TransXML: XML Schemas for Exchange of Transportation Data

DETAILS

64 pages | | PAPERBACK
ISBN 978-0-309-09872-4 | DOI 10.17226/14027

AUTHORS

Frances D Harrison; Eric Ziering; Transportation Research Board

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

NCHRP REPORT 576

TransXML: XML Schemas for Exchange of Transportation Data

Eric Ziering
Frances Harrison
CAMBRIDGE SYSTEMATICS, INC.
Cambridge, MA

AND

Paul Scarponcini
BENTLEY SYSTEMS, INC.
Exton, PA

WITH

INFO TECH, INC.,
Gainesville, FL
MICHAEL BAKER JR. INC.
Moon Township, PA
CHARLES E. CAMPBELL
Orangevale, CA

Subject Areas
Planning and Administration • Design • Highway Operations, Capacity, and Traffic Control

Research Sponsored by the American Association of State Highway and Transportation Officials
in cooperation with the Federal Highway Administration

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2007
www.TRB.org

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

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

In recognition of these needs, the highway administrators of the American Association of State Highway and Transportation Officials initiated in 1962 an objective national highway research program employing modern scientific techniques. This program is supported on a continuing basis by funds from participating member states of the Association and it receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

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

The program is developed on the basis of research needs identified by chief administrators of the highway and transportation departments and by committees of AASHTO. Each year, specific areas of research needs to be included in the program are proposed to the National Research Council and the Board by the American Association of State Highway and Transportation Officials. Research projects to fulfill these needs are defined by the Board, and qualified research agencies are selected from those that have submitted proposals. Administration and surveillance of research contracts are the responsibilities of the National Research Council and the Transportation Research Board.

The needs for highway research are many, and the National Cooperative Highway Research Program can make significant contributions to the solution of highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement rather than to substitute for or duplicate other highway research programs.

NCHRP REPORT 576

Project 20-64
ISSN 0077-5614
ISBN-13: 978-0-309-09872-4
Library of Congress Control Number 2007921915

NOTICE

The project that is the subject of this report was a part of the National Cooperative Highway Research Program conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council. Such approval reflects the Governing Board's judgment that the program concerned is of national importance and appropriate with respect to both the purposes and resources of the National Research Council.

The members of the technical committee selected to monitor this project and to review this report were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the project. The opinions and conclusions expressed or implied are those of the research agency that performed the research, and, while they have been accepted as appropriate by the technical committee, they are not necessarily those of the Transportation Research Board, the National Research Council, the American Association of State Highway and Transportation Officials, or the Federal Highway Administration, U.S. Department of Transportation.

Each report is reviewed and accepted for publication by the technical committee according to procedures established and monitored by the Transportation Research Board Executive Committee and the Governing Board of the National Research Council.

The Transportation Research Board of the National Academies, the National Research Council, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, and the individual states participating in the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

Published reports of the

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

are available from:

Transportation Research Board
Business Office
500 Fifth Street, NW
Washington, DC 20001

and can be ordered through the Internet at:

http://www.national-academies.org/trb/bookstore

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. William A. Wulf are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is a division of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The Board's mission is to promote innovation and progress in transportation through research. In an objective and interdisciplinary setting, the Board facilitates the sharing of information on transportation practice and policy by researchers and practitioners; stimulates research and offers research management services that promote technical excellence; provides expert advice on transportation policy and programs; and disseminates research results broadly and encourages their implementation. The Board's varied activities annually engage more than 5,000 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. **www.TRB.org**

www.national-academies.org

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

CRP STAFF FOR NCHRP REPORT 576

Robert J. Reilly, *Director, Cooperative Research Programs*
Crawford F. Jencks, *Manager, NCHRP*
Christopher J. Hedges, *Senior Program Officer*
Eileen P. Delaney, *Director of Publications*

NCHRP PROJECT 20-64 PANEL
Field of Special Projects

Loren Risch, *Kansas DOT* (Chair)
Steven C. Brown, *Nebraska DOR*
Roemer M. Alfelor, *FHWA*
Todd R. Bergland, *Minnesota DOT*
Terry Hale, *New York State DOT*
Maurice Harmon, *HNTB Corporation*
Liang Y. Hsia, *Florida DOT*
David C. Lively, *California DOT*
Sandra B. Salazar, *Johns Hopkins University*
Gregory T. Witt, *Kentucky Transportation Cabinet*
James Pol, *FHWA Liaison Representative*
James Ramsey, *AASHTO Liaison*
Thomas Palmerlee, *TRB Liaison Representative*
Ray Krammes, *Other Liaison*

AUTHOR ACKNOWLEDGMENTS

The authors wish to thank the following individuals who provided valuable input and perspectives on transportation data exchange needs during the course of this effort:

The NCHRP Project Manager, Christopher Hedges; the members of the NCHRP Project Panel, Jim Ramsey, Loren Risch, Steven Brown, Roemer Alfelor, Todd Bergland, Lyman (Terry) Hale, Maurice Harmon, Liang Hsia, David Lively, Thomas Rioux, Sandra Salazar, and Gregory Witt; the FHWA liaison, James Pol; the AASHTO liaison, Mary Beth Leaf; Nathan Crews of LandXML; Chuck Conley of AASHTO; Jennifer Whitted of Natek, Inc.; Carl Gonder of Alaska DOT; Michael Griffith of FHWA; Steve Jackson of Idaho DOT; and the many individuals who actively participated in the TransXML community website.

FOREWORD

By Christopher J. Hedges
Staff Officer
Transportation Research Board

This report develops a proposed common framework for exchange of transportation data in eXtensible Markup Language, known as TransXML. This framework can be used for developing, validating, disseminating, and extending current and future schemas. The research team conducted an extensive review of current XML efforts in the transportation sector, identified specific gaps and opportunities, and developed a plan to address the highest priority needs. The team developed a number of pilot schemas and applications, as well as recommendations for tools and platforms that can simplify XML coding for other transportation applications. The report summarizes the benefits that can be achieved by the adoption and expansion of TransXML, and outlines future efforts that will be needed to ensure its success.

The planning, design, construction, maintenance, and operation of transportation infrastructure all require exchanging large volumes of data. Until recently, transportation agencies have been hindered by the lack of common data formats that would facilitate the exchange of data across different platforms and applications. XML (eXtensible Markup Language) has proven to be a universal structured data-transfer methodology with great potential for the transportation sector. Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the web and elsewhere. XML data structures, known as schemas, provide a mechanism to develop and adopt common formats for data exchange.

The XML schemas that are of interest to AASHTO and its member departments typically support transportation infrastructure-related business processes in local, state, and federal departments of transportation (DOTs) as well as the work of their partners. These schemas are of equal interest to vendors who develop software products to support the transportation industry. There are currently several open consortia of private- and public-sector organizations working to create schemas to support the transportation industry, including LandXML, aecXML, ITS XML, and OpenGIS. For some transportation applications, XML schemas do not yet exist, but their availability would facilitate the exchange of transportation data broadly across multiple business areas in a format independent of the software that produced it. However, there existed no formal mechanism or framework within the transportation community to develop and maintain XML schemas, promote schema consistency and acceptance, solicit broader input, and build consensus.

There is a need for development and long-term support of XML schemas for exchange of transportation data that are widely accepted, thoroughly documented, and published on the Internet for access by any organization or individual. The long-term vision is an overall set of XML schemas for transportation applications in a framework to be called *TransXML*.

Under NCHRP Project 20-64, a research team led by Cambridge Systematics developed XML schemas in four pilot business areas in the transportation sector: roadway design, construction/materials, bridge structures, and transportation safety. Sample applications were developed for each of the schema to demonstrate their use. Data models were developed for each schema using Unified Modeling Language (UML). Geographic Markup Language (GML) was used as a consistent framework for XML coding across the four areas.

This final report is accompanied by several appendices in electronic format that can be downloaded at http://trb.org/news/blurb_detail.asp?id=7338. These include (1) a detailed review of the XML schema and relevant data standards, (2) a summary of an experiment that led to a recommendation to use a common features profile within GML to simplify the generation of XML code, (3) UML models used as the primary design tool for the TransXML schemas, and (4) the actual TransXML schemas and applications developed in this project.

This project should serve as the beginning of an effort that will become a broad umbrella for a wide variety of interrelated data exchange formats in all areas of transportation data exchange. The success of TransXML will rely on a sustained effort over time by the transportation community. Communication must continue among all of the stakeholders involved in business processes that require the exchange of data; those who will be the primary beneficiaries of common data exchange formats.

CONTENTS

SUMMARY

TransXML: XML Schemas for Exchange of Transportation Data

Project Overview

NCHRP Project 20-64 was undertaken to develop TransXML—a family of transportation data exchange formats using XML—the eXtensible Markup Language. These data exchange formats are intended to make it easier for transportation agencies to share information within and across the different stages of the transportation facility life cycle—spanning planning, design, construction, maintenance, and operations. XML is being used in a variety of other industries to enable data sharing. However, XML in and of itself is not a "silver bullet"; its value depends on getting agreement on data exchange formats from data owners, data users and software vendors. TransXML responds to the urgent need for the transportation community to work together on common data exchange formats. TransXML provides an open, vendor-neutral format for storing, exchanging, and archiving data. It therefore allows agencies to have more control over their data, and to get more value out of it over the long haul.

The justification for the TransXML project was to save agencies money. Currently, countless hours are wasted reentering or recreating information that already exists in electronic form. Data quality suffers in the process, and additional resources must be invested to ensure that data are correct with each transfer. Some transportation agencies have already recognized the value of XML, and are developing their own XML formats or "schemas" to facilitate data exchange across applications. A collective, broad-based approach to this problem reduces the need for multiple efforts of this nature. More importantly, it enables and encourages software vendors to build in standard data import and export features consistent with the agreed-upon formats. This will give agencies more choices, and will reduce the costs of configuring and implementing new applications. TransXML also provides a way for agencies to archive their valuable data in a text format that is human-readable and independent from the software used to create and store it.

Scope and Products

NCHRP Project 20-64 was designed to start development of TransXML by focusing on four transportation business areas: (1) Survey/Roadway Design, (2) Transportation Construction/Materials, (3) Highway Bridge Structures, and (4) Transportation Safety. XML Schemas were developed within each of these areas, along with sample applications that demonstrate how the schemas can be used. The initial set of TransXML schemas are described in Table ES-1.

For each business area, key stakeholders were identified and input was solicited via e-mail and via a collaborative website established for the project. Key milestones in the project were as follows:

- Definition of the scope for schemas based on an analysis of data exchange needs and gaps,
- Development of data models for each schema using Unified Modeling Language (UML) class diagrams,

Table ES-1. TransXML schemas.

Business Area	Schemas
Survey/Roadway Design	• **Area Features (AF) Schema –** Allows data from GIS to be overlaid on design drawings in CAD systems. • **Geometric Roadway Design (GRD) –** Subset of LandXML adopted into TransXML – allows for sharing of roadway alignment, cross sections, geometry across members of a design team, between designer and surveyor, and from design into machine controlled excavation equipment. • **Design Project (DP) –** Allows design project pay item data to be exchanged across design, cost estimation and bid preparation systems.
Transportation Construction/ Materials	• **Bid Package (BP) –** Supports exchange of construction bid package data between agency systems and contractor bid preparation software. • **Construction Progress (CP) –** Supports exchange of information about partial pay item quantities placed from field data collection systems to construction management systems. • **Materials Sampling and Testing (MST) –** Allows exchange of construction site installed quantities and materials used and tested information from field data collection systems to laboratory systems, central construction progress tracking and contractor payment systems. • **Project Construction Status (PCS) –** Allows exchange of construction project status information from construction management systems to stakeholder information systems (e.g., project web sites).
Highway Bridge Structures	• **Bridge Design and Analysis (BDA) –** Allows for analysis of the same structure in multiple structural analysis software packages.
Transportation Safety	• **Crash Report (CR) –** Allows exchange and sharing of crash records data. TransXML adopted the NHTSA/JusticeXML crash records XML Schema that is based on the Model Minimum Uniform Crash Criteria (MMUCC). • **Highway Information Safety Analysis (HISA) –** Allows for exchange of highway information between inventory systems and safety analysis software.
All	• **Linear Referencing (LR) –** An XML schema for linear referencing information consistent with ISO 19133 – used by the other TransXML schemas.

- Selection of the Geography Markup Language (GML) to provide a consistent framework for XML encoding across the new XML schemas to be developed, and
- Development of the XML schemas and sample applications.

Future of TransXML

NCHRP Project 20-64 has provided an initial set of TransXML schemas and has established an umbrella framework for development of additional schemas. In order to reap the full benefits of this work, there is a need to educate potential users about how to use the schemas, and to provide technical assistance for agencies and vendors choosing to modify their systems to read and write data in TransXML format. There is also a need to extend the current set of schemas to address other data exchange needs both within and outside of the four initial TransXML business areas.

Future efforts should recognize the key lessons learned during the course of this project:

- The TransXML concept of a broad umbrella for a set of interrelated transportation data exchange formats has the advantage of supporting interoperability of data across different business processes and life-cycle phases of transportation facilities. However, this umbrella concept creates challenges in that there are several distinct stakeholder communities within transportation. This means that extra effort is needed to (a) provide sufficient depth of communication and technical understanding for multiple different stakeholder groups, (b) ensure technical consistency and coordination across individual schema, and (c) maintain broad interest and support for the project.
- For TransXML to succeed, there needs to be a substantial and sustained effort to raise awareness of the project. Significant resources must be allocated for communication with stakeholders—through multiple channels—in order to ensure that the XML schemas are widely understood and perceived to have clear value.
- Resources must be devoted to coordinate TransXML efforts with related standards efforts since TransXML touches many areas where schema and standards already exist or are under development. This will maximize the potential for synergies and harmonization across efforts.

Based on these lessons learned, a model for future TransXML stewardship was developed, including a mission statement and a set of recommended functions and roles. Several different stewardship models were evaluated based on an examination of other XML efforts and consideration of key criteria—including neutrality, stability, agility, technical expertise, marketing capability, and administrative infrastructure.

Because it is critical to maintain momentum for the TransXML project while a more permanent stewardship arrangement is being established, the project team recommends that AASHTO take on temporary stewardship for the project for a period of up to 12 months. During this transition period, AASHTO or a designated contractor would maintain the TransXML website, continue with a skeletal level of outreach and liaison activities, and initiate one or more pooled fund projects to demonstrate implementation of selected TransXML schema. During this time, discussions would be pursued to explore—and ideally obtain commitment for—ongoing sponsorship for the project.

SECTION 1

Introduction

1.1 Research Context

Need for Transportation Data Exchange Standards

The need for development of common data standards and simple data exchange mechanisms within the transportation domain has long been recognized. Transportation agencies make use of a variety of data sets and software tools to support planning, design, construction, maintenance, and operations activities. Tools within each area are specialized, with data requirements tailored to the specific function being performed. However, there are natural connections across the life cycle of activities that create the need for information flows across individual applications. From the perspective of an individual project, information developed in one phase becomes a starting point for the next. Designers build on the efforts of planners, construction managers begin with the work of the designers, and then maintenance and operations managers begin with as-builts and other information produced in the construction phase. Information also needs to flow in the reverse direction—for example, designers should ideally learn about the constructability, maintenance requirements, and safety (via crash records) of a particular design approach.

There also are important horizontal flows of information within particular phases of the transportation facility life cycle. For example, design of a roadway alignment involves consideration of bridge geometry, retaining wall design, drainage structures, traffic volumes, and right-of-way acquisition impacts. The efficient integration of all these facets of roadway design can provide early feedback to designers, which in turn can have profound effects on the cost and timing of the project. New approaches like design-build can fast-track the process, putting even greater demands on the efficient flow of information. Tighter budgets necessitate better design efficiency to minimize construction costs. Increased reliance on software in many areas begs for more automated approaches for moving data from one application to the next.

Ideally, there would be seamless connections across systems to facilitate data exchange, but this is the exception rather than the rule. For example, planners assemble a vast amount of information about a project as it moves from the Statewide Transportation Improvement Program (STIP) to preliminary design. Instead of building upon this information, the designer begins with a clean sheet of paper. One reason for this is the incompatibility of the planner's Geographic Information System (GIS) and the designer's Computer-Aided Design (CAD) system. As the facility moves to construction, information is manually extracted from the design plans into the construction management system, wasting time and inevitably introducing errors.

Most agencies rely on commercial software tools for a wide spectrum of functions, including highway and bridge design, construction management, maintenance management, field inspection data collection, and traffic forecasting. When a mix of tools from different vendors is used, when an agency chooses to switch to a new tool in the same functional area, or when an agency's partners (e.g., contractors, other jurisdictions) attempt to share data created with disparate tools, they face a host of data compatibility and consistency issues.

Lack of interoperability across systems means that agencies must expend considerable resources to build custom interfaces, or live with duplicative data creation processes—with associated inefficiencies and loss of data consistency and quality. In many instances, the result is that information that could be valuable for decision-making is simply not available. Commonly agreed-upon data standards would yield substantial benefits, including improved efficiency, better information quality, and increased flexibility to make use of emerging software that best addresses particular business requirements.

XML: An Enabling Technology for Data Interoperability

In recent years, XML (an acronym for eXtensible Markup Language) has become a near-universally accepted and supported mechanism for data exchange across platforms and applications. XML data structures, known as *schemas*, provide a mechanism to develop and adopt common formats for universal data exchange, thereby, allowing separate information systems to communicate. XML provides a formal, self-documenting structure to share data, independent of the software that produced it. It also provides a mechanism for long-term archiving of data that might otherwise be difficult to access when host applications are retired. XML schemas have been developed in many industries, including publishing, insurance, education, and electronics.

XML is not a silver bullet to the problem of standards development, but it is a maturing technology that promises to improve the flow of information. With XML-based interfaces between software applications, and between people and those applications, the flow of information can become more efficient and effective. This can lead to improvements in the facility life-cycle processes, and, ultimately, to improvements in the facilities themselves.

There currently are several open consortia of private and public sector organizations working to create transportation XML schemas. However, until this project there has been no formal framework within the transportation community to sponsor, develop, and maintain XML schemas; promote schema consistency and acceptance; solicit broader input; and build consensus. For many transportation applications, XML schemas do not yet exist.

1.2 Research Objectives

The objectives of NCHRP Project 20-64 were to develop broadly accepted public domain XML schemas for exchange of transportation data and to develop a framework for development, validation, dissemination, and extension of current and future schemas. The ultimate goal of this effort is *TransXML*—a family of XML schemas for transportation applications, and a recommended institutional structure and process for implementing and sustaining its use. This project focused on four business areas within transportation (1) Roadway Survey/Design, (2) Transportation Construction/Materials, (3) Highway Bridge Structures, and (4) Transportation Safety. However, it is envisioned that TransXML will ultimately encompass a broader set of schemas for other transportation business areas.

The research objectives recognized the importance of establishing an umbrella framework for future transportation XML schema development, in order to avoid overlapping and inconsistent efforts. They also recognized that transportation data standards are only effective if they are widely adopted. The technical development of TransXML is not sufficient to achieve the desired result of better data exchange within and across transportation business areas. An institutional framework or "roadmap" is also necessary to encourage the continued development and adoption of TransXML. The institutional structure and processes for TransXML are essential for achieving TransXML's long-term goal of easy transportation data exchange.

1.3 Research Approach

This research was undertaken in two major phases. In Phase I, the project team documented relevant existing XML efforts; identified opportunities in surface transportation where use of XML could yield significant benefits; identified specific gaps and opportunities in the four identified key business areas; and mapped out a plan for development of schemas to fill the highest priority gaps. Phase I also established a website for information dissemination and for collaborative schema development among TransXML stakeholders, and identified potential lead organizations for stewardship of TransXML.

In Phase II, the project team designed and developed a set of XML schemas and associated sample applications. Phase II also developed recommendations for institutional structures, processes, and funding mechanisms for continued schema development and for sustaining the TransXML effort.

The research work program had the following tasks:

- **Task 1. Develop a Project Website**—Establish initial website for the TransXML project to post goals, plans, status, work in progress, final deliverables, names of participants and partners, outcomes, and other pertinent information.
- **Task 2. Identify Surface Transportation Business Areas for TransXML**—Provide a high-level view of surface transportation data exchange requirements by identifying which are presently being addressed using XML schema, and which would benefit from development of new XML schemas.
- **Task 3. Review Existing XML Schemas**—Document the status of existing XML efforts in transportation, and identify specific public domain schemas that can be incorporated into TransXML.
- **Task 4. Identify Potential Organizations for TransXML Stewardship**—Identify candidate organizations and agencies that could take on long-term responsibility for TransXML, and document the relative advantages and disadvantages of each.
- **Task 5. Business Area Gap Assessment**—Assess the four business areas identified for initial investigation (see Table 1)

Table 1. Business areas for TransXML.

Highway Bridge Structures	Transportation Safety	Roadway Survey/Design	Transportation Construction/ Materials
• Bridge Analysis and Design • Bridge Load Rating • Bridge Construction • Bridge Inspection • Bridge Management • Bridge Operations • Bridge Maintenance	• Crash Reports • Crash Location • Roadway Inventory • Citations • Driver Information • Vehicle Information • Emergency Medical Info • Federal Motor Carriers Info • Crash Analysis • Work Zone Safety	• Design Surveys • Parcel/Boundary Surveys • Geometric Design • Pavement Design • Right-of-Way • Construction/ Stake-out • Survey Feature Codes • Cross Section • Pay Item Quantities	• Estimates • Proposals • Letting and Award • Construction Management • Materials

to identify and prioritize where XML schemas would be beneficial, and where activities are already supported by existing XML schemas.

- **Task 6. Schema Development Plan**—Based on the priorities established in Task 5, develop a plan for full development of the initial set of TransXML schemas that can be successfully accomplished within the confines of the budget and schedule.
- **Task 7. Develop Moderated Web-Based Collaboration Tool**—Develop a website which will serve as a forum for collaboration across a wide spectrum of stakeholders on the development of TransXML schemas.
- **Task 8. Phase 1 Interim Report**—Prepare a report documenting work completed in Tasks 1 through 7.
- **Task 9. Develop Business Area XML Schemas**—Design and develop XML Schemas using an open process that encourages wide participation by stakeholders.
- **Task 10. Develop Web-Based Validation Software**—Provide web-based software to check submitted TransXML schemas for compliance with TransXML schema requirements (and W3C specifications).
- **Task 11. Software Documentation, Source Code, and Sample Data**—Develop sample software and datasets demonstrating application of the new TransXML schemas.
- **Task 12. Recommend Implementation Framework**—Recommend a framework for supporting continued schema development and improvement.
- **Task 13. Recommend TransXML Stewardship Model**—Develop recommendations and an implementation plan for future funding, management, and maintenance of TransXML to encourage its sustained development and use.
- **Task 14. Final Report**—Prepare final report documenting the results of Tasks 1 through 13, and a companion CD-ROM including the text of this report, the TransXML schemas and documentation, the Tasks 10 and 11 software programs and data, source materials for the Tasks 1 and 7 websites, and an archive of all postings made to the collaboration tool.
- **Task 15. AASHTO IS Presentation**—Develop and deliver a summary presentation for the TransXML project at a meeting of the AASHTO Administrative Subcommittee on Information Systems.

1.4 Contents of Report

This report summarizes the results of Tasks 1 through 13. It is organized in the following sections:

- **Section 1** list sources cited in this report.
- **Section 2** provides a high-level view of surface transportation data exchange requirements, providing a context for formulation of long-term goals for TransXML.
- **Section 3** reviews prior XML schema and data standards relevant to the four focus areas of roadway survey/design, transportation construction/materials, highway bridge structures, and transportation safety. It identifies prior schema which could be incorporated into TransXML.
- **Section 4** identifies and prioritizes gaps between key business areas where XML schema would be beneficial, and business areas already supported by prior XML schemas.
- **Section 5** describes the TransXML schema development process and work products.

- **Section 6** discusses future stewardship of TransXML. It examines options and presents an implementation plan.
- **Section 7** lists sources cited in this report.
- **Appendix A** includes detailed information on existing XML standards that were evaluated for relevance to TransXML. These reviews were current as of October 2004.
- **Appendix B** is the summary report documenting the investigation of the use of the Geographic Markup Language (GML) as the basis for TransXML.
- **Appendices C and D** present the UML models developed in this project, which represent the results of the design phase of XML schema development.
- **Appendix E** documents the XML schema and applications developed for this project that were delivered to NCHRP in electronic form.
- **Appendix F** contains applications and source code.
- **Appendix G** provides feedback on schemas and UML models.

SECTION 2

Surface Transportation Data Exchange Needs

2.1 Overview

The research statement for Project 20-64 identified four business areas within surface transportation as the initial focus for TransXML: roadway survey/design, construction/materials, bridge structures, and transportation safety. The different elements of these business areas are shown in Table 1 (which appears in the preceding section). While these selections cover only a subset of the surface transportation domain, they provided a representative cross section of transportation facility life-cycle phases as well as transportation facility components.

This section introduces the four business areas from the perspective of potential data exchange requirements. Then, it provides a high-level view of other areas of surface transportation that should be considered as part of future work on TransXML.

2.2 Roadway Survey/Design

Roadway design information includes horizontal and vertical centerline alignments, other aspects of the roadway geometry such as cross section, subsurface, and superelevation; and information about ancillary components such as pavement, shoulders, curbs, sidewalk, drainage pipes, and structures.

There are three important sets of data exchange requirements for roadway survey/design: information sharing among different members of a design team; information sharing between the road surveyor and designer; and finally, information transfer from design into the construction phase.

Information about the roadway alignment from roadway designers needs to be shared with numerous other disciplines, such as bridge engineers, right-of-way specialists, drainage experts, soils engineers, traffic operations, and traffic safety analysts. This communication begins during preliminary design as the alignment first begins to take shape and progresses as the detailed design evolves until a final design is approved. Thus, design data needs to be exchanged across multiple software tools that support individual specialties on a repetitive basis. These needs include exchange of information between geographic information systems and CAD tools.

A two-way data exchange between the surveyor and designer is required. The surveyor provides the designer with information about the existing characteristics of the construction site. Once the design is complete, the designer provides the surveyor with information required for construction stake out.

On completion of the design, there is a need to communicate information about the design to the prospective construction team. The project can be viewed as a set of components or construction items, each with a requisite amount of relevant information to be tracked and communicated. The component items determined for bidding during the design phase transition into the components required for tracking construction progress and payment during the construction phase. Quantity takeoff, estimating, and specifications represent data exchange opportunities in preparation for bidding and contracting. In addition, geometric descriptions used in engineering calculations during design provide the basis for stake out calculations during construction.

2.3 Transportation Construction/Materials

Data in the transportation construction/materials area includes both physical and business representations of a construction project. The physical view incorporates plans and specifications created in the design process and adds more specific information on materials placement and testing. The business view overlays information needed to bid, schedule, monitor, inspect, and manage the work.

As noted above, there are significant opportunities for efficiency improvements by automating transfer of information from design into the construction phase. Use of surface models produced in design for grading equipment machine control is an example of this—where considerable cost savings are already being achieved.

Other data exchange needs during the construction phase include support for electronic bidding processes; transfer of bidding information into contractor payment systems; transfer of work tracking and payment information between central and field offices; transfer of materials testing information among the construction site, laboratories, and central offices; and information sharing about construction project status.

2.4 Highway Bridge Structures

This business area spans bridge planning, design and analysis, construction, inspection, management, operations, and maintenance. Thus, bridge data includes information required for initial design of the structure, information about the physical design of the structure as a whole and its individual elements, information about the loads carried by the structure, condition and load rating information required for maintenance and permitting, and information about the function of the structure in the road network.

Key data exchange processes in the bridge structures area include transfer of highway geometry parameters from highway design to bridge design; transfer of bridge design information to highway engineers; provision of coordinates and station information to surveyors; loading of bridge design inputs (geometry, materials, loading information) to multiple design software packages; transfer of design information (geometry, quantities, digital terrain model information) to estimating systems; transfer of as-built information to inspection, rating bridge management and maintenance management systems; and transfer of bridge design and inspection/condition information to permitting and routing systems. Federal reporting of bridge inventory and inspection information represents another data exchange need.

2.5 Transportation Safety

Transportation safety data includes information about crashes that occur (where, when, why, who, how), information about their consequences (emergency medical information, insurance information) and information about the characteristics of facilities, vehicles and drivers that pose safety risks. This latter set of data includes highway design and operational characteristics, vehicle registration information, vehicle inspection information, motor carrier inspection and driving records, and citations. Raw crash report data is linked to roadway, vehicle, motor carrier, and driver information to yield additional information—for example, highway crash rates used for analysis of the need for safety projects are derived from crash reports and traffic data linked by location.

Crash data need to be transferred from the collection point (e.g., a police report) to processing point(s) (for validation and linkage with other data) and then on to archive point(s). This process commonly involves exchange of the data across multiple agencies. Once processed, crash reports and linked crash information are shared across a wide variety of audiences (public agencies at state and local levels, interest groups, insurance companies, etc.). Given consistent key information required for linking across safety data sets, different types of safety data can be gathered from their respective sources and joined together for a variety of reporting and analysis purposes.

Linkage of transportation safety-related information depends to a large extent on location identification. Unfortunately, different data sources typically use different location referencing methods and different roadway system models. Emerging software standards in these areas may provide a common understanding and representation scheme for XML encoding.

Transportation safety has a significant real-time component as well, including activities such as roadside inspection of commercial vehicles, emergency response, incident management, driver information, and work zone management to ensure safety of road users. These activities rely on exchange of real-time data on traffic, incidents, response status, road conditions, and weather. These types of data exchange requirements are encompassed within the ITS area (see below).

2.6 Broader Framework for TransXML

The four business areas selected for initial focus of TransXML cover an important, but relatively small portion of the universe of surface transportation data exchange needs. A broader framework for TransXML is illustrated in Figure 1. The first row in this diagram shows the initial four business areas; the second row shows potential future business areas. The final row of the figure shows areas where there are already active XML and standards development efforts that TransXML should link to or coordinate with.

The broader framework for TransXML includes the following distinct but overlapping viewpoints:

Geospatial View—Transportation data is fundamentally geographic in nature, and therefore there has been considerable work to date on development of data standards and models focused on the geospatial representation of transportation information. These allow for sharing, linking, and integrating a variety of geographic data sets (including those from nontransportation domains). For example, the

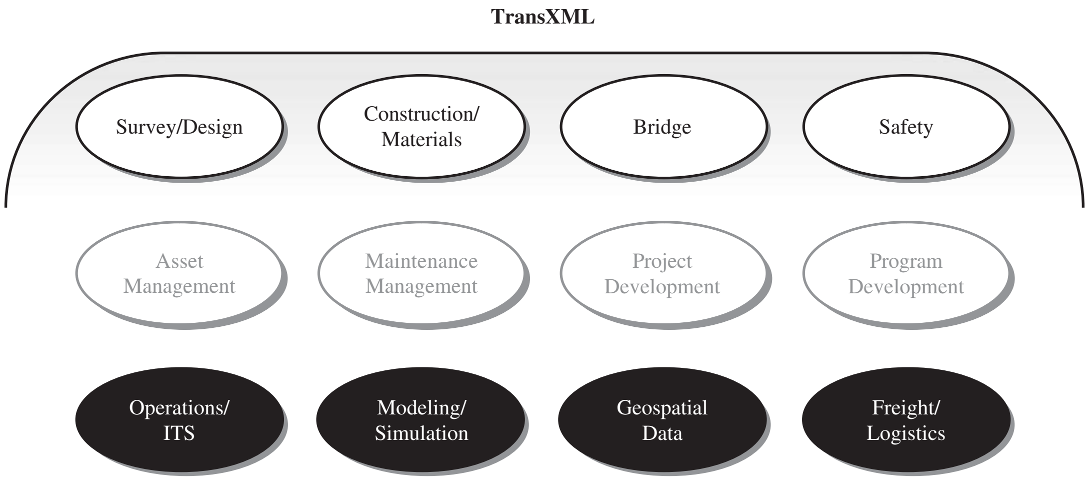

Figure 1. TransXML current and potential future scope.

Geospatial One-Stop (GOS) data content specification defines a standard representation of data related to road, rail, and transit modes. Feature models have also been developed by GIS vendors that can be used to develop geodatabases providing enterprise maintenance of and access to geographic data in support of multiple applications. The Open Geospatial consortium (OGC) is supporting continued development of Geographic Markup Language (GML), which provides an abstract model for representing geographic information.

Infrastructure View—This view is concerned with data exchange within and across different phases of the life cycle of planning, designing, constructing and maintaining transportation infrastructure. It incorporates planning, engineering/design, and business perspectives and is the primary focus of the initial TransXML effort. However, this view goes beyond the four selected business areas, extending to activities including planning and project development (environmental assessments, permits, right-of-way, etc.), program development/budgeting, asset management (inventory, inspection, performance monitoring, life-cycle cost modeling, selection and prioritization of treatments), and maintenance management (identifying, scheduling and managing maintenance activities). A recent TRB paper (*1*) proposed a semantic architecture to represent the highway construction domain using an integrated supply chain model, encompassing planning, design, field construction, and maintenance. Elements of ebXML (*2*) are used for electronic business transactions related to highway construction. This paper can be revisited at a later date as TransXML's scope is broadened.

Safety View—This view (to be addressed in the initial TransXML effort) is concerned with collection, reporting, analysis and use of information needed to reduce transportation crashes and fatalities.

ITS/Operations View—The National ITS Architecture (*3*) represents an important reference model for surface transportation processes and data flows. This Architecture defines functions and data flows for real-time operation and management of the surface transportation system. It encompasses eight different "bundles" of user services: travel and traffic management, public transportation management, electronic payment, commercial vehicle operations, emergency management, advanced vehicle safety systems, information management (including archiving real time data for use in other applications), and maintenance and construction management. As data interoperability is a key objective of the ITS architecture effort, there are a number of active ongoing ITS data standards efforts, including NTCIP (National Transportation Communications for ITS Protocol), CVISN (Commercial Vehicle Information Systems and Networks), Archived Data, Incident Management, Traffic Management, ATIS (Advanced Traveler Information Systems), and TCIP (Transit Communications Interface Profile). XML data formats have or are being developed for several ITS-related areas, including motor carrier safety data, transportation management center to center communications and traveler information exchange.

Travel and Traffic Modeling and Simulation View—This view is concerned with exchange of data inputs and outputs of transportation modeling and simulation tools—including socioeconomic data, traveler characteristics, network representations, and network characteristics (travel times and costs). The Traffic Software Data Dictionary (TSDD) provides an example of data exchange needs from the traffic modeling point of view. A recent paper (*4*) reported development of TrafficXML to support exchange of data between simulation programs of different vendors. FHWA's Next Generation Simulation Models (NGSIMS) effort is currently developing a data dictionary to support new microsimulation algorithms being developed under that project.

Freight Logistics View—This view is concerned with the intermodal freight supply chain—sharing of information on shipment and equipment status and location. The TransXML developed by Transentric covers this area; the standard is now being extended by the Open Applications Group.

SECTION 3

Current Practice Review

This section reviews prior XML schema and data standards in surface transportation in order to

- Identify gaps to be addressed in TransXML,
- Identify XML schema that could be immediately incorporated into TransXML,
- Document XML schema and standards that could be incorporated into future schema development efforts under the TransXML umbrella, and
- Identify independent standards and XML schema development efforts that TransXML development should be aware of and/or coordinate with.

3.1 Overview of Existing Schema and Standards

There have been a number of efforts, on the part of individual agencies and vendors and involving consortia of public and private organizations, to develop and use transportation-related XML schemas. While significant progress has been made, particularly in the area of geographic data, the transportation arena is still at a relatively early stage of XML schema development and adoption.

Transportation data standards, which may evolve or be incorporated into XML schema, also are of relevance to this research effort. Transportation data standards are already being used throughout the industry. Often these standards are in the form of data dictionaries that contain standard codes for specific events. These standards are useful for aggregating data at the state and national levels in order to make sense of data coming from disparate sources. They are also essential in organizing and analyzing large volumes of data.

This section provides an overview of standards and XML schema efforts in the four TransXML business areas. In addition, the related areas of geographic data and ITS are included here, since these areas may provide important building blocks for TransXML. The schema and standards efforts reviewed are listed in Table 2 and summarized below.

Roadway Survey/Design

LandXML provides at least partial support for the Design Surveys, Parcel/Boundary, Geometric Design, Construction/Stakeout, Survey Feature Codes, and Cross Section topics within this business area. LandXML includes specifications for raw and reduced surveying data, surface data, parcel data, and a 3D road model. LandXML also includes a standard format for official electronic design submission. aecXML supports the Pay Item Quantities topic. Two XML initiatives in the Geotechnical area are relevant to the pavement design area (though by no means provide complete coverage for that area). The Traffic and ITS design topic is partially addressed by the Traffic Software Data Dictionary.

Construction/Materials

The aecXML schema and Trns•port data model address each of the topic areas within the Construction/Materials business area. In addition, XML initiatives sponsored by the American Institute of Steel Construction and the American Iron and Steel Institute address specific materials. The aecXML Infrastructure effort produced a draft schema that addresses communication of pay items from design to estimation systems. This was based on the aecXML common object schema, which provided a broad range of contract, project, and organization elements relevant to construction. The AASHTOWare Trns•port suite uses XML schema for exchange of information across its modules related to construction contracts, project cost estimates and daily work reports. In addition, the Trns•port data model includes additional content that could serve as the basis for XML schema in the construction business area.

Table 2. XML schema and data standards efforts.

Schema/Standard (Links)	Content	Developer/ Participants	Screening Results
LandXML www.LandXML.org	Focus is the exchange of civil design information, including raw and reduced surveying data, surface data, parcel data, and 3D road model. Includes standard format for official electronic design submission. Most recent version incorporates mechanisms for interoperability with the Federal Highway Administration (FHWA) Interactive Highway Safety Design Model (IHSDM). Includes content of AASHTOWare SDMS.	Derived from earlier ASCII-based Engineering and Surveying-Exchange (EAS-E) initiative. Schema is supported by over a dozen commercial applications	Detailed evaluation
Geographic Markup Language – GML http://www.opengeospatial.org/specs/	A comprehensive XML schema for encoding both spatial and nonspatial geographic information. Feature-centric model, defining abstractions of real-world phenomenon (e.g., roads) with properties having names, values, and types. Includes Rules for Application Schemas.	Open Geospatial Consortium (OGC)	Detailed evaluation
XGDF XML http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=30763	The ISO TC204 (ITS) GDF standard supports location-based services, with a current focus on car navigation systems. Data model includes features (e.g., roadway, structures, and railways), relationships between features, and attributes of features or relationships. Includes roadway features, other transport modes, area features. Focus of attributes is on navigation needs. The next version of GDF is being called XGDF. Work has begun on adding both an SQL and XML encoding.	ISO TC204 (Intelligent Transportation Systems)	Detailed evaluation
Geospatial One-Stop http://www.geo-one-stop.gov/Standards/index.html	As part of the National Spatial Data Infrastructure, the Geospatial One-Stop (GOS) data content standard formalizes how geographic information in any of seven themes can be represented for transfer between government agencies. The transportation theme includes modes of road, rail, transit, air, and navigable water. It is their intent to submit this specification to ANSI for standardization. The road mode of the transportation theme is consistent with constructs found in GDF. It includes the ISO 19133 linear referencing clause. An implementation specification based on OGC GML is anticipated.	Federal Geographic Data Committee. Modeling teams had representatives from government, industry, and academia. The U.S. Department of Transportation Bureau of Transportation Statistics led the transportation theme model development.	Detailed evaluation
ISO 19133 http://www.iso.ch/iso/en/stdsdevelopment/techprog/workprog/TechnicalProgrammeProjectDetailPage.TechnicalProgrammeProjectDetail?csnumber=32551	Linear Referencing data standard. Provides a standard, generalized content format for specifying a location, applicable to most any linear referencing method.	International Organization for Standardization (ISO) Technical Committee TC211 (Geographic information/Geomatics)	Detailed evaluation

Table 2. (Continued).

Schema/Standard (Links)	Content	Developer/ Participants	Screening Results
aecXML – common object schema http://www.iai-na.org	Architecture, engineering, and construction industry schema. Provides a content format for specifying building, plant, infrastructure, and facility information.	aecXML	Detailed evaluation
aecXML Infrastructure Project (unpublished)	Infrastructure project schema. Provides a content format for specifying infrastructure projects.	aecXML	Detailed evaluation
AASHTOWare Trns•port	XML Schema exists for construction contracts, project cost estimates, and daily work reports. Trns•port includes a broader construction management data model from which additional schema could be generated. Specific evaluations were completed for: • Bid subcontract commitment; • Construction project and contract information; • Construction contractor employee payroll submission format; • Construction project payment item cost estimates; • Construction project daily work report; • Construction project payment financial format; • Construction project proposal elements (including payment items); • Construction project proposal bid format; and • Construction project subcontract format.	American Association of State Highway Transportation Officials (AASHTO)	Detailed evaluation
AASHTO BRIDGEWare	No XML Schema, but standard data structures have been established that could provide a basis for a schema. Data domain for bridge design/rating (Virtis/Opis) covers superstructure description for common types of girder bridges. Work on substructure domain is in progress. Data model for bridge inspection, needs analysis and project planning is part of the Pontis portion of BRIDGEWare.	American Association of State Highway and Transportation officials (AASHTO)	Detailed evaluation
NCHRP 12-50 XML standard http://www4.trb.org/trb/crp.nsf/All+Projects/NCHRP+12-50	Data structure (including an XML representation) for bridge analysis and rating. Includes specification information for steel girder and pretensioned prestressed concrete girder type bridges.	National Cooperative Highway Research Program (NCHRP)	Detailed evaluation

(continued on next page)

Table 2. (Continued).

Schema/Standard (Links)	Content	Developer/ Participants	Screening Results
NCHRP 20-7, Task 149 SteelBridge XML (proposed) http://www4.nas.edu/trb/crp.nsf/All+Projects/NCHRP+20-07#149	The data model provided in the appendix of this report starts with the AASHTO Virtis/Opis data model and adds on specific classes related to fabrication and construction of steel bridges.	National Cooperative Highway Research Program (NCHRP)	Detailed evaluation
American Institute of Steel Construction (AISC) CIS/2 http://www.aisc.org/	Standard for steel construction data exchange across CAD and analysis programs. This standard was referenced in the draft final report for NCHRP 20-7, Task 149. Since this standard is related more to buildings than transportation, a detailed review was not conducted.	American Institute of Steel Construction (AISC)	Reference
American Iron and Steel Institute XML standard http://xml.coverpages.org/aisi.html	Effort to develop standardized XML terminology to be used throughout steel-related transactions documents. Because there appears to be limited recent development of the guidelines for this workgroup and it does not appear to be implemented a full evaluation was not conducted.	Workgroups at the AISI web site	Reference
Global JusticeXML http://it.ojp.gov/jxdm/3.0/index.html	Comprehensive schema and data dictionary for data exchange among justice and public safety communities. Includes incident reports, driver histories, arrest warrants.	U.S. Department of Justice Office of Justice Programs (OJP), Global Justice Information Sharing Initiative (GLOBAL), Georgia Tech Research Institute (GTRI)	Detailed evaluation
AASHTOWare – TSIMS http://tsims.aashtoware.org/ContentManagement/PageBody.asp?PAGE_ID=3&TAB_ID=4 http://tsims.aashtoware.org/ContentManagement/PageBody.asp?PAGE_ID=5&TAB_ID=8	The Transportation Safety Information Management System (TSIMS) is a proposed AASHTO project, and AASHTO will be proposing the use of XML in the upcoming solicitation release as the primary data exchange interface between all subsystems. The goal of TSIMS is to develop a uniform approach to management of traffic safety information. Guidelines are being developed, so that any vendor or agency can interface existing systems with it. The Object Broker at the core of the TSIMS will use XML for data interchange with outside systems. The entire TSIMS data dictionary will also be XML-compliant. (Please note: As of the time NCHRP 20-64 was completed, the TSIMS project had been discontinued. It is expected that this project will be replaced by a new, reduced scope project entitled “Safety Management System.”)	AASHTO/FHWA	Reference

Table 2. (Continued).

Schema/Standard (Links)	Content	Developer/ Participants	Screening Results
FMCSA http://cvisn.fmcsa.dot.gov/ Documents/ Document_Nav_Frame_Page_ documents.shtml http://infosys.fmcsa.dot.gov/ provu.asp	In the area of transportation safety, the Federal Motor Carrier Safety Administration (FMCSA) has adopted XML as the standard data format for moving data between applications. Specific XML applications include Query Central, the Motor Carrier Profile report, the Inspection Selection System (ISS) data refresh, the Safety Audit upload process, and an interagency data exchange to automate driver/vehicle/carrier clearance along the border.	FMCSA	Detailed evaluation
Crash Records Markup Language (CRML)	XML tags for crash records information.	University of Florida Transportation Research Center	Detailed evaluation
FARS ftp://ftp.nhtsa.dot.gov/ FARS/FARS-DOC	National fatal accident reporting system – coding manual published with all data elements.	NHTSA	Reference
FRA http://safetydata.fra.dot.gov/ officeofsafety/Downloads/ Default.asp	Standard reporting data formats for railroad accidents/incidents.	FRA	Reference
Traffic Model Markup Language – TMML http://www.ce.ufl.edu/trc/ research/tmml.htm	TMML is a mechanism to share data among traffic modeling software products. Envisioned as an XML-compatible language prescribing class structure and data element tag names required to represent traffic model data, and create output data in format easily rendered by office productivity software. A limitation of TMML is that data must be specified in a separate Traffic Software Data Dictionary (TSDD), which serves as the source for the vocabulary and tags identifying classes and attributes. McTrans has adopted TMML for all of the HCS software modules.	University of Florida Transportation Research Center	Reference
SAE ATIS http://www.sae.org/	The SAE Advanced Traveler Information Systems (ATIS) standards committee is developing an eXtensible Markup Language (XML) vocabulary for traveler information exchange. This work builds on the existing data dictionary (SAE J2353), message sets (SAE J2354 and J2369) and other related standards work.	SAE	Reference
Traffic Management Data Dictionary (TMDD) http://www.ite.org/tmdd/ index.asp	Traffic Management Center-to-Center Communications data dictionary and messages – includes XML schema representation.	AASHTO/ITE	Reference

(continued on next page)

Table 2. (Continued).

Schema/Standard (Links)	Content	Developer/ Participants	Screening Results
Traffic Software Data Dictionary (TSDD) http://www.tfhrc.gov/its/its3.htm#traffic http://www.signalsystems.org.vt.edu/documents/Attach/Leonard_tsddtsom.pdf	Traffic engineering vocabulary, with many terms and definitions from the Highway Capacity Manual, FHWA documents, and CORSIM manuals. The TSDD is documented in P1489 format and was developed in parallel with the TMDD. An associated effort, the Traffic Software Object Model (TSOM), contains Unified Modeling Language (UML) object model diagrams that describe object classes, their attributes, and their relationships in the traffic simulation domain.	FHWA-funded effort	Reference
IEEE P1512 http://www.standards.its.dot.gov/fact_sheetp.asp?f=12	Common Incident Management Message Sets for Use by Emergency Management Centers.	IEEE SCC 32 ITS Incident Management Working Group	Reference
UTDF – Universal Traffic Data Format http://www.trafficware.com	Data exchange between signal controller systems and other software.	TrafficWare	Reference
National Transportation Communications for ITS Protocol (NTCIP) Standards	Family of communications standards for message and data transfer between ITS control devices. Includes standards for “Center to Center” communications (e.g., between weather monitoring systems and freeway management systems), and “Center to Field” communications (e.g., between a traffic management center and individual signal controllers).	AASHTO/ITE/NEMA	Reference
COSMOS/PEER-LL Geotechnical Data XML Schema http://geoinfo.usc.edu/gvdc http://dmrl.usc.edu/pubs/sci2003-web.pdf	XML schema for geotechnical information.	Consortium of Organizations for Strong Motion Observation Systems (COSMOS)	Detailed evaluation
Highway Performance Monitoring System (HPMS) http://www.fhwa.dot.gov/ohim/hpmsmanl/hpms.htm	Standard reporting format for highway performance information – reported annually by all state DOTs to FHWA.	FHWA	Reference
NCHRP 20-57 http://www4.nationalacademies.org/trb/crp.nsf/All+Projects/NCHRP+20-57	This project developed generic tools to support asset management, along with a draft XML schema for the transportation asset performance tradeoff domain.	NCHRP	Reference

Table 2. (Continued).

Schema/Standard (Links)	Content	Developer/ Participants	Screening Results
Logistics XML/TransXML http://www.openapplications.org/wg/LogisticsXML.htm http://xml.coverpages.org/tranXML.html	XML standard for e-commerce activities of shippers and carriers.	Open Applications Group (OAGi)	Reference
RecML – Recreation One-Stop XML http://www.xml.gov/presentations/itpioneers/RecML_files/frame.htm	XML specification that defines terms for recreation areas (parks), facilities (trails, campgrounds, etc.), activities (hiking, wildlife viewing, etc.), alerts (temporary closures), events, and similar recreation elements.	Recreation One-Stop Initiative	Not directly related to TransXML

Bridge

The AASHTO BRIDGEWare data model addresses the Analysis and Design, Rating, Inspection, Management, and Maintenance topic areas in the Bridge Structures area and could serve as the basis for development of XML schemas in the bridge area. NCHRP 12-50 included development of an XML representation of a data structure for bridge analysis and rating; NCHRP 20-7 Task 149 addresses the bridge construction area—it included a proposed XML standard for steel bridges based on the Virtis/Opis data model. This standard referenced an American Institute of Steel Construction standard (CIS/2) for exchange of steel construction data across CAD and analysis programs. Another related effort, still in its early stages, is ongoing at the American Iron and Steel Institute.

Safety

The topic areas listed in Table 1 for the safety area have been addressed to some extent by a number of different efforts. Important standards and guidelines related to transportation safety crash reports include ANSI D16 (Classification of Motor Vehicle Traffic Accidents), ANSI D20 (Data Element Dictionary for Traffic Records Systems), the Model Minimum Uniform Crash Criteria (MMUCC), and standard reporting formats defined by NHTSA for the Fatal Accident Reporting System (FARS), and by the Federal Railroad Administration (FRA) for its Railroad Accident/Incident Reporting System.

There are several XML efforts related to transportation safety that are related to the scope of TransXML as follows:

- The Global JusticeXML effort provides XML schema relevant to transportation safety, with an emphasis on Citations, Driver Information (arrest warrants, and driver history) aspects of safety (as opposed to traffic analysis and crash reporting aspects).
- NHTSA developed MMUCC XML as a subset of the Global JusticeXML Data Model (GJXDM)—this XML schema was released late in the TransXML project and therefore was not included in the Task 3 literature review.
- The Traffic and Criminal Software package (TraCS) developed by the State of Iowa in partnership with FHWA includes the MMUCC data elements and an XML data export routine.
- The Federal Motor Carrier Safety Administration (FMCSA) uses XML as a standard data transfer format, addressing the Federal Motor Carrier topic area within the Safety business area.
- A portion of the LandXML schema addresses Roadway Information for safety analysis—it is being used in conjunction with FHWA's Integrated Highway Safety Design Model (IHSDM), to bring in grade information needed for roadway design safety analysis.
- The AASHTOWare TSIMS effort produced a preliminary draft data dictionary with roadway characteristics relevant to traffic safety analysis. The TSIMS functional design envisions extensive use of XML as the means of data exchange across the different modules of that system.
- The IEEE 1512 family of standards cover incident management, traffic management, and hazardous materials incident response.
- The COMCARE Alliance and the Emergency Interoperability Consortium have produced several XML data standards related to emergency response, including the Common Alerting Protocol (CAP), the Emergency Data Exchange Language (EDXL), and Vehicular Emergency Data Sets (VEDS).

Other Relevant Efforts

XML efforts that are related to the safety area but are cross-cutting in nature were also reviewed. These reviews are summarized below.

Spatial Information Standards. Three standards bodies, ISO TC211 (GIS), OGC, and ISO/IEC SC32 JTC1, have developed standards for representing spatial information, much of which has relevance for the Safety business area (Crash Location). TC211's 19107 Spatial Schema is a comprehensive specification for a robust set of abstract geometry and topology types. OGC has selected a subset of geometry types (e.g., points, polylines, polygons), referred to as simple feature geometries. Their specification addressed implementations based upon CORBA, SQL, and later Java. The JTC1 standard (SQL/MM Part 3 Spatial, 13249 3) covers an SQL'99 object relational implementation of simple feature geometries plus circular and compound curves and curve polygons. Though initially independent efforts, the three groups decided to harmonize their standards based upon a common geometry object model. This model is now supported by most of the major GIS vendors. Topology standardization is less mature. TC211 addresses topology at an abstract level. An effort to expand OGC simple feature geometries to include topology failed, though GML has been extended to include topology.

Location Referencing Standards. At least three standards currently are under development for standardizing the manner by which a location can be specified, which is of relevance for the Crash Location topic of the Safety business area. The farthest along is within the ISO TC211 19133 Location-Based Services, Tracking, and Navigation. The ITS Location Referencing Message Specification (LRMS) has just been reissued for comment as SAE J2266. WG 3.3 of ISO TC204 (ITS) has been struggling with achieving consensus in their location referencing standard, due to disparities in location referencing in America, Europe, and Asia.

Roadway Information Standards. At least two standards address roadway information, which is of relevance to the Roadway Inventory, Incident Management, Work Zone Safety, and Emergency Evacuation topics within the Safety business area. The ISO TC204 (ITS) Geographic Data Files (GDF) 14825 standard evolved from its European CEN counterpart. It specifies how roadways are represented for the in-car navigation systems market and is supported by roadway map vendors, such as Navigation Technologies and TeleAtlas. As part of the National Spatial Data Infrastructure, the Geospatial One-Stop (GOS) data content standard formalizes how geographic information in any of seven themes can be represented for transfer between government agencies. The transportation theme includes modes of road, rail, transit, air, and navigable water. It is their intent to submit this specification to ANSI for standardization. The road mode of the transportation theme is consistent with constructs found in GDF. It includes the ISO 19133 linear referencing clause. An implementation specification based on OGC GML is anticipated.

Traffic/ITS Standards. There have been several efforts to establish data standards in the area of traffic management and ITS. The Traffic Management Data Dictionary (TMDD) defines the specific data elements exchanged between Advanced Traffic Management System (ATMS) and other ITS applications, such as Advanced Public Transit Systems, Advanced Traveler Information Systems (ATIS), and Commercial Vehicle Operations. Its goal is to provide nationally accepted definitions to consolidate, resolve, and facilitate data exchange. Similar to TMDD, the P1512 Incident Management Data Dictionary defines data elements for information generated and transmitted between the emergency management subsystem to all other subsystems and providers. The ATIS Data Dictionary, developed by the Society of Automotive Engineers (SAE), contains a minimum set of medium-independent messages and data elements needed to deploy ATIS services, and provide the basis for future interoperability of ATIS devices. The TSDD was developed in parallel with the TMDD, and standardizes data used by traffic analysis software. Finally, the Universal Traffic Data Format (UTDF) establishes a standard data format for signal and traffic-related data at intersections.

3.2 Screening and Evaluation of Schema and Standards

The XML schema and data standards reviewed in Section 3.1 were screened to determine which should be evaluated in detail for potential incorporation in TransXML. The screening criteria applied were as follows:

- **Relevant** to the surface transportation business focus areas of bridge structures, safety, survey/design, and construction/materials;
- **Beneficial** to address a real business need and support data exchange across disparate applications for transportation planning, design, construction, maintenance and operations; and
- **Nonproprietary** namely, in the public domain.

Table 2 indicates which schema and standards were selected for detailed evaluation. The detailed evaluations were designed to assess the suitability of the schema or standard for inclusion in TransXML. Evaluation criteria were as follows:

- **Extensible**—Reusable elements;
- **Technically sound**—Well-formed, well-structured, simple/ clean design, use of elements as opposed to attributes;
- **Practical**—As demonstrated by the level of existing use, level of support from users and vendors, documentation, existence of sample schema and programs using those schema, ease of creating instance documents, ease of parsing; and
- **Compatible**—With related data and XML standards that have been adopted by the user community—e.g., a schema with a crash type element should use compatible coding with existing ANSI standards for traffic data records.

The detailed evaluations of the schema/standards are included in Appendix A.

SECTION 4

Gaps and Opportunities for TransXML

4.1 Criteria for Identifying XML Schema Candidates

XML Schemas provide the mechanism for exchanging information (data) in order to carry out various transportation business functions. Rather than using a data-centric approach to identify TransXML schema opportunities (e.g., what are all of the things about a bridge that anyone would ever need to know?), the approach taken for this project was to use a process-centric view (e.g., what information about a bridge is required to evaluate load restrictions?). The latter question is more focused, and easier to get agreement on.

Not all sets of data that feed transportation business processes are good candidates for XML schema. A *good* candidate for an XML schema is one which has the potential to save time and money, or facilitate improved access to information by

- Allowing information produced in one process to be used in other(s);
- Eliminating the need for information to be entered from scratch (e.g., getting information from GIS systems into CAD systems or using information from road designs to control GPS-guided excavation equipment);
- Allowing agencies to analyze the same set of data in several different software tools, to easily port data from one tool to a new tool, or to share input data across different agencies which make use of varying toolsets;
- Eliminating or reducing the need for agencies to build custom applications and interfaces to connect legacy systems; and/or
- Enabling reuse of the same information for multiple purposes (e.g., sharing information about a scheduled bridge construction project with the DOT permitting office, maintenance staff, and traveler information systems).

The potential benefits of an XML schema increase as the number of different users of and uses for the same "packet" of information increases, and as the complexity and criticality of the information increases (since these factors affect the costs of duplicate data and the impact of errors that can result from duplicate data entry processes).

The ability of a candidate XML schema to achieve these benefits depends on the likelihood that it will be broadly accepted and put into practice. This in turn depends on the ease of getting consensus on a standard data structure, the business case that can be made for the schema, the incentives and disincentives for adoption across the stakeholder community, and the level of advocacy and assistance that is provided to overcome initial barriers to adoption.

Therefore, *poor* candidates for XML schema have the following characteristics:

- There is wide variation in data content across agencies and no mandates or incentives to standardize the data;
- Information is shared across a small number of identifiable systems or individuals, mechanisms for data transfer are already in place and functioning well, and it would be costly to retool systems to read/write another format;
- The structure of data content is highly dynamic in nature and therefore a schema could be obsolete by the time it is put into practice;
- The data content is highly detailed and expression in the verbose XML format would result in performance problems for the intended applications; and/or
- The data content is so simple or trivial that it is not worth the effort to pursue XML encoding.

With these considerations in mind, candidate business processes for XML schemas were identified within each of the four business areas identified in Table 1. Then, the highest priority candidates were selected based on an assessment of the potential payoff from a technical standpoint, the likelihood of adoption from an institutional/business case perspective, and the level of effort that would be required to develop and gain stakeholder agreement given the work that has already been accomplished by the previous efforts identified in Section 3. The selections also took into account the fact that several schema development efforts are ongoing (or planned) within other organizations (e.g., LandXML), and therefore any improvements would need to be pursued in coordination with these other organizations.

The priority business processes identified in this section are the basis for the selection of the initial set of schemas that were developed for TransXML.

4.2 Roadway Survey/Design

In order to explore schema opportunities in the roadway survey/design business area, the research team drew upon a business process modeling exercise conducted as part of a prior project (conducted by Bentley Systems, Inc. for the Minnesota Department of Transportation [MnDOT]). This project produced a set of data flow diagrams which depict the preliminary and detailed design of roadways and related structures. The diagrams show the design activities (functions) performed as well as the data that is exchanged between these functions.

The data flow diagrams were reviewed from a TransXML perspective. Each data flow was evaluated to determine if it is an appropriate candidate for a TransXML schema. The following subprocesses emerged as candidates for XML schema:

- Sharing of roadway geometric design information (horizontal and vertical alignments, pavement section, superelevation, cross sections, geometrics) across design team members as it evolves throughout the design process;
- Utilizing information produced in the design phase as the basis for developing as-built information from the construction phase, and then making the as-built information available for use during the maintenance and operation phase;
- Sharing of surface models across design, survey and construction, including use of this information for automated machine control;
- Transferring pay item information in the design phase for further development in the construction phase; and
- Transferring GIS-based planning information on area features into CAD-based design software.

These subprocesses cover all of the specific topic areas listed in Table 1 under the Roadway Survey/Design business area, with the exception of Pavement Design and Traffic and ITS Design (Right-of-Way may be partially addressed by LandXML, and the GIS-into-CAD area). These two topic areas are fairly large in scope, and involve distinct stakeholder communities (pavement designers and traffic engineers), and should be addressed once the initial set of schemas are established.

The first three subprocesses have been addressed by the LandXML schema. The following discussion identifies gaps in the existing schema that were considered within the scope of TransXML (in coordination with LandXML.org).

Geometric Design/Surface Model Information (LandXML)

The most significant information exchanged during design and carried forward into subsequent life-cycle phases is the geometric design of the roadway. Beginning with preliminary design, the roadway design evolves through the refinement of the horizontal and vertical alignments and the addition of pavement section, superelevation, cross sections, and geometrics. This is an iterative process based upon project constraints, codes and other regulations, budget, and time. It is done in conjunction with other disciplines which impact or are impacted by the design, including but not limited to right-of-way, drainage, utility, hydraulic, traffic, environmental, and aesthetic concerns. XML can provide a method of sharing the roadway design as it evolves during the design process. It can also provide the basis for capturing as-built information as the roadway is actually constructed. This information can then be utilized during the maintenance and operation of the roadway.

As each member of the design team develops a design solution for their particular part of a project (drainage, roadway, bridge, traffic), they are dependent upon and have influence upon the designers of the other parts of the project. For example, a roadway designer begins with a rough alignment from the preliminary design phase. As this alignment is refined, the results need to be communicated with the hydraulics engineer to complete the site drainage; with the bridge engineer to develop the detailed bridge geometry; with the traffic engineer to begin staged construction planning; and eventually with the surveyors for stake out (see Figure 2).

The information being exchanged which constitutes the roadway design includes the horizontal and vertical alignment, cross sections, superelevation, and geometric information.

This area has been a focus for LandXML, which is already in widespread use, supported by LandXML.org and yielding substantial benefits. However, there are areas of LandXML which can be improved:

- Provide a semantic model and documentation in order to increase clarity, as certain areas are open for interpretation by users, thereby increasing the usefulness of the schema for interoperability. For one-off data transfers between two specific processes, this interpretation is acceptable as long

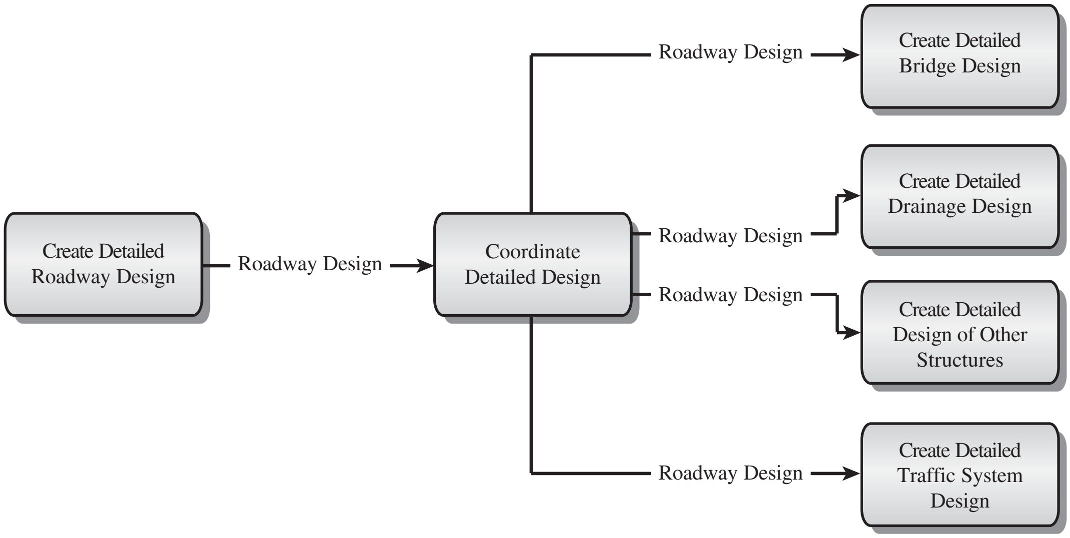

Figure 2. Data flow diagram: coordinate detailed design.

as the two processes make the same assumptions. As the use of the schema broadens (e.g., extending into construction and maintenance and operation), the ambiguities can become more problematic.

- Support the feature-based approach of most spatial standards, including the ISO TC211 191xx geospatial standards, OGC specifications including GML, and Geospatial One-Stop. LandXML supports the CAD notion that geometric representation is paramount and that attributes or feature properties can be hung off of these, almost as an afterthought. An enterprise view would be better served by an object or feature-based approach, where the geometric representation is merely one more attribute of a feature. This would accommodate multiple geometric representations for the same road feature, such as a 1:24,000 scale GIS linestring approximation and a 1″=50′ engineering true curve representation.
- Expand to include general purpose topological primitive constructs, i.e., link-node linear networks. Currently LandXML supports linear topology with feature to feature relationships between pipes and structures. TransXML needs to support other linear networks, such as the roadway network itself.
- Provide additional enhancements (e.g., the ability to break by lane for superelevation, the need to transition superelevation nonlinearly, and the featurization at individual points of cross sections).

As noted above, because LandXML is already widely used, and there is an established mechanism for improving this schema, the NCHRP 20-64 project did not pursue the above improvements independently from LandXML.org. Rather, NCHRP 20-64 provided a mechanism for developing technical content for proposed improvements that were submitted to LandXML.org for consideration.

Contract Pay Items

Contract pay items span the design and construction business areas. Pay items are the basis for estimating the cost of the project, comparing alternative design solutions, obtaining a contractor to construct the roadway, assessing the progress of the construction, providing the basis for partial (progress) payments during construction, billing the work completed, and potentially feeding maintenance and operation systems such as roadway inventory and asset management. Though contract pay items predominate in the construction phase, they begin during design and are therefore proposed as a design business area schema which can evolve into a construction schema as more information is added to the pay items.

Pay items provide the basis for estimating, bidding, and construction management. The design engineer determines what pay items make up the project and determines how much of each pay item will be required to complete the job. This is based on a standard set of contract pay items with predefined units of measurement. The pay items and their quantities are then passed to the estimator to predict the cost of the project. Often the items and quantities are included in the contract documents for the project (see Figure 3).

The information being exchanged includes the standard contract pay items with their units of measure, and the quantities of each that will be used on this project. The

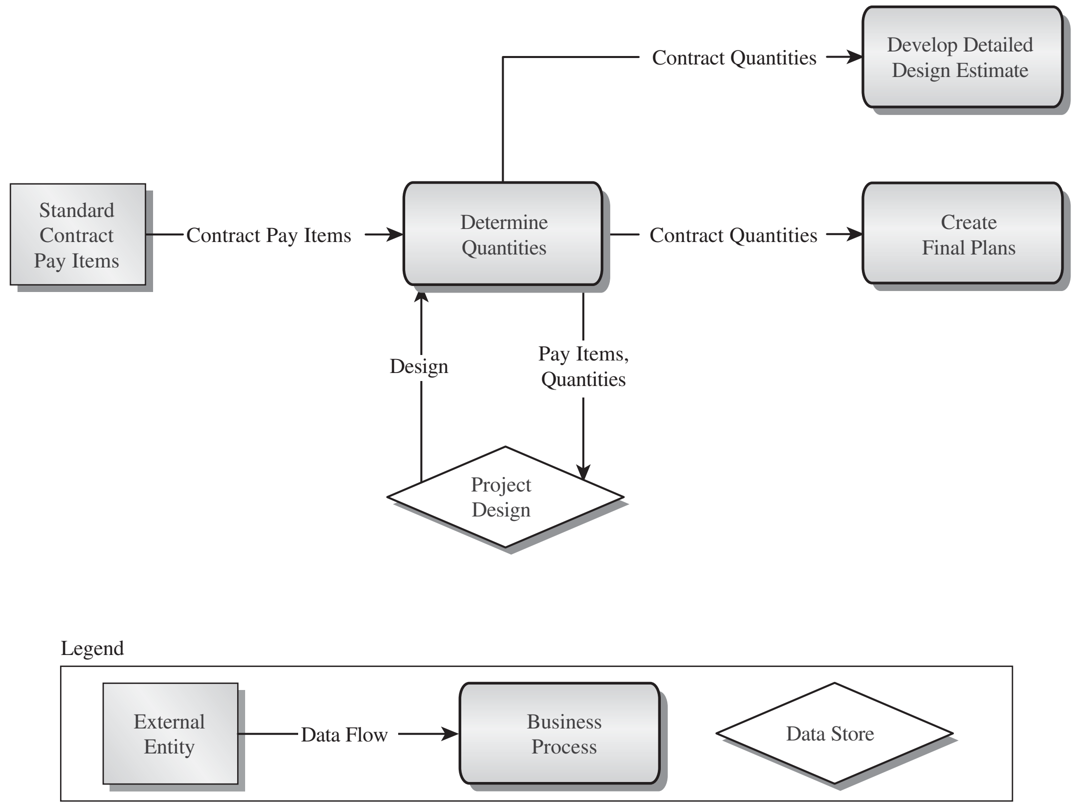

Figure 3. Data flow diagram: determine quantities.

aecXML infrastructure effort produced a good first cut at such a schema, which has been implemented in product at the MnDOT. However, this schema does not completely meet the needs presented here because it doesn't address quantities or prices. A schema that provides a single, consistent list of pay items on a project, with quantities and prices would better support the needs of the designer, estimator, plans developer, contract administrator, contractor, and inspector.

The design-to-construction process emerged as a high-priority candidate for TransXML. The model and resultant XML schema developed here were further augmented within the construction business area in support of pay item data exchange for construction phase activities.

The aecXML infrastructure break out was done under the IAI aecXML Domain Committee umbrella, but it was largely an independent effort by the aecXML Infrastructure Working Group. While the IAI aecXML effort is still active, the aecXML Infrastructure Working Group is now dormant. Given that the focus of the larger aecXML effort has been on buildings which have substantially different requirements than roadway projects, it was appropriate for NCHRP 20-64 to take on further enhancements to this schema.

Area Feature Support

Instead of starting with a clean sheet of paper, it would be advantageous if designers could capitalize on the information collected during the planning phase. Much of this information is available with GIS software which is often incompatible with engineering design software.

A significant number of data flows involve polygonal data, including environmental areas, soils, wetlands, land use, flood plains, site improvement areas, right-of-way, and cadastral information. Though less significant than the previous two schema opportunities, it does represent a significant gap in the design area. It also is exemplary of the gap between GIS and CAD, wherein traditional systems of each type have provided significant frustration for users in their inability to exchange data. It also is an opportunity for incorporating the common geometry model adopted by the three leading international GIS standards.

An example of a design activity requiring such a data exchange is creating preliminary design (see Figure 4).

LandXML does support a survey view of a land parcel element. However, its geometry does not provide support for holes and islands and constraints such as closure, nonoverlapping and nonintersecting rings as specified in other spatial

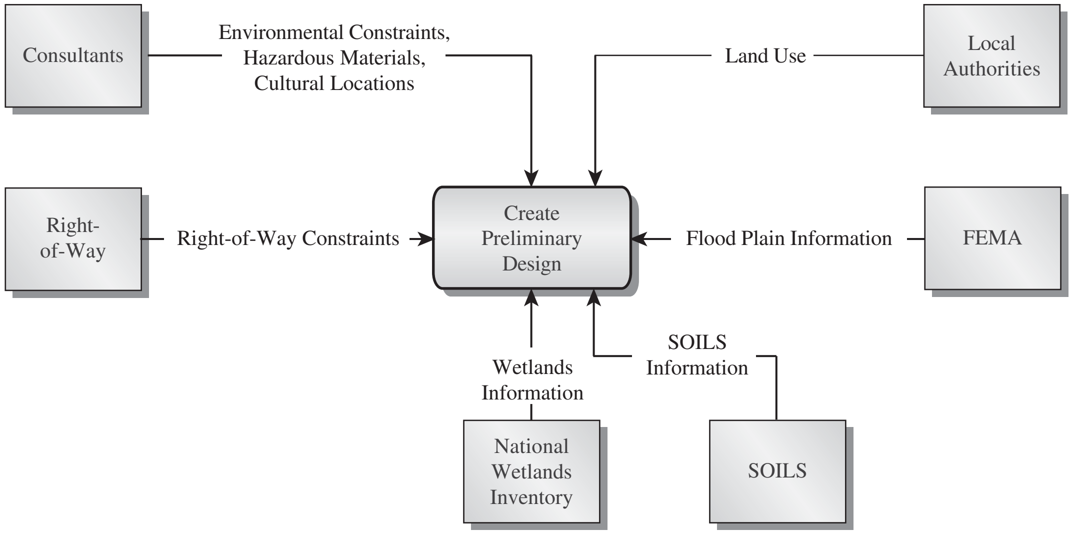

Figure 4. Data flow diagram: create preliminary design.

standards. In addition, there is also no support for general purpose area feature geometry usable for other, nonparcel, area features. General polygon support is not necessarily roadway-specific.

4.3 Transportation Construction/Materials

Several candidate business processes for XML schemas were identified for the construction/materials business areas in Table 1, i.e., estimates, proposals, letting and award, construction management, and materials. These processes are listed below roughly in chronological order based on the life cycle of a construction project:

- Acquire reference information for project cost estimates;
- Submit project definition of work to oversight institution;
- Bid package preparation;
- Submit project bid;
- Track installed quantities, materials used and tested during construction;
- Publish construction status to stakeholders;
- Change project scope;
- Monitor civil rights, on-the-job training (OJT), labor requirements compliance; and
- Monitor subcontract progress.

These subprocesses cover the topic areas listed in Table 1 for the Design/Construction business area. Three of these candidates were identified as the highest priority candidates that could be addressed within the scope of the TransXML project: Bid Package Preparation; Track Installed Quantities, Materials Used and Tested During Construction and Publish Construction Status to Stakeholders. These areas were judged to have the greatest potential payoff given the number of different users of the information. They also include foundation elements that can be later reused in the other areas—e.g., pay items. The three selected candidates are discussed below, followed by brief descriptions of the other candidates which could be considered for future extensions to TransXML.

Bid Package Preparation

Thousands of transportation proposal bid packages are published each year, often in paper form, to a community of tens of thousands of contractors, subcontractors, and suppliers. A standard transportation Bid Package XML schema would enable agencies to publish bid packages electronically in a standard form that can be directly loaded into bid preparation systems. As a result, information flows will be streamlined, and redundant data entry and the associated opportunity for error will be substantially reduced.

Following the creation of pay items in design, a proposal including the pay items is assembled into a bid package that is published to the contracting community. Primary contractors, subcontractors, and suppliers utilize this information to prepare their bids and quotes. The bid package is often published in paper form, requiring manual loading of the pay items into the tools contractors use to prepare bids. TransXML can provide a standard form for publishing the quantitative elements of bid package information in electronic form, enabling contractors and suppliers to directly load the information into bid preparation systems.

In preparation for letting a proposal, transportation agencies publish a proposal bid package. Contractors use this information to prepare their bids. Subcontractors and suppliers use this information to identify potential business opportunities and submit quotes to primary contractors. In the event that project changes occur after publication but prior to the letting, the agency publishes these changes in proposal amendments.

The bid package contains the following:

- Letting location and date;
- General proposal information including the project location, type of work, and vendor qualification requirements;
- Contract time information including the completion date and liquidated damages rate;
- Pay item information including the item description, quantity, and unit of measure;
- Design drawings; and
- Miscellaneous “boilerplate” information.

Proposal amendments can include changes to any element of this information. Agencies publish bid packages on a monthly or semimonthly basis.

With the exception of the graphical information (drawings) and boilerplate, the bid package information lends itself well to representation in XML form. The aecXML Infrastructure Project schema already provides some of the general proposal information (project location, type of work, etc.) and pay item information required for a bid package. TransXML can augment the Infrastructure Project schema to support the letting process requirements for publishing bid packages. This can be done by supplementing it with proposal information including the letting location and date, vendor qualification requirements, contract time information, and the additional attributes required for amendments.

Track Installed Quantities and Materials Used and Tested During Construction

A broad range of field devices are used to measure construction progress and track material use, sampling, and testing. Various elements of this information are communicated frequently among field, project office, test lab, and central office personnel throughout the construction project. A standard XML schema for this information would enable integration of the diverse data collection and data management systems utilized to track this information, thereby streamlining information flows and reducing the opportunity for error. TransXML can provide a standard form for communicating progress information between systems that utilize the physical project view of items installed and materials sampled and tested and the construction progress payment systems that utilize the business project view. This information lends itself well to representation in XML form.

The information being exchanged includes pay item descriptive information, partial quantities placed, placement locations, materials samples collected, the field tests performed, and the outcome of those tests. (Note that this candidate process does not address the actual test measurement and process details of the tests that are performed. Developing a consensus on that aspect of materials testing could not be accommodated within the scope of NCHRP 20-64.)

Key data flows in the daily project work tracking process are as follows:

- At the project construction site, inspectors track partial quantities placed for pay items, project locations where those quantities were placed, the frequency of samples taken for component materials, tests performed and the extent to which the materials meet established specifications. Physical measurements are taken using various types of field equipment. Once collected, this physical view information must be translated into the business view of the project (pay items), and recorded manually or with data collection systems such as handhelds, laptops, or personal computers in the field office. This translation and subsequent entry into the construction progress tracking system or central lab system is typically a manual process that has an associated risk of translation and data entry error.
- Materials sampling and testing data generally is exchanged between field offices and central laboratories on a daily basis.
- The inspectors communicate the progress information to the project engineer, who monitors the overall project status.
- On a periodic basis, the project engineer submits a progress estimate to the central office. The progress estimate consists of the aggregated partial quantities placed for the pay items and the record of the materials sampling and testing that were performed during that pay period.
- The central office processes the progress estimate, triggering a payment through the financial system.

The aecXML Infrastructure Project schema provides a good starting point for this TransXML candidate—it already has a pay item complex type that includes the required unique item identifier, description, unit of measure, and one or more associated costs consisting of a quantity, price, and a cost type. This quantity can be a partial quantity placed. The schema also already has a location complex type. However, the following gaps were identified:

- Partial quantities placed with a location;
- Material samples collected; and
- Field tests performed and outcome of those tests.

Publish Project Construction Status to Stakeholders

External stakeholders such as the general public, elected officials, oversight or regulatory agencies, and other institutions such as utilities and railroads require or can benefit from access to timely information about the status of a transportation construction project. This information is managed within the agency in their construction management system and is provided to different stakeholders in different forms. TransXML can provide a mechanism for communicating construction project progress information from the agency's construction management system to external stakeholders. This information can be presented in a variety of forms appropriate for the individual target audiences.

The information being exchanged includes project description, location, and fiscal, schedule and progress information, including milestone dates and those affecting traffic. The frequency of publication depends on the practices of the agency.

The aecXML Infrastructure Project schema provides a ProjectType complex type (imported from the aecXML Common Object Schema) that includes required general project information (project description, location, etc.) and project status information, including current phase, percent complete, begin date, and several completion dates. This schema requires supplementation with additional milestone dates and fiscal and progress information to be useful for status tracking.

Acquire Reference Information for Project Cost Estimate

Estimators use production, labor, and equipment rates and historical bid and as-built information to estimate labor, equipment, and material costs for the project pay items. An estimator often gets this reference information from a variety of internal and external sources, including data warehouses and commercial and government publications and services.

The information being used includes production, labor, and equipment rates, historical bid information, and as-built information.

Submit Project Definition of Work to Oversight Institution

At various stages in the project life cycle, the transportation agency must submit the definition of project work to various oversight institutions. For example, in the United States, state transportation agencies must submit their construction plans, specifications, and estimates (PS&E) to the Federal Highway Administration (FHWA). When the project is awarded, they must submit the awarded contract to FHWA. Similar requirements apply for other oversight institutions such as the Army Corps of Engineers, the Coast Guard, and, for international projects, financial institutions such as the World Bank.

The information being provided includes project descriptive information, pay items with quantities and either estimated or contract prices. Depending on the requirements, it may include additional information such as funding sources and projected funding allocations.

Submit Project Bid

At or before the time of the bid letting, contractors submit their bid for a project to the transportation agency. The bid document has the bid price for each pay item. Depending on agency requirements, it may include additional information such as subcontracting commitments and bid bond information.

The information being submitted includes a bid price for each pay item, and depending on agency requirements, additional information such as subcontracting commitments and bid bond information.

Change Project Scope

As project construction proceeds, changes to the scope of the project are identified including adjustments to authorized pay item quantities, addition of new work, and changes to contract time requirements. Project staff define the changes to be made and submit them for review and approval. On approval, the changes are communicated to the contractor.

The information being exchanged can include contract and project descriptive information, pay item quantity adjustments, new pay items, and modified contract time specifications.

Monitor Civil Rights, OJT, and Labor Requirements Compliance

Transportation agencies must monitor contractor compliance with civil rights, OJT, and labor requirements. Contractors may subcontract portions of contract work to other contractors. Agencies require contractors to submit certified contractor payroll information regarding contractors' and approved subcontractors' workforces.

The information being exchanged includes contractor, contract, and employee identifiers, employee characteristics including address, gender, and ethnicity, and daily work information including hours worked and wage rate information for the pay period.

Monitor Subcontract Progress

Transportation agencies record progress on both prime contractor and subcontractor work. Agencies require contractors to submit information on the work being performed by subcontractors.

The information being exchanged includes contract descriptive information and pay items with quantities provided by the agency, and subcontractor descriptive information and pay item subcontracting information provided by the contractor.

4.4 Highway Bridge Structures

The topic areas for bridge structures included in Table 1 spanned the full life cycle of a bridge—analysis and design, construction, inspection and load rating, management, operations, and maintenance. Given the common data exchanges that occur within and across these areas, three specific business processes within these areas were identified as good candidates for XML schema:

- Bridge Analysis and Design;
- Truck Permitting and Routing; and
- National Bridge Inventory (NBI) Reporting/Data Exchange.

Each of these candidates is discussed below.

Bridge Analysis and Design

The major candidate within the bridge structures area for inclusion in TransXML is a physical description of the bridge geometry and structural characteristics. This information is developed during the structure analysis and design phase, and then it is used throughout the bridge life cycle for development of load ratings and as input to permit and routing applications (see below).

A standard XML format for describing the structural elements of a bridge would facilitate use of multiple design and analysis packages for a given design problem. According to the AASHTO LRFD specification commentary, the verification of computational processes used for bridge analysis is the responsibility of the engineer. Because of the large amount of specification checking required for a single structure, verification of a process is not always simple. Hand checks are not always practical because of the iterative nature of many of the specification checks. Passing information, both bridge description (input) and analysis/specification results (output), is often desirable to expedite the comparison of two processes.

In addition to having a standard description of the bridge for input to design and analysis software, it would also be useful to have a standard format for representing outputs of a bridge analysis. This would facilitate comparisons of the structural analysis results across different bridge software applications which may have different output formats.

The complexity of bridge structure description information makes it a good candidate for XML; however this complexity also means that taking on the task of developing XML schema for all bridge types is beyond the scope of the current project. There will be a need to clearly define a manageable subset of structure types to focus on. In order to make this effort most productive, it is desirable to build upon existing work that has been accomplished within the AASHTOWare program on Virtis/Opis, and on prior NCHRP studies that have addressed bridge specification.

The AASHTOWare Virtis/Opis Project provides a comprehensive bridge domain which contains bridge descriptions for the purpose of structural analysis for many bridge types. These types include steel plate girder, rolled beam and built-up multigirder superstructures, reinforced concrete tee-beam and slab superstructures, sawn timber multibeam superstructures and precast, prestressed concrete I-beam and box-beam superstructures. Currently the Virtis/Opis software provides a reporting tool that produces an XML representation of the bridge domain. This XML information is used with a dynamic XSL template generator to produce user-defined reports for viewing using an Internet browser.

The NCHRP Project 12-50 process provides a start for the definition of Bridge Specifications information. This process is currently being applied in research on several NCHRP bridge-related projects.

Truck Permitting and Routing

Each state has procedures for issuing permits to vehicles which exceed established size or weight limits for travel on state highways. A number of different permitting and routing applications are in place to select a permissible route for a vehicle based on the vehicle's characteristics (length, axle configuration and weight) and the characteristics of the road network—including bridge horizontal and vertical clearances and load ratings. A standard packet of information about bridge characteristics required for permitting and routing is a logical candidate for an XML schema.

NBI Reporting/Data Exchange

The Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges defines a standard data structure for required annual NBI reporting on all public highway bridges. The current reporting format is a fixed format text file. NBI data is produced by bridge management systems (including AASHTO's Pontis system) and many state DOTs have developed NBI reporting utilities that work with their bridge databases. NBI data is shared across different agencies and is used for a variety of analyses.

Developing an XML schema for NBI data would provide a self-documenting format for NBI data and would provide a much-needed opportunity to address longstanding problems with the fixed format (e.g., constraints on the number of digits for certain data items). It would allow for development of standard web services for validation, queries, and reports.

Several changes to the NBI data items have been proposed, and an update is likely over the next couple of years. It would be best to roll out a new NBI XML format in conjunction with this update. Therefore, the timing was not right for development of an NBI XML schema as part of NCHRP Project 20-64.

4.5 Transportation Safety

The transportation safety area involves many different business processes on the part of multiple agencies. Major processes include the following:

- Collection and processing of crash reports, including work flow to validate and approve reports across multiple agencies (state and local law enforcement, state highway safety office, state DOT);
- Federal reporting of safety data from state agencies and operators to NHTSA (Fatal Accident Reporting System, State Data System), FMCSA (SAFETYNET) and FRA;
- Linking crash data with other data sources (highway inventory, vehicle registration, driver licensing, enforcement, medical/injury control, road weather information) to provide information needed to analyze causal factors and outcomes;
- Sharing of crash reports within and across multiple agencies and organizations;
- Querying, reporting and mapping crash data based on a variety of criteria;
- Determining high-accident locations by conducting statistical analysis of crash data and vehicle miles of travel on different types of roadway sections;
- Analyzing crash information to develop strategic safety plans;
- Identifying and evaluating safety countermeasures for specific locations;
- Evaluating the safety implications of alternative facility designs and of alternative construction and maintenance practices, (including work zones);
- Commercial vehicle licensing, inspection, and permitting processes, including those related to hazardous materials transport;
- Real-time (ITS) management of incidents and emergency response activities; and
- Real-time (ITS) work zone safety management processes.

Real-time exchange of incident information is already being addressed in the ITS arena; and exchange of motor carrier safety information is being addressed within FMCSA. The two highest priority remaining opportunities for XML schema in the safety arena are for crash reporting and highway safety analysis. These are discussed below.

Crash Reporting

Crash records are an obvious candidate for an XML schema. Crash records are used by many organizations and individuals for a variety of different purposes. They are sufficiently complex to make manual reentry of data costly. Their accuracy is critical. Even though crash data are not currently standardized, the Model Minimum Uniform Crash Criteria (MMUCC), the FARS reporting requirements, and ANSI D20 and D16 provide a solid base of widely adopted standards on which to build an XML schema. Many states have uniform crash reports and are in various stages of implementing electronic crash reporting systems; several are already using XML. The AASHTO TSIMS design (*5*) recommends XML as the primary data exchange format for import and export functions, passing data to and from external legacy systems and the TSIMS data warehouse, and between TSIMS and external safety analysis applications. (Please note: As of the time NCHRP Project 20-64 was completed, the TSIMS project had been discontinued. It is anticipated that this project will be replaced by a new, reduced scope project entitled "Safety Management System.")

A standard crash record document (even a base that could be adapted by individual states) would facilitate implementation of electronic crash reporting systems, reducing the time and cost of crash records processing. It would allow for development and sharing of software (e.g., web services) for

- Validating individual crash reports,
- Aggregating crash reports from different sources,
- Querying and reporting of crash data,
- Mapping of crash data, and
- Analyzing crash data.

During the course of this project, NHTSA released an XML schema for crash records based on the MMUCC. This schema was developed in coordination with the Global JusticeXML effort, which had previously included some material for the driver history and citation elements of a comprehensive crash records schema, but not most of the MMUCC elements.

Highway Safety Analysis

While the MMUCC elements include some information on the characteristics of the crash site, it is impractical to collect detailed highway inventory information as part of a crash report. In order to identify high-accident locations, evaluate potential countermeasures and conduct safety analyses correlating highway characteristics to crash risks, it is necessary to link crash data to highway inventory data via a common location reference. Therefore, in addition to an XML schema for crash data, it would also be beneficial to define a schema for a standard set of highway inventory elements that could be extracted from existing inventory systems for use in safety analysis. FHWA's Safety Analyst software database is a good first cut at defining the inventory elements required for such analysis. The draft TSIMS data dictionary (which focused on highway data elements) and relevant location and linear referencing elements needed to link crash records to highway inventory provide additional sources for such a schema.

Functions/processes that could make use of this schema include the following:

- Get highway information for a crash and store it in an archive for statistical analysis and queries;
- Get highway information for a crash and use it to validate information on the crash report (e.g., consistency of different location attributes of the crash record, pavement surface type, median type);
- Get safety-related highway inventory information for an identified high-accident location in order to identify and evaluate countermeasures; and
- Given a set of crashes over a given time period, establish peer groups based on highway characteristics, calculate average crash rates and/or hazard indices and identify high-accident locations.

SECTION 5

TransXML Process and Products

This section describes both the process used to develop TransXML schema, and the work products of the effort.

5.1 Schema Development Process

A schema development plan was prepared based on the analysis of gaps and opportunities presented in Section 4. This plan identified the schema and sample applications to be developed, defined the technical standards and processes to be followed in order to provide consistency across the schema, and established mechanisms for stakeholder involvement.

Table 3 summarizes the work included in the schema development plan. An initial activity in the plan was to conduct an experiment to determine whether the XML encoding for TransXML should be based on the Geography Markup Language (GML). This would allow the TransXML effort to take advantage of a rigorous XML foundation, a structured approach that enhances interoperability and tie-in to open geographic data standards. This experiment was conducted, and after considerable debate, the decision was made to adopt GML for new schema but to allow for inclusion of pre-existing, non-GML-compliant schema into TransXML. This approach struck a balance between establishing a consistent framework for TransXML on the one hand, and avoiding duplication of existing, accepted XML schema. Section 5.2 discusses the technical standards established for TransXML and the results of the GML experiment.

A total of 10 data exchange topic areas were selected for detailed data modeling. For one of these areas—Geometric Roadway Design—the schema development plan called for data modeling only, with no XML schema development. This decision was made because of the large and well-established user base of LandXML. In lieu of developing a new TransXML schema, the project team recommended enhancements for consideration by LandXML.org. For a second area—Crash Records—data modeling was completed, but a decision was made not to implement a TransXML schema. Instead, the project team elected to adopt the NHTSA crash records schema that was released during Phase II of the TransXML project. Thus, eight new schema were developed for TransXML, and two external schema were adopted (NHTSA MMUCC XML and a subset of LandXML).

The schema development process involved the following steps:

- Solicit participation in schema development by key stakeholder groups for each of the schema areas through networking, e-mails, and establishment of a collaboration website for TransXML. Stakeholder involvement is discussed in Section 5.3.
- Develop UML models for each of the schema areas in order to gain consensus on data content and structure, and revise based on stakeholder comments. The UML modeling effort is described in Section 5.4.
- Develop XML schema and sample applications. The process used to generate the schema is described in Section 5.5; the content of the schema and applications is described in Section 5.6.

5.2 Technical Framework for TransXML

Schema Development Standards

Schema development standards were established to ensure consistency across the different schema efforts undertaken as part of TransXML. They also provided for adherence to W3C standards for XML schema and for the use of best practices which are generally accepted in the broader community of XML schema developers.

In order to define development standards for TransXML, the following sources were consulted:

- AASHTOWare Standards and Guidelines Notebook: http://aashtoware.org/docs/020820_Complete_Viewable_S&G.pdf (or see AASHTOWare.org);

Table 3. TransXML schema development.

Item	Scope	Comments
A. GML Experiment		Experiment done in order to make an informed decision about basing TransXML on GML.
B. Geometric Roadway Design (GRD)	Exchange of roadway geometric design information across design and survey software packages, and for machine control.	No XML produced – enhancements suggested to LandXML.org based on UML modeling.
C. Design Project (DP)	Exchange of design project pay item data across design, cost estimation, and bid preparation software.	Renamed from "Contract Pay Items" in original Schema Development Plan.
D. Area Features (AF)	Exchange of area features data (e.g., wetlands locations) across GIS and CAD systems.	
E. Bid Package (BP)	Exchange of construction bid package data between agency systems and contractor bid preparation software.	
F. Construction Progress (CP)	Exchange of information about the progress of a construction project in terms of partial pay item quantities placed.	This was originally combined with materials testing in the original Schema Development Plan.
G. Materials Sampling and Testing (MST)	Exchange of construction site installed quantities and materials used and tested information from field data collection systems to laboratory systems and central construction progress tracking and contractor payment systems.	Renamed from "Construction Quantities and Materials" in original Schema Development Plan.
H. Project Construction Status (PCS)	Publication of construction project status information to stakeholder information systems (e.g., project web sites, partner agencies).	
I. Bridge Design and Analysis (BDA)	Conduct bridge structural analysis in multiple software packages without the need to reenter data; compare analysis results across systems.	
J. Crash Report	Exchange of crash data from police reports to validation/processing systems and archives; enables standard queries of crash databases.	(NHTSA/Justice MMUCC XML Schema adopted as part of TransXML).
K. Highway Information Safety Analysis (HISA)	Exchange of highway inventory information between inventory systems and safety analysis software.	

- Draft Federal XML Developers Guide (2002) http://xml.gov/documents/in_progress/developersguide.pdf;
- aecXML Style Guidelines: http://www.iai-na.org/aecxml/documents.php;
- LandXML Detailed Documentation: http://www.landxml.org/spec.htm;
- JusticeXML naming conventions training slides: http://justicexml.gtri.gatech.edu/workshop/Day2/9_Naming.pdf; and
- GML: http://www.opengeospatial.org/docs/02-023r4.pdf—Section 8.

The following guidelines were established:

- The Schemas will conform to the W3C XML Schema Recommendations: http://www.w3.org/XML/Schema;
- Use of attributes within TransXML schema will be limited to conveying metadata about elements that are relevant only within the local scope of the element and are not likely to require further extension or parsing;
- New TransXML schema will avoid duplication of existing elements of related schema that are currently in widespread use through use of inclusion methods;

- The most appropriate and restrictive data types will be used for element types and attributes;
- A default namespace will be used to avoid the need to qualify all element types and attributes with namespaces;
- The temporary namespace placeholder: "http:/www.transxml.net/schema" will be used as the default namespace in the development of TransXML schema;
- All element names will be self-explanatory and meaningful to subject matter experts;
- Plural case shall only be used for element names that represent lists. Otherwise, singular case shall be used;
- Upper Camel Case will be used for schema elements and data types (e.g., "UpperCamelCaseElement");
- Lower Camel Case will be used for properties (e.g., "lowerCamelCaseProperty");
- Schema component names will not contain Underscores (_), Periods (.), and Dashes (—);
- Schema element and attribute names will not use abbreviations, unless there is agreement among stakeholders that the abbreviations are widely understood and that they are necessary for conveying the element meaning; and
- Schema element and attribute names should avoid the use of acronyms, unless the resulting names would be unreasonably verbose and/or the acronyms are nationally recognized. If acronyms are used, they will be spelled out within schema annotations.

Even with adoption of the schema development standards described above, the W3C XML Schema specification leaves considerable room for structural and stylistic differences across schema developers. Also, it was apparent the schema in the transportation field needed to make use of data elements that are required in other application areas—most notably for location, but for other elements as well. The research team felt strongly that associating TransXML with an existing XML effort that provided a consistent technical framework, including foundational elements, would provide a leg-up for TransXML.

The GML was identified as the most promising base for TransXML. The other option would have been the Global JusticeXML Data Model (GJXDM), which has been emerging as a foundation for XML efforts in the Justice, and more recently, Homeland Security communities. While either choice would have been beneficial for TransXML, GML was considered to be better aligned with the types of schema to be produced under TransXML. In particular, it provided better treatment of information describing physical objects and their properties (including geographic location), as well as consistency with adopted open geospatial standards.

Geography Markup Language

The GML is a set of XML building blocks designed to be a foundation on which specific industries, like transportation, can develop domain-specific application schemas. A GML schema follows the W3C XML schema standard so all GML-compliant schema are also XML-compliant and can be validated with available XML tools. GML was developed by the Open Geospatial Consortium (OGC) and is a draft ISO standard within the TC 211 191xx family of GIS-related standards. GML is the format for request and response messages for OGC Web Feature Services (WFS), which means that a variety of web applications for querying and serving geographic data from a variety of sources are being designed to understand GML.

Even though GML is clearly a product of the geospatial data community, it is not only a markup language for geographic data. It provides an equally valid approach to XML encoding for applications that have no geographic content. It claims to be an "XML-based mark-up language used to encode information about real-world objects [features]." These features can be concrete (tangible) like roads and bridges or abstract like property or jurisdictional boundaries. GML provides a standard way to represent features, and properties of these features, which can include (but does not need to include) information about their location and geometry. For example, GML can be used to provide multiple representations of a roadway—its abstract characteristics (e.g., functional classification), its location in space and along a longer route, and its horizontal and vertical alignments. The feature-centric approach supports data exchange across the enterprise because it allows for objects (such as roads) to have geometrical representations of different types and at different levels of precision, supporting a variety of applications (including inventory, GIS, and CAD).

The guidance and structure provided by GML results in more uniform schemas which are more human readable and predictable for software processing. The tradeoff to be made in adopting a foundation such as GML (or GJXDM for that matter) as the basis for TransXML is the additional complexity for schema developers and overhead imposed on the XML files themselves. In order to explore this issue, an experiment was conducted which involved developing parallel XML and GML schema, instance documents, and extraction programs for a candidate application area. The results of this experiment are provided in Appendix B. Key conclusions of this experiment were as follows:

- Use of GML required less effort in developing schemas than straight XML because standard, predefined GML types provided a foundation upon which to build TransXML-specific types. These include features, xlinks, basic types (for measurements, codes, etc.), geometry, topology, reference systems, temporal, units of measure, and styles.
- The GML application schema was easier to read than the straight XML schema. This is because the GML schema was less cluttered with basic type definitions. The GML object-

property approach encouraged stronger typing, which reduces the ambiguity evident in less structured approaches.
- There was no significant difference in the instance documents (.XML files) from the XML and GML schemas. GML has a few extra lines for tags due to the object-property distinction. The difference in case (UpperCamelCase convention for objects, lowerCamelCase for properties) for element tags clarifies the difference. The XSL documents which generate sample reports from XML and GML instance documents are virtually identical due to the similarity in instance documents.

In order to reduce overhead in processing of GML-based TransXML schema, a decision was made to use a subset of GML for TransXML. This subset of GML 3.1.1 is based on the current OGC proposal for Simple Feature GML (SFGML), in OGC 05-033r24, GML simple features profile, Appendix D, March 7, 2006.

5.3 Stakeholder Involvement

Stakeholder Identification and Communication

Stakeholders for each of the business areas were identified, and an e-mail list was developed to inform these individuals and groups about the TransXML project and encourage their participation. Information about the project was also sent to the AASHTO and TRB liaisons at all state DOTs, and to relevant TRB committee chairs for further distribution. The May 2005 TRB e-newsletter included a blurb on the project with a link to the TransXML website. Presentations on TransXML were given at the September 2004 International Highway Engineering Exchange Program conference in Lincoln, Nebraska; the January 2005 TRB annual meeting in Washington, D.C.; the May 2005 AASHTO Information Systems Subcommittee meeting in Baton Rouge, Louisiana; and the August 2005 Traffic Records Forum in Buffalo, New York. Two press releases for the project were issued—one in the spring of 2005 and a second in January 2006. Articles on the project were published in the *AASHTO Journal* in May 2005; and in the Spring 2005 edition of the AASHTO Trns•port News.

Key stakeholder groups identified for the different business areas were as follows.

Roadway Survey/Design

- DOT design divisions,
- Civil design software vendors,
- Engineering firms,
- LandXML.org,
- aecXML Infrastructure Working Group participants, and
- Current IAI aecXML Domain Committee members.

Construction/Materials

- AASHTO,
- FHWA,
- Construction management software vendors,
- Construction contractors,
- aecXML Infrastructure Working Group participants, and
- Current IAI aecXML Domain Committee members.

Bridge Structures

- AASHTO Virtis/Opis Panel,
- Bridge software developers,
- Bridge design firms, and
- DOT bridge engineers.

Highway Safety

- NHTSA,
- FHWA Safety Office/Turner Fairbank Highway Research Center,
- The Association of Traffic Safety Information Professionals (ATSIP),
- National Safety Council,
- Iowa DOT TraCS Consortium,
- AASHTO TSIMS representative(s),
- GLOBAL JusticeXML, and
- COMCARE Alliance.

Collaborative Website

An interactive website was developed to enable a collaborative process of schema development. The "transxml.net," "transxml.org," and "transxml.com" domain names were reserved and were directed at this site. The website included the following features:

- Background information on the TransXML project, including an overview presentation on the project that can be downloaded, a list of contacts, and schedule information;
- Links to external sources of information about related XML and data standards efforts;
- A section for each business area, with capabilities for stakeholders to review sources and resource materials, download draft documents, submit comments and proposed revisions, and participate in threaded discussions on relevant topics;
- Automated notification for interested registered users when new material is added to specific areas of the site; and
- A protected area for the development team to share documents and links to external resources.

A moderator was assigned for each of the four business areas. Moderators were responsible for posting relevant materials to the website, and monitoring and responding to stakeholder comments.

An administrator for the entire website managed general project content (not specific to a business area) and user privileges. The website was set up with public areas containing general project information and a working group section for each of the four business areas. Stakeholders interested in joining working groups were able to register on the site. The site administrator then granted them access to the relevant working groups, allowing them to read and download materials, and participate in discussions. Registration was required to prevent spamming, but it may have presented an obstacle to achieving greater stakeholder participation—particularly since there was a time delay between registration and granting of access.

The initial website went live on March 26, 2004. As of the end of June 2004, over 100 individuals had signed up on the website to be on the project's mailing list and/or participate as schema developers or reviewers. At the close of the project, 376 people had registered on the website.

While several general comments were made at the start of the project, and a few individuals did perform a detailed technical review of UML models and submitted comments, the overall level of participation in the TransXML project was considerably lower than anticipated.

The collaborative website was developed using open source. NET software. The entire site, including the supporting database with all site content is being provided to NCHRP as part of the deliverables for this project.

5.4 UML Modeling

Unified Modeling Language (UML) class diagrams (ISO Standard 19501) were used as the primary design tool for the XML schemas developed for TransXML. UML modeling is standard practice—several standards bodies (OpenGIS, TC211, TC204), have adopted UML as their modeling language of choice. While XML Schemas are humanly readable, it is extremely difficult to grasp the data relationships being represented in an XML schema (.xsd) document. UML models present this information graphically and symbolically, facilitating the review process. Just as it is always easier to modify software (or any constructed item) in the design stage than after it is under construction, it is much easier to modify UML models than completed XML schema. In addition to facilitating the design review process, UML diagrams also serve as useful documentation for the XML schema and are invaluable for supporting maintenance of the schema over time.

Figure 5 shows a sample UML class diagram. This diagram conveys the following information:

- A road project is composed of 0 or more road project pay items.
- A road project has three attributes: a project number, a project name, and a project engineer. The project number has a special ID data type; the project name and project engineer attributes are character strings.
- A road project pay item has two attributes: quantity and estimate. Both quantity and estimate are numerical data types.
- A road project pay item must correspond to an entry on a standard contract pay item list.
- A standard contract pay item is described by four attributes: an item number (ID), an item name (character or string), an optional description (character string), and units of measure (which is a special measurement units data type).
- Measurement units must be selected from a code list, which has four fixed choices: linear feet, cubic yards, tons, or each as well as an option for a user-specified other value.

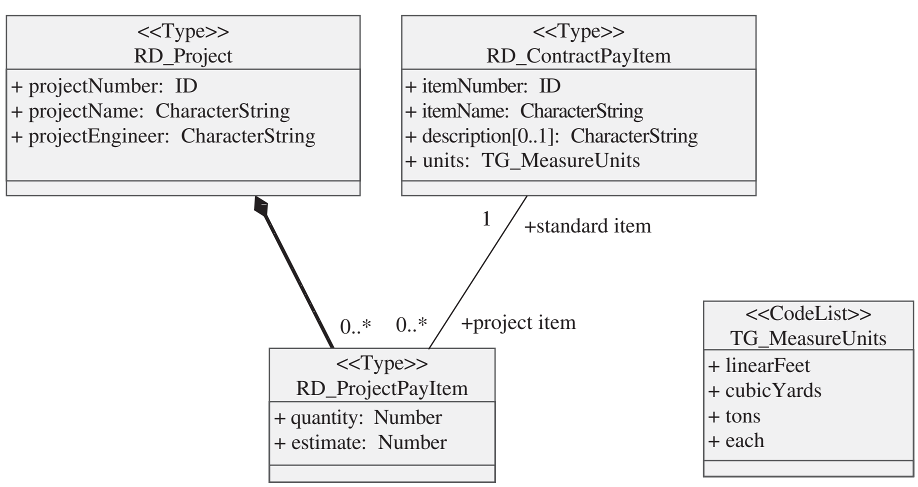

Figure 5. Sample UML class diagram.

Two UML modeling tools were utilized within the development team: (1) IBM's Rational Rose (an "industrial strength" tool) and (2) Sparx Systems Enterprise Architect (a lower-cost UML modeling tool). Appropriately, an XML-based modeling data exchange format (XMI) was used to transfer model information across the two tools.

Standard conventions for the TransXML UML models were established and documented for use by the data modelers from the four business areas. A presentation was developed on how to read UML class diagrams, and posted to the TransXML website as a resource for stakeholders wishing to comment on the diagrams.

UML models for the selected TransXML schemas are documented in Appendix C.

5.5 GML Encoding and Validation

Creation of GML Application Schemas for TransXML

The TransXML schemas make use of the Open Geospatial Consortium GML simple features profile, which is a subset of GML. This subset was developed to simplify the process of writing software that generates and parses the GML/XML documents. The simple features profile is described in OGC document [OGC 05-033r24] (Copyright ©2005 Open Geospatial Consortium, Inc. Used with permission. All Rights Reserved. To obtain additional rights of use, visit http://www.opengeospatial.org/legal/ which explains what needs to be in the schema.)

The Open Geospatial Consortium has made available a program called ShapeChange which converts UML into GML. Documentation on the ShapeChange tools can be found at: http://www.interactive-instruments.de/ugas/. The ShapeChange tool was used to create an initial or partial cut at each of the TransXML schema except for the bridge design and analysis schema. This last schema was created through a separate process to ensure consistency with the existing Virtis/Opis XML report generator output (see Section 5.6 for further information).

Experience with the ShapeChange tool was mixed. The tool makes certain assumptions about element namespace prefixes which did not always agree with the conventions established for TransXML. For example, we use "xs" for all W3C elements and nothing for the namespace of the schema being defined. ShapeChange does the opposite. TransXML also has adopted a uniform order to elements in a schema file: root element, alphabetized base property types, other property types (also alphabetized), and finally alphabetized enumerations and code lists. It was not apparent what order ShapeChange used. ShapeChange was useful in generating the enumerated and code list types, to reduce manual entry. It also provided verification of the structure and content of GML constructs.

In order to guide development of future TransXML schema, a template UML model and a companion template.xsd file were developed.

Future TransXML Schema Validation

For the purposes of future development of XML schemas under the TransXML umbrella, the following requirements are recommended:

1. The schema should conform to the rules for the GML simple feature subset; and
2. The schema should utilize the linear referencing standards outlined in ISO 19133 (Clause 6.6) for all linear referencing.

OGC is currently preparing an abstract test suite that will enable TransXML schema developers to ensure that the first of these two requirements is met. Because of their focus on GML, their test suites are likely to be more thorough and complete than could possibly be developed under this project. In addition, OGC is considering and is likely to recommend incorporating the relevant portions of ISO 19133 as part of GML, and is likely to maintain its test suites as GML evolves. Therefore, it was determined that the TransXML project should not develop redundant validation software. Developers of future TransXML schema should make use of the OGC test suites for schema validation.

The GML application schema validator (Version 2.1.2) checks the schema's validity against Version 1.0 of the W3C XML Schema specification and Version 2.1.2 of the OGC GML rules. The schema must be available on a web server via a HTTP URL. The following checks are included in the validator:

- All feature type definitions must extend xmlns(gml=http://www.opengis.net/gml)AbstractFeatureType or form a complex type that extends that type;
- A GML feature definition must not have a direct child element that derives from xmlns(gml=http://www.opengis.net/gml)gml:AbstractFeatureType;
- A GML feature definition must not have a direct child element that derives from xmlns(gml=http://www.opengis.net/gml)gml:AbstractGeometryType; and
- The schemas must define the schemaLocation for all import and include statements that are resolvable from the source schema URL.

The OGC test suite also provides a GML instance document validator. This validator checks a GML instance document against the GML 2.1.2 schemas and against the application schemas that are defined in the instance document. The instance document being validated can be made available on a web server, or alternatively, its body can be pasted in to the validator form.

The OGC validators can be accessed at: http://cite.opengeospatial.org/gmlTools/index.jsp.

5.6 TransXML Schema and Sample Applications

The schema and sample applications developed for TransXML are briefly described below. See Appendix E for documentation of the schema, and Appendix F for electronic versions of schema and applications.

As part of the UML modeling process, some common building blocks were identified and packaged into three separate components for inclusion in multiple schema. These components are as follows:

- A linear referencing schema based on ISO 19133,
- A reference schema that can be used to store master lists of project pay items or funding sources, and
- A TransXML base schema with various shared elements (e.g., addresses, phone numbers, project identifiers).

Geometric Roadway Design Schema

Description

The Geometric Roadway Design (GRD) schema defines the design control elements of a roadway, including horizontal alignments, profiles, cross sections, superelevation, and geometric control features. This information is critical throughout the design process and must be communicated clearly between various design stakeholders and design data consumers. The information is passed to construction so the road can be built.

The part of LandXML that addresses geometric roadway design is the adopted TransXML schema in this area. This part of LandXML was adopted because it provides good coverage of the important information elements and already has an established base of user and vendor support. Work was conducted within the TransXML project with the aim of making a contribution to future development of LandXML. This work focused on the development of a UML model based on LandXML Versions 1.0 and 1.1. The purpose of this model was to provide a well-documented and unambiguous view of the data contained in the XML schema (in order to reduce the occurrence of inconsistent interpretations for element and attribute meanings) and to identify and recommend improvements to extend the usefulness of the LandXML schema.

The semantic model developed for this project deviates from the existing LandXML model where schema change recommendations are proposed for consideration by LandXML in order to reduce ambiguity. In addition to these detailed recommendations, the following general recommendations were made:

- **Future Extensions**—Rather than adding new elements to future versions of LandXML, extend existing elements where possible. Where there are multiple ways of representing the same data (using different elements), the chances are increased that a LandXML data set produced by civil design software application will not be suitable for consumption by another application that reads LandXML files. For example, one application that imports/exports alignments as CoordGeom elements and another application that imports/exports alignments as series of AlignPI elements will not be able to exchange alignment data even though they are both LandXML-compliant. If the intent of adding the PI-based alignment element is to facilitate reporting, a viable alternative would be to introduce a provision whereby the LandXML 1.0 CoordGeom components of an alignment may be optionally grouped. This grouping element may then carry optional attributes such as PI coordinates and stations. The current LandXML 1.1 AlignPIs element handles some common combinations of spiral transitions, but was not designed to model three centered curves and tapered curves. The same goes for cross sections. Currently CrossSectSurfs models both design and existing cross sections. With the introduction of DesignCrossSectSurfs, cross section data may be represented using two different methods, when there is no practical reason to do so. The state attribute may be used to delineate between existing versus proposed cross sections, and it is not uncommon to encounter man-made existing sections such as existing pavement, existing sidewalks etc . . . that require the same flexibility as design sections.
- **Coordinate Encoding**—Implement stronger typing to facilitate validation, change ordering of coordinates to be consistent with the standard XYZ representation, make explicit whether the coordinates are 2D or 3D.
- **Units**—Standardize and consolidate types of units used and improve documentation of which units are expected for the different elements. Rely on style sheets to render data in user preferred units.
- **State Attributes**—Review and revise use of state attributes to eliminate the potential for conflicting values between container elements and their members.
- **Stations**—Improve documentation for use of station attributes to distinguish between internal stations (beginning station plus cumulative length) and user station values.

Consider adding qualifiers to distinguish the type of station value provided.

- **Naming Conventions**—Improve consistency in naming conventions for data types and code list elements.

Resource Documents

- Crews, N., E. Hall, and D. Rebolj, LandXML Schema Version 1.0 Reference, 2002;
- Crews, N., LandXML-1.0.xsd; and
- Crews, N., LandXML-1.1.xsd.

Design Project Schema

Description

The aecXML Infrastructure schema provided the means for roadway designers to obtain information about possible pay items available for use on a design project, and allowed them to specify which of these will actually be used on the project. The TransXML Design Project schema builds on this and adds pay item quantity and cost information. This extends the value of the schema, allowing for transfer of data to the following:

- **The Estimator**—Who will use it to determine the estimated cost of the project,
- **The Plans Developers**—For preparation of the contract documents, and
- **The Contract Administrator**—For preparation of the bid proposal.

The Design Project schema was developed in coordination with the Bid Package schema so that information created in the design phase can be augmented (not recreated) for use in the construction phase—both for electronic bidding, and for construction inspection and tracking. A separate reference schema was created for master lists of pay items and funding sources to provide a common link across the Design Project and Bid Package schemas.

The Design Project schema includes information about pay items (ID, type, description, units of measure) as well as a list of contract pay items and their respective funding sources on a given project. It also supports the concept of multiple design alternates for different aspects of a given project.

The schema is designed to support evolution of cost information from the initial estimate through multiple iterations of the design, and on to bidding and letting. Various costs are included, including the designer's initial estimated cost on pay items for an individual project, an estimators estimated cost per item on the project and for a proposal which may include several projects, each contractor's pay item bid price, the awarded contractor's price, and any price adjustments made by change orders after the project is let.

Finally, the schema also allows for inclusion of location referencing for the design project. A shared TransXML linear referencing schema is used for this purpose.

A sample application of the Design Project schema was developed that demonstrates:

- Retrieval of a list of master pay items (display and query from an instance document for the reference schema with pay items);
- Creation of a subset of pay items from this master list for use in a design project (creation of a new reference instance document with pay items); and
- Creation of a design project from the project-specific pay item list, including quantities and unit prices (creation of a design project instance document).

The application was developed based on the JavaServer Faces technology, using the Sun Java Studio Creator 2 as an IDE. This application runs on any servlet container that implements the JavaServer Pages specification 2.0, such as "The Apache Tomcat 5.5 Servlet/JSP Container."

Resource Documents

- IAI aecXML Domain Committee, aecXML.xsd.

Area Features Schema

Description

The Area Features schema represents information about area features such as environmental areas, soils, wetlands, land use, flood plains, site improvement areas, right-of-way, and cadastral information. Of primary concern to the designer is the location of these areas with respect to the roadway project being designed. This information is typically stored in a GIS but would be helpful if it could be included as a backdrop to a CAD design drawing.

Currently LandXML includes land parcels as the only area features and it approaches these from the perspective of surveying. The Area Features schema was developed to provide a more general purpose area feature capability. It supports the transfer of features information from agencies outside of the design office, such as planning, National Wetlands inventory, counties and municipalities, FEMA, right-of-way, soils, and consultants.

Area features are common to most GIS software and have been standardized by ISO TC211. The TransXML Area Features schema is consistent with TC211, and makes use of the native GML elements for spatial representation of surfaces. The schema supports XML documents containing a single area feature as well as documents containing collections of features (e.g., all of the lakes and ponds within a county).

A sample application (Import TransXML Area Features) was developed to demonstrate use of the Area Features schema. This application allows the user to import and display GIS area features from an XML instance document into a Microstation CAD design drawing. Sample data files are provided representing seeding, erosion, and pond areas. The application was developed on Microstation Version 08.05.01.xx Windows x86 using MDL (Microstation Development Language).

Resource Documents

- Crews, N., E. Hall, and D. Rebolj, LandXML Schema Version 1.0 Reference, 2002;
- Crews, N., LandXML-1.0.xsd;
- Burggraf, D., LandGML0.6.xsd;
- ISO, ISO/TC 211/WG 4/PT 19136 Geographic Information—Geography Markup Language (GML), ISO CD 19136, February 7, 2004;
- Ron Lake, David S. Burggraf, Milan Trninic, and Laurie Rae, Geography Mark-Up Language, John Wiley & Sons, Inc., San Francisco, California, 2004; and
- ISO 19107, Geographic Information—Spatial Schema, International Organization for Standardization, Geneva, 2001.

Bid Package Schema

Description

In preparation for letting a construction contract, transportation agencies publish a proposal bid package. Contractors use this information to prepare their bids. Subcontractors and suppliers use this information to identify potential business opportunities and submit quotes to primary contractors. In the event that project changes occur after publication but prior to the letting, the agency publishes these changes in proposal amendments.

The bid package includes general proposal and pay item information. Proposal amendments can include changes to any element of this information. The Bid Package schema builds upon the Design Project schema to support the letting process requirements for publishing bid packages. The Infrastructure Project schema is augmented with additional proposal information including the letting location and date, vendor qualification requirements, contract time information, and the additional attributes required for amendments. Proposal milestones are also included to support Cost Plus Time bidding and other situations where contractors bid on time to complete milestones.

Thousands of transportation proposal bid packages are published each year, often in paper form, to a community of tens of thousands of contractors, subcontractors, and suppliers. The standard transportation Bid Package XML schema will enable agencies to publish bid packages electronically in a standard form that can be directly loaded into bid preparation systems. As a result, information flows will be streamlined, and redundant data entry and the associated opportunity for error will be substantially reduced.

The TransXML Bid Package schema package incorporates the TransXML Design Project, Reference, and Linear Referencing schemas.

A sample application was developed using XSLT to create an HTML Bid Package report from an XML instance document based on the TransXML Bid Package XML Schema.

Resource Documents

- IAI aecXML Domain Committee, aecXML.xsd and aecXML_infra_v33.xsd; and
- AASHTO, AASHTOWare Trns•port product documentation.

Construction Progress Schema

Description

At the project construction site, inspectors record daily construction progress on pay items. This information is gathered by project inspectors and communicated to the project engineer. The project engineer prepares a progress estimate using this information and submits it to the central office to trigger a progress payment to the contractor.

The information being exchanged includes pay item descriptive information, and partial quantities placed and placement locations. The Construction Progress schema builds upon the Design Project schema to enable association of a location with a partial quantity placed.

A broad range of field devices are used to measure construction progress. Various elements of this information are communicated frequently among field, project office, test lab, and central office personnel throughout the construction project. A standard XML schema for this information will enable integration of the diverse data collection and data management systems utilized to track this information, thereby streamlining information flows and reducing the opportunity for error.

The TransXML Construction Progress schema package incorporates the Materials Sampling and Testing schema, the Bid Package schema, and the Reference schema.

An application was created to generate an HTML webpage showing a Daily Construction Diary for a project as described in a TransXML Construction Progress Report XML instance document. This web page can be used to review daily construction progress. The daily construction diary web page is generated by applying an XSLT stylesheet to a TransXML Construction Progress Reporting XML instance document. The colors and styles of the web page are defined using cascading style sheets.

Resource Documents

- IAI aecXML Domain Committee, aecXML.xsd and aecXML_infra_v33.xsd; and
- AASHTO, AASHTOWare Trns•port product documentation.

Materials Sampling and Testing Schema

Description

At the project construction site, inspectors record daily construction progress on pay items. Associated with pay item progress is a parallel tracking of the component materials of the pay item and the extent to which these materials meet the agencies materials testing requirements. This information is gathered by project inspectors and laboratory personnel and communicated to the project engineer. The project engineer prepares a progress estimate using this information and submits it to the central office to trigger a progress payment to the contractor.

The information being exchanged includes materials sampling and testing requirements, materials samples collected, field tests performed, test acceptance methods, and the outcome of those tests. The Installed Quantities and Materials Used and Tested schema builds upon the Design Project schema and encompasses material samples collected, field tests performed, and the outcome of those tests.

A broad range of field devices are used to track material use, sampling, and testing. Various elements of this information are communicated frequently among field, project office, test lab, and central office personnel throughout the construction project. A standard XML schema for this information will enable integration of the diverse data collection and data management systems utilized to track this information, thereby streamlining information flows and reducing the opportunity for error.

The TransXML Materials Sampling and Testing schema package incorporates the TransXML Design Project, Construction Progress, and Reference schemas.

An application was developed using XSLT to produce a webpage showing the sampling and testing activity for a material sample as described in a TransXML Material Sampling and Testing XML instance document. This webpage can be used to review the sampling and testing status and results for a material sample.

Resource Documents

- IAI aecXML Domain Committee, aecXML.xsd and aecXML_infra_v33.xsd; and
- AASHTO, AASHTOWare Trns•port product documentation.

Project Construction Status Schema

Description

Project stakeholders including the general public, elected officials, oversight or regulatory agencies, and other institutions such as utilities and railroads require or can benefit from access to timely information about the status of a transportation construction project. This information is managed within the agency in their construction management system. The information is provided to different stakeholders in different forms.

The information being exchanged includes project description, location, and fiscal, schedule and progress information including milestone dates and those affecting traffic. The Project Construction Status schema builds upon the Design Project schema to support publication of project construction status information. This schema was augmented with additional milestone dates, and fiscal and progress information.

This XML schema will enable the automated publication of transportation project status information in a standard format that can be presented in a variety of forms appropriate for the individual target audiences.

A sample application was developed using an XSLT stylesheet to generate an HTML project construction status web page for the projects in a contract as described in a TransXML Project Construction Status XML instance document. This web page can be published on a public website to provide project status information to interested stakeholders.

Resource Documents

- IAI aecXML Domain Committee, aecXML.xsd and aecXML_infra_v33.xsd; and
- AASHTO, AASHTOWare Trns•port product documentation.

Bridge Design and Analysis Schema

Description

The purpose of the TransXML bridge design and analysis schema is to allow for transfer of bridge description information across bridge structural analysis packages to allow for comparative analysis of the same bridge using multiple bridge analysis processes.

The AASHTOWare Virtis/Opis bridge domain was used as a starting point for the development of the TransXML Bridge Analysis/Design schema. The AASHTO Virtis/Opis bridge domain provides a comprehensive description for the purpose of analyzing many bridge types. These types include steel plate girder, rolled beam and built-up multi-girder superstructures, reinforced concrete tee-beam and slab superstructures, sawn timber multibeam superstructures and precast, prestressed concrete I-beam and box-beam superstructures. Due to the complexity of this domain and limited resources, the following subsets of the Virtis/Opis structure types were included in this initial bridge structure schema for TransXML:

- Multigirder steel rolled beam and steel plate girder girder-line structures,
- Multigirder prestressed concrete I-beam and box-beam girder-line structures,
- Multigirder reinforced concrete tee-beam and I-beam girder-line structures, and
- Reinforced concrete slab-line structures.

This subset covers roughly 75 percent of bridge structure types—it excludes only timber multigirder, floor systems (girder-floorbeam-stringer), and girder-system definitions.

The Virtis/Opis software provides a reporting tool that produces an XML representation of the bridge domain. Before being implemented into the Virtis/Opis software, the XML structure was reviewed and approved by AASHTO Technical Advisory Group members from nine different states. This XML information is used with a dynamic XSL template generator to produce user-defined reports for viewing using an Internet browser. While the output is in XML format, there is no corresponding XML schema. Based on input from the AASHTO BridgeWare community, a decision was made to constrain the structure of bridge data in developing the TransXML bridge schema so that little or no modifications would be required to make XML output from the current Virtis/Opis report writer validate against the TransXML bridge schema. This was a case in which the desire to have TransXML accepted by a major stakeholder group (AASHTO BridgeWare users) had to be traded off against the objective of a schema development process that responded to a wider spectrum of organizations. Because of the preexisting constraints, the adoption of GML was less rigorous for this schema. A bridge was at least defined as a subtype of GML feature to enable GML-aware software to recognize it and handle it accordingly, for example, display it on a map along with other features.

A sample application, the TransXML Bridge Input Converter, was developed to demonstrate the translation of a Bridge TransXML instance document produced by one piece of bridge analysis software to a format that could be interpreted by another bridge analysis software package. An example is provided with the application that uses a TransXML instance document of a prestressed concrete bridge generated from the AASHTO Virtis database to create an input file that can be processed by the Pennsylvania Department of Transportation's LRFD Prestressed Concrete Girder Design and Rating (PENNDOT PSLRFD) software. While this application demonstrates the conversion for this specific software package, the concept for the conversion is applicable for other software packages that utilize an input file/output file method of operation.

Resource Documents

- AASHTO Virtis/Opis Database and API documentation;
- M. Mlynarski, J. A. Puckett, C. M. Clancy, M. C. Jablin, and P. D. Thompson, *NCHRP Report 485: Bridge Software Validation Guidelines and Examples,* Transportation Research Board, Washington, D.C., 2003; and
- PENNDOT LRFD Software manuals.

Crash Report Schema

Description

Information about highway crashes is recorded in police reports, and then transferred to a variety of other agencies (Registry of Motor Vehicles, DOT, courts, etc.) for processing, archiving, and analysis. In many states, crash reports and crash reporting practices are not uniform across jurisdictions. However, there is some degree of commonality to crash data—crashes involving fatalities are subject to the NHTSA Fatal Accident Reporting System (FARS) requirements; and a minimum set of data elements (the Model Minimum Uniform Crash Criteria or MMUCC) have been developed by NHTSA, FHWA, and the National Association of Governors' Highway Safety Representatives in the interest of achieving greater uniformity in crash data.

The purpose of the Crash Report schema is to provide a standard data exchange format for the information recorded at the time of the crash (not the information which may be linked to the report after the crash). This format can facilitate transfer of crash data from collection systems to systems that validate the data; from validation systems to archival databases; and from archives for reporting, aggregation, and analysis applications.

A UML model for crash records was developed based on the Model Minimum Uniform Crash Criteria (MMUCC), and the FARS reporting requirements. On completion of this UML model, NHTSA released an XML schema for the MMUCC elements that was developed in coordination with GJXDM. Prior to this schema's release, detailed information about its content was not publicly available. In order to avoid duplication of resources (and production of a competing crash records schema), the research team recommended that the NHTSA MMUCC XML schema be adopted by TransXML. The TransXML UML model was provided to NHTSA along with some comments regarding inclusion of FARS elements into the schema. The research team was informed that subsequent versions of the schema would include the FARS elements that could not be translated from existing MMUCC elements.

The research team developed a sample application to demonstrate how multiple sources of crash data using different XML schemas can be combined in a single crash report. This application merges two XML data files—one using the NHTSA MMUCC XML schema, and the other using a different schema such as those that might be used by a state or municipality. XSLT allows the application to link together any crash records schemas. The columns shown in the report can be changed by modifying the XSLT files.

The MMUCC XML schema is available at: http://www.crashdata-xml.us/. This site also includes a demonstration of an Electronic Data Transfer (EDT) gateway server to have Police Accident Reports (PAR's) immediately forwarded to the FARS case management system as soon as they are entered into the state crash records database. NHTSA is funding a pilot program for states wishing to implement this EDT capability.

Resource Documents

- Model Minimum Uniform Crash Criteria Guideline, Second Edition (2003);
- ANSI Standard D20-2003 Data Element Dictionary for Traffic Records Systems, American Association of Motor Vehicle Administrators, April 2003;
- ANSI Standard D16.1-1996 Manual on Classification of Motor Vehicle Traffic Accidents, Sixth Edition;
- 2004 Fatal Accident Reporting System Coding and Validation Manual, NHTSA;
- Transportation Safety Information Management System, Phase I Consolidated Report, prepared by Littleton PRC, June 2001; and
- JXDM-3.0.2.xls (from http://it.ojp.gov/jxdm/3.0.2/).

Highway Information Safety Analysis Schema

Description

The Highway Information Safety Analysis schema describes safety-related highway inventory items that relate to a specific incident location. It provides a set of data elements that can be integrated with crash data in order to identify high-accident locations, analyze the need for engineering countermeasures, or evaluate specific countermeasures proposed for a location. The schema was based primarily on the FHWA's SafetyAnalyst data dictionary. Its design also considered the contents of the preliminary TSIMS data dictionary that defined a preliminary set of highway inventory items that are required for safety analysis.

It was originally intended to build upon either ISO 14825 Geographic Data Files (GDF) or the FGDC Framework Data Content Standard (formerly Geospatial One-Stop) as the infrastructure base. As part of the National Spatial Data Infrastructure (NSDI) initiative, the latter appeared to be more appropriate. However, it is only now becoming stable enough for consideration. What has been provided for HISA appears to align well with the Framework standard. As TransXML evolves beyond Project 20-64, it would be our recommendation to include Framework inside TransXML with appropriate connections to HISA.

A sample application was developed that allows users to search through crash records stored using the NHTSA MMUCC XML schema, and to link these crash records to related highway safety information stored using the HISA schema. Users can search based on four sample data fields, and a small set of data fields are presented for each crash record. The program can easily be extended to search on or display additional data fields.

Any XML file that conforms to the NHTSA MMUCC XML schema can be filtered and viewed by this application. Crash records stored using a different XML schema can be converted to the NHTSA schema using XSLT. XSLT also allows data from multiple XML data files to be combined. In this sample application, Highway Safety Information associated with each crash (stored in a separate XML file) is presented alongside each crash record. While the linkage between these two sample data sets is arbitrary, the application shows how these two schemas can be used together to create a complete view of a crash. An XSLT file is used to customize the formatting of each crash record. Any data item in the NHTSA MMUCC XML schema that can be expressed with XPath can be removed or added as a category in the search filter.

Resource Documents

- Model Minimum Uniform Crash Criteria Guideline, Second Edition (2003);

- ANSI Standard D20-2003 Data Element Dictionary for Traffic Records Systems, American Association of Motor Vehicle Administrators, April 2003;
- 2004 Fatal Accident Reporting System Coding and Validation Manual, NHTSA;
- Transportation Safety Information Management System, Phase I Consolidated Report, prepared by Littleton PRC, June 2001;
- Transportation Safety Information Management System Data Dictionary;
- American National Standard for Information Technology—Geographic Information Framework—Data Content Standards For Transportation Networks: Roads, Information Technology Industry Council;
- Highway Performance Monitoring System Manual; and
- FHWA SafetyAnalyst data dictionary (draft).

SECTION 6

Future Stewardship of TransXML

6.1 Introduction

One of the objectives of NCHRP Project 20-64 was to define an appropriate mechanism for continued maintenance and development of XML schemas for transportation data exchange after the project is completed. This section presents the analysis of options for ongoing stewardship of TransXML. To develop a stewardship approach for TransXML, the research team first looked at existing models for XML stewardship in other domain areas. The team looked at their organization, funding, participants, and activities. Based on the existing models, and the experience gained during NCHRP Project 20-64, the research team developed a proposed mission statement for a future TransXML project, along with a set of goals and functions. Then, four options for future stewardship were identified and evaluated. Finally, a more specific 2-year implementation plan for TransXML was developed.

In the remainder of this section, the name *NCHRP Project 20-64* is used to refer to the initial TransXML effort that is the topic of this final report. The name *TransXML Project* is used to refer to a broader continuing effort that will ideally evolve out of the work from NCHRP 20-64.

6.2 Existing Models for XML Stewardship

The following existing efforts were identified in order to understand the range of existing models for stewardship that could be considered for TransXML:

- **OASIS**—The Organization for the Advancement of Structured Information Standards (OASIS) is a global consortium founded in 1993 (originally as SGML Open) that works toward adoption of e-business standards. OASIS has established, well-defined mechanisms to support open standards development processes, including formation of technical committees, membership and participation requirements, voting procedures, and standards approval processes. Membership in OASIS is required in order to participate in technical committees. Membership fees in OASIS are $250 (for individuals), and between $1,000 and $5,750 for organizations (higher-cost sponsorship memberships are also offered). OASIS membership composition is currently 15 percent government and academic, 51 percent technology providers, and 34 percent users. Current government members are primarily from the legal/justice, homeland security, and defense communities. Examples of currently active OASIS technical committees include: election and voter services, electronic procurement, emergency management, and product life-cycle support.
- **Justice**—The JusticeXML data model was developed jointly by the U.S. Department of Justice (DOJ)—Office of Justice Programs (OJP), and the Global Justice Information Sharing Initiative (Global). The U.S. Attorney General approved establishment of the Global Advisory Committee (GAC) in 2000, to "ensure appropriate input from local, state, tribal, and Federal agencies regarding information sharing and integration within the justice community." In accordance with the Federal Advisory Committee Act (FACA), the GAC's charter is for a 2-year period (the current charter started October 2002). The U.S. DOJ provides all support services for the GAC, and its operating expenses (for the 2-year charter period) were estimated at $2 million. The GAC appoints subject-matter Working Groups to research specific topics and prepare recommendations. The Global Infrastructure/Standards Working Group (ISWG) created the XML Structure Task Force (XSTF) to develop the Global JusticeXML Data Dictionary (GJXDD). This body involved a wide variety of stakeholders. The bulk of the technical development work was done by the Georgia Tech Research Institute (GTRI).
- **LandXML**—LandXML is a nonprofit industry consortium formed in 2000 to support continued improvement to and use of the LandXML schema. There are about 300 members

representing government agencies, software vendors, engineering firms, and academic institutions. There is no membership fee, and LandXML operates on volunteer labor. An employee of Autodesk has been providing significant leadership and support for the effort.

- **GML/OGC**—GML was authored by a private company (Galdos Systems, Inc.). The Open Geospatial Consortium (OGC) currently maintains the specification through an open working group process. Galdos staff lead this working group and continue to make significant technical and marketing contributions. However, the OGC's intellectual property policy requires that all contributors of technical material for adopted specifications provide a royalty-free license to any party that wants to use the specification (both OGC members and nonmembers). OGC is a not-for-profit organization founded in 1994, with the mission of advancing interoperability among IT systems that process geo-referenced information. OGC currently has roughly 250 members, from private industry, government, and academia. It has a staff of 13 and a 14 member board of directors. Membership in OGC ranges from $300 per year for local government staff (nonvoting membership) up to $50,000 per year for Principal Members (who can chair technical committees and receive support services from OGC staff). The OGC operates three programs: (1) the Specification Program, (2) the Interoperability Program, and (3) the Outreach and Community Adoption Program. The Specification Program houses working groups that develop Implementation Specifications (e.g., GML). The Interoperability Program complements the Specification Program by organizing experiments, test beds, and pilot projects (e.g., Geospatial One-Stop, LandGML). The Outreach and Community Adoption Program undertakes activities to support widespread use of OGC specifications (e.g., strategic alliances, conferences, and seminars).
- **aecXML**—The aecXML schema was originally developed by Bentley Systems, who spearheaded the formation of the aecXML industry consortium. Several working groups were formed to focus on topic areas including design, projects, procurement, and catalogs. The International Alliance for Interoperability (IAI) adopted aecXML in 2000. IAI is a program of the National Institute of Building Sciences, a nonprofit organization. The purpose of IAI is to facilitate information exchange within the building industry. IAI has a staff of six, and a five-member management committee, and a 31-member board, consisting of government and private industry organizations. IAI's membership fees range from $1,000 to $10,000. Technical activity is performed in working groups by members on a volunteer basis. Momentum for maintaining and improving the standards depends on the sustained energies of these volunteers.
- **TranXML/LogisticsXML**—TranXML was developed by Transentric, a company specializing in outsourced rail car tracking. The intellectual property for TranXML was acquired in 2002 by the Open Applications Group (OAGi), a nonprofit industry consortium. OAGi was formed in February 1995 by a group of enterprise software vendors (including SAP and PeopleSoft). Its initial work was funded by these founders and focused on specifications for financial transactions and supply chain integration. OAGi's membership currently consists primarily of software vendors and manufacturing companies. A working group within OAGi is now using TranXML as an input to the development of a broader LogisticsXML schema.
- **ITS Standards Development**—While not necessarily related to XML schema, mechanisms for ITS standards development also provide useful models for the TransXML stewardship investigation. Based on the national ITS architecture, several standards development organizations (SDO's) are responsible for developing ITS standards through a consensus process. These SDO's include: the American Association of State Highway and Transportation Officials (AASHTO), the Institute of Transportation Engineers (ITE), the Institute of Electrical and Electronics Engineers (IEEE), the Society of Automotive Engineers (SAE), the American National Standards Institute (ANSI), the National Electronic Manufacturers Association (NEMA), and the American Society for Testing and Materials (ASTM). This work is done in a collaborative fashion. For example, there is an IEEE Incident Management Working Group which has been responsible for the 1512 family of standards. Participants in this group have included FHWA, ITE, AASHTO, several state DOTs, public safety agencies, and ITS industry vendors. Cooperative agreements are also a frequently used mechanism; for example, a joint ITE/AASHTO steering committee was responsible for the development of the Traffic Model Data Dictionary (TMDD) and message sets for external traffic management centers (MS/ETMC2) standards.

6.3 Lessons Learned

It is useful to summarize some key observations related to stewardship based on the experience gained during NCHRP Project 20-64.

Focus on Value Added

It was recognized from the project's inception that simply producing an XML schema will not guarantee that the schema will be adopted or that it will provide value, and substantive stakeholder involvement in schema development is essential to success (where success is defined as a critical mass of adopters resulting in collective time and cost savings). The stewardship function must therefore ensure that resources are being devoted to the most promising schema development opportunities—where value can be clearly demonstrated and widely understood. It must also devote considerable resources to communication—through multiple channels.

Need to Coordinate with Related Standards Efforts

The second observation is that the scope of TransXML touches many areas where schema and standards already exist, and therefore the future TransXML project must work within the environment of existing schema and standards development efforts that are now occurring within various communities:

- **Design/Survey**—LandXML is an established interchange standard, with an established national (and international) stakeholder community. TransXML can participate as a stakeholder in LandXML, but cannot and should not control schema development in this arena.
- **Location Data**—The OGC's GML effort and the Geospatial One-Stop Initiative are defining open standards and XML encodings for geographic information. TransXML should be a consumer of these efforts, and not attempt to duplicate or conflict with the work that is occurring.
- **Safety**—The Global JusticeXML effort is gaining acceptance and visibility, and is being used as the basis for the National Information Exchange Model (NIEM), jointly sponsored by the Departments of Justice and Homeland Security. For the safety business area, NCHRP Project 20-64 adopted the NHTSA MMUCC XML schema based on the GJXDM. Future TransXML efforts in the safety arena will need to stay coordinated with related JusticeXML (GJXDM) developments and NIEM developments.
- **ITS**—ITS standards work is ongoing within numerous standards development organizations (SDO). Some of this work is highly relevant to the longer-term objectives of TransXML. A future TransXML steward must determine its appropriate role with respect to ITS XML schema and its relationship to efforts being coordinated by the ITS Joint Program Office.

Even though the intent is for TransXML to be an umbrella for development of XML schema for transportation applications, it cannot fully control its world. Several of the logical building blocks for transportation XML schema are and will continue to be developed through outside efforts. This implies that part of TransXML's role might be to fill gaps that are not being addressed elsewhere, to serve as a "skunk works" for developing schema and applications that then get turned over to appropriate groups for further development and certification, and to serve as a voice of coordination within the transportation sector and provide a liaison function across different schema development efforts.

Recognize Distinct Communities Within Transportation

The scope of TransXML includes a collection of schema that have distinct (though in some cases overlapping) sets of stakeholder communities. While schema for highway design, bridge design, construction management, and safety may all be relevant to transportation agencies, they are of interest to very different groups/individuals within those agencies, and there are distinct private sector and academic communities across these areas. If and when the scope of TransXML broadens further (as intended), this disparity will increase. Therefore, the nature of TransXML may be more one of a coordinated federation of communities rather than a single cohesive group working towards a specific product. This is an important characteristic to keep in mind in evaluating stewardship options.

Marketing and Communication Are Key

A key finding of NCHRP 20-64 was the need to aggressively address the marketing and promotional aspects of TransXML. The TransXML website presence that was established and maintained in NCHRP 20-64 proved to be reasonably effective in promoting collaboration among willing and interested parties, but recruiting and encouraging the participation of such parties was more difficult than anticipated. A substantial effort (well beyond that which was originally budgeted) was required to raise awareness of the project to a point at which members of the stakeholder community began to participate in the technical aspects of the project, and the resulting participation was generally neither intensive nor maintained over time.

A major reason for this is that there are two different populations of individuals who need to be addressed in outreach and communications: industry and agency management and IT leaders, and technical staff. The former category of individuals has the broader interests of their agencies at heart, and is most influential in decisions about investments in software and information technology. Communications to these individuals must focus on these overall organizational benefits, and drive the commitment of resources and attention to important technical issues. The latter individuals are the ones who are able to deal with XML schema at a detailed technical level. This is a relatively small population; only a limited subset of software and IT professionals in transportation agencies (a) are technically qualified to contribute meaningfully to schema development efforts; (b) have a personal or professional interest in the development of data interchange standards; and (c) have the time and resources to commit their efforts to a project such as TransXML. Communications to these groups must have a very different emphasis.

Role of AASHTO

A final observation is the recognition of the key role of AASHTO in the genesis of TransXML, and the logic of AASHTO continuing to play a key role in this effort. The AASHTOWare Standards and Guidelines Notebook includes an "AASHTOWare XML Implementation and Migration" specification. This specification has the objective of providing a framework to (1) develop XML schema for incorporation into AASHTOWare products and (2) participate in joint development and maintenance of XML schemas with public and private sector partners. It states that XML schema can facilitate data exchange between AASHTOWare products, between AASHTOWare products and third-party products, and between other products utilized by AASHTO member organizations. It also recognizes the ongoing schema development efforts at LandXML, aecXML and OGC, and states that "AASHTOWare task force members, contractors, and AASHTO staff will participate in these consortiums, where appropriate, and support the development and enhancement of these schemas."

6.4 Goals and Mission Statement for the TransXML Project

Goals of the TransXML Project

Based on the above observations, the following goals are proposed for the future TransXML Project.

Establish Needs and Demonstrate Value

NCHRP 20-64 has established a process for developing consistent, coherent XML schemas in predefined business areas. Several schemas have been developed accordingly, and the process is repeatable for future schemas. However, the key to the long-term success of TransXML is ensuring that the schemas are actually **used.** The best way to do this is to focus schema development on areas where they are most needed. Interoperability must start with a practical understanding of what data needs to be shared, and how sharing of these data will benefit the transportation community. The criteria stated in Section 4.1 of this report for what constitutes a good candidate for an XML schema can provide high-level guidance for this. Specific opportunities need to be carefully evaluated to establish a clear vision of where interoperability is most needed and where it could generate the greatest benefits.

Broaden the Business Area Focus

NCHRP Project 20-64 was designed to focus specifically on four business areas within the much broader surface transportation domain. As noted above, some (though not all) of the impetus and support for the project was related to the desire to enhance data interoperability for existing and proposed AASHTOWare products, which correspond to the four selected business areas. The project was designed to identify and fill gaps in these business areas in which schema were required but did not yet exist. This proved to be an effective means for concentrating the initial project efforts on well-defined application areas and user communities.

At the same time, this approach limited the topical breadth of NCHRP 20-64, and concentrated work in technical areas in which there were a limited number of participating vendors. As a result, adoption of the initial XML schemas is expected to be somewhat limited in breadth. While future TransXML efforts should certainly continue to support and enhance the value of AASHTOWare product investment in these four business areas, the TransXML Project should also attempt to

1. Address data exchanges and overlaps with other well-established standardization efforts such as those pursued by FMCSA in the commercial vehicle information systems area, and those pursued in the ITS community. This will help to broaden the scope of TransXML and start the process of building working relationships with these other vital communities.
2. Play a proactive role with new or proposed AASHTOWare efforts, to ensure that data exchange and support for interoperability are fundamental to the initial releases of new AASHTOWare products.
3. Identify and address other major transportation topic areas which are not covered by AASHTOWare or other existing standardization efforts. Many standards bodies, such as OGC, ISO TC211, and ANSI L1 have been focused on foundation issues, and are only now getting involved in transportation-specific applications on top of those foundations. TransXML can provide transportation expertise to support them in these efforts.
4. Promote the importance of data interchange and interoperability to other agencies that sponsor research in software and information technology (including, for example, the NCHRP and SHRP II research programs).

Formally Embrace Coordination with Other XML Schema Efforts

NCHRP Project 20-64 incorporated an effort to identify and evaluate existing XML schema development and other standardization efforts. This was done in part to avoid overlaps with existing work, and in part to identify workable stewardship models for TransXML. The experience of NCHRP Project 20-64 and, in particular, the decisions made in that effort to adopt certain technical components of other schemas—or in one case a complete schema—as part of TransXML, underscore the importance of developing a deeper and continuing relationship with these other efforts.

The TransXML Project should therefore emphasize coordination with other larger-scale, relatively well-established XML Schema standardization efforts in the transportation arena (ITS, commercial vehicles, crash records, etc.) so as to avoid proliferation of incompatible standards. While this sounds straightforward, it represents a substantial practical challenge. Other standards efforts have different goals and objectives and serve different audiences, and often rely on very different technical foundations. For the TransXML Project to influence existing schema standardization efforts, it will need to build up credibility through meaningful participation. In doing so, the project should be able to leverage its adoption of GML as a unifying framework. Participation in these other efforts requires familiarity with the work performed to date and the needs and other characteristics of participants in the relevant standardization community. This, in turn, will require time and financial resources.

Balance Schema Development, Advocacy, and Industry Coordination

The TransXML Project will need to strike a careful balance between different levels of activity: providing a technical foundation and infrastructure (both administrative and collaborative) to support schema development activity; addressing specific data exchanges through schema development; and working with other transportation standards bodies to help establish an overarching architecture for transportation data. The first of these is essential to yield a coherent and coordinated set of work products and to simplify and streamline schema development efforts; the second is essential to provide a continuous demonstration of successful efforts and value being derived from the project effort; the third is in many respects the "heart of the matter" with respect to the ultimate objectives of the effort.

While coordination of existing XML schema development efforts (and, where possible, providing support for these efforts via TransXML standards and infrastructure) is extremely important, it should not dominate the project. The construction and materials business area clearly needs a stewardship body to facilitate development of schemas. The safety business area also needs a stewardship body to coordinate a wide variety of schema, despite the fact that there are important and relatively well-established schema development efforts (specifically, NHTSA's crash records schema) already going on in this business area.

Beyond Data Exchange; Towards Interoperability

NCHRP Project 20-64 emphasized data exchange among applications—defining a common format for data to allow one application to export a data set that can then be imported into a variety of other applications. Current information systems technology is increasingly directed towards service-oriented architectures (SOA), in which a variety of information systems operate in parallel, providing services to each other in an "on call" environment. This approach (often referred to as "web services" in the context of systems that interoperate over the World Wide Web) often leverages XML as a communications medium that allows information systems to operate in tandem, packaging services provided by a number of different systems into applications that meet specific end-user requirements. The TransXML Project should make an effort to promote these potential benefits of XML in addition to the potential data exchange benefits. Examples of potential web services in transportation might include searches for facilities (rest areas, transit stations) by location, queries of traffic conditions on a route, validation of standard data sets (e.g., NBI, HPMS), or price lookups for standard pay items. In each of these examples, the web service represents a modular function that consumes and produces a standardized XML data set, and can be called from a variety of web-based applications.

TransXML Project Mission Statement

The NCHRP 20-64 project team recommends the following Mission Statement for the TransXML Project:

> The TransXML Project promotes data exchange and interoperability of software applications and information systems used by transportation agencies in areas where such interoperability is most needed and will generate the greatest benefits. The Project intends to increase the utility of these systems and enable them to work together more effectively, overcoming existing data communication obstacles, empowering the industry to operate more efficiently and, ultimately, improving transportation in the United States.

6.5 Function and Roles of a TransXML Stewardship Organization

The TransXML stewardship organization should provide the following functions:

- **Continually develop and refine a vision of where interoperability is most needed and where it can generate the greatest benefits.** The steward must not only devise a clear vision of practical, beneficial interoperability, but must then frame that vision in a way that enables stakeholders to perceive the potential gain. The greater the need, the more likely it will be that interested parties will participate in the process. This will be an ongoing effort as needs are likely to change over time.

- **Develop, maintain, and promote a family of XML schemas that target these transportation business areas.** The TransXML project can perform these functions in-house using project staff, or oversee work that might be performed through coordinated voluntary efforts or via paid contractors or subcontractors (academic, nonprofit, or commercial). As the business areas addressed by schema become more diverse, it becomes increasingly unlikely that a single organization will hold the required technical expertise, and increasingly likely that contractors or third parties will need to apply their specialized skills and services. The implication is that the TransXML Project must have the flexibility to arrange for technical work to be performed by others.
- **Participate in and/or endorse other XML schema and standardization efforts that support its mission, advocating on behalf of public sector transportation agencies to encourage those efforts to meet the needs of the TransXML community.** The TransXML Project should establish and invest in maintaining formal relationships with the following XML schema and standardization efforts: LandXML, OGC's GML and CAD-GIS efforts, the FGDC Framework Data Content Standard for NSDI, NHTSA crash records schema, the family of ITS standards efforts being coordinated by the ITS Joint Program Office, and the Volpe Center's efforts to establish XML schema in the commercial vehicle arena on behalf of FMCSA. The TransXML Project must also monitor the evolution of new standardization efforts and make decisions about what position to take with respect to these efforts.
- **Provide industry-wide coordination to ensure coverage across all important segments of the transportation industry.**
- **Establish and promote technical standards to ensure technical compatibility, uniformity, and nonredundancy in the development of XML schema for transportation.** These two oversight functions, taken together, are absolutely crucial for TransXML to achieve its mission and fulfill the longer-term goal of creating a library of consistent, compatible schemas that serve the industry. Gaps in coverage of key transportation business areas will inhibit unified support from agencies and the vendor community; the lack of standards will guarantee confusion and inefficiency as poorly coordinated schema efforts result in conflicts among members of the stakeholder community. These functions must be executed by committed, long-term core TransXML Project staff, and it is a requirement that the project steward have such staff available.
- **Promote the benefits of TransXML to transportation agencies, software and information technology firms, industry research organizations, and the consulting community; champion the adoption and development of TransXML schemas through advocacy, technical assistance, and selective financial support.**
- **Provide communication mechanisms and infrastructure to enable appropriate participation of the various categories of industry representatives in the TransXML schema development process.** For the TransXML Project to succeed, the steward will need to aggressively lead outreach and marketing activities targeted to both the management and technical communities through: information dissemination (including costs of attendance) at relevant industry conferences (e.g., HEEP, AASHTO IS, TRB, ITE, NEMA, ATSIP); preparation and distribution of regular updates on the project through press releases and website updates; proactive electronic communications, including e-mail, newsletters, and participation in relevant news and discussion groups; and preparation and distribution of white papers and other educational materials.

6.6 Criteria for TransXML Stewardship

Seven core criteria were identified to evaluate candidate organizations' capabilities to carry out the functions defined above:

1. **Leadership**—Established credibility and relationships with the TransXML stakeholder communities, degree of synergy between the organization's established mission and the objectives of TransXML, impetus and incentives in place for the organization to serve as a champion, ability to build consensus within the stakeholder community(ies);
2. **Neutrality**—Perceived degree of neutrality across vendors, and perceived degree of commitment to open standards;
3. **Stability/Sustainability**—Availability of a stable source of funds; likelihood that the organization can sustain itself over a 5- to 10-year period;
4. **Agility**—Ability to productively and expediently make decisions and move forward with initiatives within a reasonable timeframe;
5. **Technical Expertise**—Availability of in-house staff with technical expertise needed to provide support (training, respond to questions, correct errors) for the schemas developed under this project and for further schema development and application, and/or ability to access these resources;
6. **Marketing Capability**—Ability to mount an effective marketing effort to promote adoption and use of the schemas—

even the most sound technical efforts do not flourish without a dedicated marketing effort on behalf of the steward organization; and

7. **Administrative Infrastructure**—Availability of administrative staff and IT support resources that can take on additional incremental responsibilities.

6.7 Recommended Model for TransXML Stewardship

Four options were developed for stewardship of TransXML:

Option 1: Manage within the AASHTOWare program. This is the status quo "path of least resistance" approach. The premise is that current AASHTOWare subscribers would be willing to have a portion of their license fees go to support incorporation of XML into the AASHTOWare products, and for continuing to support the collaborative schema development process launched within NCHRP 20-64. It is also possible that a separate TransXML Task Force within the AASHTOWare umbrella could be established to solicit interest in a TransXML subscription. This would support continued schema development to provide interoperability across non-AASHTOWare applications. Further discussions are required to determine whether an appropriate package of incentives could be provided to encourage a sufficient level of interest (since the premise of TransXML is that it is freely distributed, unlike AASHTOWare product licenses).

Option 2: Establish a cooperative agreement including AASHTO, U.S. DOT, and others to support TransXML. This could take the form of a new joint task force (e.g., like the AASHTO/FHWA/TRB Task Force on Asset Management), or depending on interest, it could be a more ambitious new partnership entity with an executive director and a board of directors (e.g., like the National Partnership for Highway Quality, which involves AASHTO, FHWA, APWA, and several industry groups). Member agencies would provide funding and in-kind services. Technical work would be accomplished through contracts and/or volunteer working groups. These working groups could pursue collaborative schema development via established mechanisms provided by OGC, OASIS, and others.

Option 3: U.S. DOT-funded competitively awarded multiyear contract with a university research center, nonprofit, or commercial organization (or team) to house TransXML. This option presumes that one or more agencies within U.S. DOT would provide a multiyear commitment of funding for TransXML (analogous to the U.S. DOJ's commitment to JusticeXML). The contract would be written to include all of the desired functions of TransXML, including stakeholder liaison, schema development, application development, and technical support functions.

Option 4: Establish a new independent nonprofit organization that relies on grants, memberships, and voluntary contributions from the stakeholder community. This is the OGC model. The success of this approach would be very dependent on public agency grants (at least initially), or the development of a business model that provides incentives for membership or revenue-generating products and services.

The research team identified a number of factors or criteria for evaluating the most appropriate stewardship option:

1. The TransXML mission statement should be most closely aligned with the mission statement of the coordinating body.
2. Long-term, sustainable funding is essential for the success of TransXML.
3. The administrative structure and funding mechanism of the coordinating body must lend itself to the activities described in the TransXML mission statement: stakeholder liaison, schema development, application of development, and technical support.
4. The coordinating body must have credibility and perceived neutrality among the stakeholders.
5. TransXML must be able to reach beyond the traditional boundaries of AASHTOWare applications.
6. TransXML must have broad vendor support.
7. Decisive, central leadership is required. The management structure must be efficient and able to make decisions and achieve consensus quickly.

6.8 Work Plan for the TransXML Project

The recommended work plan for the TransXML Project is divided into two phases. The initial phase covers the transition to the new steward organization. The second phase covers the ongoing operation of the project under its new stewardship arrangement.

Transition Phase

There will inevitably be a gap between the end of NCHRP Project 20-64 and the full transition of TransXML to its future home. In order to ensure a smooth transition and maintain momentum, the researchers recommend that a temporary home be established for the project. Specifically, they recommend that AASHTO take on responsibility for the TransXML project's transition period. The Subcommittee on Information Systems could take the lead, with participation from the Special Committee on Joint Development (SCOJD). The objectives of this transition period are (1) to maintain momentum for TransXML implementation and (2) to obtain commitment by a unit within U.S. DOT to take on stewardship of the project.

The transition phase should begin as early as possible and end when a commitment for longer-term project stewardship has been secured. After a total of 12 months have elapsed, if no workable stewardship arrangement has been identified, then the temporary stewards may conclude that the time is not right for the project to continue in its current form.

The following activities are recommended for the transition phase:

1. Appoint an AASHTO staff person to coordinate the effort.
2. Form a task force for the TransXML transition project. This project task force should include some NCHRP 20-64 panel members to provide continuity. Other desirable members include representatives of state DOTs that have expressed an active interest in TransXML (Nebraska, Idaho, Florida, Minnesota, New York), and representatives from one or more of the following units of U.S. DOT that have an interest in transportation data interoperability: the Intelligent Transportation Systems Joint Program Office, the FHWA Office of Asset Management, and the Office of the Chief Information Officer (CIO).
3. Ensure that the members of the task force are familiar with the results of the TransXML project. The NCHRP Project 20-64 Final Report and project briefing can be used as a resource for this.
4. Agree on and implement an action plan for continuity and continued progress during the transition period. This agenda should consider the following elements:
   - **TransXML Website Hosting**—Identify an agency or organization willing to take on temporary hosting of the TransXML project website. This will ideally be either AASHTO or a member of the TransXML Transition task force. Website hosting is not likely to require much effort; the goal would be to maintain the project's presence on the web, including access to the UML models and schema developed as part of the project. The TransXML domain names should be transferred from the NCHRP Project 20-64 contractor to the new host. The new host will be able to obtain all of the website files from the CD-ROM provided to NCHRP at the close of NCHRP Project 20-64.
   - **Conferences**—Ensure that the work conducted as part of TransXML is promoted at key conferences, including international HEEP, the Traffic Records Forum, the annual TRB conference, and the annual AASHTO Information Systems conference. The sample applications developed as part of NCHRP Project 20-64 can be used to communicate what can be done with the new XML schema.
   - **Support for Early Adoption Projects**—Early adoption of TransXML schemas by state DOTs is critical. State DOT implementation of the initial TransXML schemas would demonstrate technical commitment by the parties most affected by the project. This in turn would present a more persuasive case to potential stewardship organizations and funding agencies of TransXML. One of the NCHRP 20-64 panel member states (Florida) has already begun implementation of the TransXML construction and materials schemas. This state DOT should be encouraged to share its experience with peer agencies to establish a base of credibility and support. In addition to offering assistance and encouragement to states taking the initiative to implement TransXML, a pooled fund project (e.g., using State Planning Research (SPR) funds) should be pursued to demonstrate implementation of the TransXML schema, and prepare additional web-accessible demonstrations of how these schemas can be used, and what their benefits are.
   - **Liaison with Other Related XML and Data Standards Efforts**—Establish relationships with external groups in order to ensure that the work done by TransXML is considered (i.e., not duplicated) by similar efforts, and also to stay abreast of developments which may be incorporated into future TransXML schema. These groups include the Open Geospatial Consortium (OGC); the U.S. DOT ITS Standards Program; the Geotechnical Management System (GMS) effort (see FHWA Pooled Fund Project TPF-5 [111]); and the Department of Justice/Department of Homeland Security National Information Exchange Model effort. Other important stakeholder groups within each business area were listed in Section 5 of this report.
5. Pursue discussions with units within U.S. DOT that could potentially take on stewardship responsibility (primary or shared) for TransXML: the office of the CIO, the ITS Joint Program Office, the Volpe Transportation Center, and the FHWA Office of Asset Management. Use the materials provided below (under Ongoing Operation) to structure these discussions and provide a starting point for identification of resource requirements.

Ongoing Operation

Given the proposed mission statement (see Section 6.4) and functions (see Section 6.5) for the TransXML stewardship organization, the following staff roles and activities are recommended for consideration. They provide the basis for establishing a rough estimate of resource requirements for the effort. It should be kept in mind, however, that there are many possible variations on how this effort can be accomplished. Other XML efforts have been able to successfully leverage external resources (most notably voluntary labor).

Project Lead—Provides overall project direction, advocates for continued flow of resources to support the project, manages the contractor, and provides project liaison and communication functions.

Project Manager—Manages and coordinates all work in coordination with the Project Lead. With input from the technical architect and the evangelist (see below) develops a recommended work plan on a semiannual basis. This work plan includes an appropriate mix of promotional/educational activities, liaison activities, new schema development work (both to broaden the base of schemas and to extend existing schemas to better serve user needs), and implementation assistance work.

Technical Architect—Provides a technical focal point for schema development and liaison with other XML and standards efforts. Works to ensure reuse of existing TransXML components. Continues to extend the base of shared components that span all TransXML schemas. Works with external standards efforts to minimize duplication of effort and to foster consistency in definition of data concepts. Identifies externally developed schema for incorporation into the TransXML umbrella. Manages modifications to existing schema with consideration of backward compatibility issues. Coordinates with the Open Geospatial Consortium on GML issues and evolution. Provides technical assistance to TransXML schema developers to ensure consistency and compatibility across new schema efforts. Maintains a consolidated set of UML models for TransXML schema.

Programmer/Analyst—Supports the technical architect in maintenance and development of the TransXML code base, consisting of UML models, schema (.XSD), sample instance files (.XML), and sample applications. Develops utilities as needed for XML validation and sample applications. Maintains the TransXML website.

Evangelist—Serves as primary liaison with various stakeholder communities, including AASHTO committees, state DOTs, MPOs, and software vendors. Manages stakeholder forums on the TransXML website. Develops and delivers conference presentations on TransXML schema, applications, and value added. Identifies opportunities for collaboration with external schema development efforts. Promotes and provides assistance for TransXML implementation. Develops and disseminates case studies of TransXML implementation. Identifies needs for new schema or schema extensions based on stakeholder input.

Administrative Assistant—Provides clerical and production support to the project.

Table 4 presents a rough estimate for the minimum level of funding necessary to support the above positions. This estimate does not include the Project Manager position, and it assumes that all of the other positions are full-time and provided by a contractor. An additional $150,000 is added to fund a grant program that could be administered by the Project Manager in coordination with the Contractor. The idea behind this grant program would be to provide seed funds for projects to develop additional sample applications, and import/export routines for existing systems in order to help implement existing TransXML schema.

While it is certainly possible to define a “minimalist” approach for TransXML that would cost less, it is the research team’s view that to credibly address the mission statement that has been defined above, there needs to be a critical mass of visibility and effort and the resources to support that.

A scaled-down version of a continued TransXML project would need to operate under a more modest mission statement that emphasized either maintenance and support for the existing schemas or a single coordinator/architect position combined with matching or grant funds to continue schema development in specific areas.

Table 4. TransXML contract budget estimate.

Budget Item	Cost Estimate (Annual)
TransXML Project Manager	$200,000
TransXML Architect	$200,000
Evangelist	$120,000
Programmer/Analyst	$100,000
Administrative Assistant	$50,000
Direct Costs (including travel for 5-10 conferences per year)	$50,000
Demonstration Project Funding	$150,000
Total	**$870,000**

SECTION 7

References

1. El-Diraby, T. A., *A Semantic Architecture for the Management of Highway Construction Supply Chains,* November 15, 2003, TRB 2004 Annual Meeting CD-ROM.
2. Electronic business XML initiative, established by the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) and the Organization for the Advancement of Structured Information Standards (OASIS). See: http://xml.coverpages.org/ebXML.html.
3. National ITS Architecture Executive Summary, prepared by the Architecture Development Team for Federal Highway Administration, October 2003. AASHTO Highway Subcommittee on Maintenance, Resolution to Define and Adopt Common Maintenance Performance Measures, Resolution 01-07 (July 19, 2001).
4. Daiheng Ni and John D. Leonard II, "Development of TrafficXML: A Prototype XML for Traffic Simulation," *Transportation Research Record 1879,* TRB, Washington, DC, 2004.
5. Litton PRC, *Transportation Safety Information Management System Phase I Consolidated Report,* prepared for the American Association of State Highway and Transportation Officials, June 2001.

Appendices A Through F

The researchers submitted Appendices A through F on a companion CD-ROM that contains a variety of supplemental information and digital data files. The CD-ROM contents are available for downloading at http://trb.org/news/blurb_detail.asp?id=7338. The list below describes each Appendix to this report submitted on the CD-ROM.

- Appendix A: Detailed Review of XML Schema and Data Standards
- Appendix B: GML Experiment Summary Report
- Appendix C: UML Models
- Appendix D: UML Models (in native XMI format)
- Appendix E: XML Schema files (in native XSD format)
- Appendix F: Sample applications, source code, and data

APPENDIX G

Additional Feedback on Schemas and UML Models

Comments from the TransXML community were sought on the UML models through August 30, 2005, and on the TransXML schemas through December 31, 2005. Comments received beyond those dates could not be incorporated into schema and sample applications developed in the course of this project. This additional feedback has been collected in this appendix so that it can be addressed in future TransXML development.

G.1 Safety Business Area

Additional Comments from Reviewer #1

The existing schema need to be supplemented in the area of site information. There is little allowance for information that is needed for accident reconstruction and safety analysis. The length, location, and orientation of skid marks are crucial to estimating the speed of a vehicle and the driver's actions at the time of the accident. The type of any barrier that was struck is essential to studies aimed at evaluating the performance of those barriers. While accident reports may be available, a mechanism is needed to sort to the ones that involve the specific barrier system of concern.

There is still a lot to be studied in regard to acceptable widths of clear zones and clear areas. Accident data is key to those studies. The schema needs to allow the lateral travel of errant vehicles to be documented, as well as the sideslopes of the clear area. Quality of the clear areas is also an open question, as higher center of gravity vehicles are less likely to perform well on sideslopes that have been considered acceptable for traditional personal vehicles. The schema needs to allow recording of both sideslopes and vehicle rollover information.

The amount of damage to a barrier system is a key piece of the accident record and also valuable to the maintenance effort to repair that guide rail.

The safety schema should include these and other data about the physical condition of the roadside and the evidence of the crash dynamics.

The TransXML effort should develop draft schema additions for these needs rather than just passing on the suggestion to NHTSA.

Additional Comments from Reviewer #2

The comments below refer to the UML models documented in the section labeled "Highway Information Safety Analysis Package" which can be found on pages 73 through 100 of the UML Model Appendix to this report.

- Page 83, 7.2.23: urbanRural (Area Type) could be expanded to include suburban

With respect to Intersections:

- The actual geometry of an intersection cannot be determined/reconstructed from the data elements in the schema (e.g., skew is not represented).
- The "Intersection" is a sub-element of "RoadLocation" and has two attributes named "locationReferenceMinor Road" and the "minorRoadName." This implies that the road that is identified by "RoadLocation" is the major road for all of its intersections, which is not necessarily true. The model cannot represent situations in which the roadway under analysis is the minor road at an intersection. This could be acceptable if an assumption is added to the model that an intersection is counted as a sub-element for a roadway only if the roadway is the major road.
- "Corner" data are not in the model.
- "Turn Speeds" are not in the model.

With respect to Roadway Segments:

- The following elements used in IHSDM are not explicitly modeled:
  - Roadside hazard rating;
  - Ditches (defined implicitly via the cross-section elements);
  - Obstruction offset;
  - Shoulder width and slope;
  - Curve widening;
  - Bridge presence/width;
  - Speeds: design and 85th percentile; and
  - Percent RVs.
- "Average lane width" is provided (which includes the auxiliary lane width), but the width of individual lanes cannot be defined.
- Turn lanes are not explicitly defined, except for TWLTLs. Under the Auxiliary Lane Type code list, there are "accelerationLane" and "decelerationLane"—are these meant to represent turn lanes?
- Driveway Density is provided, but the locations of individual drives cannot be specified (which might be needed for future IHSDM/HSM models).
- It is unclear whether more than one auxiliary lane can be modeled per direction. Also the relative placement of the auxiliary lane with respect to the thru lanes cannot be modeled.
- Only one shoulder type per direction can be modeled. Also, the "Composite" shoulder type that is included in the schema was eliminated from the IHSDM roadway model.
- The "Bikeway" attribute for each direction of a Road Segment seems to duplicate the "bicycleLane" attribute in "Auxiliary Lane Type."
- Curvature, superelevation, and grade data duplicate elements in the Geometric Roadway Design section.

Additional Comments from Reviewer #3

The developed Safety Schema do not cover crash dynamics or roadside geometry concerns in any place close to the level of detail that I was hoping for. My major motivation far participating was to ensure that the developed schema adequately addressed sideslopes, backslopes, guide rail types, terminal and attenuator types, etc. That did not happen.

G.2 Survey/Design Business Area

Additional Comments from Reviewer #1

The comments below refer to the UML models documented in the section labeled "Geometric Roadway Design, 2nd Draft," which can be found on pages 172 through 254 of the UML Model Appendix to this report. The page number references shown below refer to the numbering scheme used within that section of the appendix, not to the page numbers of the appendix itself.

- Intersections are not modeled (but intersections are covered in the "Highway Information Safety Analysis (HISA) Package").
- Page 38, "Line," "dir[0. .1]" attribute: How is the direction of the line defined (e.g., using N/S/E/W, azimuth, etc.)?
- Page 41, "Spiral," "Recommendation," line 7: Should the references to "begin length" and "end length" of a spiral instead be to "begin radius" and "end radius?"
- Page 55, Superelevation>Carriage Way>Lane: Lane width is entered indirectly via offsets, but there does not appear to be a way to identify lane type (thru, turn, climbing, passing, etc.).
- Page 55, The proposed Superelevation model cannot model a break in cross-slope within a lane. (Not sure if this is important.)
- Page 56, "Superelevation": For "standard AASHTO" transitions, it appears that the "beginStation" and "endStation" attributes are redundant with the Critical Transition stations related to Transition Types "entryNormalCrown" and "exitNormalCrown."
- Page 59, "Transition Type": The "specialTransition" attribute is defined as "Any special transition location." It is unclear whether the type (e.g., beginning of alignment) is to be specified for the transition location, or just labeled as "special transition" regardless of type?
- Page 64, "CrossSect": What does the "name" of a cross section refer to?
- Page 69, "DesignCrossSectSurf": What does "typical width" refer to?
- Taper locations (e.g., begin/end of turn-lane taper) are not explicitly modeled. However, they could be modeled using cross-sections, if the "critical" points are captured.
- Shoulder width and slope are not modeled. Only one shoulder section can be modeled per side.

Acronyms

AAMVA	American Association of Motor Vehicle Administrators
AASHTO	American Association of State Highway and Transportation Officials
AASHTO IS	AASHTO Information Systems Subcommittee
AEC	Architecture, Engineering, and Construction
AISC	American Institute of Steel Construction
AISI	American Iron and Steel Institute
ANSI	American National Standards Institute
APWA	American Public Works Association
ASTM	American Society for Testing and Materials
ATIS	Advanced Traveler Information Systems
ATMS	Advanced Traffic Management Systems
ATSIP	Association of Traffic Safety Information Professionals
BDA	Bridge Design and Analysis
BRASS	Bridge Rating and Analysis of Structural Systems
CAD	Computer-Aided Design
CD	Committee Draft
CEN	European Committee for Standardization
CES	Cost Estimation System
CORBA	Common Object Request Broker Architecture
COSMOS	Consortium of Organizations for Strong Motion Observation Systems
CRML	Crash Records Markup Language
CVISN	Commercial Vehicle Information Systems and Networks
DIS	Draft International Standard
DOJ OJP	U.S. Department of Justice Office of Justice Programs
DOM	Document Object Model
DOT	Department of Transportation
EAS-E	Engineering and Surveying-Exchange
EDT	Electronic Data Transfer
ER	Entity-Relationship
FACA	Federal Advisory Committee
FAQ	Frequently Asked Questions
FARS	Fatal Accident Reporting System
FGDC	Federal Geographic Data Committee
FHWA	Federal Highway Administration
FIPS	Federal Information Processing Standards
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
GDF	Geographic Data Files
GIS	Geographic Information System
GJXDM	GLOBAL JusticeXML Data Model
GJXDD	GLOBAL JusticeXML Data Dictionary
GLOBAL	Global Justice Information Sharing Initiative
GML	Geographic Markup Language
GOS	Geospatial One-Stop
GPS	Global Positioning System
GRD	Geometric Roadway Design
GTRI	Georgia Tech Research Institute
HEEP	Highway Engineering Exchange Program
HISA	Highway Information Safety Analysis
HPMS	Highway Performance Monitoring System
HTML	Hypertext Markup Language
IAI	International Alliance for Interoperability
IEEE	Institute of Electrical and Electronics Engineers, Inc.

IHSDM	Interactive Highway Safety Design Model
INCITS	International Committee for Information Technology Standards
ISO	International Organization for Standardization
ITE	Institute of Transportation Engineers
ITS	Intelligent Transportation Systems
LAS	Letting and Award System
LRFD	Load and Resistance Factor Design
LRM	Linear Referencing Method
LRMS	Location Referencing Message Specification
MMUCC	Model Minimum Uniform Crash Criteria
MS/ETMC2	Message Sets for External Traffic Management Centers
NBI	National Bridge Inventory
NEMA	National Electrical Manufacturers Association
NGSIMS	Next Generation Simulation Models
NHTSA	National Highway Traffic Safety Administration
NSDI	National Spatial Data Infrastructure
NTCIP	National Transportation Communications for ITS Protocol
OAGi	Open Applications Group
OASIS	Organization for the Advancement of Structured Information Standards
OGC	Open Geospatial Consortium, Incorporated
OJT	On-the-Job Training
PAR	Police Accident Report
PDF	Portable Document Format
PDI	Pontis Data Interchange
PES	Proposal and Estimates System
PSLRFD	Prestressed LRFD Bridge Design and Rating Program
RDF	Resource Description Framework
RDL	Report Definition Language
SAE	Society for Automotive Engineers
SDMS	Survey Data Management System
SDO	Standards Development Organization
SFXML	Simple Feature XML
SQL	Structured Query Language
STIP	State Transportation Improvement Program
STLRFD	PennDOT Steel LRFD Bridge Design and Rating Program
TAG	Technical Advisory Group
TC	Technical Committee
TCIP	Transit Communications Interface Profile
TraCS	Traffic and Criminal Software Package
TMDD	Traffic Management Data Dictionary
TSDD	Traffic Software Data Dictionary
TSIMS	Transportation Safety Information Management System
TSOM	Traffic Software Object Model
UFTRC	University of Florida Transportation Research Center
UML	Unified Modeling Language
UTDF	Universal Traffic Data Format
U.S. DOT	United States Department of Transportation
VEDS	Vehicular Emergency Data Sets
WG	Working Group
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language
XSLT	Extensible Stylesheet Language Transformations
XSTF	XML Structure Task Force

Abbreviations and acronyms used without definitions in TRB publications:

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