AEGIST

Measure once.
Use many times.

Applications of Enterprise GIS for Transportation (AEGIST) Guidebook

U.S. Department of Transportation
Federal Highway Administration

2019

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-HEP-20-014		

4. Title and Subtitle	5. Report Date / 6. Performing Organization Code
Applications of Enterprise GIS for Transportation: Guidance for a National Transportation Framework (AEGIST Guidebook)	5. Report Date: October 2019 6. Performing Organization Code: HEP
7. Author(s) Jack (Al) Butler, Robert Scopatz, Steve Anderson, VHB, Inc. Joseph Hausman, FHWA, Office of Planning	8. Performing Organization Report No.
9. Performing Organization Name and Address Federal Highway Administration Office of Planning, Environment, & Realty 1200 New Jersey Ave, S.E. Washington, DC 20590	10. Work Unit No. 11. Contract or Grant No. Task Order Proposal Request HPPI180003
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Policy and Governmental Affairs 1200 New Jersey Ave, S.E. Washington, DC 20590	13. Type of Report and Period Final Report, March 2018-September 2019 14. Sponsoring Agency Code HPL-1

15. Supplementary Notes

Input and guidance on the development of the Guidebook was provided by a Technical Advisory Group of industry subject matter experts.

16. Abstract

The objective of the AEGIST Guidebook is to help build the National Roadway Base Map by constructing a consistent, enterprise-level deployment of a spatial data infrastructure at State departments of transportation. The AEGIST Guidebook shows how agencies at all levels of government can work together to develop a consistent means of collecting, maintaining, and publishing spatial data. At a detailed, technical scale, the AEGIST Guidebook addresses the specific needs for defined objects and their attributes, including the required accuracy, resolution, and precision. At a higher, organizational scale, the AEGIST Guidebook shows how to construct a more reliable and sustainable spatial data governance and management structure based on civil infrastructure management.

17. Key Words	18. Distribution Statement
Geospatial, GIS, GIS-T, AEGIST, CIM	No restrictions. This document is available to the public through the Federal Highway Administration website.

19. Security Classification. (of this report)	20. Security Classification. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	148	Free download.

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

ACKNOWLEDGMENTS

Input and guidance on the development of this guidebook were provided by a Technical Advisory Group (TAG) of subject matter experts. The Federal Highway Administration expresses appreciation to the following State and local agency representatives for participating in the TAG and acknowledges this guidebook would not be possible without their valuable contributions:

Eric Abrams, Iowa Department of Transportation

Chad Baker, California Department of Transportation

Jeromy Barnes, Alabama Department of Transportation

Jared Causseaux, Florida Department of Transportation

Sean Diehl, Timmons Group, Raleigh, NC

Will Holmes, Kentucky Transportation Cabinet

Erin Lesh, North Carolina Department of Transportation

Jim Mitchell, Louisiana Department of Transportation & Development

Margaret Pridmore, Idaho Transportation Department

Chade Saghir, Southeast Michigan Council of Governments

Derek Troyer, Ohio Department of Transportation

Unless otherwise noted, FHWA is the source for all figures contained in this guidebook.

FOREWORD

The **Applications of Enterprise GIS for Transportation (AEGIST) Guidebook** is the latest in a series of publications designed to help the Federal Highway Administration (FHWA) and the States migrate to the enterprise level for creating, maintaining, and governing data related to roadways and their characteristics, elements, and events.

Its immediate predecessor, the All Road Network of Linear Referenced Data (ARNOLD) Reference Manual, set the direction for adding detail and uniformity to the roadway feature geometry included in transportation datasets. The ARNOLD Reference Manual was endorsed by the Highway Performance Monitoring System (HPMS), which specifies the data States must supply annually to FHWA. The AEGIST Guidebook is intended to help the States implement the guidance in the ARNOLD Reference Manual.

The idea for the AEGIST Guidebook originated with a proposal by James E. (Jim) Mitchell, Ph.D., of the Louisiana Department of Transportation and Development, and Jack A. (Al) Butler, PMP, AICP, of the City of Ocoee, FL, to a group of professionals in a geographic information system (GIS) user community to develop a uniform set of business rules and resulting sample database design. The user group members, who were primarily from State departments of transportation (DOTs), had discovered that disparate business rules among the States were making it very difficult for software vendors to develop a more robust product. The problem was that each DOT had its own set of business rules that imposed differing product requirements: What helped one State could hinder the product's use in another. After working for years to find a way forward, the user group recognized it was in their best interest to develop a shared set of business rules that could serve as a national specification for a GIS for transportation (GIS-T) product to support the data collection and editing process. A significant output of that process is the annual HPMS submission and its mandatory inclusion of ARNOLD-compliant centerline features supporting linear referencing.

In addition, States recognized the need to better support movement toward meeting the Federal requirements for collecting and reporting the fundamental data elements (FDEs) of the Model Inventory of Roadway Elements, Version 2.0 (MIRE 2.0) in a manner consistent with HPMS.[1] The FDE represent attributes and usage data associated with roadway segments, intersections, and interchanges. Database objects are needed to attach and organize these data elements—objects that need to be contained within a structure adopted by the States. Any such structure needs to be consistent with the linear referencing framework within which the States currently organize their roadway data.

Parallel with these data content and performance needs were the movement to building information modeling (BIM) for project design and advances in GIS technology and spatial data, and GIS development focused on the States and their needs for data maintenance tools. The integration of civil infrastructure management (CIM), a framework for enterprise-level data governance and management that builds on BIM and GIS, is rapidly advancing. Spatial data collection techniques—along with the volume of data sources they make available—have advanced and combined to make many new work processes possible. Digital orthophotographs and light detection and ranging (LIDAR) data are now mixed with terrestrial laser scanning to support augmented reality platforms that allow users to move seamlessly from looking down at a bridge from an aerial view to exploring its internal structure. It was in this context that FHWA funded the AEGIST Guidebook project.

---

[1] See USDOT FHWA, *Model Inventory of Roadway Elements: MIRE 2*, Report No. FHWA-SA-17-048, July 2017, and 23 CFR § 924.17 – MIRE fundamental data elements.

At its core, the AEGIST Guidebook is a proposed roadway data publication specification for consideration by the States as a means of providing data to a number of user groups and by FHWA as a data reporting format. Significantly, however, the AEGIST Guidebook is not simply a proposal for showing how the States may create and maintain roadway data to be submitted to FHWA in HPMS and MIRE. Among the many conclusions reached by the large group of States participating in the AEGIST Guidebook's development was that adopting enterprise data governance in each State is key to achieving the internal changes needed to adjust each agency's business rules to be consistent with those contained in the AEGIST Guidebook.

One way to do so is to adopt an asset-centric organizational structure for data management. Today, a typical State DOT is organized by workgroups that perform a single mission within the overall agency. Each of these workgroups is responsible for one task and the collection, maintenance, and publication of the data related to that task. Rather than focusing on the individual tasks to be performed by each functional workgroup within a State DOT, AEGIST project participants found that focusing on the assets being constructed, operated, and maintained provided a better framework for managing the data about those assets.

The logical extension of having a shared set of business rules was the realization that it would then be possible for a national roadway dataset to be generated by combining the individual contributions of each State, thereby producing a national roadway base map (NRBM) through a decentralized process. If the States were to go further and adopt the proposed AEGIST publication format as their internal data structure, they would realize numerous internal efficiency benefits. It would also be possible for software vendors to develop products that potentially would appeal to scores of customers, rather than just the handful of States with that particular set of business rules.

Collectively, the many proposals and recommendations contained in the AEGIST Guidebook seek to move States forward on many fronts, both technically and organizationally. Endorsing enterprise data governance and its implementation through enterprise data management is perhaps the most effective way for any State to begin the process of deploying the operational recommendations of the AEGIST Guidebook.

Enterprise data governance identifies the stakeholders who need to make decisions about data: who does what and according to what schedule. Once these managerial decisions have been made and tasks assigned, data management, with its focus on data accessibility, can be implemented. Data management allows each workgroup to be secure in having its data needs met in a timely manner, regardless of which workgroup is assigned the duty to provide that data in accordance with a clear set of expectations.

There are, however, several AEGIST Guidebook recommendations that can be adopted by key individual workgroups that can provide them with substantial benefits. These include changes to the structure and workflow used to compile, edit, and publish data in the many forms expected by existing users. Relatively few people are involved in these activities in each State. Because of its focus on data delivery—a focus expressed succinctly in the project's vision, "Measure once, use many times"—the AEGIST Guidebook demonstrates techniques that will allow these few workgroups to make internal changes that can benefit them and their users without the users themselves being aware of the changes. This is because the data user groups will continue to get the same products they have always received. What will be different is how that data is created and maintained.

It is important to note that the AEGIST Guidebook is one of several parallel efforts being pursued by FHWA. There is also a project being pursued under the auspices of the Federal Geographic Data Committee (FGDC) called the United States Road Specification (USRS). The transportation community has functional responsibility for the Transportation Theme of the national spatial data framework. A working group of multiple Federal agencies is seeking to develop a spatial data standard for how cartographic features related to roadways will be structured. At the present time, they are exploring endorsement of an international standard: ISO 14825, Intelligent Transport Systems – Geographic Data Files (GDF).

While the AEGIST data structure includes cartographic objects intended to store geometric features that represent roadways, it does not actually specify the internal structure of those features. HPMS currently accepts many standard GIS file formats. In the future, it may restrict such submissions to the single format endorsed by the USRS working group and the FGDC. The USRS may also be able to provide some guidance on the taxonomy of roadway features, characteristics, elements, and events needed to further refine the content of an NRBM conforming to any structure.

Publication of the AEGIST Guidebook is simply another step along the evolutionary path of managing the Nation's transportation infrastructure. It is notably one of the few such steps in the spatial data field taken by FHWA in response to a need identified by the States and one of few with such strong involvement by so many States working together to find a way forward. The decisions made represent a shift in perspective. Rather than continuing their disparate paths for data compilation, maintenance, and delivery, the States chose to work together to find a common set of business rules and implementation mechanisms that would provide internal and external benefits. The continuing support provided through the Transportation Pooled Fund Program demonstrates the State and Federal partnership in both testing the proposals contained in the AEGIST Guidebook and in facilitating the various data production mandates.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

1D	one-dimensional space; in the context of this document, each roadway establishes a 1D datum where the only coordinate is a measure (m) that represents a location along the roadway
2D	two-dimensional (planar) space; defined by *x* and *y* coordinates
3D	three-dimensional space; defined by *x*, *y*, and *z* coordinates, also known as physical space
4D	four-dimensional space where time is added to the three physical dimensions
AADT	annual average daily traffic
AASHTO	American Association of State Highway and Transportation Officials
AEGIST	Applications of Enterprise GIS for Transportation
ARNOLD	All Road Network of Linear Referenced Data (note that HPMS has adopted the practice of omitting "Referenced" in the product name)
ATIS	advanced traveler information system
ATMS	advanced traffic management system
BIM	building information modeling
BIMI	BIM for Infrastructure
BTS	Bureau of Transportation Statistics, a part of the Research and Innovative Technology Administration in the Office of the Assistant Secretary of Transportation for Research and Technology
CAD	computer-aided design; formerly known as computer-aided drafting
CDO	Chief Data Officer

CIM	civil infrastructure management; also known as civil/construction/city information model
CIO	Chief Information Officer
CFR	Code of Federal Regulations
CMM	capability maturity model
DOT	Department of Transportation (this is a general term for all State agencies involved in building or maintaining transportation facilities and services)
ERD	entity-relationship diagram
FDE	fundamental data element (a primary subset of MIRE)
FGDC	Federal Geographic Data Committee
FHWA	Federal Highway Administration (part of USDOT)
FIPS	Federal Information Processing Standard
GIS	geographic information system
GIS-T	GIS for transportation, which is distinguished from general purpose GIS by its support of linear referencing
GUID	globally unique identifier
HPMS	Highway Performance Monitoring System, FHWA's premier highway data system
IT	information technology
LPA	local public agency (a general term for all local government agencies involved in building or maintaining transportation facilities and services)
LRM	linear referencing method(s)
LRS	linear referencing system(s)
MIRE	Model Inventory of Roadway Elements
NBI	National Bridge Inventory
NCHRP	National Cooperative Highway Research Program
NGDA	National Geospatial Data Asset
NHPN	National Highway Planning Network
NHS	National Highway System
NRBM	National Roadway Base Map
NSDI	National Spatial Data Infrastructure
NTAD	National Transportation Atlas Database
NG911	Next Generation 911, an address data standard for emergency response; also abbreviated as NextGen911
OGC	Open Geospatial Consortium
RDBMS	relational database management systems
SLD	straight-line diagram
TAG	Technical Advisory Group

TFTN	Transportation for the Nation
TRB	Transportation Research Board
UNETRANS	Unified Network for Transportation
URISA	Urban and Regional Information Systems Association
USDOT	U.S. Department of Transportation
USRS	U.S. Roadway Specification

Executive Summary

The **Applications of Enterprise GIS for Transportation (AEGIST) Guidebook** is the latest in a series of publications designed to help the Federal Highway Administration (FHWA) and the States migrate to the enterprise level for creating, maintaining, and governing data related to roadways and their characteristics, elements, and events. It is based on a consensus best practices approach to the management of technology, data, and transportation system assets. This initial publication is Version 1.0 of what should be followed by a series of revisions reflecting lessons learned from the associated FHWA Transportation Pooled Fund (TPF) Program and other State implementations.

The AEGIST Guidebook contains several core elements:

1. *An asset-centric approach to creating, editing, transferring, and publishing data about transportation facilities and services.* This approach is called civil infrastructure management (CIM). CIM is a framework for data governance and management at the enterprise level. Through CIM, an organization can manage application-essential operational components, such as policies, processes, equipment, data, and human resources. CIM builds on existing, task-focused organizational structures and the concepts of building information modeling (BIM) for Infrastructure (BIMI) to show how to migrate data assets from one workgroup to the next as a transportation asset moves through its life cycle. The AEGIST Guidebook shows how to construct a more reliable and sustainable spatial data governance and management structure based on CIM.
2. *Information on how to implement enterprise data governance.* The project participants recognized that the most organizationally efficient and lowest cost means for the States to provide data in conformance with internal and external needs is to adopt an enterprise perspective and to focus on transportation assets as the organizing element for data governance and its implementation as data management. As a result, they concluded that the most important AEGIST Guidebook content was guidelines for moving to enterprise data governance.
3. *Defined standard terms and business rules.* The project participants worked to define standard terms and business rules that should be adopted by the States to move toward a common foundation for roadway data collection, editing, and publication for use. This work was conducted using several virtual and in-person peer exchanges, along with document draft reviews at critical points in the development process.
4. *A database design that implements the business rules and standard definitions.* Presented as a series of entity-relationship diagrams (ERDs), the database design process is explained in a step-by-step manner. In addition to being a guide for the States in developing their own data collection, management, and publication processes, the database design, in concert with the standard definitions and business rules, is a guide for software developers to create products with a wide market appeal. Although the database design is substantially abstract in its overall structure, many parts have been prototyped or functionally included in existing commercial products. Multiple software vendors were provided copies of draft documents and given an opportunity to send comments and questions to the project team.
5. *A way to implement and integrate existing Federal data standards.* One of the AEGIST project's key tasks was to demonstrate ways the States can implement the requirements of the ARNOLD Reference Manual. ARNOLD is the All Road Network of Linear Referenced Data, the principal guidance for supplying cartographic features depicting public roads to FHWA as part of the

annual Highway Performance Monitoring System (HPMS) submission. Another key document is the Model Inventory of Roadway Elements (MIRE) 2.0, which sets forth the fundamental roadway characteristics and elements that must be considered in conducting highway safety analyses. The proposed database design includes all roadway elements needed for full implementation of MIRE 2.0.

6. *A proposal to create a National Roadway Base Map (NRBM) specification founded on the database design.* The NRBM is to be a roadway map of the United States showing all public and private facilities along with basic descriptive data. Its presentation will be extendable to allow the States to provide access to all their available roadway data. The States are currently required to annually submit a roadway feature base map as part of the HPMS. At the present time, a number of geographic information system (GIS) file formats are supported, and each State is free to adopt whatever data conventions it may desire. The result is that each State's data may be directly examined cartographically, but in isolation.
7. *A proposal for the States and their local government, Tribal, and regional agency partners and contractors to adopt the database design for internal purposes.* This part of the project will be continued in TPF Program. The next chapter in the AEGIST Guidebook, Chapter 6, will be written during the TPF Program. It will describe best practices in implementing the proposed database design. Additional revisions may be made to this original AEGIST Guidebook to reflect lessons learned during the pooled-fund study.

Efforts underway to define a national transportation data standard include the United States Roadway Specification (USRS) being pursued by the USRS Working Group within the Federal Geographic Data Committee (FGDC). (Notably, within the FGDC, the U.S. Department of Commerce, Bureau of the Census, and not FHWA, has lead responsibility for the national roadway dataset.) The primary focus of the USRS Working Group has been the cartographic aspect of the roadway dataset. As of this writing, the USRS Working Group is exploring adoption of an international standard: ISO 14825, Intelligent Transport Systems – Geographic Data Files (GDF).

While the AEGIST data structure includes cartographic objects intended to store geometric features that represent roadways, it does not actually specify the internal structure of those features. This omission leaves regulatory space for any future roadway feature file format requirement that may be adopted by FHWA for HPMS and, by extension, the NRBM.

The NRBM dataset specification contains an attributed roadway map of the United States in two forms: a collection of inventory routes supporting linear referencing and a collection of roadway segments spanning from intersection to intersection for use by local governments and others who do not use linear referencing. The database design accommodates both forms. Each State would continue to be responsible for roadway mapping within its borders, as it is currently. Internally, the States would be free to continue to use whatever database design and work process they currently employ. The proposed data publication process would create a national dataset in a decentralized manner. A cross-boundary matching mechanism is also provided to knit together a national product.

The vision of AEGIST is seamless integration of descriptive and performance transportation asset data—accessible to all user groups via CIM—to lower operational costs within the States.

Just publishing roadway data in a manner conforming to the NRBM specification is a significant step forward, but the AEGIST Guidebook goes further. The objective is to build the NRBM by constructing a consistent, enterprise-level deployment of a spatial data infrastructure at State transportation agencies. Because the States must work with local, Tribal, regional, and other partners to conduct their spatial data operations, the AEGIST Guidebook shows how agencies at all levels of government can work together to develop a consistent means of collecting, maintaining, and publishing spatial data. At a detailed, technical scale, the AEGIST Guidebook addresses the specific needs for defined objects and their attributes, including the required accuracy, resolution, and precision. At a higher, organizational scale, the AEGIST Guidebook shows how to construct a more reliable and sustainable spatial data governance and management structure. That structure is based on CIM.

The CIM process of transportation asset management starts with operational analyses of existing transportation system components to find areas in need of improvement. A key example that informed development of the AEGIST Guidebook is advanced safety analysis of intersections. The identified areas for improvement then go through project development, leading to construction—a series of tasks increasingly aided by BIM technology and virtual reality enhancements. Once opened to traffic, asset management, maintenance, safety data collection, and performance monitoring begin. These are the tasks where GIS technology has dominance. The BIM-to-GIS interface is undergoing rapid development by multiple vendors, with some taking GIS inside structures to give a three-dimensional (3D) view, and others allowing GIS users to visualize structures stored in 3D BIM files.

Technology has pushed forward over the years and enhanced the abilities of many State department of transportation (DOT) functional areas. Paper records kept in filing cabinets became electronic files stored on mainframe computer disks. Hand drafting turned into computer-aided design (CAD) on workstation computers. Manual mapping migrated to computer-based GIS. A second wave followed these initial introductions, where desktop computers replaced mainframes for all but the largest applications. The desktop computers themselves motivated technological changes, such as the migration to GIS as a platform for transportation system analysis and the evolution of CAD to BIM.

Traditionally, each of these migrations began at the workgroup level, as early adopters brought in tools and techniques that helped their workgroup perform. These advances moved through the organization as other workgroups found useful applications of the technology. Recommendations in the AEGIST Guidebook seek to leverage these advances to move from a workgroup focus on technological innovation to an enterprise level of technology management.

The AEGIST Guidebook aims to help States discover how to cut windows into the functional silos that have individually deployed GIS, BIM, and similar tools, so that the whole enterprise may reap the benefits of these achievements. The AEGIST Guidebook provides a technical specification for transportation datasets and a practical guide for organization change through enterprise data governance. Implementing these two components will allow the entire organization to evolve to the next level of operation, which is represented by CIM.

With all the technical and organizational tools now available, the technical obstacles precluding enterprise management have fallen away and enterprise data governance can begin using the asset focus of CIM. As an asset moves along its life cycle, so, too, should the data about that asset move through the functional workgroups of the agency.

The AEGIST Guidebook covers nine main subjects, with a focus on State DOTs:

1. A CIM framework for asset management and related data flows in the States.
2. Data governance policies and practices for spatial data and systems.
3. Linear referencing systems (LRS) and linear referencing methods (LRM) for all public and private roads.
4. Data interoperability for LRS, including roles, responsibilities, and coordination among agencies.
5. Spatial data business rules.
6. Model enterprise implementation plan with a self-assessment tool.
7. ARNOLD-compliant cartographic solutions for creating and maintaining statewide maps of all public and private roads in concert with partner agencies.
8. Temporality of data and its use.
9. Application-specific guidance for affiliated data areas, such as MIRE 2.0 for the traffic safety community.

The AEGIST Business Case

AEGIST builds on a sequence of key decisions to present a set of definitions, business rules, and data elements that can be deployed by all States. Currently, each State has its own set of definitions, business rules, and data elements that have evolved organically within the organization. Thus, the question for each State to answer for itself is, "Why should we adopt the business rules and database design presented in the AEGIST Guidebook?"

At its core, the AEGIST Guidebook is a recipe for enterprise GIS. The usual answer for a GIS-based work process is, "To make better decisions," on the premise that more information will result in different answers to traditional questions. However, that response suggests the best decisions are not being made now. This implied criticism does not make a good first impression. The short answer to the question should be, "To make different decisions." Said another way, it is to answer different questions than have been posed in the past. The traditional questions are being answered by the data currently being collected and presented to users, but this data is generally collected to meet past needs for information. The world is changing. Simply building more roads or widening existing ones will be increasingly difficult options. What are the new options? Right now, all evidence points to a stronger focus on asset management. More efficiency in using the transportation system is needed.

Including "Enterprise" in the name provides the strongest foundation for making the business case for State endorsement of the proposals and recommendations contained in the AEGIST Guidebook. The traditional silos of a State DOT—such as planning, design, production, and maintenance—cannot individually answer enterprise questions, such as the following:

- How can autonomous vehicles be accommodated?
- How can agencies compile and use federally required MIRE data on highway safety?
- To what extent are mobile devices with real-time traffic feeds altering traffic patterns?
- Is there a better way to respond to traffic congestion than by increasing roadway capacity on key routes?
- How can agencies respond to increasing citizen data requests?

The States, along with their local, regional, and Tribal partners, manage transportation assets. Efficient utilization of these assets is a growing need that no single silo (workgroup) can meet. Workgroups involved in policy, planning, safety, pavement management, traffic operations, and bridge management will have increasing decision-participation demand. Strict observance of the State/local jurisdictional boundary for transportation facilities is being replaced by partnerships formed to solve a joint problem. The complexity of many problems often mandates a multi-agency solution. The use of BIM and other information-centric tools in design and construction makes the transfer of data from production to maintenance workgroups even more difficult. Just passing along a set of redlined as-built plans will not get the job done. Data update cycles spanning months or years also will not be acceptable to a public consumer who has gotten used to the instant gratification of the Internet. Citizens expect data to be available anywhere, at any time, on whatever device they are using.

Important use cases for AEGIST can be found throughout transportation. The focus on data-driven decisions in transportation safety relies on the ability to jointly analyze crash data, roadway attributes, and traffic volumes—all of which are included in the enterprise GIS. New questions and applications for roadway inventory data are all external forces for change, but there are also internal forces. At the top of the list is the cost and time required for data collection and inventory updates. AEGIST describes how States can continue to supply data to users in the form they want while simplifying the data editing process and building links to functional area datasets that expand the roadway inventory to serve as a data access point for all users—including those asking new questions. This is where the AEGIST project's vision, *"Measure once, use many times,"* originated. Adopting the AEGIST Guidebook's technical proposals requires changes to the work process of only a few people: the staff and contractors who collect and edit basic roadway data. These are the very people who have been involved in developing the AEGIST Guidebook's proposed methodology and data structure.

The tasks that lie ahead are fairly well defined and fully vetted. What *should* be done will require more internal changes. Primary among these is wider deployment of GIS-based tools. Currently, most GIS users in transportation agencies are those who do mapping, such as in the planning and programming sections. This is like deploying GIS as a substitute for manual map production and is analogous to the initial deployment of CAD to replace hand drafting. However, similar to the way in which CAD became BIM, GIS has become more than just a way to make maps of the highway system. Effectively managing the transportation system is a spatial problem, one that today's GIS platforms are ready to help solve. This is why the AEGIST Guidebook's proposed specification has a national scope but a project-level capability in terms of accuracy, precision, and resolution. The time is right to move spatial platforms—BIM, GIS, and others—to other parts of the organization, to use the analytical and spatial inference opportunities provided by these technologies, to adopt temporal database structures in order to understand and manage an asset during its life cycle, and to let linear referencing serve as the gateway to all of a State's data resources.

The bottom line is that the proposed NRBM specification will not be just another report generated by the States simply to satisfy a Federal requirement. It is the actual data used within the State, regional, Tribal, and local government transportation agencies, and it is structured in a way that reduces the cost of data collection, editing, and publication. It is the foundation for all of an agency's data to be presented to internal and external users, and it is the means by which daunting transportation issues will be addressed by answering the old and new questions being asked.

In addition to the urgency created by the deployment of autonomous vehicles, the visualization tools available to designers, planners, and the public are rapidly moving to 3D, where consistency in data form and content is mandatory. Just as BIM moved CAD from 2D to 3D, the same thing is happening in the GIS world. Working with separate aerial orthophotographs, Google® StreetView™, and BIM files is being replaced with systems that integrate these and other data to produce a 3D view of urban areas. With the integration of data flows between BIM and GIS now underway, it will soon be possible to move seamlessly from the large scales of BIM to the geographic extent of GIS. Transportation data specifications must be adopted now in order to facilitate this technical migration and the organizational benefits that will flow from it.

AEGIST Guidebook Goals

The following provides a summary of the AEGIST Guidebook's four primary goals:

- Facilitate deployment of enterprise data governance.
- Develop a national specification for roadway data structure.
- Advance State capabilities for safety analysis.
- Provide general instruction on data organization.

Facilitate Deployment of Enterprise Data Governance

Establishing data governance in the States is the first goal of the AEGIST Guidebook. Data governance is a formal process of managing data and systems to meet the enterprise's needs for information to support decision-making. In practice, this enterprise approach must also look to the individual business unit's needs in defining requirements at the organization-wide level. It is specifically designed to be a joint exercise of senior-level managers (agency directors, workgroup managers, and the chief information officer), the many business units' subject matter experts (the practitioners), and the information technology (IT) experts supporting the enterprise and the business units.

Data governance is a set of processes for executing and enforcing authority over data. It is a data asset management method designed to deliver required performance of all defined and approved data-related business functions. It seeks a formal understanding of data needs and provides clear direction to the many involved workgroups regarding which group is responsible for meeting those needs by delivering particular data items in a specified manner according to a published time schedule.

Enterprise data management is possible once enterprise data governance has been established. Data management is the process for developing, executing, and overseeing system architectures, maintenance policies and procedures, and data quality improvement. This is where the workgroups establish their data structures and methods for delivering the required data to its users. It is the implementation mechanism that enforces governance agreements regarding who delivers what data in accordance with an established standard and schedule.

Develop a National Specification for Roadway Data Structure

The second AEGIST Guidebook goal is to present a national transportation data structure, starting with a proposed publication specification for roadway data, which can be used to provide data to all transportation data consumers. The proposed structure is provided as a template for the NRBM with the ability to expand it to cover other modes of travel.

It has also been designed, with strong participation by the States through an extensive peer exchange process, to become an internal operating data structure that can natively produce the existing data products and publish each State's contribution to the NRBM dataset. At its heart is linear referencing, which is used as an organizing element for data presented in a manner that simulates travel down a roadway. Linear referencing is a well-known manner of presenting data to internal users at the State and Federal level and is how driving directions and route information are presented to transportation data consumers using the distance between action points or referencing posted milepost values.

The intent is to continue to provide the data products expected by the many task workgroups within an agency, while simultaneously reducing the cost of data maintenance. This mission requirement is met by deconstructing the typical roadway database into its components, which allows the States to recombine them into many forms desired by data users.

The data structure itself, however, is not the big accomplishment of this task. That honor goes to the detailed and extended process of the States, local government representatives, and Federal stakeholders collectively deciding on key business rules and implementation mechanisms that allow a national database to be constructed by the States working in a decentralized manner.

Each State has separately developed business rules, methodologies, and processes over many years. These inconsistent institutional norms have precluded development of a standard software product that can be used by all the States to perform data maintenance. The AEGIST Guidebook encourages adoption of a single data editing process among the States, which would provide GIS-T (GIS for transportation) software vendors with clear direction on how to develop a software platform that can be marketed to all the States.

Advance State Capabilities for Safety Analysis

States rely on the collection and analysis of safety data to detect problems, identify the root causes of operational issues, identify countermeasures, develop solutions, and prioritize projects to mediate operational issues. States also use safety data to evaluate the effectiveness of their program of implemented safety projects. Crash, roadway, and traffic data should be linkable by location on a roadway base map. States should put in place methodologies to ensure that the location of crashes, roadway elements, and traffic data are consistent with the most current roadway base map. Showing how to do so is the third AEGIST Guidebook goal.

MIRE 2.0 and its future revisions specify the roadway elements related to highway safety analyses. States are required by Federal law to collect MIRE’s fundamental data elements. Although the AEGIST Guidebook’s data specification calls for a maximum display scale of 1:5,000, it also recognizes that the needs of project-level analyses, of which safety evaluations are a primary example, must also be met. As a result, the AEGIST Guidebook’s proposals also support larger-scale depictions of intersections and their components that are defined in MIRE. These internal intersection components—turn segments and navigation points—are often too small to be “visible” in the linear referencing system used for roadway data.

Provide Guidelines for the Organization

The fourth AEGIST Guidebook goal is to provide useful information to the States on data compilation, maintenance, and publication methodologies. Development of the NRBM specification motivated several States to create informal conceptual plans for deployment. The information provided herein is directed at a typical State DOT. Of course, there is no “typical State DOT.” Each State has its own existing data management environment and varying levels of technology use and understanding. As a result, the general guidance developed and included in this document will be further refined and published as future updates to the AEGIST Guidebook.

The AEGIST Guidebook shows how States may more efficiently compile, edit, and publish transportation data. The first implementation step is for the States to agree on the roadways, elements, and services that will be included. This agreement is established through adopting the AEGIST Guidebook’s proposed dataset design for roadways. This is one of several thematic layers to be provided as the national inventory of transportation assets. Other top-level groupings include “performance and utilization” and “projects.” Collectively, these would form the main part of the National Transportation Atlas Database (NTAD).

The following is an example of some basic components:

- Inventory
    - Roadways (NRBM)
    - Transit Services
    - Railroads
    - Pipelines
    - Airports
- Performance and Utilization
    - Traffic Counts
    - Crashes
    - Pavement Condition
    - Bridge Condition
- Projects
    - Transportation Improvement Program (TIP)/Statewide Transportation Improvement Program (STIP)
    - Other Federal-aid Projects
    - State and Local Work Programs

The proposed AEGIST Guidebook's data structure provides a place for all these thematic datasets to go. It has demonstrated that the technical issues that have historically precluded enterprise data governance and asset management through CIM have been removed as impediments. What remains to be addressed are organizational impediments. The AEGIST Guidebook presents solid information on how to move a State to enterprise data governance and asset management in a way that lowers the cost of managing transportation assets at all points in their life cycle. It also shows how to undertake this process with minimal impact on data users. It is now up to the States to use these guidelines as they deem appropriate.

Chapter 1. An Introduction to AEGIST

The **Applications of Enterprise GIS for Transportation (AEGIST) Guidebook** offers both a proposed product specification for a national transportation dataset and a recommended operational guide for how the States could publish a national dataset. Doing so is intended to help the States meet Federal data delivery requirements, principally those set by the Highway Performance Monitoring System (HPMS).[2] The proposed product specification and operational guide are two applications for an enterprise approach to spatial data management in a geographic information system (GIS) focused on transportation data (GIS-T).[3]

The product specification is for a national transportation data model and resulting dataset containing the most essential attributes and features of U.S. roadways: the National Roadway Base Map (NRBM). The NRBM should become part of the National Transportation Atlas Database (NTAD).[4] The proposed NRBM specification is also a prototype for a uniform NTAD presentation. It includes the ability to expand into other modes of travel and to support the discovery and transmission of additional roadway data elements.[5] In addition to providing maps and descriptive facts about roadways, the NRBM contains a routable network. The intent of this network is to provide the basic ability to determine whether a path can be constructed from a point of origin to a destination as a means of routing a specified vehicle. This capability is especially targeted at overweight and over-dimension vehicles, which generally require State permits for travel.

Transportation data is, by its nature, spatial; i.e., it includes location references. Contracts, bid documents, and other materials exist in a non-spatial form, but they must be put within their location context for full comprehension. To answer “Which roadway?” and “Where on that roadway?” GIS-T has traditionally limited itself to three dimensions. However, in this document, spatial also refers to the fourth dimension: time. So “When was it done?” and “When will it be done?” are other spatial questions that must be addressed. The specified data structure provides the means for data about past and future conditions to be as readily accessible as data about present conditions.

The AEGIST Guidebook’s operational guide component offers States advice on how they may better organize their spatial data production and publication processes to produce data that conforms to the national specification more cost effectively. This organizational direction will help State departments of transportation (DOTs) more fully realize the benefits of enterprise-level information governance and data management.

---

[2] For information on the history, purpose, and structure of HPMS, go to https://www.fhwa.dot.gov/policyinformation/hpms/2010/ch2.cfm.

[3] Information regarding GIS-T at the USDOT is available at https://www.transportation.gov/gis. The current draft *GIS Strategic Plan for the U.S. Department of Transportation* is available for download from a link at this site.

[4] For more information about the NTAD, go to https://www.bts.gov/geospatial/national-transportation-atlas-database.

[5] The potential uniformity in NTAD data assets primarily comes from the use of a single base map on which to place other data. To the extent possible, the uniformity also comes from internal and external design consistency. As an example of external design consistency, the draft USDOT GIS Strategic Plan proposes the agency explore adopting the National Information Exchange Model (NIEM), “which is being examined for its appropriateness” as a data publication standard. NIEM is an object-oriented data exchange standard using XML as the means of communication. The data structure concepts presented in the AEGIST Guidebook are consistent with those of NIEM and, where possible, NIEM definitions are adopted so the NRBM may be integrated into the NIEM structure at a future date. A less speculative example is offered by the common use of certain terms by this document, the National Bridge Inventory (see, for example, the definitions provided in Report No. FHWA-PD-96-001, *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation’s Bridges*), and MIRE 2.0.

These benefits include:

- A peer-based national data schema and format for publishing essential transportation data for consistent delivery to users.
- Better coordination of activities across all levels of government.
- Improved internal workflows.
- Increased efficiency in building, maintaining, preserving, and evaluating assets.
- Enhanced capability to cost effectively meet the growing need for data to support highway safety analyses.
- Clear lines of responsibility for data collection and publication.
- Reduced cost of meeting local, State, and national transportation data needs.

The AEGIST Guidebook also provides introductory material on topics such as linear referencing for agencies and staff new to the field of transportation data. The States use linear referencing to structure most roadway inventory databases. Linear referencing is a one-dimensional (1D) means of answering the two basic questions posed previously: "Which roadway?" and "Where on that roadway?" It seeks to replicate the real-world experience of someone driving along a roadway. The principles and practices of linear referencing as a data organization and discovery mechanism are described in Chapter 3.

There is a difference between the real world and the abstract one presented by the GIS user interface. GIS uses graphical abstractions that reflect the essence of a real-world entity within a visual presentation in the form of a map. Typically, roadways are shown as lines, intersections as points, and right-of-way parcels as polygons. The power of GIS, however, is not its ability to organize and display spatial data—maps have always done that. It is the ability of GIS to reveal geographic relationships between entities that may not be apparent, to go beyond simply mimicking the actions of a human drawing a map. Still, GIS is just a tool. The context in which the tool is used is what provides the true value of GIS. The AEGIST Guidebook demonstrates how GIS-T data can be organized to facilitate the discovery of spatial relationships, to go beyond just storing and displaying the facts of a highway inventory.

That context is provided by data governance, which establishes the framework for data flows and responsibilities. The primary role of governance is to ensure user access to reliable data suited for the intended purpose. Thus, advice on establishing and maintaining effective data governance is a key component of the AEGIST Guidebook. Data governance sets the policy foundation and establishes the roles and responsibilities of all involved parties. Self-assessment mechanisms, like capability maturity models, enable an organization to judge the sufficiency and resiliency of its data governance.

The AEGIST Guidebook's recommendations represent the evolution of spatial data management toward methods that go beyond the manual processes today's computer-based production mechanisms replaced. The deployment of GIS-T moved roadway map production from manual to computer-based methods. This followed the same process as deploying computer-aided design (CAD), where the new tool was used in the old way to produce paper construction plans. Now, CAD is evolving to building information modeling (BIM), which takes us beyond the simple production of 2D drawings to a 3D world that can feature augmented reality. BIM technology is becoming widely deployed in the States. As its use matures, the next evolutionary steps are being taken to better integrate BIM technology and tools with GIS and asset management as both inputs and outputs to the design-build process.

To do the same with GIS-T requires States to start treating spatial data—whether held in a BIM or GIS-T environment—as a vital asset central to the operation of the agency. Information is the product that flows from one task to the next along the transportation facility life cycle. The fundamental parameters remain: what, where, and when. State DOTs are typically organized around functional workgroups, such as safety, policy, planning, design, construction, maintenance, and operations. Each of these workgroups develops and deploys specialized tools that help them complete their function. BIM and GIS-T are two of these tools, all of which have to grow beyond the current workgroup setting for a true enterprise approach to be taken. Enterprise data governance is the key to allowing this evolution to occur and for allowing agencies to realize the full benefits of their investments in personnel and equipment.

Nature of the Document

The AEGIST Guidebook builds on prior programs and Federal requirements for data production, management, and analysis. In particular, it expands the principles of the All Road Network of Linear Referenced Data (ARNOLD) Reference Manual.[6] The AEGIST Guidebook is instructional and advisory. It offers information regarding the technical and organizational foundation of GIS-T, in addition to showing States how to construct a national transportation dataset. This dataset will initially contain essential information about all public roadways but will be extendable to other modes of travel. The AEGIST Guidebook is also advisory to the HPMS reassessment process[7] and the continuing refinement of MIRE. As a result, it may generate changes in the State-supplied deliverables mandated by HPMS or addressed in MIRE.

A national transportation dataset already partly exists in a different form. This dataset is NTAD, which is available online from the Bureau of Transportation Statistics (BTS) in the U.S. Department of Transportation (USDOT).[8] The NTAD currently includes 44 publicly available roadway databases dealing with such topics as freight movements and motor vehicle crashes. It includes HPMS, the National Highway Planning Network, and the National Highway System. Each of these datasets contains its own roadway base layer. What the proposed new national data specification will supply is the NRBM, which will be a universal roadway framework with essential data elements for the NTAD. The NRBM could be used to display the attributes contained in the existing NTAD roadway datasets, along with the essential roadway data elements supplied by the States. Users would be able to combine the data from one or more layers on a common roadway base map, while data suppliers would no longer need to produce application-specific roadway centerlines.

To help States uniformly produce the NRBM, the AEGIST Guidebook provides business rules and implementation guidelines for enterprise spatial data management at transportation agencies. The objective is to build the NRBM by constructing a consistent, enterprise-level deployment of a spatial data infrastructure at State DOTs. Because the States must work with local government, Tribal, and regional agencies; contractors; and other partners to conduct their spatial data operations, the AEGIST Guidebook shows how agencies at all levels can work together to develop a consistent means of collecting, maintaining, and publishing spatial data. At a detailed scale, the AEGIST Guidebook addresses the specific needs for defined objects and their attributes, including the required accuracy, resolution, and precision. At a higher level, the AEGIST Guidebook shows how to construct a more reliable and sustainable spatial data infrastructure. That infrastructure is civil infrastructure management (CIM).

---

[6] All Public Road Geospatial Representation Study, DOT Contract No. GS-35F-0001P, September 2014.

[7] FHWA – HPMS 9.0: Reassessment and Modernization Study Draft, July 2018.

[8] BTS is part of the Research and Innovative Technology Administration under the Office of the Assistant Secretary of Transportation for Research and Technology. BTS is the agency responsible for supplying the transportation layer of the National Spatial Data Infrastructure (NSDI). Go to https://data-usdot.opendata.arcgis.com/search?tags=Roads for the available NTAD roadway databases.

The AEGIST Guidebook covers nine main subjects, with a focus on State DOTs:

1. A CIM framework for asset management and related data flows in the States.
2. Data governance policies and practices for spatial data and systems.
3. Linear referencing systems (LRS) and linear referencing methods (LRM) for all public and private roads.
4. Data interoperability for LRS, including roles, responsibilities, and coordination among agencies.
5. Spatial data business rules.
6. Model enterprise implementation plan with a self-assessment tool.
7. ARNOLD-compliant cartographic solutions for creating and maintaining statewide maps of all public and private roads in concert with partner agencies.
8. Temporality of data and its use.
9. Application-specific guidelines for affiliated data areas, such as MIRE 2.0 for the traffic safety community.

The following discussion of CIM is not a substitute for information related to BIM, which is a technology-based holistic design methodology. As BIM technology was migrating from the vertical construction world to that of roadway and bridge construction, the acronym was modified from building information model to civil information model. This later transitioned to civil integrated management. FHWA now seeks to use the term BIM for Infrastructure (BIMI).[9] None of these terms or abbreviations embraces the full meaning of CIM as it used by the AEGIST Guidebook.

CIM is the organizing concept for implementing the recommendations contained in the AEGIST Guidebook. It is the conveyor belt that moves a transportation asset through its life cycle, with each workgroup station adding its contribution to the final product—a well-operating transportation system for the public. Adopting CIM represents a fundamental change in how a State looks at its data, which becomes both an asset on its own and the key to understanding and managing transportation assets in the real world. In doing so, the State will enhance its return on investment in BIM, GIS, and other technologies. So, while BIM is not an explicit subject of this document, it benefits greatly from the shared governance structure the AEGIST Guidebook seeks to encourage the States to adopt.

BIM – building information modeling: a technology-based holistic design methodology.

BIMI – BIM for infrastructure: BIM technology from the vertical construction world applied to roadway and bridge construction.

CIM – civil infrastructure management: a framework for data governance and management at the enterprise level. CIM can be used to manage application-essential operational components, such as policies, processes, equipment, data, and human resources. CIM builds on the task-focused organizational structures and concepts of BIM and BIMI.

[9] See, for example, FHWA Project *Identifying Data Frameworks and Governance for Establishing Future CIM Standards*, September 2017, which says, "To be consistent with global terminology, FHWA is transitioning from 'CIM' [civil integrated management] to 'BIM for Infrastructure'." This project is to identify data standards that can facilitate the movement of data from construction to operations and maintenance within an agency's asset management system. The proposed outcome is "to move towards data standards and overall data governance."

AEGIST Guidebook Development

The AEGIST Guidebook was developed through a peer-based process involving face-to-face and online meetings to propose, review, and revise/endorse its content. This interactive process was supplemented by a peer review conducted by a Technical Advisory Group (TAG) composed of key State, metropolitan planning organization, local, and Federal experts experienced in the application of linear referencing, GIS-T, data requirements, and analytical data needs at the Federal, State, and local levels. Several content proposals were substantially based on a thorough literature review. Core concepts and early drafts of key chapters were subjected to discussions during four virtual peer exchanges lasting 4 hours each and three multi-day, face-to-face peer exchanges. In addition, the drafting team conducted workshops at the Transportation Research Board Conference in January 2019 and the American Association of State Highway and Transportation Officials (AASHTO) GIS for Transportation Symposium (GIS-T 2019) in April 2019. These were followed by a final peer exchange held in Columbus, OH, in May 2019. The peer exchanges and literature search, in addition to interviews conducted with State agency personnel, produced a number of conclusions and recommendations for action. Those conclusions and recommendations are presented in the definitions and business rules described in Chapter 3.

It is important to acknowledge that the AEGIST Guidebook did not spring whole from this single effort. It builds on numerous prior attempts to develop a national standard for transportation data, as implied by the numerous footnoted references. These prior efforts notably include National Cooperative Highway Research Program (NCHRP) Project 20-27(2), the National Spatial Data Infrastructure (NSDI) Framework Transportation Feature Identification Standard proposed by the Federal Geographic Data Committee (FGDC), and the Unified Network for Transportation (UNETRANS) developed at the University of California at Santa Barbara through corporate sponsorship. Some States used the outputs of these earlier projects to inform internal efforts to improve their LRM data structure.[10] These major projects and other available documents found through the extensive literature search or previously authored by project team members formed a solid starting point for drafting the initial proposals and concepts.

None of the prior attempts produced the desired outcome: a national standard implemented by the States to produce uniform transportation data that would supply a national roadway inventory. The primary reason appears to be resistance by potential implementing agencies to changing their existing standards and work processes. Why should the AEGIST Guidebook produce a different result? There are two primary reasons. First, times have changed. In particular, States are motivated to adopt shared definitions and data structure in order to entice software vendors to develop the tools they need to create, edit, and publish data in multiple forms for a growing body of users who are utilizing increasingly sophisticated analytical tools. Second, HPMS reassessment and the desire for the HPMS dataset to support more uses while migrating to modern methods of data collection and distribution strongly suggest the AEGIST Guidebook's proposals and recommendations will be considerations in any revisions made to HPMS. The previous attempts to create and deploy a complete roadway data standard produced valuable proposals but lacked any means of mandatory implementation. As a result, adoption of the proposals was voluntary. However, less ambitious efforts, such as migrating the NBI to an XML data exchange standard for structure elements, have been developed successfully and made part of the mandatory reporting process.[11] Key parts of the AEGIST Guidebook could be voluntarily implemented by the States simply for the benefits they offer without regard to their potential for supporting a future HPMS or other Federal data reporting process.

---

[10] Notably, Iowa made changes following the guidance provided by NCHRP 20-27(2), and Virginia moved toward the UNETRANS data model. Many AEGIST data concepts came from the NSDI Transportation Framework effort.

[11] From https://www.fhwa.dot.gov/bridge/nbi/140725.cfm.

There is also the practical matter of needing well-supported software tools to implement any proposed specification. The large variation in business rules among the States has made it difficult for software vendors to develop a single product that could be broadly marketed to them all. As a result, specialized GIS-T software is limited in functionality and variety. A prerequisite for such software is a national market made possible by the States adopting common business rules. To ensure the proposed rules and their implementation are technically and economically feasible, software vendors have been informed participants in the AEGIST Guidebook's development and vetting process.

Support for a national roadway database began with the requirement for States to supply LRM-based data with compatible geometry covering all public roads as part of their HPMS submission. However, the States are able to provide their own version of the LRM-based dataset, not one conforming to a single definition or structure.[12] Many States are finding it difficult to modify their data structures and business practices to produce and maintain this extensive data, and existing GIS-T software lacks the level of sophistication to fully support the work. Adopting the AEGIST Guidebook's proposals both creates a marketable target for software vendors and offers a clear path for States to follow in modifying their internal processes to more easily create the required roadway dataset. In short, there has never been a better time for a proposed national specification such as the one contained in the AEGIST Guidebook to be adopted by the States and used by the public.

Intended Audience

The AEGIST Guidebook's primary audience is State DOTs. It is intended to aid their development of an enterprise GIS-T infrastructure and creation of an NRBM that contains roadway centerlines and essential data elements conforming to a published specification. However, the target "enterprise" is really the entirety of the many levels of government focused on transportation facilities and their use. In other words, a robust State GIS-T infrastructure must provide the means for interacting with all stakeholders, including State, Federal, Tribal, regional, and local agencies. The result is a decentralized form of national transportation data governance achieved through the shared deployment of a common set of business rules and adherence to a transportation data publication standard based on those rules.

Reaching the ultimate goal of building a national multimodal map and the data it supports starts with adoption of shared definitions and business rules. To be an effective management tool, transportation data must be uniformly available across the country at all levels of government. Each workgroup must do more than simply take care of its internal data needs. Data governance best practices clearly demonstrate that data collection, interpretation, and analysis should be built on a management system founded on shared business rules and standards to support the principle of "measure once, use many times."

The AEGIST Guidebook's content is organized to serve four basic audiences:

- Senior managers, who must implement the organizational guidance by establishing the rules of data governance and facilitating data management.
- Middle managers, who will apply the rules of data governance in organizing and directing their technical staff; i.e., leading data management.
- Technical staff, who will implement the technical recommendations and produce the data.
- Software vendors and developers.

The content of the AEGIST Guidebook was either directly developed or extensively vetted by middle managers and technical staff in State, regional, and local governments to ensure it will work. The

[12] HPMS is implementing a quality control process to check for internal consistency within each State's submittal.

inclusion of key FHWA stakeholders was also part of the process of ensuring the agency's needs would be met. It has been demonstrated that the recommendations are technically feasible, but they must also be organizationally acceptable by the senior managers who will have to make the decision to adopt CIM and endorse the recommendations contained herein. The main recommendation and content focused on senior managers is that addressing the need for enterprise data governance. A complete governance structure assigns duties and responsibilities to specific workgroups, sets production schedules, and makes investments in people and technical resources to make it happen. Only senior managers can make the decisions needed to put enterprise data governance in place.

The intent of data governance should be to provide data interoperability. It is not enough for data to be shared. It must be the responsibility of the supplier to provide the information expected by the user accurately, on time, and in the proper form. This level of data management is called data interoperability, and it establishes the obligation of the data supplier to meet the needs of the data user. It will no longer be acceptable for a data supplier to say, "Here, I made this data to meet my needs. Maybe you can use it." Data interoperability means being able to say, "Here, I made this data for both of us to use." This is the highest expression of the AEGIST project's vision, "*Measure once, use many times*." The only way this can happen is for each user's data source to be well defined and reliably available. Reliability is founded on obligation. It has to be part of the data supplier's job to provide accurate and timely information to designated users in the specified form and in accordance with an agreed upon delivery schedule. Enterprise data governance establishes these obligations, and data management provides the mechanism for implementation and enforcement.

Content Organization

The AEGIST Guidebook has a structured presentation of information intended to make it easy for a reader to access the desired content quickly. This intent is expressed using multiple sub-headings and topical presentations.

This introductory chapter, Chapter 1, builds the foundation for each of the following chapters. Chapter 2 provides advice on effective data governance. The policies of data governance, and their implementation as data management, form the skeleton for CIM and enterprise adoption of the AEGIST Guidebook's recommendations. Chapter 2 is the primary chapter directed at senior managers. It presents practical advice on governance and describes a sequence of capability maturity levels that can exist in successful enterprise GIS-T implementations. This sequence assumes a DOT will produce the State's portion of the NRBM and, ideally, adopt enterprise data governance.

The following are key steps:

1. Establish the initial data governance. This step is described separately from the others, but in reality, data governance will be active throughout the implementation and transition. The information in Chapter 2 will explicitly show where, once enterprise GIS-T data management is established, new activities will be required for the data governance group. This is likely to grow over time as stakeholders are identified.[13]

[13] A common decision made when establishing data governance at a State agency is to wait for the system to stabilize before including stakeholders from outside the organization, particularly local governments. The intent is to "get our house in order" before trying to add external participants. While logical, such a decision hurts the long-term success of governance, which is to get all stakeholders together for joint decision-making. Decisions made with only a subset of all involved parties may need to be undone or changed as new stakeholders are added to the governance structure. It is preferable to start with at least some representatives from each type of stakeholder. For a State DOT, this means including Tribal, metropolitan planning organization, and local government representatives from the beginning.

2. Identify the existing data structure(s). This step is required to establish a starting point for data migrations. It can interact well with a data governance effort that would also identify data gaps and develop plans for filling those gaps.
3. Map the existing data structure(s) to the NRBM data structure. This is primarily a step for the IT and GIS-T practitioners to establish links from old to new data structures and plan for the data transfers. A place in the new data structure needs to be established for each piece of information the organization needs, and a production mechanism must be devised for each new piece of information that is required. Since many legacy business systems will be maintained as they are, a continuing extract, transform, and load (ETL) process needs to be established that will handle the flow of data from workgroup to workgroup and from source to application.
4. Transition to the nationally uniform business rules. This step incorporates the business rules into the enterprise GIS-T and agency procedures so that data meet the new specifications and the system functionality is established as required. This step also includes the data governance group.

Chapter 2 concludes with a brief overview of methods and tools that States can use to measure their status and progress toward enterprise GIS-T. The basis of all the methods is the choice of one from among many capability maturity models (CMM) that are either already published or are being developed now. The models discussed herein are directly applicable to enterprise GIS-T. The scoring will be presented as the typical 5-point scale, where 1 is the lowest level and 5 represents a fully mature, institutionalized system or process. These include the full Urban and Regional Information Systems Association (URISA) CMM provided by the GIS Management Institute and its simplified implementation by Slimgim.[14]

Next, in Chapter 3, the AEGIST Guidebook dives into the world of linear referencing and related business rules. The origins and intent of linear referencing and its theoretical foundation are provided to serve as a primer for the new GIS-T data user. More advanced topics, such as the application of linear referencing and common technical issues, are built on this foundation. The chapter includes LRM-related business rules developed through a series of peer exchanges and other stakeholder mechanisms. These numerous business rules establish the operational framework for the other proposals provided in the AEGIST Guidebook and form the most critical content. These business rules are the core elements of nationally consistent enterprise GIS-T data management. By adopting a standard set of business rules, States may collectively produce a consistent roadway dataset that can be readily combined with similar products from other States to produce the NRBM. In addition, it may induce software vendors to enhance existing products and build new ones given the increased marketability of a common rule-based platform.

Chapter 4 presents the specifications and technical foundation for the NRBM product. The business rules of Chapter 3 are brought into reality through the NRBM design elements covered in Chapter 4. Many figures are included in this chapter to explore complex topics in a way that makes them readily understandable to the GIS-T novice. Chapter 4 describes how the NRBM and its implicit database design implement linear referencing while adding temporal support.

Chapter 5 describes ways to implement the NRBM database design by applying well-known principles expressed through entity-relationship diagrams (ERDs). While it is true that all States have a roadway

[14] For more information on the URISA GIS Management Institute and its CMM, go to https://www.urisa.org/main/gis-management-institute. The development of Slimgim-T was sponsored by FHWA and involved four U.S. States and one Canadian province. For more information on Slimgim-T, go to https://www.gis.fhwa.dot.gov/documents/Slimgim-T_GIS_Capability_Maturity_Model.htm.

inventory and mapping system, many are looking to migrate from legacy systems that are less able to meet today's needs. This implementation chapter aids software developers in understanding how their products can meet the proposed defined terms, business rules, and database design.

The data maintenance and publication process will be covered in Chapter 6, when that chapter is written. Chapter 6 will be developed during a Transportation Pooled Fund (TPF) Program as participating States implement the AEGIST proposal. The chapter will describe workflows for creating, editing, and retiring data records of all types. In the spirit of *measure once and use multiple times*, this chapter will demonstrate how the data publication process can generate the data products users want.

A Brief History of GIS-T

The Federal Aid Road Act of 1916 required the States to establish highway departments that were focused on those roadways under State jurisdiction. With the subsequent advent of the Federal-Aid Highway Program, the States were required to expand their scope of interest to all roads eligible for Federal aid. Biennial roadway condition reports were required from the States beginning in 1956.[15] These were replaced by the HPMS in 1978, which has a combination of sample and universe data requirements on the public roadway system.[16] This was the initial environment for deploying computer-based GIS-T when it arrived on the scene in the 1980s. "Manual" GIS-T, though, had been around almost since the beginning of State Highway Agencies. The States have always made maps of their highways. They also quickly adopted the well-established railroad practice of recording the location of assets and characteristics along the highways using linear referencing and straight-line diagrams (SLDs). Paper and pencil was the GIS-T technology of the time. Linear referencing is based on a few simple business rules.

County and State maps were generated for use by the public. Line symbology told the user what kind of surface was present. Highway inventory mapping consisted of drawing SLDs for each roadway, with characteristics and elements placed along the roadway centerline using text and symbols. Positions were stated in terms of a distance in decimal miles from the route's point of origin or from a reference object. The direction of measurement is called the cardinal direction and has traditionally been from south to north and west to east. Each roadway was an entity unto itself within the context of the LRMs used by the States.

To use linear referencing, one must specify the route of interest and the distance from its starting point or an intermediate reference object to find a location along the roadway. Linear referencing provides an almost direct translation from a position in the database to a location in the real world and requires only the technology provided by a vehicle's odometer to implement. Because the measurements are taken in the real world, the results are inherently 3D in that they take into account the effects of vertical curvature.

The application of computer-based GIS-T in its early days was for the 2D cartographic display of data about roadways under State jurisdiction, with extensions for other Federal-aid highways. The use of GIS-T by Federal, State, and local transportation agencies has been underway for many years. During this time, the technology has evolved along with the number of applications to which it has been applied. The problem has been that each agency's use of the technology evolved fairly independently of others, with the result that there is limited ability to share data across agencies or to compile a single dataset from multiple agencies.

---

[15] From https://www.nssga.org/advocacy/grass-roots/reauthorization-roadmap/history-federal-aid-highway-program/.
[16] From https://www.fhwa.dot.gov/policyinformation/hpms.cfm.

Because of the need to support linear referencing, GIS-T has emerged as a specialty area in the general field of geographic mapping and analysis. The ubiquitous nature of roadway mapping—everyone with a smart phone has one—hides the reality that transportation data is often both hard to find and difficult to understand. The proposals in the AEGIST Guidebook use linear referencing to make transportation data easy to find and show how to present the data in multiple forms so that it can be easily understood. It also puts the basic information about roadways in the public domain, normally a very difficult and expensive option for commercial datasets.[17]

The level of abstraction needed to display highway inventory data cartographically was usually satisfied by creating a centerline feature for each roadway. Each State chose how to do so. The result was 52 isolated State highway maps, data models, and datasets.[18] Federal-aid programs eventually led to a standard set of data to be reported to the USDOT as a condition for States to receive funds. HPMS presented the standard in 1978, and it has remained the primary reporting mechanism for information about roadways and their use, performance, and condition. Mandatory reporting by the States of the Fundamental Data Elements in MIRE 2.0 will soon be added to meet the requirements of more recent Federal-aid legislation.

Currently, HPMS requirements include a geospatial route file for all public roads following the guidance in the ARNOLD Reference Manual and the requirements of the HPMS Field Manual. ARNOLD represented a fundamental change in both extent and abstraction. The extent is now all public roads—expanded by the AEGIST Guidebook to include all roads, public and private, in the NRBM. The level of abstraction now requires directional centerlines for divided roadways, plus connectors that represent limited-access highway ramps at grade-separated interchanges, channelized turn movements, and other elements of the roadway system that support pathfinding. Collectively, these new requirements allow ARNOLD to become the implementation mechanism for the Transportation for the Nation (TFTN) Strategic Plan adopted by the USDOT in 2011.[19]

Implicit in this new environment is the requirement for all 50 States, plus the District of Columbia and Puerto Rico, to jointly produce a single, national transportation map and dataset.[20] The ARNOLD Reference Manual was created as a guide to the States for completing this task. Since it was developed, however, the differences in roadway inventory business rules among the States and difficulties with vendor products implementing some functional requirements led to a realization that additional information was needed. Something more fundamental was required to provide a complete data structure for essential transportation asset data, especially highway safety data described in the Model

---

[17] See, for example, the requirements of FAR 52.227-14 for standard Federal contracting language regarding the rights to Federal contract deliverables.

[18] As used herein, the term 'State' includes the District of Columbia and the Commonwealth of Puerto Rico.

[19] Transportation for the Nation (TFTN) was a 2008 initiative of the National States Geographic Information Council (NSGIC) that was formalized in a strategic plan developed by the USDOT in May 2011. The vision is for the Federal Government to coordinate development of a nationwide dataset of addressable roads that is built in a collaborative and shared environment. The idea actually predates TFTN, with an origin in the 1990 OMB Circular A-16 that was subsequently expressed by Executive Order 12906 (1994) as the transportation layer of the National Spatial Data Infrastructure (NSDI). A revision to OMB Circular A-16 in 2002 assigned the responsibility for developing the transportation theme (layer) of the NSDI to the USDOT's Bureau of Transportation Statistics.

[20] Although commercially available transportation system datasets exist, they do not meet the Federal requirement for the information to be held in the public domain. Nor do commercially available products contain the specific data needed for transportation system analysis and management by government agencies.

Inventory of Roadway Elements – MIRE 2.0 (July 2017). The AEGIST Guidebook is intended to meet this need.[21]

Civil Infrastructure Management Sets the New Context

The context for enterprise data governance is CIM. CIM originated with the presentation of infrastructure asset management concepts more than 15 years ago.[22] The concept is known by many names, including enterprise asset management and BIM for Infrastructure or BIMI, and has many variations. In the AEGIST Guidebook, CIM is an asset management system that starts with planning the construction or modification of a transportation asset, moves through design and construction, and, once built, transitions to deployment, operations, maintenance, and performance evaluation. Like cars moving down a factory assembly line, each functional workgroup adds information to the database being assembled about a transportation asset. Each preceding workgroup must have added their pieces of information properly to the asset database in order for the next workgroup to be able to understand what has happened and be in a position to complete its tasks. AEGIST supports all these workgroups by demonstrating best practices for data management, which is the implementation mechanism for data governance. Asset-centric CIM is the overarching organizational mechanism used in AEGIST to provide data management.

The integration framework presented in the AEGIST Guidebook is an extended version of CIM. As with BIM, the initial focus of CIM was project delivery; i.e., facility design and construction. Roadway inventories describe the present state of the transportation system. They are commonly updated as transportation assets are put into use or modifications are completed. CIM embraces the complete facility life cycle, as shown in Figure 1. The purpose of this figure is to both illustrate the life cycle and to show which utilize GIS and BIM. Of course, CIM utilizes many tools besides these. The figure indicates those four tasks where better integration of BIM and GIS is warranted for individual functional workgroups. There is also a clear need for data transfer from BIM to GIS at the end of construction to support facility inventory updates.

Many States still rely on legacy, text-based inventory database systems. Constructing an interface between BIM and these legacy systems may be difficult. GIS was initially deployed as a means of illustrating the content of those databases and is increasingly becoming the native roadway inventory database. Some States are using GIS as an input mechanism to provide environmental and roadway data to the design process, which uses BIM tools. GIS users are also consumers of BIM-derived data. As a means of providing roadway inventory updates when a construction project ends, it would be beneficial for the BIM platform to then send the as-built information to the GIS platform. This ideal data flow is shown in Figure 2, which generalizes the more detailed presentation of Figure 1.

---

[21] The official name for the AEGIST effort is "Standards and Governance for Enterprise Geospatial Systems in Transportation," which is a project being pursued by FHWA's Office of Policy and Government Affairs under TOPR HPPI180003, with funding also being provided by other offices.

[22] See, for example, Cagle, Ron F., *Infrastructure Asset Management: An Emerging Direction*; AACE International Transactions, 2003; available through http://connection.ebscohost.com/c/articles/14699075/infrastructure-asset-management-emerging-direction. See also, Halfawy, Mahmoud, D.S. Pyzoha, and T. El-Hosseiny, *An Integrated Framework for GIS-based Civil Infrastructure Management Systems*, presented at the Annual Conference of the Canadian Society for Civil Engineering in June 2002; available through https://www.researchgate.net/publication/267160088_An_integrated_framework_for_GIS-based_civil_infrastructure_management_systems.

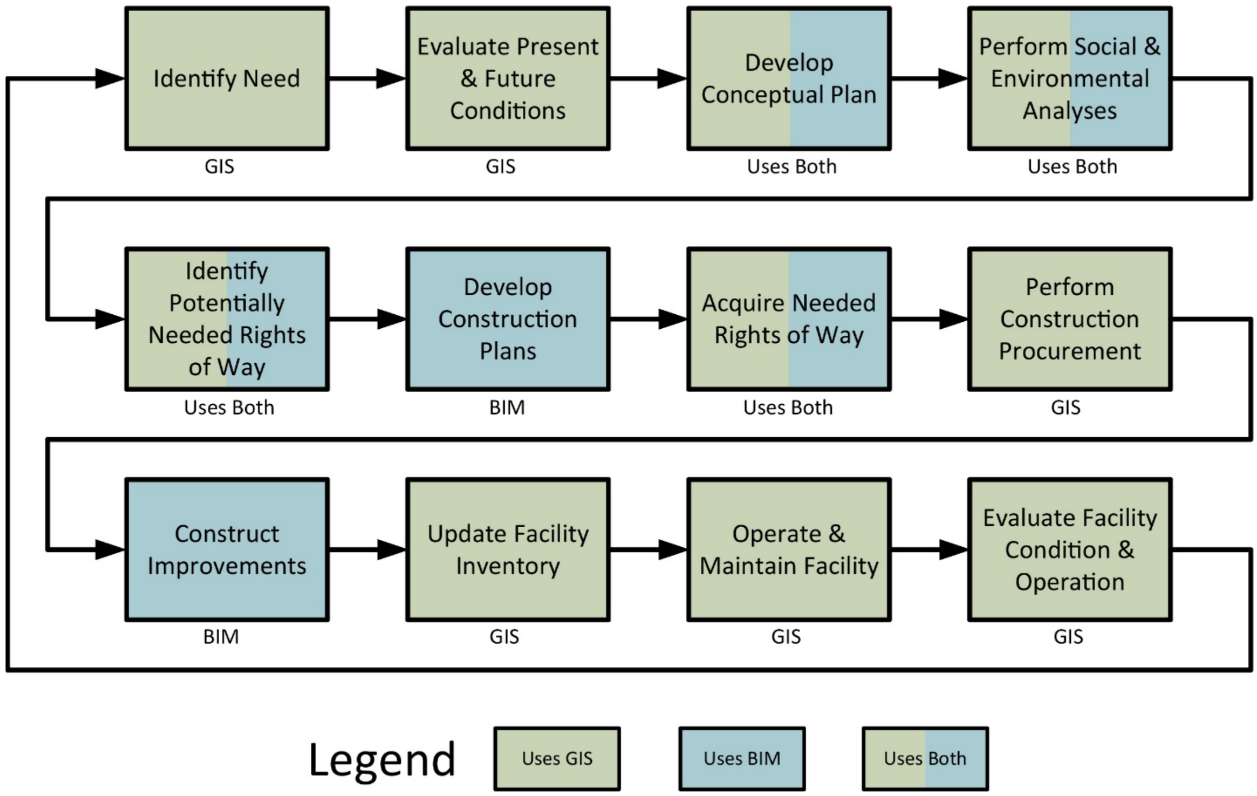


Figure 1. Typical transportation facility life cycle workflow.

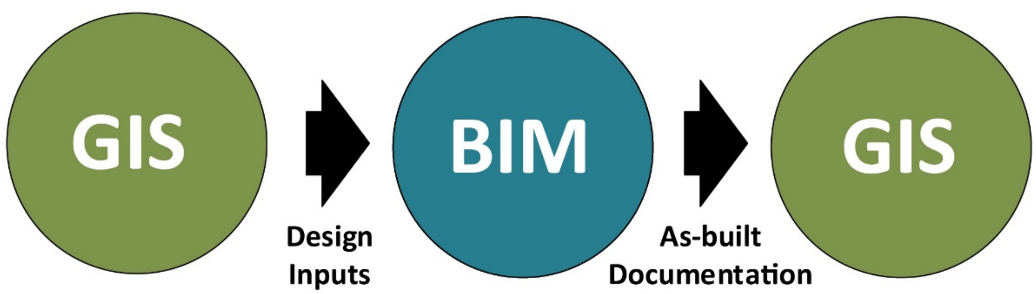


Figure 2. GIS-BIM data flows.

As shown in Figure 3, CIM has the ability to develop an all-encompassing context for everything that a State DOT does by presenting a focus on transportation assets rather than individual job functions. In this way, CIM breaches the functional silos of a typical task-oriented State DOT. CIM can provide an asset-centric data governance and data management environment.

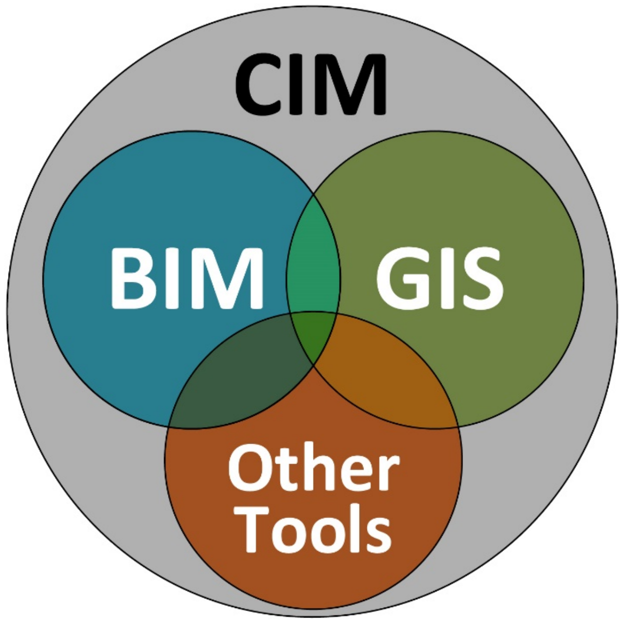


Figure 3. CIM can be a container for other asset production and management tools.

FHWA, AASHTO, and other key stakeholders have agreed to a common definition of CIM: "the technology-enabled collection, organization, managed accessibility, and use of accurate data and information throughout the life cycle of a transportation asset. The concept may be used by all affected parties for a wide range of purposes, including planning, environmental assessment, surveying construction, maintenance, asset management, and risk assessment."[23] This definition does not say that data management stops at the end of construction. It continues through the complete asset life cycle. GIS is a data management tool integrated into classical CIM, and software vendors of BIM and GIS products are already establishing data transfer mechanisms between these platforms. The technological issues have been largely solved. The real obstacle to full deployment of CIM is organizational.

Managing Change through Enterprise Governance

The purpose of the AEGIST Guidebook is to motivate change. Organizations naturally resist change. Linear referencing was deployed many years ago—generations ago in terms of employee career cycles. Many of the work practices in place today are much like verbal histories. "That was the way I was taught to do it," is often the response to the question, "Why do you do it that way?" Institutional memories fade or are lost through retirements of key personnel. Adopting nationally uniform business rules, as proposed by the AEGIST Guidebook, is especially difficult when an organization does not fully understand the reasons behind its existing business practices.

[23] Taken from the executive briefing presentation of key results for *NCHRP 10-96, Guide for Civil Integrated Management (CIM) at Departments of Transportation (DOTs)*.

A good way to get started with enterprise data governance is to identify the underlying business rules upon which the current roadway data collection and publication effort are based. Knowing why the existing process is the way it is will help the organization identify what pieces need to be updated and what changes will impact data collectors, editors, and users. *Knowing where you are will help you get to your destination.*

That proposed destination is a data management system that applies a nationally consistent set of data business rules that result in the production of a nationally consistent data product. The objective is to develop a single, shared specification for transportation data that can be adopted by all Federal, State, local, regional, and Tribal agencies. In other words, the objective is to develop a national transportation dataset through a decentralized production system. A secondary objective is to provide the software industry with a common product specification that should lead to development of new tools designed to implement the technical guidance provided.

This Guidebook demonstrates that adhering to a national specification is consistent with the need for flexibility and support for legacy processes. The guidelines provided herein can be adapted to the States' differing situations and past decisions on some of the core definitions and business rules while still producing a uniform product that can be used to assemble the national transportation dataset. To be sure, the task will be different for each State. Some States have business rules that are fairly well aligned with those proposed herein. Other States will need to make many changes to work practices and data structures to be able to adopt the proposed business rules.

One of the great benefits of adopting CIM is that it can be done incrementally. It is not necessary for it to emerge fully formed in any specific agency. Build the skeleton first and add the flesh later. That skeleton is enterprise data governance. While CIM has many technological components, at its heart, it is a structure to manage organizational workflows. Asset management fundamentally demands data management, which is the implementation mechanism for the policies established through data governance. The enduring question is, "Who did what and when?" The proposed NRBM specification seeks to answer that question, and the related advice shows how a State DOT can most effectively organize itself to deliver that data product while meeting its internal needs for asset management.

States and their partners need to find ways to work together as they adopt the guidelines presented herein. Data governance is the recommended way to develop and enforce data specifications. HPMS, ARNOLD, and MIRE are just three of the many product specifications being imposed on spatial data within the States. Each of the many workgroups or business units found in a typical DOT has its own data specifications. Bridge management staff needs to know about bridges, overhead signs, and other structural elements of the transportation system, with such information reported to FHWA through the National Bridge Inventory (NBI). Pavement management staff needs to know about the concrete or asphalt that form the linear facility itself. Traffic operations staff needs to store data about traffic signals, stop signs, speed limits, school crossings, and other regulatory aspects of the transportation system. Safety management professionals must conduct safety analyses with integrated datasets containing crash, roadway, and traffic information. All of these workgroups and others inside and outside the DOT need information on facility usage, such as traffic counts, and on maintenance work and construction projects that may impact the facility.

It is in this context of multiple—often somewhat isolated—user groups that the problem moves from being a technological one to an organizational one. Historically, it has been common for each interest group to meet its own needs, with data sharing between groups being, at best, an afterthought. This "silo" approach to data management was relatively benign when tabular data prevailed, as each group could maintain their own files. The advent of GIS has coincided with a growing need for data and problem analysis and a rapid increase in the amount and complexity of data being demanded by users.

Data costs money to compile and maintain. Data also has a shelf life and must be updated. Whether using in-house staff or contractors to compile and publish data, a typical State DOT has scores of people involved in the process. Such data collection efforts are typically distributed among the organization's various functional workgroups, each following its specific mandate.

Duplication of effort is the direct result of each workgroup meeting its own needs. It is reasonable to assume that some part of each workgroup dataset could be supplied by another workgroup. Sometimes, data sharing can provide this information through informal channels, but data sharing is inherently unreliable since it is, by definition, informal and not an intrinsic deliverable to the other workgroup. Data interoperability is the most efficient means of distributing the responsibility for data production and delivery across an enterprise, thereby reducing data duplication. To construct an enterprise dataset, an organization's management must span workgroups to coordinate their efforts. Data delivery by one workgroup to another must be part of the supplier's functional obligations in order for the recipient to depend on it as a reliable source for vital information. Absent such a formal obligation on the part of the supplier, the recipient has no choice but to duplicate the data collection effort to ensure its needs are met.

Data duplication increases the cost of information to the organization. To the extent that these data production efforts generate slightly different results, the cost of data duplication includes the cost of different parts of the organization having different views of the same transportation facilities. There is also the cost of lost information, as one workgroup fails to pass along what it has to the next workgroup along the asset life cycle, whether for organizational or technical reasons.

Enterprise-level management of data through an organizationally endorsed governance process is an effective means of reducing the cost of data. Enterprise data management produces data integration to support data interoperability across internal workgroup boundaries. A key prerequisite to reaching this level of operational governance is documenting the data being compiled and used by each workgroup. With this information, the organization can find common elements and seek ways to combine resources across internal boundaries. However, the document must go beyond merely describing what is being collected and used to include why and how. It must also address the organizational issue of creating the quilt of GIS-T data described in the TFTN Strategic Plan.

The problem is that organizational issues require extensive interaction between agencies that may not share a common vision. Many non-State organizations see production of the ARNOLD dataset as a State problem and are not likely to participate in a process that requires agreement by all local governments. In other words, there are two kinds of problems to be solved: technical problems and organizational problems. Although this Guidebook contains instructions for organization governance that facilitate reaching its objective of producing a national transportation dataset, it is fundamentally a technical document. Thus, it is necessary to turn an organizational issue into a technical one for it to be reliably solved. In other words, the organization must adopt a formal workflow process. In this case, that process is how to compile and move data throughout an organization in the most efficient manner. Enterprise data governance is the proposed means of devising and enforcing that process, which is really data management. The CIM concept of asset life cycle management and the data flows it depends on provide the rough outline for data management.

All the foregoing are good reasons for change, but if everything is working fine now, why change? That is not the first question to ask. The first question is, "How do we know whether things are working fine for everyone today?" This is actually a question that needs to be answered on a regular basis, as users' needs change over time due to external forces. The best way to answer the question is to establish enterprise data governance. This is also the best way to anticipate, detect, and respond to change.

Change imposed by outside forces is disruptive. Change imposed by an organization on itself is transformative. Complacency—the decision to do nothing—will inevitably lead to disruptive change.

The transformative changes proposed by AEGIST have been developed through a comprehensive vetting process and will be tested by several States through a subsequent pooled-fund process for deployment assistance. AEGIST sets a destination of creating and maintaining the NRBM and shows a path to getting there. It is for each State to decide for itself how it will travel to that destination.

An early decision made during the AEGIST Guidebook's development was to omit address data from the NRBM specification. This decision was partly based on the principle that the initial NRBM dataset should primarily contain that information for which the State DOT or one of its partners is the authoritative source. State DOTs are not the authoritative source for addresses, which are created and managed by local governments, 911 centers, and other local and regional entities. There is also the matter of how to structure address data. Traditionally, address ranges have been assigned to roadway segment centerlines to provide a crude geocoding solution. It is becoming increasingly common to instead place address points at geographic locations.

There are additional considerations that make incorporating address data into the specification a decision for the future. Although there is a new standard for address data transmission associated with emergency response, called Next Generation 911 (NG911)[24], it is just beginning to be deployed and is for a specific application. NG911 is also not the only published address data format. There are several alternative address data format standards, including the U.S. Thoroughfare, Landmark, and Postal Address Data Standard, which was developed for Federal agencies.[25] Basic choices need to be made by the stakeholders before such information can be reliably incorporated into the NRBM.

Incorporating addresses is not the only issue for the future. As noted elsewhere, this initial publication of the AEGIST Guidebook is expected to be revised over time as it is deployed in States participating in the pooled-fund program. Some of those revisions will be to add domains for various attributes included in the product specification.

The proposed AEGIST data specification includes the 'LRM Object' as the representation of roadway characteristics, elements, and events that are positioned along roadways using linear referencing. Obvious LRM Objects include the characteristics of speed limit, functional classification, surface type, and number of lanes; the elements of bridge, tunnel, culvert, sign assembly, travel restriction, and guardrail; and the events of project, maintenance activity, and crash. These are the business data maintained and used by the States. Some of these objects, such as those representing functional classification and bridge, have existing names, descriptions, and domains provided by national standards. Some are included in HPMS, which provides a default national standard for the roadway data it includes. However, there is a clear need for an explicit effort to specify each LRM Object that will be included in the NRBM dataset or can be made available by the States by using the NRBM as a discovery mechanism. Such an effort goes beyond the scope of the pooled-fund program, which will focus on

---

[24] For more information, go to https://www.911.gov/issue_nextgeneration911.html. This website can provide a number of supporting documents.

[25] The *U.S. Thoroughfare, Landmark, and Postal Address Data Standard* was adopted by the Federal Geographic Data Committee in February 2011. The overall mission of this standard is to "meet the diverse address data management requirements of local address administration, postal and package delivery, emergency response (and navigation generally), administrative recordkeeping, and address data aggregation." One of its stated goals is to "Provide a systematic, consistent basis for recording all addresses in the United States." For more information, go to https://www.fgdc.gov/standards/projects/address-data.

deploying CIM and enterprise data governance, determining how the States can produce the NRBM dataset, and helping the States adopt the NRBM data structure for their internal use.

One possible source of a typology of LRM Object names, definitions, and domains is the U.S. Road Standards (USRS) project sponsored by the FGDC. USRS is currently looking at this issue, as well as addressing what standard to endorse for the national roadway dataset. It is also exploring the benefits of adopting the Geographic Data File specifications of ISO 14825 as the standard for submitting HPMS roadway geometry. Currently, States may submit any geographic file format compatible with the standards adopted by the Open Geospatial Consortium (OGC). HPMS staff is also considering migration to using web services as a means to provide continuous access to State HPMS submissions rather than continuing the current once-a-year physical submission process.

The States primarily provide the roadway data used by the USDOT and, through that agency's use of the data, indirectly contribute to the Transportation Theme of the National Geospatial Data Asset (NGDA). While the USDOT is the Federal agency tasked with overall sponsorship of the Transportation Theme, the All Roads Dataset is currently managed by the U.S. Department of Commerce, Bureau of the Census. As the AEGIST development process has well demonstrated, extensive involvement of the States and their local, regional, and Tribal partners is essential to the successful adoption of any specification by the States. FGDC subcommittees and the NGDA governance process are not currently structured to directly support extensive non-Federal membership. To fully realize the benefits of the AEGIST data specification, it is recommended that the USDOT take over management of the All Roads Dataset and begin the process of outreach to the broader community of roadway data suppliers and users to create the organization needed to construct the LRM Object typography. In this context, USRS could be the mechanism for propagating the decisions made regarding LRM Object definitions to the Federal agency members of the FGDC.

Chapter 2. Enterprise GIS-T Data Governance and Capability Maturity Models

The information in this chapter is derived from multiple sources, including two National Cooperative Research Program reports (NCHRP). NCHRP 666: Target-Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies (Volume I: Research Report and Volume II: Guide for Target Setting and Data Management)[26] includes sections on data governance that explain the processes and define roles and responsibilities. NCHRP Synthesis 508: Data Management and Governance Practices[27] provides a review of relevant literature and descriptions of State and local transportation agency practices in governance, data warehouses, data integration, and data quality management. This chapter also draws on past FHWA guidance on data governance and descriptions of capability maturity models (CMMs) for enterprise GIS. The reader should be aware that there are many different models of data governance, and the terminology and process descriptions may differ across the various models. This chapter aims to provide practical advice for implementing and sustaining a data governance effort for spatial data in an enterprise GIS. The advice can be generalized to other data governance efforts and specifically references efforts in data governance for asset management and safety.

As part of the discussion of data governance in enterprise GIS, this chapter incorporates descriptions of CMMs. It presents these models in two ways: first, as a way to assess and improve data governance itself, and second, as a noteworthy practice in assessing and improving enterprise GIS implementations. CMMs are a generalizable tool. The advice here applies to both uses of CMMs, and States are encouraged to make both types of CMM a part of their standard practice for systems and data management.

What is Data Governance?

In transportation, data governance is a formal process of managing data and systems to meet the enterprise's needs for information to support decision-making. In practice, this enterprise approach must also look to the individual business unit's needs in defining needs at the organization-wide level. Data governance is specifically designed to be a joint exercise of executive-level managers (agency directors, the chief information officer, and chief data officer in particular), the many business units' subject matter experts (the practitioners), and the information technology (IT) experts supporting the enterprise and the business units. Data governance is a set of processes for executing and enforcing authority over data. It is a data asset management method designed to deliver required performance of all defined and approved data-related business functions.

Data management is the process for developing, executing, and overseeing system architectures, maintenance policies and procedures, and data quality improvement.

[26] National Academies of Sciences, Engineering, and Medicine 2010. Target-Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies – Volume I: Research Report, and Volume II: Guide for Target Setting and Data Management. Washington, DC: The National Academies Press. https://doi.org/10.17226/14429.

[27] National Academies of Sciences, Engineering, and Medicine 2017. Data Management and Governance Practices. Washington, DC: The National Academies Press. https://doi.org/10.17226/24777.

Key Terminology

Data governance efforts introduce some terms and definitions that may be unfamiliar. It is useful to begin with a shared vocabulary so that the remaining discussion can focus on practical considerations for implementation. Many of the terms relate to roles and responsibilities within a data governance activity. The key terms are defined as follows:

Capability Maturity Model (CMM): A (typically) 5-point scale rating the ability of a data system or a process to deliver desired levels of performance. The five levels are usually described as (1) Initial or ad hoc, (2) Repeatable, (3) Defined, (4) Managed, and (5) Optimized. A CMM assessment helps States understand their status on an objective scale from least to most mature. For the purposes of enterprise GIS and data governance, the two CMM assessments are related, but separate. States should consider doing both assessments and developing goals and action plans to achieve their desired capability levels in both: one CMM for enterprise GIS, and one for data governance/management. States should understand that achieving level 5 in either CMM requires broad participation among all relevant stakeholders. The plans developed for improving enterprise GIS and data governance should be incorporated into the relevant sections of the State’s traffic records strategic plan and the State’s Strategic Highway Safety Plan, among others. CMMs are useful in data governance because they assess existing performance and help the data governance group set system capability goals.

Community of Interest: A very broadly defined group of people who need the data or provide services important to the data’s quality. This group may be accessed by the data governance group as part of a data gaps analysis and to gauge satisfaction with the available data and services.

Data Governance: The processes for executing and enforcing authority over data asset management and data-related business functions.

Data Governance Committee (Group, Council): A panel of key staff charged with implementing and enforcing department-wide data management and system policies. This group should involve data collectors, managers, users, and IT support staff relevant to each system and supported business area served by the data systems they manage.

Data Management: The processes for developing, executing, and overseeing system architectures, maintenance policies and procedures, and data quality improvement.

Data Quality: Data has several measurable attributes that tell users, managers, and collectors about its suitability for any intended use. The attributes are:

> **Timeliness:** measures of the duration of processes along the timeline from initial event or implementation until data availability for analysis.
>
> **Accuracy:** measures of the proportion of records with correct or erroneous data in any specified data elements.
>
> **Completeness:** measures of the proportion of records with full information for specified data elements.
>
> **Uniformity:** measures of data collectors’ adherence to established data standards for specified data elements.
>
> **Integration:** measures of the ability to merge data from two or more data sources linking records for the same people, locations, or events.
>
> **Accessibility:** measures of users’ ability to obtain and use data to meet supported business needs.

Metadata: Every data element in a well-documented system has associated information describing that data element—how it is collected, date stamps, data attributes, applicable data standards, quality measurements, validation rules, and others. The metadata is stored in the database for access by any user who needs to know the information.

Stakeholder(s): A broad group of people including those who collect, manage, or use the data. The stakeholders include members of business units as well as IT staff.

Roles and Responsibilities

Figure 4 shows a conceptual relationship of the levels at which data governance operates. The core of data governance starts with standards for system implementation set by the agency's upper management. System-level decisions (which platforms will be supported and how IT priorities are set) are typically the responsibility of the agency's Chief Information Officer (CIO). An agency's Chief Data Officer (CDO) may also take on the executive-level role of establishing the requirements for enterprise data management that all data governance groups would then implement. This is the level where leaders set policy and select the standard tools and design standards.

The next level is the data governance group, which is responsible for implementing systems according to the agency's defined standards. This group is responsible for enforcing those standards so that every system is appropriately documented, works well in the chosen software and hardware environment, and meets the architectural design standards. The data governance group is also responsible for defining data needs, establishing data definitions and quality standards, and enforcing the data quality standards. This is a cooperative process involving multiple agencies and key personnel from involved business units and the IT professionals who support the systems. The larger group made up of system stakeholders and a community of interest includes anyone who collects, manages, or uses the data and anyone with a business need for data that meets the standards. The data governance group serves as bridge between this broad group of stakeholders and the system deployment and support efforts. The stakeholders are a source of information on data gaps, business needs, analytic support requirements, and decision support.

Figure 4. The data governance group serves as bridge between stakeholders and system deployment and support efforts.

The following defined roles may carry different titles in different agencies or environments. The important thing is to identify the function of each defined role and assign the corresponding responsibilities as part of the data governance program.

Agency and IT leadership: This is typically executive-level decision makers, including the CIO and CDO. They set goals, allocate resources, and encourage data governance as agency standard practice.

Data Governance Board: The deliberative body that implements the executive-level decisions and sets data management policies for the systems under its control.

Data Stewards: The personnel who make sure data is collected, managed, and used according to the Data Governance Board policies.

Data Business Owners: The (often director-level and subject matter expert) personnel from the business units who define the business requirements for data.

Data Custodians: The IT support staff responsible for data system design, implementation, or maintenance.

Working Groups: The permanent or ad hoc committees who research technical issues and provide procedural, policy, and tool selection recommendations to the data governance board.

An individual may serve in more than one role. The important thing is to see that each role is covered in the data governance effort and those responsible know what they are required to do.

Practical Data Governance

Figure 5 shows a schematic overview of data governance, its inputs, and its products. As a noteworthy practice, a data governance effort would begin with upper-management support, including agency directors and the CIO and CDO. These provide guidance and empower the data governance group.

Figure 5. A flow diagram of data governance.

The data governance group itself arises from a list of stakeholders and functions as a set of decision makers charged with implementing data governance as defined by the agency or State. The data governance group relies on a data needs/gaps analysis and a capability assessment (CMM results) for the domains (systems and data) to be governed. This could include a CMM for each of the key data sources or business units and a CMM for the enterprise GIS.

The data governance group produces (or at least standardizes the contents of) system documentation and a set of relevant data standards and defines the quality control requirements for each governed system and dataset. These must be communicated and enforced so that all data collected is documented and meets the established standards. This process supports effective data integration, enterprise GIS implementation, advanced analyses, and more as part of the process of meeting business units' and departmental needs for data and system capabilities. Feedback from the supported functions gives the data governance group updated information in the form of new data needs and revised CMM scores. Thus, the data governance group also monitors changes in capabilities over time.

In practice, implementations and data governance may not work in this ideal manner. In some cases, the upper-level manager may not be familiar with or support the idea of data governance and must be convinced through a practical demonstration before he or she feels comfortable promoting its use. Even when there is encouragement and support from upper management, it can be difficult for a State agency to begin a data governance program because the business units may not want to give up a measure of control over their internal systems. Once established, data governance can be difficult to maintain because it is easy for key staff to be overwhelmed by the long list of tasks and associated time spent in meetings.

Sharing lessons learned from successful States may help. Learning from the experiences of peer agencies through FHWA-sponsored peer exchanges and technical assistance can help frame data governance as a more manageable and beneficial process.

Designating a data governance leader or champion helps keep the (potentially multiple) data governance groups working toward a shared standard and goal. Even for those who are not the designated champion, a change in job descriptions may be helpful so that the staff can have data governance as a part of their formal, assigned duties.

One important lesson is to recognize that data governance can be effective even if the State does not implement every aspect of a fully mature process from the start. A small, sustainable effort that demonstrates the value of data governance will be more useful than a massive, formal effort that fails. This incremental approach is common among the successful States; however, it should be recognized that the data governance effort must ultimately expand and become more formalized in order to achieve higher-level capabilities.

Examples of an incremental approach include starting data governance with just one or two key systems (e.g., asset management, safety, maintenance). This is obviously *not* an enterprise approach, but it can help to "sell" the concept of data management when the governed systems demonstrate success in meeting needs for data and system capabilities. Much of the progress made in the first few governance efforts can become the tested standard for how enterprise governance is achieved. Another common incremental approach is to start data governance with only the parts of the system controlled by the State DOT. This allows the DOT to establish and refine its processes before asking other agencies to devote their own resources.

There is a danger with any of the incremental approaches in that governance can stall at this incomplete level. For example, a DOT may take years to get around to involving local agencies in data governance.

This could create a missed opportunity because local agencies can support statewide data collection and management if they know what the standards are and have been consulted along the way.

Ultimately, formal data governance may require a complete cultural shift in a State DOT *and* in its relationships with partnering agencies. This will, by necessity, need the buy-in of executive-level staff not just in the DOT, but also in the county and municipal agencies, Tribes, and metropolitan planning organizations. To get that statewide buy-in, a DOT should start with that as the vision. Every incremental improvement in data governance should be building toward achieving the enterprise, statewide vision of formally governed transportation data and GIS. That is why CMMs are so important. Using CMMs for data governance and enterprise GIS, States can tell where they are and what is missing before they can achieve the next level of performance. CMM scores lead directly to action plans that can be incorporated into documents such as the Traffic Records Strategic Plan, the Highway Safety Plan, the Strategic Highway Safety Plan, and more.

Resource Questions

When setting up a new or expanded data governance effort, questions will arise about the number of hours required of the people involved and the balance between face-to-face meetings and individual work by the data governance group members. There are, of course, no established benchmarks for these activities; however, the following considerations should be helpful:

Start small. Setting up a data governance group and process for the first time will be time-consuming. It is a good idea to start with a pilot in one specific business domain (e.g., safety, asset management, maintenance) before tackling the larger issues represented by an enterprise GIS implementation. The experience in the pilots can help inform decision makers about the appropriate levels of staff involvement and other resource needs for the effort. Since data governance should show tangible benefits quickly, decision makers can also use the pilot implementations to judge the value of the effort and adjust their expectations.

Designate a champion. Data governance efforts go more smoothly when there is an in-house person facilitating the processes. Ideally, the agency can devote a person full time to assisting all data governance groups, especially during the early phases of implementation. Even if that level of investment is not possible, assigning a single person to lead each data governance group and providing help with meeting logistics (room reservations, scheduling, meeting preparation, etc.) will make the process more manageable.

Understand the charge. When the data governance group meets, it should take the time to acquaint all participants with purpose, mission, and vision for the effort. Next, the group should plan its activities and make assignments to schedule critical tasks. Working groups can tackle assigned tasks and then report back to the full committee. This encourages members to work in parallel and to complete their tasks on time. Through this process, it will quickly become obvious if particular members or teams are overburdened. One typical constraint is that the IT staff supports multiple systems and might be part of several data governance efforts.

Stakeholder Engagement

Data governance is designed to meet users' data needs as efficiently as possible. The data governance group must understand those needs in detail before they settle on data standards and before they can begin to set priorities for how best to fill any gaps they perceive in the data or in the systems' support for decision-making. There are several points where the community of interest should be part of the process:

Stakeholder registry. The data governance group needs to know who the stakeholders are, and how they use the data. Each system's documentation should include a stakeholder registry listing the agencies and business units interested in the data. The registry can also list major needs or any other pertinent information the data governance group desires to know about each stakeholder.

Gap analysis. System documentation lists the definitions of each data element. Data needs arise from interviews with stakeholders and a review of the data requirements for specific software tools or analytic methods in use by the business units supported by the system. A comparison of the two is a data gap analysis that the data governance group can use to identify and prioritize data collection needs.

Training. Whenever the State is relying on local agencies for data collection or quality control, they should provide training on data standards. The training will not always need to be lengthy and formal—information sharing is often sufficient. The data governance group should decide which agencies need training and help develop training content. Delivery will typically fall to the data stewards and their local agency contacts. Similarly, the data governance group can engage other stakeholders (e.g., State agency personnel) who also need the training.

Data governance group participation. The stakeholder registry is a good source of candidates for an expanded data governance group. Ideally, the group would include local agency participation from the start. It is common, however, for State DOTs to begin data governance as a purely in-house effort with a view to later expand to other agencies once the internal staff is comfortable with the process. Inviting representative local agency participants is a good way to increase stakeholder engagement.

Data Governance Group Charge, Mission, and Vision

Data governance is similar to strategic planning. It is a group process aimed at long-term management and improvement. States should begin by thinking about the domain of data governance in their operations, and what powers the data governance group will have to make decisions, implement changes, and enforce standards. Some States balk at this level of formality when they first start data governance, preferring a more grass roots versus top-down approach. The following assumes the recommended top-down approach, but brings in suggestions for the less formal implementation as well. Even if it begins less formally, however, data governance should eventually become a formal, repeatable part of how the agency manages data and systems.

Foundation. In most States, the DOT (and perhaps State IT) upper management will set the charge for data governance. That initial kickoff should tell the data governance group what it is supposed to accomplish and may define its authority and responsibilities. If a State wants to start a more grass roots form of data governance, the data governance group should spend time defining its own goals before it begins the actual governance process.

Vision. The vision for data governance should reference *how* data and systems ought to be managed. This is an aspirational statement of the way things *should* be if every system and data management effort worked according to the requirements set by upper management. If data governance is a grass roots effort, the group may need to establish their own vision in absence of upper-level requirements for system environment, documentation requirements, and data management practices.

Mission. The mission should describe what the data governance group will do. This is similar to the foundation document that may have been passed down from the agency executives. It should include an expanded description of how the group operates. Where possible, it can include quantifiable goals.

Data Governance Meetings

The data governance process requires meetings and meeting preparation. States are free to choose the frequency of the meetings, so the following are considerations for the data governance group as it works to establish its own process.

Start up. The first data governance efforts in an agency, and the first implementations of data governance over any specific system, require several meetings. The data governance group has to decide, as a group, how they will operate, who else should be invited, if working groups are helpful, and what list of issues they should address. The level of effort should settle into a less intense schedule once the data governance group is established and has gone through a few meetings together.

Continuous efforts. Like any strategic effort, follow-through is important. When agencies or individuals accept an assignment as part of the data governance group, they must complete the work and report back. The data governance group should agree to meet multiple times throughout the year so long as there is a list of items awaiting their decision. They can also form working groups to tackle specific issues (e.g., gap analyses, data integration, data collection plans). The working groups should also report back to the full committee on a regular basis. If the State designates a data governance champion, one role for that person is documenting progress on assignments and keeping the list of issues up to date. The agenda for each data governance group meeting should be drawn from its established list of issues and the work assigned to individuals and working groups.

Member satisfaction. Data governance efforts sometimes fail. They do not usually fail because of a lack of any important work that needs to be done. They fail most often because the members do not see the value of meeting again. The data governance group should take time to assess its own members and their satisfaction with the group's progress. If data governance is not delivering the promised improvements, then it should change. The group can improve its own processes, up to a point. If the group lacks authority to make changes, however, then it must go back to upper management for direction. The important thing is to establish a process that invites some internal reflection on how to improve data governance so that the meetings do not falter or the group become ineffective.

Enterprise GIS-specific Governance Considerations

Data governance for spatial data is an important part of implementing an enterprise GIS for transportation. While the majority of this guide gives examples of data specifications for transportation, the following considerations are specific to how a State can govern data no matter what standards it implements.

Focus. Enterprise GIS is intended to include data that has (typically) existed in multiple systems in multiple formats. Incorporating all the various business unit-specific systems into one enterprise GIS does not necessarily mean that governance over what are now individual parts of the GIS must be subsumed into a larger, enterprise-level governance effort. The State may decide, for example, that governance works best at the level of business needs so that it might establish data governance for safety data, separate governance for asset data, and so on. These individual efforts would still benefit from the spatial data guidelines provided in the remainder of this document. The important point is that the data (and the individual parts of the enterprise GIS) are governed consistently, at least with respect to the spatial data standards the State will adopt.

The State should decide how it can best govern each data source now and in the future. A flexible approach is advised. It may be best to start with a pilot (as mentioned at the start of this section) and expand governance to all systems and data as the processes and efforts prove their worth. If the State is confident in its ability to sustain a spatial data governance effort, it can also decide that this is the ideal place to start.

One potential long-term structure for spatial data governance is to establish working groups based on the governance groups previously established for the business-unit specific systems. Ultimately, the goal is for every system and every data source to be managed through a formal data governance process. There is no single best way to accomplish that, and each State is free to adopt the structure that works best for its purposes.

Stakeholders. Data governance in enterprise GIS requires another group of stakeholders—those who manage data in the GIS environment. This group is not necessarily the same people as the data custodians (the IT staff) responsible for system support for business units in the DOT. Where relevant to the success of the statewide GIS, the local GIS experts should become part of the data governance group.

Gap analysis. Where the success of the statewide system relies on local data input, the spatial data environment does not really impose any new requirements on the gaps analysis or the data governance group's need to understand local agency capabilities. However, it is clearly important as part of the gap analysis to address not just the business unit-specific data gaps, but also gaps in GIS capabilities and knowledge. The data governance group should include these in its prioritization efforts.

Metadata. Spatial data has an expanded set of metadata elements that the data governance group should define and require. The other sections of this guide describe the various types of important metadata. These include versioning, dates, and precision, along with the usual items related to measurable data quality. These are important parts of the data standards, and the data governance group should spend time deciding which items apply to each dataset and include those in the data standards requirements.

Non-spatial data. Enterprise GIS for transportation includes non-spatial data elements. These have different metadata and data standards requirements that are not the subject of this guide. The data governance group should still adopt standards for these elements, many of which can be obtained through discussions with business units and agency administrators. Examples include financial data and some resource allocations.

Enforcement. Spatial data governance, like all data governance, will succeed or fail based on the level of compliance with the data standards. The data governance group needs to know its own authority over enforcing the standards and then decide how best to work within that authority to encourage compliance. Where enforcement is lacking, data quality will suffer and the enterprise nature of the system will falter.

Using Capability Maturity Models and Self-Assessment

There are several CMMs describing multiple levels of enterprise GIS implementations. The Urban and Regional Information Systems Association (URISA) established the GIS Management Institute to help organizations advance GIS competency and maturity levels. On the URISA website, there are several GIS CMMs and self-assessment tools available free of charge (or free to government agency members of URISA). Slimgim offers a set of formal and repeatable tools for assessing and improving enterprise GIS. A variant tailored for State DOTs, known as Slimgim-T, is focused on State-level enterprise GIS implementations in transportation. In addition, a 2018 FHWA report[28] is available that documents noteworthy practices for completing a CMM assessment of enterprise GIS, using several State case studies as examples.

The following is a brief overview of CMM and how it might be implemented with respect to enterprise GIS. CMMs typically define five levels of capability maturity. The labels for the five levels may vary among models, but the following are typical:

Level 1: Initial or ad hoc. This is the lowest level capability. It is typically described as individual efforts for single systems or applications with no central control or enforcement of standards. Data quality is handled individually for each application or system. Spatial data systems are organized at the desktop level without shared standards even within a business unit's applications. In a CMM for data governance this level is described as ad hoc, because each system's data is managed however the business unit decides is best and there is no centralized standard or coordination applied. In a CMM for enterprise GIS, this level is characterized by stand-alone implementations of GIS within individual business units with no central coordination or requirements for shared environments or standards.

Level 2: Repeatable. This level includes process and policy definitions for data management, but the definitions are not shared across all applications and business units. Data quality is still managed at the individual system level. Spatial data standards are shared within each business unit, but they may differ across business units. In a CMM for data governance, this level would include some upper-level policy guidance on what system documentation and review processes *should* include; however, each system's data is managed separately and the policy may not be enforced. In a CMM for enterprise GIS, upper-level management will have set policies for GIS implementations; however, stand-alone systems will persist without coordination.

Level 3: Defined. At this level, high-priority projects are managed through a formal data governance process with clearly defined policies. The agency manages core data using shared standards that apply across platforms and business units. Spatial data standards apply throughout the organization for those core data sources subject to data governance. In a CMM for data governance, departmental standards are enforced for most systems; however, there are still business unit-specific systems that are not governed according to standards. In an enterprise GIS CMM, a similar pattern can be seen where high-priority, mission-critical implementations are developed using a single, enforced standard, but some GIS implementations still do not meet the defined standard.

Level 4: Managed. This level includes agency-wide processes that are applied across all systems and platforms. Data standards are enforced throughout the agency or organization. Spatial data standards apply equally to all systems managed by the agency. In a CMM for data governance, every data system in the State agency is governed with a formal, repeatable process; however, the standards may not fully apply to systems outside the agency. In an enterprise GIS CMM, this level has all internal GIS implementations on a single platform and managed in accordance with established standards and policies; however, outside agencies may have their own systems and standards.

Level 5: Optimized. At this highest level, data management policies address needs both within and outside the organization so that governance applies (and includes) all internal and external sources. Spatial data standards are set cooperatively with all stakeholders and enforced throughout the State. In a CMM for data governance, all transportation data adheres to the same standard no matter which

---

[28] Green, M. Lucivero, A., Rovan, A., and Quinton, D. (2018). ***Slimgim-T: GIS Capability Maturity Model for State Departments of Transportation.*** Federal Highway Administration, Washington, D.C. Retrieved from FHWA website 5/28/2019: https://www.gis.fhwa.dot.gov/documents/Slimgim-T_GIS_Capability_Maturity_Model.pdf.

agency collects or manages it. In an enterprise GIS CMM, there is a statewide GIS that applies to all transportation and even relevant non-transportation data across all agencies.

States can conduct a self-assessment to measure their capability maturity level for enterprise GIS and for data governance. The URISA GIS Management Institute website is a good source for the enterprise GIS self-assessment tools. In 2015, NCHRP 814, Data to Support Transportation Agency Business Needs: A Self-Assessment Guide[29] provided a tool for States to judge their data needs and plan actions to make and monitor data improvements. In 2012 and 2018, FHWA completed an assessment and CMM of all States plus the District of Columbia and Puerto Rico on safety data management. That assessment included data governance roles and responsibilities, policies, and processes. States can start with their CMM scores for data governance from that assessment and use their State-specific action plan to help them reach those goals. Information on these two assessments can be obtained from FHWA Office of Safety.

Final Considerations on Data Governance

There are many data governance models and methods. States should choose the ones that work best for them and be ready to modify methods that do not work as well as planned. There is a great deal of work to do at the start of any data governance effort. The data governance group needs to know what is expected of them and what the goals are for the systems, the data, and for data governance in general.

As agencies gain experience in data governance, the process does become easier to manage. There is some "tension" between doing governance well right from the start and starting cautiously to build the necessary trust in the process. This guide acknowledges that tension and cautions against starting with too large an effort, even if it means that spatial data governance lags (slightly) behind enterprise GIS implementation. Ultimately, though, consistent spatial data governance is crucial to long-term success of the enterprise GIS. States should move toward fully formal spatial data governance as quickly as possible. This should, also in the long term, save a lot of effort by reducing duplication. Since the standards for spatial data are shared across all business units, governing the various component parts of an enterprise GIS should be similar and result in economies of scale.

State DOTs are still in the learning stages for data governance and enterprise GIS implementation. This combination produces some reluctance by the DOTs to expand their discussions to bring in other State agencies and local agency partners. From a data governance perspective, failing to bring in other agencies right from the start is a mistake, but one that it is perhaps difficult to see before the attempt is made. The best advice FHWA can offer is to plan to bring in other agencies at the earliest possible time so that they are formally included in the data governance group. In the meantime, the data governance group must have access to the broad list of stakeholders so that they know where the data, knowledge, and capabilities gaps are. This level of engagement should be part of the effort from the start.

[29] National Academies of Sciences, Engineering, and Medicine 2015. Data to Support Transportation Agency Business Needs: A Self-Assessment Guide. Washington, DC: The National Academies Press. https://doi.org/10.17226/23463.

Chapter 3. Fundamental Rules and Concepts

This chapter presents the fundamental concepts of linear referencing systems (LRS) and the linear referencing methods (LRM) they manage and apply. It defines several terms so that the States and other entities implementing the recommendations of this Guidebook may have a common language. The chapter concludes by listing the core business rules for data contributors and managers so that States can provide a consistent product and simplify the data editing and publication process. This consistency is a prerequisite for producing a National Roadway Base Map (NRBM) using the States' data products.

Linear Referencing Fundamentals

Maps are abstractions of reality intended to express a central idea of a set of facts. Map design choices are made to emphasize some aspect of the data. Maps should be drawn to a scale that is appropriate for the spatial resolution and accuracy of the data they seek to present, the medium displaying them, and their intended use. The features on a map are simplifications of real-world entities. Linear referencing is one means of simplifying real-world entities; i.e., roadways. It offers a means to model locations in one dimension (1D), that is, to convert the two-dimensional (2D) or three-dimensional (3D) geospatial location of an entity to a 1D position along a linear transportation facility. The "thing" existing at that location, whether it is a descriptive characteristic, element of the roadway, or an event that happened at the location, is a 1D object.

These 1D objects have traditionally been referred to as "events" that occur along a specific route and can exist as point events or linear events.[30] Point events are located by a specific distance (measure) along a specified roadway. Linear events are identified by a starting distance and an ending distance (start and end measures) on a specified roadway. Linear events can also be specified using a starting measure and a length, and the sum of these two values determines the ending measure. Prior to the advent of geographic information system focused on transportation data (GIS-T), linear referencing supported the construction of data tables defining events that were measured along a roadway and identified by the distance along that roadway from a point of reference (typically the beginning of the roadway). GIS-T technology developed data structures and mathematical processes (dynseg) to support drawing point and linear events from the tabular data.

GIST-T has extended the utility of the 1D LRS data model by providing tools that create map data from geographic coordinates as points or lines that can incorporate 1D data with other information provided in 2D and 3D forms. One of the principle reasons for doing so is to analyze the spatial relationships of roadways to other phenomena. Once in the GIS-T environment, events acquire the characteristics of points or lines and can take advantage of the power of geospatial analytics and cartographic display methods. Of particular importance is the ability of GIS to combine data from disparate sources and create new information.

Another utility of using GIS-T is the ability to take data presented in a 2D or 3D coordinate system and collapse them into one dimension; i.e., to convert them into point and linear events based on how they interact with roadway centerlines. In GIS-T, the relationships between a roadway and elements that sit off the roadway, along the shoulder, or in any lane can be converted into 1D events. Methods have been derived to place such objects at a position offset from the centerline. Objects collected by different

[30] The term 'event' came into formal use with National Cooperative Highway Research Program Project 20-27(2), one of the earliest fundamental research projects dealing with the transition to GIS-T from legacy systems. This Guidebook and the NRBM limit the use of the term 'event' to its more common definition of things that happen. 'LRM object' is the general term used herein to collectively refer to characteristics, elements, and events.

programs and stored in disparate databases based on other datums and coordinate systems can be analyzed together. For example, signs and crash locations can be treated as point events that can be associated with surface types, pavement conditions, traffic volumes, and lane width, all of which are linear events.

Despite the strength of linear referencing in modeling spatial relationships, it has limitations with respect to cartographic presentation. For example, the path of linear objects that lie alongside a roadway, such as guard rails, retaining walls, and sound walls, cannot be accurately depicted using LRS. These are independent of the roadway and have completely separate geometric characteristics.[31] It is, however, appropriate to use an LRM to place the feature *along* the roadway, where its coincidence can be modeled along with other roadway data.

Typical guidelines for designing spatial datasets involve thinking of the problem as creating a number of map layers. For example, there might be layers for rivers and bodies of water, property parcels, buildings, and transportation facilities. Each map layer consists of geometric features with data attached to them. Many of these objects, like parcels and buildings, are naturally distinct and well defined. Linear entities, like rivers and transportation facilities, typically must be subdivided to create discrete features to which attributes may be attached. A new segment is created whenever one of the included attributes changes in value. Street map layers are often subdivided into blocks terminated at intersections in order to support address-based geocoding applications. A user can click on a segment to view its attributes.

This approach works well enough when there are few attributes to include, but it gets complicated to construct and maintain the data when there are numerous attributes and they change frequently or randomly vary in value. There is also the issue of how to include something that exists along the segment or when values change at mid-segment locations. The States have very large spatial extents and numerous attributes to track in their roadway inventory systems. These systems are intended to address four fundamental issues:

1. Transportation facilities are often not discrete. Yes, one can walk outside and readily identify a roadway, for example, but what are the physical limits of that roadway? More importantly, perhaps, what is the identification of that roadway? Is it defined by the name or some other characteristic? A pavement management office may want to subdivide the roadway in accordance with its pavement type or year of construction. The planning staff may need to subdivide the roadway by functional classification. A State may want to use signed route numbers, while a local government may prefer their street names to define the extent of each roadway. What happens when State route numbers overlap and are applied to a single piece of pavement? As these examples demonstrate, there is no single way to segment transportation facilities when designing a database for multiple user groups.

2. Transportation facility attributes change frequently. A single attribute at one location may not change for years, but given the spatial extent of a State roadway system, an attribute is changing somewhere every day. Editing the database to keep it updated is a daunting task. If an agency uses the block segment approach and a street is repaved for a distance of, say, 2 miles, there may be dozens of block segments that now need a new entry for "Date Last Paved." Not only is there the physical process of entering the new data, there is also the problem of getting the new data to the editors. Someone has to be given the task of telling the editor what was

[31] The NRBM treats these features as elements of the roadway and can have their own geometry independent of the roadway centerline.

changed, how it was changed, and when the change took effect. Then the editor has to provide the updated information to all users.

3. Many attributes are continuously variable. There can be no gaps, for instance, in speed limit, jurisdiction, functional classification, and many other attributes. Every piece of each roadway must have one and only one value for these attributes.

4. Attributes do not always change at the end of a segment. Some attributes also do not apply to a whole segment. A 500-foot-long block might have a 30-foot loading zone designated along its length. The parking enforcement and traffic operations divisions will need to know that it exists and where it is located. Roadway elements, like signs, are point entities that may also need to be included. How can that be done with a segment-related design?

As a result of these many issues with segment-based data structures, the States and other agencies dealing with large transportation systems adopted the linear referencing data structure to deal with these various drawbacks of the segmented approach. This data structure is based on the separation of transportation feature geometry from the business data stored in tables. Figure 6 shows how this data structure differs from the segment-based approach.

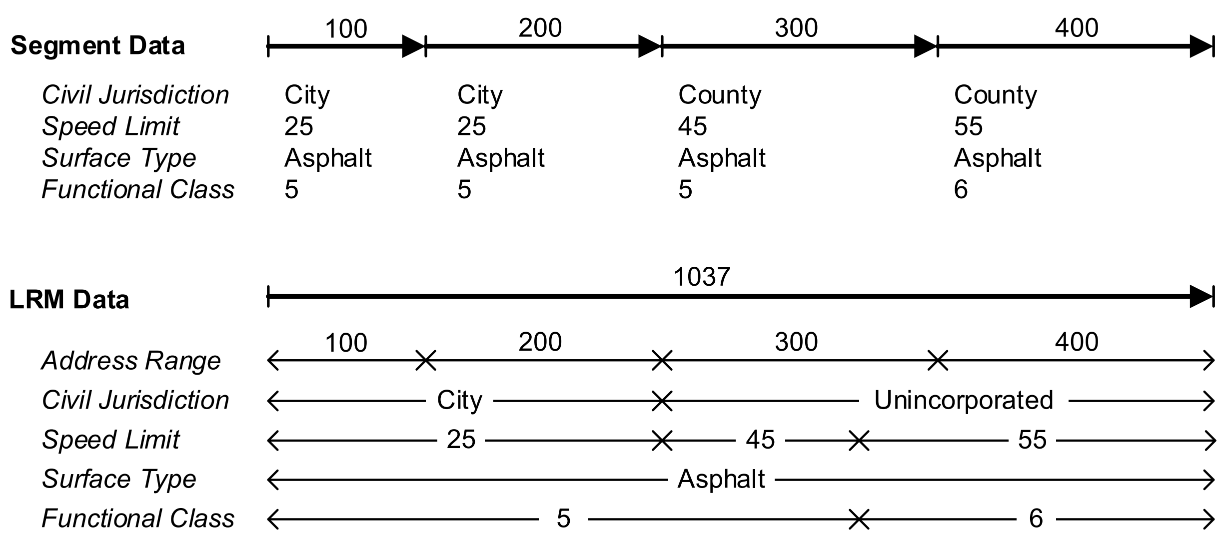


Figure 6. Segments versus LRM routes.

The segment-based design repeats the data for each segment. Every field shown has at least two segments with the same value. The **`SurfaceType`** = 'Asphalt' entry is repeated four times. The LRM-based database design has no repeating values. A change in the value of a field requires that only one or two entries be modified. More importantly, perhaps, is the ability to change values at mid-segment locations. In Segment 300, the segment-based design has a single entry for **`SpeedLimit`** and **`FunctionalClass`**. The LRM data shows that these values actually change at a mid-segment location. Thus, the segment-based data has the value that applies to most of the segment, while the LRM-based data has the actual location where the change occurs.

The segment-based data contains four segments; the LRM-based data has only one. To map, for example, **`SpeedLimit`** values, all that is required for the segment-based data is to assign a color to each defined speed limit, read the `SpeedLimit` value for the segment, and turn the segment centerline that color. It does not matter how long the segment may be or what path the centerline might follow. Each segment is a standalone entity. It takes four rows in a Segment feature class to fully describe this part of the street network. (The contents of the Shape column are not visible to the user.)

Segment Feature Class

SegmentID	CivilJurisdiction	SpeedLimit	SurfaceType	FunctionalClass	Shape
100	City	25	Asphalt	5	*
200	City	25	Asphalt	5	*
300	Unincorporated	45	Asphalt	5	*
400	Unincorporated	55	Asphalt	6	*

Things are more complicated for LRM-based data. Before it can map segments based on speed limit, the GIS software must first create the segments. It does this by using a process called "dynamic segmentation," or simply "dynseg." Dynseg depends on LRM measure values that tell the software where segments begin and end. The LRM Data example in Figure 6 omitted a key piece of information: the measure values. They have been provided in Figure 7.

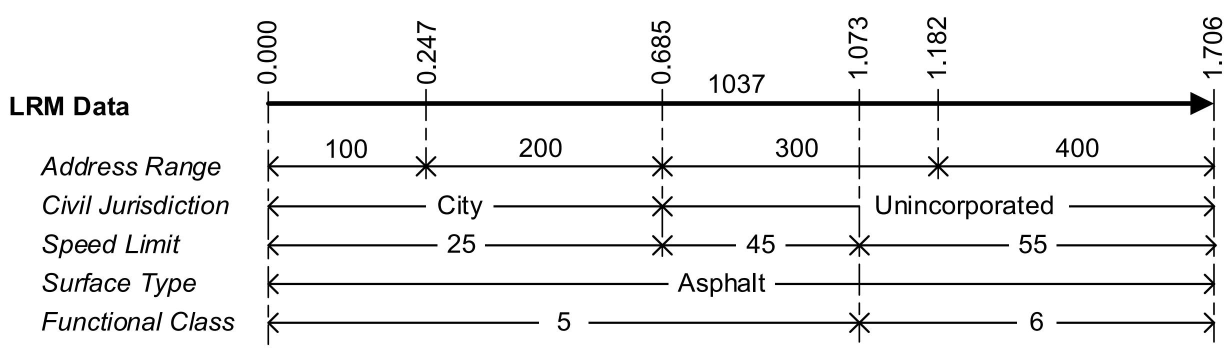


Figure 7. LRM route with measure values.

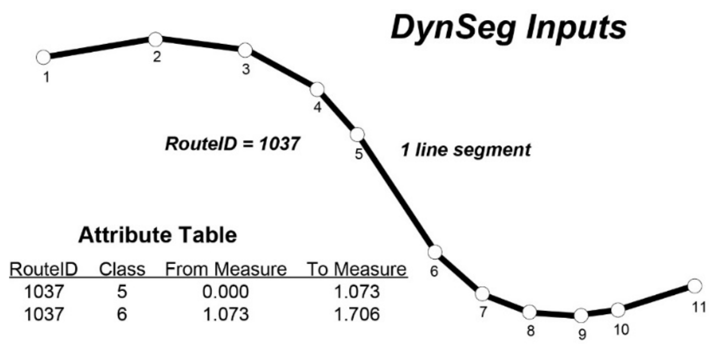
A typical vector centerline is a series of vertices, each with an *x,y* pair of coordinates saying where that vertex is located on the map surface.[32] When working with LRM-based data, a third coordinate must be added to each vertex: a measure. This additional coordinate supplies the cartographic features with the same measure values stored in the business tables, thereby facilitating the use of dynseg to produce the segments needed to display a particular attribute or set of attributes.

---

[32] One of the basic differences between roadway centerlines created in GIS and those created in computer-aided design (CAD) is the manner in which curves are constructed. Curves in GIS are commonly constructed as a series of straight line segments that provide the appearance of a curve at the desired display scale. The larger the scale, the more vertices are required since the straight line segments have to become smaller. CAD generally constructs curves using geometric and mathematical formulas, such as to define a point of curvature, a point of tangency, and an inscribed angle, degree of curvature, or radius of curvature. The difference in structure is based on the intended product's application; one approach is not inherently better than the other for creating a map at a specified scale. All current GIS platforms support the construction of curves using the CAD approach in order to support data exchange. The NRBM does not specify the manner in which curves should be constructed to produce roadway centerlines, except for the requirement that they are compatible with a map display scale of 1:5,000 and the data must be delivered in conformance with Highway Performance Monitoring System (HPMS) requirements. Most States are producing roadway centerlines that can support much larger scales, in part due to the use of larger scale aerial photos as the original source.

The segment database stores the segment identifier, the name of each attribute, and the value of each attribute. The *x,y* coordinates stored for each segment centerline say where it goes. LRM data needs to include more. LRM data tables include the route identifier, a beginning measure value, an ending measure value, the attribute name, and the attribute's value. On the geometry side, there is also a difference between segment and route features: a route feature has to establish a 1D datum. In other words, it has to include the starting and ending measure values. Thus, a feature class is needed for route centerlines and one or more business tables for route attributes—both of which have to be calibrated to the same beginning and ending measure values. The dynseg process takes the beginning and ending measures of each record in a business table and constructs a centerline segment that conforms to the linear extent of the record using the geometry provided by the route centerline. Each business table generates different record-based line segments.

Figure 8 shows how dynseg works to create segment geometry to display the attribute(s) of interest. The inputs are a feature class with measure values assigned to its vertices and an attribute table with the location of each record defined by a route identifier and one or two measure values.[33] In this single-route example, the centerline consists of a set of 11 vertices, each knowing its measure value. The dynseg process had to create vertices wherever needed to define the start/end of a segment. It does so by using straight-line interpolation to find the point along the line that matches the proportional distance implied by the measure values assigned to the two vertices.

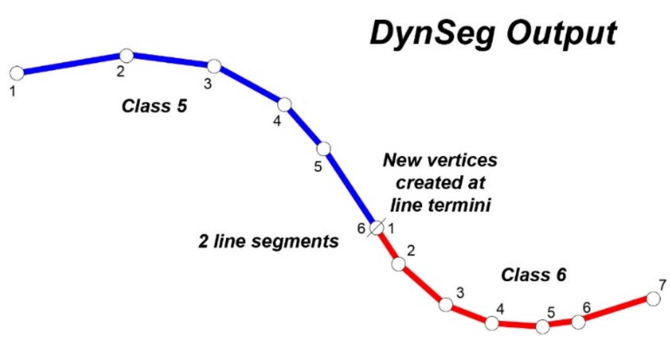


Attribute Table

RouteID	Class	From Measure	To Measure
1037	5	0.000	1.073
1037	6	1.073	1.706

Original Centerline

Vertex	Measure
1	0.000
2	0.214
3	0.589
4	0.708
5	0.821
6	1.118
7	1.196
8	1.202
9	1.329
10	1.474
11	1.706



Resulting Centerlines

Segment	Vertex	Measure
1	1	0.000
1	2	0.214
1	3	0.589
1	4	0.708
1	5	0.821
1	6	1.073
2	1	1.073
2	2	1.118
2	3	1.196
2	4	1.202
2	5	1.329
2	6	1.474
2	7	1.706

Figure 8. How dynamic segmentation (dynseg) works.

[33] LRM and dynseg support point and linear route attributes. A point attribute has a single measure value. A linear attribute has two measure values, one for the beginning point and one for the ending point.

In trying to find the starting point for the record segment it needs to create, the algorithm seeks either an existing vertex that has the correct measure value or the two vertices that define the limits of the line segment on which the vertex and measure can be inserted. It starts by finding the start of the line, which here matches an existing vertex with Measure 0.000. It then works its way along the line until it reaches the two vertices that define the segment where the end of the first row will be found at measure 1.073. These are Vertex 5, where m = 0.821; and Vertex 6, where m = 1.118. Using straight-line interpolation, the algorithm calculates that Measure 1.073 is 86% further down the line than Vertex 5 (1.118 – 0.821 = 0.255; 1.118 – 1.821 = 0.297; 0.255 / 0.297 = 86%). The algorithm creates a vertex at this location along the line to terminate the first feature. It then repeats this process to create a coincident vertex that starts the next feature at Measure 1.073, and then to find the end of the segment at Measure 1.706. A similar process is used to locate a point object, except that it only needs to find or create a single vertex.

The most important structural difference between segment and LRM databases is that a segment has a value for each route attribute built into the feature class, while LRM databases separate the business data about the routes from the geometric features used to display them. This means a segment can be a centerline with attributes attached, just like all other map layers in a GIS. A route can also be a single centerline feature on the map, but it cannot show any data values until dynseg has occurred. The user has to pick which attributes to display, and then run dynseg to display those values by subdividing the route centerline into pieces that match the spatial extent of each business data record being displayed. The end result is still a set of segments that match the business data, but each piece of business data creates a different set of segments because each attribute is allowed to begin and end where it actually happens along the route.

Figure 9 shows the inputs for this example route: one Route feature class with a geometric description of the path of Route 1037 and an Attribute table with all the data used to describe this route.

Route Feature Class

RouteID	BeginMeasure	EndMeasure	Shape
1037	0.000	1.706	*

Attribute Table

RouteID	AttributeType	AttributeValue	BeginMeasure	EndMeasure
1037	Address Range	100	0.000	0.247
1037	Civil Jurisdiction	City	0.000	0.685
1037	Speed Limit	25	0.000	0.685
1037	Functional Class	5	0.000	1.073
1037	Surface Type	Asphalt	0.000	1.706
1037	Address Range	200	0.247	0.685
1037	Speed Limit	45	0.685	1.073
1037	Address Range	300	0.685	1.182
1037	Civil Jurisdiction	Unincorporated	0.685	1.706
1037	Speed Limit	55	1.073	1.706
1037	Functional Class	6	1.073	1.706
1037	Address Range	400	1.182	1.706

Figure 9. Example route inputs. The Attribute table rows are ordered first by BeginMeasure value, and then by EndMeasure value.

There are 12 rows in the Attribute table to store the same data that the segment approach could fit into only four, and the segment-based design already had the geometry that matched the extent of each record. It looks like the LRM approach is more complex, but it actually produces a smaller file. The four segment records contain five attributes in each for a total of 20 values. The LRM database contains only 12 records. This difference is due to the lack of repeating values in the LRM database. For example, the same value for **`SurfaceType`** (Asphalt) is given four times in the segment database but only once in the LRM database. The lack of duplication in the LRM data structure means there are fewer records to change, and that means the risk of error is lower and the workload of data maintenance is reduced.

The drawback is there is not a map of roadway data sitting there with all this information. Users will still need to go through the dynseg process to create the segment geometry needed to display an attribute's value in a map. For example, to make a map of speed limits, a user would select all the records from the Attribute table that have an **`AttributeType`** value of 'Speed Limit'. The result would be the following:

Speed Limit Table

RouteID	AttributeValue	BeginMeasure	EndMeasure
1037	25	0.000	0.685
1037	45	0.685	1.073
1037	55	1.073	1.706

Dynseg would create three Speed Limit feature class segments, each matching the spatial extent defined by the beginning and ending measure values contained in the table's rows. Each segment would have a SpeedLimit value matching that of the source table:

Speed Limit Feature Class

RouteID	BeginMeasure	EndMeasure	SpeedLimit	Shape
1037	0.000	0.685	25	*
1037	0.685	1.073	45	*
1037	1.073	1.706	55	*

Figure 8 showed how dynseg worked to produce a Functional Classification feature class with two members. All these examples assume that the attributes are mixed together into a single table. Often, each attribute is in its own table.

Considering the extent of a statewide LRM database perhaps including a million or more attribute records, a lot of computing power is required to make a map. So, why would a State DOT adopt an LRM database design, create all these measure-based attribute records, and have to go through dynseg every time it wanted to turn the data into a map? The answer is because the LRM-based design fixes the basic problems imposed by a roadway inventory application:

- Records can be combined to produce whatever features are needed; users are not limited to a fixed segmentation scheme.
- Each attribute value is in a single row, which makes it easy to find and modify any value.
- Data quality control is easy to implement.
- Attribute values can change anywhere along the route.
- Business data can be edited by its authoritative workgroup without having to also maintain a geometry reference layer.

In addition, it is important to note that those measure values are based on real-world distances. It is much easier to compile and use data based on real-world distances than trying to force data into a fixed-segmentation approach. All it takes to use measure values in the field is a vehicle odometer and knowledge of where each route starts and the path that it follows.

But here is the clincher for NRBM: It is a simple matter to produce a segment-based output from an LRM-based information system—all the attribute records are joined using the measure values. A new record will be created wherever an attribute value changes. In the example, the attribute that changes the most often is **`AddressRange`**, which has four records.[34] However, this process, with an LRM foundation, will generate five rows, not four, since it will split the 300 block into two records that more correctly show each piece with different speed limit and functional class values.

Segmented data also plays another role in NRBM: It makes updates more manageable. Geometry changes will be propagated using whole segments. Each State or other NRBM provider will define fixed segments along each inventory route that will form the basis for geometry revisions. If a reshaping occurs along part of one of these segments, the old segment geometry will be retired and a new geometry feature will be provided as a complete replacement. This will allow each end of the modification to tie back to existing geometry so it can be used to construct LRM inventory routes.

Before concluding this section, there are a few more fundamental aspects of LRM to cover. First, most people use the term 'event' to talk about route attributes. This is a term that was developed many years ago and popularized by the NCHRP 20-27(2) research study, which explored how to develop a common transportation database design. (Yes, this has been attempted before.) So, instead of talking about attributes or attribute tables, most LRM language speakers will talk about events and event tables. This Guidebook explicitly avoids using the event term. Here, the general term 'LRM object' is used and is defined as including characteristics (attributes), elements (physical components of the roadway), and events (things that happen). Each of these LRM objects has its own data structure and operational considerations. Since route identifiers and measures can be used to say where they are located, the general term 'LRM object' is more inclusive.

Second, the routes used so far to explain how LRM and dynseg work are not the signed or named routes that populate the real world. These are inventory routes. They may have initially been created by following the path of a signed route, but once created, they no longer have any connection to the signed route. This distinction is necessary to avoid problems when the path of a signed route changes—and these changes happen with remarkable frequency. If measure values are derived from the real-world distances defined by the path of a route, all the measure values downstream of the realignment must be modified. This is a lot of work, and it makes it difficult to compare locations over time. Instead of signed routes that change over time, the AEGIST Guidebook is based on inventory routes that do not change when the path of a signed route changes. The only thing that happens when a signed route changes its path is the Signed Route characteristic changes for the affected portion of the route.

Third, the Guidebook refers to the old records as "retired" rather than deleted. This is to retain the historical record to provide a temporal aspect to the database. Any analysis that compares the past to the present (or the future) needs the old information. The time that the system changed and how it changed needs to be recorded, but this revision cannot erase the past; it must be preserved. Few State

[34] Using address range data in this example may appear to conflict with the prior discussion about omitting address data from the initial NRBM specification, but such information is common in roadway databases constructed with fixed segmentation conforming to blocks. How the data is stored is specific to each database creator.

roadway inventory databases are structured to do this, but it is a primary functional requirement of the NRBM dataset.

Lastly, it is important to talk about the difference between accuracy, precision, and resolution. The ARNOLD Reference Manual specifies a spatial baseline scale of 1:5,000. A spatial scale of 1:5,000 is the same as 1 inch equates to 417 feet on the ground (5,000 / 12). Using the traditional map accuracy standard of 1/30-inch error,[35] the related spatial accuracy would be ±14 feet and the minimum distance that can be detected is ±7 feet; i.e., objects in the real world must be at least 7 feet apart to be shown as existing at separate locations. This is the maximum resolution.

Figure 10 illustrates the difference between accuracy, precision, and resolution. Accuracy is the measure of how close a sample value is to the true value. The 1:5,000 scale requirement of ARNOLD can be converted to an accuracy standard of ±14 feet by using the traditional map accuracy standard of 1/30-inch error. Precision describes the repeatability of a measurement. Accuracy and precision work together to determine data quality. Good repeatability allows correction of systemic errors, much as one might adjust the sights on a rifle. A rifle with a tight clustering of shots on the target can be adjusted to achieve higher accuracy.

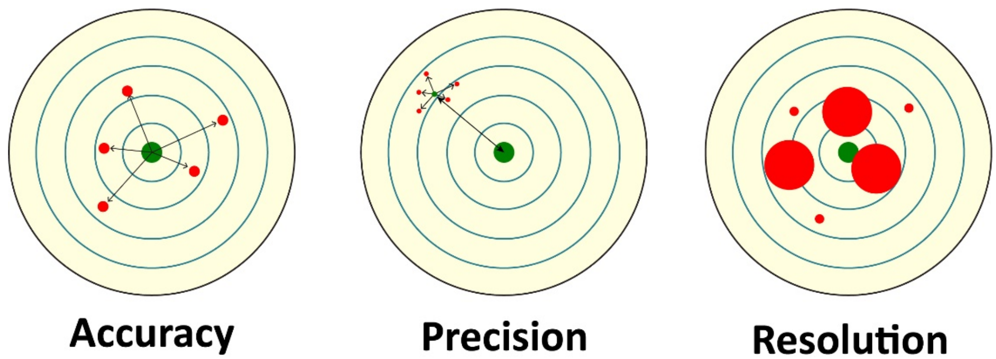


Figure 10. Accuracy, precision, and resolution.

Resolution is a little different. Resolution depicts how big the bullet is. In the case of ARNOLD, the spatial resolution of 7 feet is determined by the accuracy requirement of ±14 feet; i.e., two objects separated by 7 or fewer feet will appear to be in the same location. This is slightly less restrictive than the 5.28 feet (0.001 mile) LRM measure resolution that is in common use by the States today. The 0.001 mile standard has been adopted for LRM-based data published as part of NRBM.

The deployment of GIS software to work with LRM data using decimal values has illustrated a problem with digital fractions. A common practice is to store the data with a precision of seven decimal places. Working with data at this level of precision can often induce small changes in the value being stored. To eliminate the problems caused by decimal measurements within the digital computing environment the Guidebook has adopted a standard LRM measure unit of M (the Greek letter *mu*), which is equivalent to 0.001 mile. This standard will restrict measure values to whole integer numbers. It is a simple matter to convert to decimal values by dividing a value by 1,000 at the time of publication, for those users who

[35] The traditional, scale-dependent national map accuracy standard says that the minimum accuracy of a dataset should be 90% of the points being within 1/30-inch of their true horizontal position at a publication scale of 1:20,000 or larger. This standard is based on the resolution of the human eye. Applying this standard, the permissible error is ±14 feet on the ground at a scale of 1:5,000 (5,000 / 12 = 416.67 feet for 1 inch of map; 416.67 / 30 = 13.89 feet for the accuracy standard).

need this format. Some GIS platforms currently use this convention by internally converting measure values to integers. (See Measure Unit #49.)

Defined Terms

In any effort of this type, starting with a common language is a core requirement. Many of the following definitions are derived from or are taken directly from various Federal publications. Where there is a conflict or inconsistency between such sources, the AEGIST Guidebook provides a single definition.

1. All Public Roads – A public road is any road or street owned and maintained by a public authority and open to public travel. [23 U.S.C. 101(a)].[36]
    a. The term "maintenance" means the preservation of the entire highway, including surfaces, shoulders, roadsides, structures, and such traffic-control devices as are necessary for safe and efficient utilization of the highway. [23 U.S.C. 101(a)]
    b. To be open to public travel, a road section must be available, except during scheduled periods, extreme weather or emergency conditions, passable by four-wheel standard passenger cars, and open to the general public for use without restrictive gates, prohibitive signs, or regulation other than restrictions based on size, weight or class of registration. Toll plazas of public toll roads are not considered restrictive gates. [23 CFR 460.2(c)]
    c. A public authority is defined as a Federal, State, county, town or township, Indian tribe, municipal or other local government or instrumentality with authority to finance, build, operate, or maintain toll or toll-free facilities. [23 U.S.C. 101(a)][37]

   Public roads include those on Federal lands, such as Indian reservation roads, the Land Management Highway System, forest highways, and Forest Service development roads.
2. Anchor Point – A formally defined point of origin or terminus for an anchor section. Since an anchor point may be somewhat ambiguously defined as the center of an intersection or the point where a boundary crosses a route, the location of an anchor point is typically established through the use of a reference object and statement of distance and bearing from that object to the anchor point.[38]
3. Anchor Section – A formally defined path and length for a route that is used to create a 1D datum for linear referencing. Each anchor section begins and ends at an anchor point.
4. Approach Segment – The portion of an LRM inventory route leading to an intersection; a.k.a., approach. It is defined by a linear LRM object that begins at a specified distance from the intersection and ends at the outer navigation point of an intersection.[39] In the case of a divided highway, only the directional centerline(s) that take(s) traffic toward the intersection will contain an approach segment. An approach is one of the five objects listed in the Model

[36] Although the certification requirement of ARNOLD is for all public roads, the facility extent called for is all public and private roads; see, e.g., p. 49 in the *ARNOLD Reference Manual*. From a practical perspective as it is related to this *Guidebook*, there is no functional difference between public and private roads.

[37] *Op. cit*, Ch. 2.

[38] Anchor points, anchor sections, and reference objects are all derived from NCHRP 20-27(2), *Development of Systems and Application Architectures for Geographic Information Systems in Transportation* (April 2001), which provided them as a means of formally defining the beginning, ending, and official length of what is called an inventory route in this document. See *NCHRP Report 460: Guidelines for the Implementation of Multimodal Transportation Location Referencing Systems* (2001) for the results of the research effort. Anchor points and anchor segments became Framework Transportation Reference Points and Framework Transportation Segments in the proposed NSDI Framework Transportation Feature Identification Standard.

[39] When shown at smaller scales, the approach segment will terminate at the intersection point.

Inventory of Roadway Elements (MIRE) for attaching highway safety data. MIRE 2.0 lists 38 potential attributes for each intersection approach; not all are applicable at each intersection.

5. Attribute – Any characteristic that is represented by a column in a table. A table may have any number of columns (attributes). A record identifier column where each record has a unique value is a mandatory attribute for every table.
6. Barrier – A physical element or characteristic along a route that limits its use by specified vehicles. A barrier may be defined by a bridge or overhead structure, which may limit the width, height, or weight of a vehicle that may pass over or under the bridge, or by a traffic regulation. A barrier is represented along an inventory route using a linear LRM instance.[40]
7. Bike Lane – A designated portion of the travel surface that is reserved for use by persons riding a bicycle. A bike lane is part of a motor vehicle route. A bike lane may define navigation points but must not create additional intersection points. An exclusive facility for use by cyclists is a bike trail.
8. Bridge – Generically, a structure for the conveyance of vehicular or pedestrian traffic across an obstacle; a type of structure. Within the context of the NRBM and its expected use of the National Bridge Inventory (NBI), a bridge is “a structure including supports erected over a depression or an obstruction, such as water, highway, or railway, and having a track or passageway for carrying traffic or other moving loads, and having an opening measured along the center of the roadway of more than 20 feet between undercopings of abutments or spring lines of arches, or extreme ends of openings for multiple boxes; it may also include multiple pipes, where the clear distance between openings is less than half of the smaller contiguous opening.” A bridge is present on both the route going across (on) the bridge and the one going under it.
9. Calibration – The process of adding measure values to a route centerline. This is typically done by using a calibrate feature operation in a GIS platform using at least two calibration points, one for each end of the centerline.
10. Calibration Point – A point feature that provides a fixed measure value along a centerline for the purpose of controlling measure location error. A calibration point applies to a single route, but may include measure values for multiple LRMs existing at that location. Coincident calibration points may exist at some locations, such as at intersections, where calibration of multiple routes may be needed.
11. Cardinality – The relationship of two object classes to each other. Cardinality is characterized as one-to-one (1:1), one-to-many (1:m), many-to-one (m:1), and many-to-many (m:m). Cardinality may also mean the direction of increasing measure values along a roadway. The specific meaning of the term at any point in this document will have to be gleaned from the context.
12. Carriageway – An individually defined roadway surface. A divided roadway has two carriageways. The Highway Performance Monitoring System (HPMS), implementing the ARNOLD Reference Manual guidance, requires a separate centerline for each carriageway. The business rules contained herein are consistent with that guidance. Each side of a divided roadway (carriageway) defines its own inventory route, with measure values increasing in the direction of travel on that carriageway. The term ‘directional centerline’ is accordingly used in this document rather than ‘carriageway’.
13. Centerline – A linear geometric feature that defines the path of a route in 2D or 3D space. A centerline extends from an origin to a terminus. It is an abstract representation that seeks to emphasize the general path of a linear transportation facility across the surface of the Earth.

[40] Barriers represent an enhancement to the basic topological structure required for pathfinding and represent an ARNOLD enhancement; see, for example, p. 50 in the ARNOLD Reference Manual.

Such facilities naturally have a width, but that aspect of the real-world facility may only be treated as an attribute of the facility. The location of the centerline within the facility cross-section is subject to multiple considerations, the most important of which is the intended maximum map scale for which it will be used. If an inventory route is a divided roadway, then it will need two centerlines, one for each direction of travel. There are three centerline types:

a. *Segment Centerline:* the geometric feature associated with a route segment.
b. *Route Centerline:* the geometric feature associated with an inventory route. A route centerline may represent an inventory route, signed route, trail, or some other logical path (traversal) within a transportation system that uses linear referencing to denote the position of objects.
c. *Ramp Extension:* the geometric feature that visually extends the inventory route that represents a directional ramp from the end of that ramp to an apparent intersection with a limited-access highway's route centerline. A ramp extension is a purely cartographic object provided for map appearance. It does not carry specific attributes, nor does it include LRM positions. A ramp extension does not participate in route topology.[41] Figure 11 shows how the ramp inventory route centerline extends to the physical gore and ends in a ramp terminus. An at-grade intersection is placed at the equivalent location of the physical gore on the related limited-access road directional centerline. A topological connector links the ramp terminus to the intersection in order to provide continuity. This "kluge" in the basic model is one negative side-effect of representing roadways using centerlines when they are actually area features. The ramp extension is needed to provide a more "natural" appearance for map users but violates the abstraction rule for inventory routes.

14. Channelized Turn Movement – A path through part of an intersection's navigation space that is reinforced with physical barriers as a means of positive guidance for motorists. In this Guidebook, a channelized turn movement is represented by a turn segment; i.e., a directional feature provided for travel continuity and improved pathfinding algorithm performance. A channelized turn movement is not an LRM inventory route; it exists only within the navigation space of the related intersection. Channelized turn movements are typically provided for larger scale mapping products.
15. Characteristic – An attribute of a transportation facility, such as functional class, speed limit, travelway width, and number of lanes; a type of LRM object. A characteristic is cartographically represented by symbolizing or annotating the centerline feature.
16. Class – A defined type of object in a database structure; a.k.a., object class. A class is the representation of a group of similar objects conforming to a shared definition within a data model; i.e., an abstract representation of a group of similar objects that includes defined attributes and may include specified behaviors. In a relational database implementation, a class is a table consisting of rows and columns, with behaviors embedded in software that acts on the data.
17. Column – A means of expressing an object's attribute using a vertical list of values stored in a table consisting of rows and columns.

[41] The ramp's inventory route ends (or begins) at the physical gore. The ramp extension goes from that point to the appropriate limited-access highway's directional centerline following the length and path of the deceleration or acceleration lane with a diverge or merge segment to get from the end of that lane to the limited-access highway centerline. There is no LRM applicable to the ramp extension because it does not represent a separate inventory route. Everything from the limited-access highway centerline to the physical gore is part of the limited-access highway.

18. Coordinate System – A defined framework for stating the position of objects relative to each other using at least two location references (*x,y*). There are geographic and projected coordinate systems.
19. Datum – A set of rules and reference points that are used to establish the surface used by a coordinate system. The AEGIST Guidebook discusses three datum types:
    a. Cartesian – A 2D planar (*x,y*) or 3D (*x,y,z*) non-planar framework for stating a location by giving the appropriate value for two or three coordinates.
    b. Linear – The 1D reference framework for controlling the production of LRM positions.[42] Each inventory route creates its own linear datum. Positions along a route are determined by distance from an origin, as that distance is measured along the path of the facility. A linear datum may be formally defined by a State through anchor points and anchor sections.
    c. Temporal – A datum based on units of time with a beginning point and possibly an ending point. The proposed NRBM data structure is bi-temporal in that it tracks two time datum, one for the validity period of information in the database and another for the time when changes are made to the database. LRM-based data in the NRBM has one spatial dimension and two temporal dimensions.
20. Departure Segment – That portion of an LRM route leading away from an intersection; defined by a linear LRM instance contained within the navigable space of the intersection. Note that a given portion of a route may serve as both an approach segment and a departure segment. Departure segment is not a MIRE 2.0 entity and is not included in the NRBM database.
21. Divided Highway – A multi-lane facility with a curbed or positive barrier median or a median that is 1.2 meters (4 feet) or wider;[43] a.k.a., divided roadway or divided facility. A divided highway is represented by directional centerlines.
22. Domain – The range of valid values for a given attribute. A domain may be a list of valid choices or a range of numeric values.
23. Dynamic Segmentation – The process of calculating the location of LRM instances along the relevant inventory route. The inputs to dynamic segmentation, commonly abbreviated as 'dynseg', are an LRM instance table and an inventory route centerline after it has been calibrated so as to match the LRM used in the LRM instance table.
24. Edge – The topological connection between two nodes. An edge may be represented by a line; however, the spatial relationship between two adjacent nodes need not be instantiated using a geometric object. Each route segment must be represented by one or more edges.
25. Element – A component of the transportation facility, such as a bridge, sign, or guardrail; a type of LRM instance. An element may be cartographically illustrated using symbology applied to the route centerline or by using its own independent geometry.
26. Entity-Relationship Diagram (ERD) – A graphical means of describing object classes and their relationships to form a data structure. ERDs come in a variety of forms, but typically object classes are rectangles and relationships are lines connecting related object class rectangles. The object class rectangles may contain only the name of that class in a conceptual data model or be expanded to include the names and descriptions of the included attributes, along with specified behaviors and interfaces in a physical data model. ERDs are a means of expressing the organization of data and behaviors within an application but do not necessarily represent the

[42] Derived from *ARNOLD Reference Manual*, September 2014, p. 41.
[43] Ibid.

actual structure of the data in that database. As used herein, ERDs will not include behaviors and interfaces; they will serve only to express the organization of data.

27. Event –Something that happens to or along a transportation facility, such as a vehicle crash, work program project, or maintenance activity; a type of LRM instance analogous to an LRM Object.[44] An event may be cartographically illustrated using symbology applied to the route centerline through an LRM position reference or may be shown using its own independent geometry. If an actual event occurs in more than one location, it will need to be subdivided into discrete LRM objects for each location-specific component.
28. Feature – A geometric object used for cartographic display in a GIS. In a vector object class, features can take the form of a point, line, or polygon defined by vector graphics. Some features may be multi-part objects.[45] Centerlines, calibration points, navigation points, and element geometry are all features.
29. Feature Class – A group of features having a common definition and treated as a cartographic layer of similar objects that may be added or removed from a dataset using a single action.
30. Field – The intersection of a row and a column in a data table. Each field stores a value describing an attribute of the object. These values are constrained by the attribute definition, such as the data type or maximum number of alphanumeric characters, and may additionally be constrained by a stated domain of valid values.
31. Geometry – The cartographic feature created to represent an entity in the database. All geometries supported in this Guidebook are vector graphics; they include point, line, and polygon and may be multipart. Geometries are abstract and cannot reflect all aspects of the entities they seek to represent.
32. Globally Unique Identifier (GUID) – A random 32-digit hexadecimal number that has a very small chance of being duplicated. A hexadecimal number is one using Base-16, where the decimal digits of 0-9 are supplemented by the capital letters A-E, which represent the decimal equivalents of 11-15. Counting in hexadecimal, the number 10 = 16 in decimal (Base-10). A GUID contains 128 bits of information. The possible number of GUIDs is 10 to the 38th power. The 32 hexadecimal digits are grouped into segments of 8, 4, 4, 4, and 12 digits. Most database management systems can generate GUIDs.
33. Graphical Connector – A linear feature constructed to provide a more realistic appearance, typically created to connect the end of a ramp inventory route with the appropriate limited-access highway directional centerline. LRM measures are not provided for graphical connectors.
34. Highways – Roads, streets, and parkways and all their appurtenances (23 U.S.C. 101).[46]
35. Identifier – The name given to an object so as to uniquely identify it. In a database, the identifier is called the primary key and may consist of multiple fields.
36. Instance – A single object; a single row in a table.
37. Instantiation – The act of creating an example of a real-world or conceptual entity, such as an object or a row in a table.

[44] The term ‘event’ has traditionally been used to describe all data elements that are tied to a route using a measure value; i.e., everything was an event. This practice was formalized by NCHRP 20-27(2), which is a widely used foundation for LRM and transportation database design. This Guidebook substitutes the term ‘LRM Object’ to refer to the collection of all entities and attributes in the database that may be given an LRM position description.

[45] This document limits itself to the types of geometry typically found in highway inventories; i.e., vector graphics. GIS also supports raster, or surface, data structures. Data in such a form is to be converted to its vector equivalent for use in the ARNOLD dataset; i.e., to a point or linear event located along a route. As noted elsewhere, facility elements may additionally have their own geometry.

[46] Ibid.

38. Interchange – A point on one or more limited-access highways where traffic may move from one route to another. An interchange is an inventory route element. Each interchange must include at least one ramp (acceleration or deceleration) and two intersections (both ends of the ramp). MIRE 2.0 lists 25 attributes for interchanges. All intersections and ramps within the interchange should be treated as internal components in an object-oriented implementation of the NRBM database design.
39. Interchange Point – A geometric point feature that is used to represent the general location of an interchange. The location of the interchange is defined by an LRM position for only the limited-access highway, but the actual interchange point feature is not cartographically placed on a route centerline. When more than one limited-access highway is involved at a single interchange, each such highway will have an LRM position recorded for the interchange's location on that highway. The LRM position description of the interchange is used by many States to create an interchange number within the context of the milepost signage LRM.
40. Intersection – A point where vehicles and/or pedestrians may move from one route to another. Intersection is an inventory route element and a MIRE-defined object to which up to 28 attributes may be specified. An intersection is represented in the transportation network's topology by a node. Intersections may be characterized as being one of the following types:
    a. *At-grade:* where roadways are all in the same plane and movement may occur between routes at one or more internal navigation points.[47]
    b. *Continuity:* where a jurisdiction may create an intersection to represent the transition from one LRM inventory route to another, such as in a curve where route identifiers change, or to link a deceleration/departure route to an LRM inventory route at the virtual gore point. A continuity intersection is also used to mark the divergence or convergence of a centerline and two carriageways. Some continuity intersections mark the boundary of county/parish jurisdictions,[48] which form the limits of inventory routes. When a county/parish boundary is also a State boundary or a supplier boundary, then the appropriate intersection type (i.e., State boundary match point or supplier boundary match point) should be used instead of a continuity intersection.
    c. *Median cut:* a location along a dividing median where traffic may cross the median and/or execute a U-turn. A median cut intersection does not subdivide a route segment; it occurs along a route segment. A median cut intersection is located on each route segment of a divided roadway. A topological connector is not provided to link the two median cut intersections for use by routing applications.[49]
    d. *Multimodal interchange point:* a connection from a roadway to another mode of travel, such as a vehicle ferry loading dock. A multimodal interchange point is used when

---

[47] There was some consideration given to creating two at-grade intersection subtypes, simple and complex, in order to recognize the special case of offset intersections as defining the complex subtype. Such an intersection uses a single traffic control mechanism to manage traffic flows from two Tee intersections on opposite sides of a main road that are far enough apart to be considered as separate LRM objects but not far enough apart to utilize independent traffic control. It was ultimately decided that these intersections actually represent a traffic control strategy; i.e., one traffic signal may control multiple intersections. The object-oriented implementation of the NRBM model does include simple and complex subtypes, with the complex intersection class also having a roundabout subtype.

[48] Some States do not have counties or parishes that may serve as logical limits for inventory routes. Each State is therefore empowered to select the political and other jurisdictions that will result in subdivisions in inventory routes, or to decide that sub-State breaks in inventory routes will not be imposed.

[49] If a path is not provided to nearby intersections or is not possible except through the median cut, then the median cut will need to be treated as an at-grade intersection and become a route segment terminus.

continuity of travel is provided for the same vehicle. A multimodal terminus type is appropriate when the nature of travel requires a change in the transporting vehicle, such as a bus terminal or airport.

e. *Railroad grade crossing:* a point of interaction between railroads and other modes of travel; a.k.a., highway-rail crossing.[50] A divided highway will cause two adjacent intersection points to be created so that information about the shared railroad grade crossing may be discovered.[51]
f. *Ramp junction:* the point at which the inventory route representing the ramp topologically connects to the limited-access highway. The location of the ramp junction corresponds with the ramp's physical gore. A topological connector links the ramp junction intersection on the limited-access highway to the ramp merge/diverge point terminus for routing applications.
g. *Roadway access:* such as needed for a major driveway that has the characteristics of an at-grade intersection but involves only one route. Traffic control, crash instances, and other traffic data are typically provided. Intersections of this type do not automatically break route segments.
h. *State boundary match point:* serves as the outer boundary of a given State's roadway system.
i. *Supplier match point:* provides a boundary connector to an adjacent data partner's centerline feature to provide path continuity across the boundary.

41. Intersection Point – A geometric point feature that represents the general location of an intersection. An intersection point need not be located on an inventory route but its position within the LRM of all involved inventory routes will be recorded as an LRM object. An intersection point defines the end or beginning location of at least two route segments.
42. Inventory Route – A linear transportation facility defining a 1D LRM datum. The only means of stating a location along the inventory route is by giving its distance along the route as measured from the point of route origin. An inventory route is composed of one or more whole route segments. The route for which the applicable inventory data is to be recorded, as opposed to a named or numbered route. In the case of a bridge, intersection, or other object that may involve more than one inventory route, information about said object must be related to all relevant inventory routes.
43. Linear Referencing Method (LRM) – A means of establishing a position description along a transportation route using a combination of a route identifier and a measure. In most instances, the measure represents the distance from the route's point of origin to the location in some specified unit of measure, such as miles, feet, or meters. Typical methods used are milepoint,

[50] MIRE 2.0, available at from FHWA at https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa17048.pdf, implicitly defines highway-rail crossings as a type of intersection by including Railroad Crossing Number as an intersection attribute, but does not explicitly list it as an intersection type (pp. 87-88). It also includes railroad crossing warning devices as intersection traffic control types (pp. 99-100) and as intersection approach control types (p. 118), but fails to include any examples of highway-rail crossing geometry for intersections or approaches. The distance from a highway-rail crossing to a roadway intersection may be a useful future addition to MIRE data elements.

[51] Guidance regarding the number and location of railroad grade crossing intersections is found in USDOT Publication DOT/FRA/RRD-23, *Federal Railroad Administration Guide for Preparing U.S. DOT Crossing Inventory Forms*, March 2015. In general, a highway-rail crossing is defined by the location of warning devices; i.e., all tracks located within the protected area are considered to be a single crossing. Note that a grade-separated railroad crossing provided by one or more bridges is still treated as a railroad grade crossing intersection.

milepost, reference point, and link-node.[52] The LRM chosen for NRBM is milepoint, which means that LRM position measures are stated in terms of a distance, in units of 0.001 miles, from the route's point of origin, and that the distance between two points on a route is the mathematical difference in the measures defining those positions in the LRM. However, a milepost LRM is typically used for Interstate highways and other limited-access roadways. The milepost values generally begin at a State boundary or signed route origin. Milepost LRMs should also be provided as a separate linear datum.

44. Linear Referencing System (LRS) – A rule-based framework for managing and applying LRMs. Such rules include how to calibrate field equipment, the various workgroup responsibilities for each LRM and their use, and the governance structure that manages changes in organizational LRMs. The LRS should align LRM positions in all measures so that information from various business groups utilizing different LRMs can be accurately mapped.[53]
45. LRM Object – Any characteristic, element, or event that has an LRM-based position description. An LRM object is an example of an LRM instance, which is anything that can be given a location description using linear referencing. LRM objects are those LRM instances that are included in a database. LRM objects come in point (one LRM-based position) and linear (two LRM-based positions) types.[54]
46. LRM Position – A description of a location using a linear reference consisting of a defining inventory route identifier, a measure value in whole *M* units, and, if there are multiple LRMs, an indication as to which 1D reference system applies to the LRM position. An LRM position may be stated as an "at" LRM location using a single measure or as a "from" and "to" LRM location using a pair of LRM measures.
47. Location – The place where an object exists in the real world. An object has one location but that location can be described by multiple positions.
48. Measure – A number that represents a position along a route, as defined by the applicable LRM and its 1D reference datum; a.k.a., a route measure. A measure value cannot be stated without also providing the related route identifier and, if there is more than one, the applicable LRM datum. A measure value is not absolute and represents only a relative position found between two calibration points. A measure (*m*) may be stored for each centerline vertex as one of several coordinate variables; e.g., a GIS software shapefile allows the coordinates of *x,y,m,z.*[55]
49. Measure Unit – LRM measures are stated in terms of a unit of measure, which is 0.001 mile (5.28 feet). This measurement unit has been converted to integer values and is symbolized by the Greek letter M (*mu*). LRM measures stated using an integer M value can be converted to miles by dividing by 1,000; i.e., 1,000 M = 1 mile. Using integer values eliminates the difficulty of working with floating-point decimal numbers and the resulting potential for changes in precision (number of decimal places) between applications. Measure values may continue to be stored using decimal values, but the resulting mathematics will not produce aberrant answers.
50. Median – Any physical or regulatory barrier, except a single- or double-painted centerline, that subdivides a roadway into directional flows. Median types recognized by the NRBM include

[52] *HPMS Field Manual*, Ch. 2.

[53] Paraphrased from a quote taken from the AASHTO Innovation Initiative-Line Referencing System in the *ARNOLD Reference Manual*, p. 41.

[54] This Guidebook uses 'LRM object' as a preferred term over the more historically common 'event'.

[55] This is the order of coordinates in the shapefile specification; i.e., *xy, xy, xy, xy, …, m, m, m, m, …, z,z,z,z….* There is also a vertex ID field available that follows the *z* coordinate string.

‘Unprotected’, ‘Curbed’, ‘Positive Barrier’, and ‘Continuous 2-way Left-turn Lane’. [56] Each median should be represented in the dataset by a median element with an LRM position defined on the directional inventory route located on each side of the median.

51. Multiplicity – A relationship descriptor on ERDs that says how many of the object class may be associated with a single instance of a related object class. Multiplicity is typically expressed using a single number or a range of numbers with ‘..’ placed between the lower and upper limits of the range; when there is no upper limit to the number, it is shown as *m*. Examples include 1, 3..9, and 0..*m*. When the number 0 is shown as the lower limit, it means the existence of an object of the class is optional for any given instance of the related class. Some texts use an asterisk (*) instead of *m* to show there is no upper limit, as in 4..*.
52. Navigation Point – A junction within a navigation space where travel choices are made by a driver or pedestrian. Some navigation points mark the location of entry into an intersection and its navigation space. Turn segments must terminate at a pair of navigation points. An at-grade intersection of two or more bi-directional route segments has a single navigation point coincident with the intersection point.
53. Navigation Space – The area within an intersection where vehicles and/or pedestrians may select from multiple paths through and out of the intersection. A single intersection may contain only one navigation space. Due to inherent limitations of the LRM concept and the related scale of abstraction, the NRBM does not see into navigation spaces. To the LRM and the NRBM, an intersection is a singularity with no internal structure. However, MIRE and other safety and operational analyses have to look at the internal structure of an intersection. To meet this need, the NRBM supports the use of separate cartographic treatments for revealing this internal structure, which may include a navigation space containing navigation points and turn segments for all but the simplest intersections.[57] Intersection analyses may link approach segments to the navigation space.
54. Network – A topological representation or model of the transportation system. The NRBM topology requirement is provided as a means of providing quality assurance for the accurate representation of route segments using segment centerlines and forms a rudimentary pathfinding structure for LRM objects.
55. Node – The topological termination of an edge. A node may represent a decision point in a transportation network; i.e., an intersection or a terminus.
56. Object – Any real world or conceptual entity.
57. Object Class – See Class.
58. Object-Oriented Programming – A means of constructing databases and applications in the form of discrete object classes that have defined behaviors expressed to the outside world by one or more interfaces. The internal structure of data and software is not visible to the user. An interface specification describes the inputs to the class, called arguments or messages, and the output of the class, called results. What happens inside the class is called the method. The three

---

[56] This list of median types differs from that currently contained in the *HPMS Field Manual*, which explicitly says that “a continuous turning lane is not a median” (Item 35 – Median Type). However, participants in the various AEGIST peer exchanges supported the change in definition to include continuous two-way left-turn lanes, which meet many of the listed purposes of a median found in the *HPMS Field Manual*. These include minimizing interference with opposing traffic, providing a recovery area for out-of-control vehicles, offering a stopping area in case of emergencies, and serving as a storage area for left-and U-turn vehicles. HPMS also supports four versions of the Positive Barrier option: ‘Unspecified’; ‘Flexible’; ‘Semi-rigid’; and ‘Rigid’. However, these are optional values and may not appear in all State datasets; therefore, the NRBM will use the domain containing only the four mandatory values. The default value of ‘Positive Barrier’ is equivalent to ‘Unspecified’.

[57] A simple at-grade intersection formed by two bi-directional roadways has a single navigation point and no turn segments. The navigation space is defined as the polygonal area of the intersection, which includes pedestrian crosswalks.

principles of object-oriented design are encapsulation, polymorphism, and inheritance. Encapsulation means that everything regarding the data and behavior of the class is inside the class. Polymorphism means that the user can send the same message to two object classes and get different results due to the unique methods of each class. Inheritance means the programmer can create a “child” class from a “parent” class, where the child class inherits all the interfaces and methods of the parent, and then adds new interfaces and/or methods. The creation of a class is called instantiation. A class may be abstract in that it can serve as a concept for a class but is not actually instantiated itself. “People” is an object class. ”Person” is an abstract object. “You” are an instantiated object in the People class. What you do is your method. How you interact with the outside world, both in perception and expression, is through interfaces. The memories you contain and the logic process you use to act in response to inputs are the data encapsulated in your brain. How your body works is not exposed to the outside world; it is encapsulated. You inherited certain data and behaviors from your parents.

59. Position – A location stored in a database using a description conforming to specified rules. For example, a bridge’s location may be described using geographic coordinates or a route/LRM measure. A position is a relationship between a location and a datum. Time may also be a part of the position description, as in when the object was at this location.
60. Primary Key – The attribute(s) that are used to define a single instance of an object class. Primary keys are generally generated by the database management software and serve only to define a single object in the class. A single attribute is used to form a simple primary key. Multiple attributes are used to form a complex primary key. All object classes included in this document use complex primary keys in order to separate records that differ by time but describe the same object. Minimally, both the object identifier and the record date are needed to select a single record.
61. Projection – A mathematic process of converting a spherical coordinate system to a planar system, such as may be used to convert locations on the surface of the Earth stated using latitude and longitude values to locations on a flat piece of paper for mapping purposes. Projections typically seek to preserve shapes and/or spatial relationships.
62. Public Key – An attribute that meets the general requirements of a primary key but is more readily usable by humans; a.k.a., a candidate primary key. Typically, a public key has “intelligence” or a means of being read or decoded by humans so as to both identify the record and the entity it represents.
63. Ramp – A special inventory route that represents a path of entry to or departure from a limited-access highway at an interchange.[58] Typically, one end of an inventory route that serves as a ramp will coincide with a ramp junction point feature. The related intersection point feature is placed on the related directional centerline at right angles to the physical gore. The LRM measure for the ramp intersection is determined by this location. See Figure 11 through Figure 13 for more information.
64. Ramp Extension – See Centerline, Ramp Extension type.
65. Raster – A model of a surface where the area of the surface is subdivided into equal subareas, called pixels, that contain the attributes of that subarea. Pixels are arranged into rows and columns. The position of a single pixel is defined by its row and column identifiers.
66. Record – An object stored in the database, such as a row in a table.

[58] Any ramp conveying traffic in two directions shall be treated as a pair of directional routes and must be represented by directional centerlines.

67. Record Date – The date a record was created. Every time the value of an attribute is changed in a given record, a new record date value is written.
68. Reference Object – A readily recoverable object in the real world that can be used to establish an LRM position.

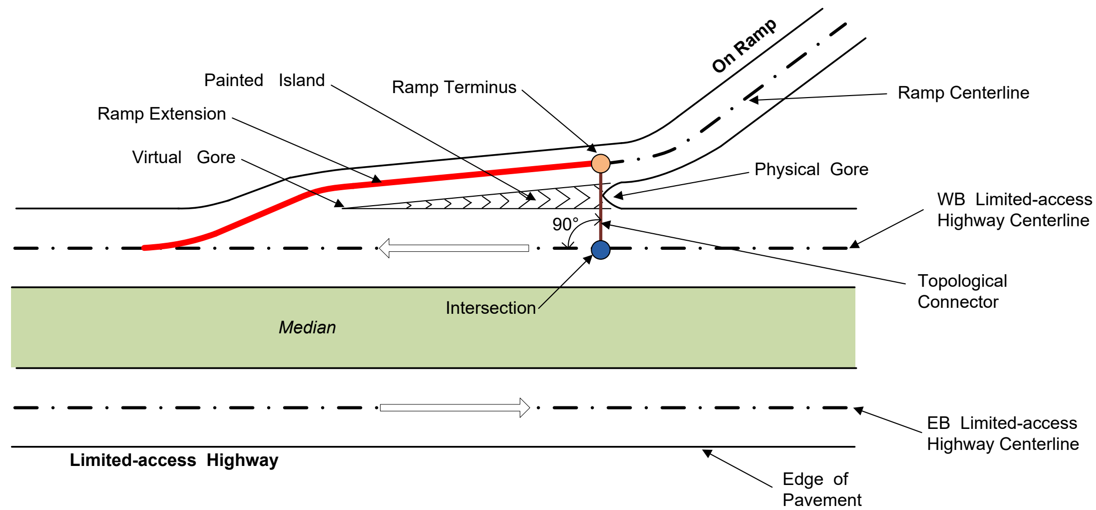


Figure 11. Comprehensive view of interchange objects. Of these several objects, the topological connector linking the ramp terminus with the ramp junction on the limited-access highway centerline is the one that would normally not appear on a roadway map.

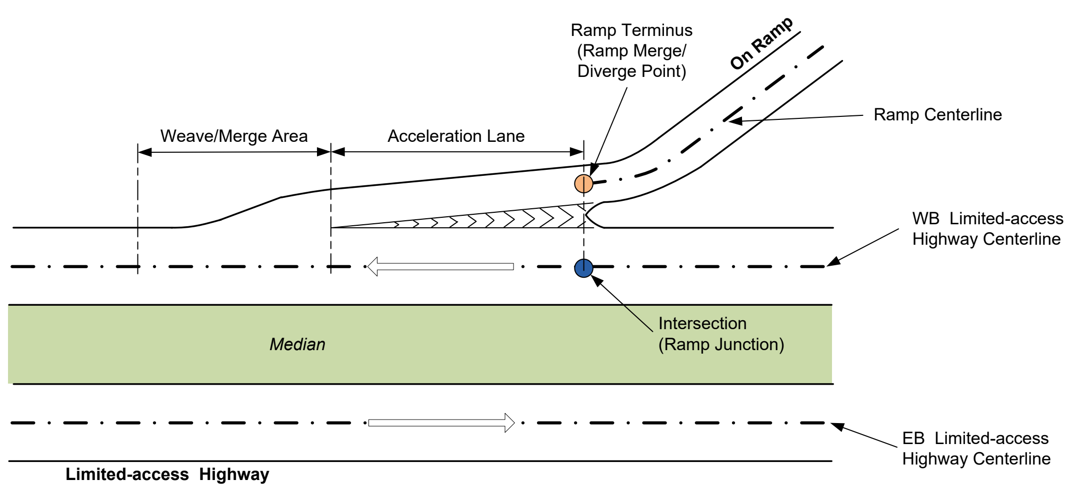


Figure 12. LRM interchange objects from Figure 11: one ramp centerline and two limited-access highway directional centerlines illustrating the path of three inventory routes, along with the ramp terminus point (merge/diverge type) and intersection point (ramp junction type).The weave/merge area and acceleration lane labels show possible linear LRM instances that would be characteristics of the westbound limited-access highway route.

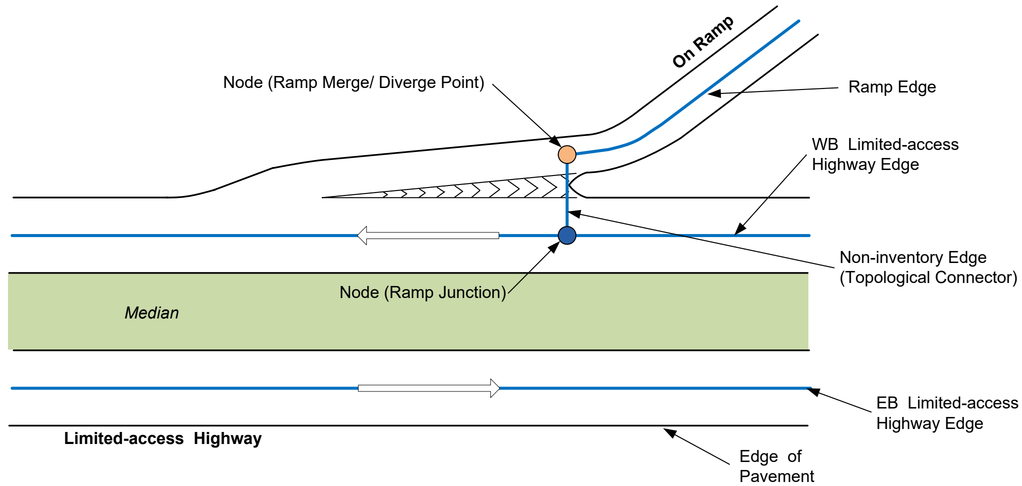


Figure 13. Topological objects from Figure 11. Route segment and edge breaks occur at the ramp junction node on the westbound limited-access highway inventory route centerline.

69. Restriction – A potential impediment to travel. Examples include bridges, tunnels, gates, and time-of-day limitations to travel for all or some vehicles. The routable aspect of topology in the NRBM includes only physical limitations to travel, such as may be presented by weight limits on bridges.
70. Roadway – The portion of a highway intended for vehicular use.[59] The term is all encompassing and does not imply the presence of a prepared surface; a.k.a., road and street.
71. Roundabout – Any of several configurations of an at-grade intersection where vehicles operating along a directional rotary may move from one route to another; a.k.a., traffic circle or rotary. A roundabout will contain multiple internal navigation points and turning segments but is represented by a single intersection point (LRM instance) in the NRBM.[60]
72. Routable – The characteristics of a topological network with the ability to determine whether any two points—an origin and a destination—are physically connected by a continuous sequence of route segments. A routable network has the ability to determine whether a path is possible and to identify those route segments that form the path. Travel cost is determined by the length of each route segment and (optionally) other user-selected attributes. Travel restrictions, such as bridges and tunnels, may be present on various route segments that can cause them to be eliminated from the transport network prior to conducting the routing analysis. This is the simplest aspect of pathfinding, which includes the ability to provide turn-by-turn directions to a traveler as to how to navigate the network to get from the origin to the destination. Routability depends on the route segment choices available at each intersection. As a result, internal components of an intersection, such as navigation points and turn segments, are not included in the routable network. The routing application should be responsive to restrictions present on route segments.

[59] Ibid. The *HPMS Field Manual* interchangeably uses the term 'road' and 'roadway' but provides a definition only for 'roadway'.

[60] Some States have segmented and digitized their base inventory routes to reveal what the NRBM specification treats as basic or foundational objects contained within the point intersection. To accommodate these States, an intersection may be treated as a linear LRM instance that spans the distance along an inventory route between the outer navigation points for an intersection.

73. Route – The linear transportation facility defining the datum for determining measure values along the facility by applying an LRM. A route is known by its identifier. A route, as defined here, is not synonymous with a numbered or named route, which may be treated as a static traversal. A route may be illustrated cartographically using a single centerline and, in the case of a divided highway, by two or more carriageways. There are the following recognized route types, although others may be added for specific subsets of the NRBM dataset:
    a. *Inventory:* the basic entity for storing roadway data. Inventory routes must include an LRM. An inventory route consists of one or more whole route segments. It defines a path through part of a transportation system, including such transitory traversals as a bus route, where time, rather than geographic distance, is the means of defining a location as a 1D LRM position.
    b. *Named Route:* a published traversal that follows the path of a linear transportation facility identified by name or assigned route number. The path of a named route is created by selecting all linear LRM instances of one or more inventory routes with the desired name or route number. A named route should consist of a sequence of route segments.
74. Route Centerline – See Centerline, Route type.
75. Route Segment – A transportation facility path extending from an intersection to an intersection, from an intersection to a terminus, or from a terminus to an intersection. The ordering of termini defines the topological direction of the route segment. A route segment is an atomic element of an inventory route; i.e., an inventory route consists of one or more route segments. Each route segment must be represented by one edge and terminate at two nodes. Cartographically, a route segment is represented by a segment centerline, which may be part of a directional pair. A route segment is represented in the transportation network topology by an edge. When representing the path of a divided roadway, one or two topological connections may be constructed and added to the end of the route segment (as defined by the inventory route's LRM measures) to the intersection point.
76. Segment Centerline – See Centerline, Segment type.
77. String – A data type that allows alphanumeric characters to be used to provide a value given for a defined attribute. The maximum number of such characters is usually specified in parentheses.
78. Table – An organization of data presented as a set of rows and columns. Each row describes a single object. Each column describes a single attribute of such objects. Note that the actual storage of data need not conform to the table structure, which is a means of showing data to a user.
79. Terminus – A route segment beginning or ending point that is not a roadway intersection. The following are terminus types:
    a. *Cul de sac:* the end of a roadway with an optional area for vehicles to turn around. The terminus object class includes an attribute for describing the additional pavement area available for vehicle maneuvering stated as a radius from the end of the centerline. If the roadway terminus is a dead end without additional pavement width, then the cul de sac radius will be zero.
    b. *Map limit:* the edge of the NRBM dataset. Primarily, these are the end of a roadway centerline at an international boundary when that roadway continues into the adjacent country. A map limit does not represent the end of the roadway, only the limit of the NRBM dataset's ability to describe the roadway.
    c. *Ramp merge/diverge point:* the point where a limited-access highway ramp begins or ends at the physical gore. A topological connector links this terminus to the related ramp junction intersection on the limited-access highway.

80. Terminus Point – A geometric point feature that represents the location of a terminus. A terminus point may only be coincident with a single route segment.
81. Topological Connector – A linear feature that is created to connect the end of a route segment's centerline to the actual intersection point that represents the intersection and the point where all connecting route segments must terminate. Topological connectors are needed only on divided roadways. One or two topological connectors can be used to extend the centerline of a route segment to the intersection point in order to provide topological connectivity. Doing so observes the general rule that all route segments begin and/or end at an intersection.
82. Topology – Within the context of the NRBM standard, topology is the connectivity of linear facilities at points of intersection or termination. Topology is provided only for route segments, which may be defined by the path between two intersections or one intersection and one terminus. Topology is not provided for inventory routes or ramp extensions.
83. Trail – A linear transportation facility devoted to travel by unlicensed vehicles, ridden animals, and/or pedestrians. Trails may be limited to use by pedestrians alone or to specified subsets of motorized and non-motorized vehicles. Route segments and inventory routes may be defined for trails. Trail heads and crossings involving one or more inventory routes should be noted by an LRM object and may define an intersection.
84. Transformation – The mathematical process of converting positions from one reference framework to another.
85. Travelway – That portion of a transportation facility that is intended for the movement of people and/or vehicles. The travelway does not include shoulders but does include bike lanes and other portions of the facility dedicated to the movement of designated vehicles.
86. Traversal – A logical path through a transportation system. A traversal may be stable, as in the case of a numbered route or named street, or it may be transitory, as in the case of a path used once for a specific journey. An LRM may be created for any traversal by transforming the original LRM positions using a mathematical process.
87. Tunnel – A structure conveying a traveled way beneath an obstacle. In the NBI, a tunnel is recorded as a bridge, with the route going through the tunnel begin listed as being under the bridge.
88. Turn Segment – A path between navigation points within a navigation space. Movements along turn segments may be restricted by vehicle type, time of day, etc. Turn segments have no LRM length. Most turn segments are short pathways internal to traditional intersections.
89. Vector Graphics – An abstraction of entities in the real world using simple geometric objects (features), typically points, lines, and polygons. A point feature is defined by a single set of coordinates. A line and the boundary of a polygon are described using one or more ordered coordinate sets. The actual structure of vector graphics is specific to each vendor, although some general standard methods exist. Typically, the coordinates of the first vertex (origin) in a sequence for lines and polygons is fully stated in terms of the applicable coordinate system and subsequent vertices are stated as offsets from the origin. Some GIS platforms support multipart vector objects. Attributes are associated with each vector feature.
90. Vertex – A single coordinate set describing a geometric location in a coordinate system (plural – vertices). A vertex must be stated using all the required coordinate axes. A point consists of a single vertex. A line consists of two or more vertices. The boundary of a polygon must include at least three vertices.

Business Rules

The intent of the AEGIST Guidebook is to show agencies how to construct a national transportation map and dataset (NRBM) in a decentralized manner by using contributions from the States, Tribes, and Federal agencies that conform to the specification provided herein. Local and regional government agencies may serve as original sources of information supplied by the States. The best way for these different data sources to provide a consistent product is for data providers to adopt a common set of business rules and implementation policies.

Through an extensive literature review, three online peer exchanges, a face-to-face peer exchange, and other mechanisms, a series of core business rules have been developed that represent a foundation for States and their partners to build the NRBM dataset in a manner compliant with ARNOLD, MIRE 2.0, and other Federal requirements. The business rules will be used to make database design decisions and to specify software functional requirements.

Following are the core business rules and implementation policies:

1. Although the initial implementation is expected to be through HPMS data submittals by the States and other contributors, the long-term goal is to provide the information via web services using an online portal supported by the USDOT.
2. The proposed NRBM data structure must—
    a. Support facility inventory using linear referencing.
    b. Provide basic topology (routability).
    c. Be vendor independent.
    d. Provide the means for multiple modes of travel to be included.
    e. Be constructible over time in a modular fashion.[61]
3. The data editing and data usage environments must be separate—
    a. Data editing must include data checking and version certification capabilities and practices.
    b. Data users should be able to subscribe to a custom update cycle (e.g., daily, monthly, once per year) that conforms to the needs of their application.
    c. Automated data extraction, transformation, and loading (ETL) should be provided in the data publication process.
    d. Users should be able to download only those changes that were made to the dataset since their last download; i.e., data transfers should be updates that alter only the records added, deleted, or modified by source data edits during the period.
4. The enterprise LRS should be extensible to include all entities in the facility inventory.
5. An LRS must provide four essential functions:
    a. Locate – establish the location of a point in the field by reference to identifiable objects.
    b. Position – translate or express the location of a point to a database position. There can be many position descriptions in the database for a single real-world location.
    c. Place – convert a database position description into a real-world location.
    d. Transform – convert a position description in one LRM to a position description for another LRM.[62]

[61] The initial focus is on roadways carrying motorized vehicles, but the database needs to be extendable to include other modes of travel.

[62] Derived from Vonderohe, Alan, et. al, A Generic Data Model for Linear Referencing Systems – Findings of NCHRP Project 20-27(2), September 1997; p. 6.

6. Organizational issues, such as creating match points and assigning globally unique route identifiers, must be addressed by technical solutions that do not depend on inter-agency communications. Each State will have to be the arbiter of boundary issues and the location of match points within its territorial jurisdiction. FHWA will be the arbiter of State boundary match points.

Each object in the NRBM will be given a nationally unique identifier as a public key. An example for creating nationally unique identifiers using Federal Information Processing Standard (FIPS) codes is shown in Figure 14. It concatenates the two-digit State FIPS identifier and the three-digit county/parish FIPS identifier to create a geographic location at the county/parish level.[63] The State DOT will assign an agency identifier to each data supplier, including itself. The data source will assign an internally unique identifier for each data object it supplies. This example assumes the proposal to subdivide inventory routes at county/parish boundaries is implemented. Note that these identifiers are public keys for data publication and use, not primary keys for data creation and maintenance.[64] Source datasets should use GUIDs generated by the data editing software as their internal identifiers.[65]

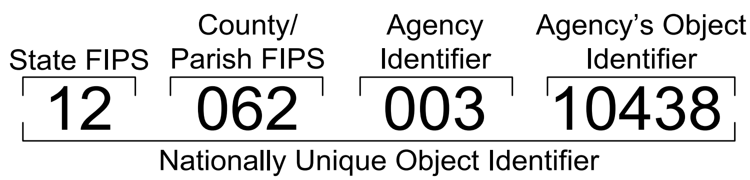


Figure 14. Intelligent key components of the proposed generic public key.

7. Any data provided by a source must be returnable to that source without alteration of any value contained in the source dataset.
8. The data structure must support multiple ways of cartographically presenting transportation data, such as at different scales or levels of abstraction, without requiring the duplication of the related business data.
    a. Linear facilities, such as named routes and bike lanes, may be displayed using one or more centerlines.
    b. At larger scales, a single bidirectional centerline may be decomposed into directional centerline and/or mode-specific centerlines.
    c. Elements that are affiliated with a route, such as bridges and noise walls, may need their own geometric representations that are independent of the route centerline.

---

[63] The need for a break at a county boundary, or its functional equivalent particular to a given State, is to support several needs, particularly those related to administrative units and their impact on data production and usage. The first is to make it possible for metropolitan planning organizations, local governments, and other sub-State groups to be able to readily extract a subset of the State dataset by simply selecting one or more county FIPS identifiers. The second is to facilitate route segment maintenance by the States, which often use county boundaries for defining their administrative units and the responsible unit for data maintenance. The third is to more readily support census applications and others that are organized along county boundaries. The fourth is to provide a public key creation process that incrementally produces a globally unique value.

[64]The first three identifier components are fixed length. The leading zeroes in the example show that these various components and the resulting national identifier are string data values; i.e., they are alphanumeric characters, not numbers.

[65] An agency's object identifier is not limited to the five characters shown. The current NRBM specification suggests the agency's object identifier be up to 32 characters in length, with shorter identifiers being right-justified and zero-filled. This is because a GUID requires 32 hexadecimal digits. Any data supplier using GUIDs will need 32 characters to store their internal identifier. Support for hexadecimal numbers is why the public key attribute must be a string data type.

   d. Intersections and interchanges may be shown at a small scale as a point feature, and at a larger scale as a collection of component turn segments and navigation points.
9. Temporality is a vital component of spatial data that must be provided in the data structure through record-level metadata. Two temporal vectors are needed, one for the time when records are created and another for the validity period of each record. These temporal vectors should provide transactional update capabilities, wherein a user may download only those records that have changed since the last update, and allow users to track the following:
   a. Edits made to the data and by whom they were made.
   b. Changes over time, such as to show historical views of the transportation system or to present future views resulting from planned projects.
10. The state of the transportation system, as presented in the database, must be recoverable for any point in time contained within the historic and future scope of the dataset. Active records may reflect future changes to the roadway system, such as the result of planned construction projects. An analysis of future conditions and changes over time will need to access data about past and future roadway conditions.[66]
11. Updated information may be delivered to the editor in random manner, with earlier updates being received after later updates have been made. The data editing mechanism must allow asynchronous edits; i.e., edits should be allowed out of chronological order.[67]
12. Accuracy, precision, and resolution standards in a dataset must be well defined and enforced.
13. Each piece of information should be contained only once in the editing environment, but may be published in multiple forms.
14. The database design must provide network topology at the route segment level. Such topology is provided for quality control purposes and is intended to serve basic applications for routability; i.e., is it possible for a given vehicle to traverse the system between the proposed origin and destination? More refined applications of topology, such as to provide a turn-by-turn narrative of a shortest-time route, will require commercial datasets.
15. In order to support basic routability applications, data related to bridges, tunnels, and other restrictions to travel must include the route traveling across (on) the bridge, the route traveling under the bridge, the maximum weight for travel across the bridge, and the maximum vehicle dimensions (width and height) for traveling over and under the bridge.[68] The location of the bridge relative to the involved routes must be defined by an LRM instance for the route traveling across the bridge and the route traveling under the bridge. The bridge structure itself should be illustrated by a polygon (independent element geometry).
16. Each agency implementation of a transportation facility database must contain a data dictionary that can be conveyed to any user as a means of describing the database contents.
17. The data dictionary must include record-level metadata so the user can properly interpret the data contained in any temporal view of the database.
18. The data dictionary must include the members of any defined-value domain and the valid domain range for a numeric field. This data must be temporal so that the correct values apply to any time-specific view of the transportation system.
19. A record may have one of the following three record status values:

---

[66] Temporality is applied by using the From and To dates for record validity, whether a currently active record showing the latest information or a retired record that was valid for a past period of time. The From date for the replacement record should match the To date for the older record.

[67] To allow asynchronous editing of records, data maintenance staff will need to have access to retired records so that the from and to dates may be modified to fit the more recently received information.

[68] These attributes, which may impact the routability of over-dimension and overweight vehicles, are not directly available in the NBI but may be initially derived from that source pending field measurement.

   a. *Active:* the record describes the entity as it exists today; a From Date value must be provided and the To Date value is normally null. The From Date value may be in the past or the future. The To Date value may be non-null when the record reflects the planned future end of life for the entity. There is normally a single active record for each entity in the database. Multiple active records may be in the database for a given entity when a chronological sequence of future states is being described.
   b. *Replaced:* the original record was subsequently found to be in error and does not represent the real world at any time in history; both the From Date and To Date values must be null.
   c. *Retired:* the record validly represented the state of the entity for an historical period of time, but is no longer the current state of the entity; valid entries are required for both the From Date and To Date attributes.
20. The status of the record, described above, is distinct from the status of the entity it describes. For example, there may be an active record that describes a future roadway segment. Entity status domains vary by entity type.
21. Edits that must be supported by software tools include the following:
   a. Add a new record (includes marking the record as active with a starting date of validity).
   b. Delete a record (the record is marked as replaced and the time period of applicability is deleted; the record itself is not actually deleted, but it will not be returned in a result set when searched by date).
   c. Modify an existing record (includes retiring the old record and adding a new active record with appropriate time stamps and editing information).
22. Modifying an existing record may require the following:
   a. Changing the value of a field in the record.
   b. Changing the LRM-based position of the described entity.
   c. Subdividing a linear entity.
   d. Extending a linear entity.
   e. Shortening a linear entity.
23. Within the LRM, each inventory route must be assigned a unique identifier, a starting measure value, an ending measure value, and an official length. The enterprise data system must organically provide a means of making inventory route identifiers globally unique. Additional points may be defined along the route by stating a declared measure value, along with optional information, such as a geographic coordinate location, for calibrating measures.
24. Inventory routes will be subdivided at county/parish boundaries for consistency with other datasets, to allow ready selection of subsets by political jurisdiction, and to provide reasonable inventory route lengths in rural areas. Other points of subdivision may be determined by the States, such as to reflect the use of a municipal data source. A State's NRBM submission would consist of a number of whole counties/parishes.
25. Route segments are linear elements on inventory routes as a means of providing limited roadway data in discrete linear pieces, supporting topology, and controlling the extent of cartographic edits to route centerlines.
26. An inventory route consists of one or more whole route segments. Cartographic editing of route inventory centerlines will result in one or more whole route segments being retired or replaced.
27. With the limitations described below, an inventory route's path must be arbitrarily determined. While an initial path may be defined using various attribute values, such as the traversal of a numbered route across a county, once the path is determined, it is completely separate from the means by which it was initially defined. Changes in the attributes used to initially define a route's path must not alter the path of the route over time. Physical changes to the route, such

as may occur due to construction or abandonment, will generally result in retirement of a portion of the route and the creation of a new route.

28. An inventory route must not overlap another inventory route; i.e., each route segment can be part of only one inventory route. Signed routes may overlap, but they are constructed from a sequence of inventory route segments formed by linear LRM objects.
29. Inventory routes must not bifurcate; i.e., a single inventory route identifier cannot be assigned to divergent paths. Each path must be its own inventory route.
30. Inventory routes may be self-intersecting.[69]
31. In the NRBM, measure values on an inventory route must be monotonic; i.e., the numbers must increase proportionally and continuously, with no gaps and duplications, along the length of the inventory routes. A data source's internal LRS may support one or more LRMs with stationing, formulas, offsets, and other non-monotonic structures; however, the LRM used for the published NRBM dataset must be monotonic.
32. A point location in an LRM is described by a combination of a route identifier and a measure value representing the distance of the location (offset) of the described point from the origin of the route. A single, real-world entity may be represented by multiple point-like LRM objects, one on each applicable route (such as an intersection of two roadways) or to represent multiple aspects of a single entity (such as a bridge that is also a jurisdictional boundary).
33. A linear location in an LRM is described by a combination of a route identifier, a measure value representing the distance of the starting location from the origin of the route, and a measure value representing the distance of the ending location from the origin of the route. A single linear LRM instance cannot bridge a junction between inventory routes. Zero length linear LRM instances are permitted.
34. Elements, characteristics, and events occurring along a route must receive measure values that place them in the correct linear sequence and relative distance apart.
35. Inventory routes may be mode- or function-specific, depending on the nature of the linear facility they represent; however, the inventory routes defined for NRBM and described in this Guidebook are currently limited to roadways for motor vehicle travel.
36. The basic geometric representation of an inventory route is a centerline. When necessary to more fully describe a linear facility that follows multiple paths (i.e., a divided highway), a pair of directional centerlines may be created to geometrically represent the path of each carriageway in the NRBM submission. A single centerline may additionally be required for other users who need a different level of abstraction or a smaller scale product.
37. Directional centerlines diverge or converge at a transition point defined by a continuity intersection, which serves to provide topological continuity and a point of reference for LRM measures. The continuity intersection is illustrated by an intersection point that actually serves to terminate the three connecting centerline features.

[69] This business rule is thought by some project participants to violate the general requirement for federally supported datasets to conform to the OpenGIS specifications of the Open Geospatial Consortium (OGC), which was believed to prohibit self-intersecting lines. According to the Consortium's *OpenGIS Simple Feature Specification for SQL* (June 1999), a self-intersecting line string (a sequence of three or more vertices) is classified as a non-simple line string. A line string (LineString) is a curve with linear interpolation between vertices. Each adjacent pair of points defines a straight line segment. A self-intersecting line is non-simple because it includes a single location within two or more line segments. [pp. 2-5] Thus, a self-intersecting line string is included in the OGC specification and this business rule does not appear to violate the general Federal requirement for consistency with the OGC specification.

38. A mode-specific centerline may be physically located within the cross-section of a multimodal facility, but may still be treated as a topologically separate linear facility; e.g., a bike lane running along the shoulder of a motor vehicle highway. In other words, routes supporting specific modes of travel may overlap or be located on the same roadway in the real world.
39. The only geometric representation of an LRM facility that is cartographically consistent with the 1D datum it creates is a straight line.
40. Route data is independent of the geometric representation provided; however, route data shall be recorded by side of route when the route is divided and directional centerlines are provided. Information describing the median shall be provided for the full extent of any carriageways. Data shall be published in the direction of travel when directional centerlines are used to display the data.
41. Within the published data, traversals may be defined by applying specific rules, such as to follow the path of a named street or designated route. An LRM may be created for each published traversal that is derived from the underlying inventory route LRM measures.
42. A match point may be declared wherever connectivity in the transportation network is required at agency boundaries. A match point created at a boundary is a continuity intersection. Although cartographic coincidence is desirable for edge matching the cartography provided by adjacent jurisdictions, it is not a prerequisite to constructing the NRBM dataset. What is a prerequisite is the ability to stitch adjacent maps together for topological continuity. To provide such continuity for pathfinding applications, each jurisdiction should include match points (boundary termini are spatially identical for each jurisdiction) or boundary connection points (link termini where zero-length segments can be created by any party to provide continuity) as a means of relating its spatial data to those of surrounding jurisdictions.
43. Data positions stored in an LRM database may be projected onto a 2D or 3D map, but doing so must not alter the LRM position description.
44. Map-defined LRM positions (measures) may be used when they represent the only source of such data, but are always inferior to measures determined in the real world.
45. Addresses may be provided as point features along a route, if available, and as block face address ranges on a centerline or carriageway feature. [70] Within the highway inventory, a block face address range is stored as a linear LRM object located using an LRM position with a side-of-route attribute.[71]

[70] The original specification for the AEGIST dataset included addresses as a mandatory component in furtherance of this requirement as contained in the USDOT's Transportation for the Nation Strategic Plan and predecessor documents. However, in the course of developing the proposed specification, it was recognized that the States are not an authoritative source for addresses, which are assigned and managed by local governments.

[71] Selecting the best structure for address information is somewhat controversial. The simplest approach is to attach it to the route's centerline, ignoring side-of-road considerations. A more sophisticated method is to split address ranges into separate ranges for each side of the road; i.e., as block face records. However, numeric address ranges are often inadequate to express addressing irregularities, such as out-of-order addresses, overlapping ranges, and multiple addressing authorities within a single block. There is also the simple fact that structures along the roadway are being addressed, and not the roadway itself. As a result, address points are the preferred means of less ambiguously conveying information about specific addresses. Such points are not located on the roadway centerline but are near the physical location of the building or unit to which the address is assigned. There is also the matter of which address format to support. There is an official Federal Geographic Data Committee-endorsed standard developed by the Urban and Regional Information Systems Association and a new address standard called NextGen 911 developed by the National Emergency Number Association. The various commercial GIS-T platforms have their own addressing formats and methods. Therefore, even if the States were an authoritative source of address information rather than a consumer of it, there would still be the issue of how to present such information. This problem exceeds the capabilities of the current AEGIST project.

46. For consistency with the NBI, a tunnel is a type of bridge where there is only a route going under the structure.
47. Bridges and tunnels are presented in the NRBM dataset as restriction elements.
48. Bridges often bisect geographical and manmade features that serve as political boundaries that may require route segmentation to occur; e.g., at State and county boundaries. In such cases, the bridge should be discoverable on all applicable routes and route segments, with the portion of the bridge on each route being listed as a linear LRM instance with the appropriate extent. The related bridge element record will contain the full description of the structure.
49. An intersection is conceptually composed of the following three atomic components:
    a. One or more approach segments.
    b. A navigation space where vehicles and/or pedestrians may make choices regarding direction of travel at defined navigation points.
    c. One or more departure segments.
50. Each intersection must be represented as a single point. Some applications, such as safety analysis, may need the internal geometry of an intersection. Such internal components are navigation points and turn segments. The LRM scale of 1:5,000 does not apply to the internal structure of intersections. The area containing the internal components of an intersection is a navigation space and may be defined by a polygon.
51. Within an intersection, navigation points and turn segments connect approach and departure segments to provide topological continuity through the intersection.
52. Approach and departure segments are functional assignments based on spatial association. Only approach segments are included in the NRBM dataset.
53. Each position within an intersection's navigation space where one centerline intersects another defines a location where a pathfinding decision may be made; i.e., creates a navigation point.
54. A navigation space may contain internal turn segments that connect navigation points. Some of these segments may represent channelized turn movements. Such internal segments are "owned" by the intersection, and are not assigned to a specific route. Navigation points and turn segments are not part of the LRM dataset.
55. Each intersection on a public road must be described by at least one LRM-based position.
56. Each intersection of two or more public roads contained within the NRBM dataset shall have at least one LRM-based position description for each such public road.
57. Although an intersection is minimally illustrated by a point feature and typically has a point LRM object describing its presence on each involved inventory route, some States have enlarged the scale of their base LRM to include navigation points and turn segments as visible features of the base roadway network. For these States, the location of an intersection on a roadway may be stored and published as a linear LRM object defining the length of the intersection's navigation space on that roadway.
58. An interchange is a means of logically grouping intersections and roadway segments for the purpose of treating them as one complex object. Each such interchange shall be given a GUID that may be used as a foreign key for relating components of the interchange.

59. A limited-access highway ramp is an inventory route representing a departure/deceleration "exit" roadway, an acceleration/merge "entrance" roadway, or a connector of multiple directional ramps within an interchange. Most ramps are paired with a graphical connector or ramp extension that serves to complete the path from the end of the ramp as an inventory route to a point of apparent intersection with the related directional centerline. The transition from inventory route to ramp extension occurs at the physical gore. From this point, the ramp pavement serving as a deceleration/acceleration space and a weave/merge space is subsumed by the limited-access highway and its presence becomes a component of the limited access highway's cross-section. (See Figure 13 through Figure 15.)[72]

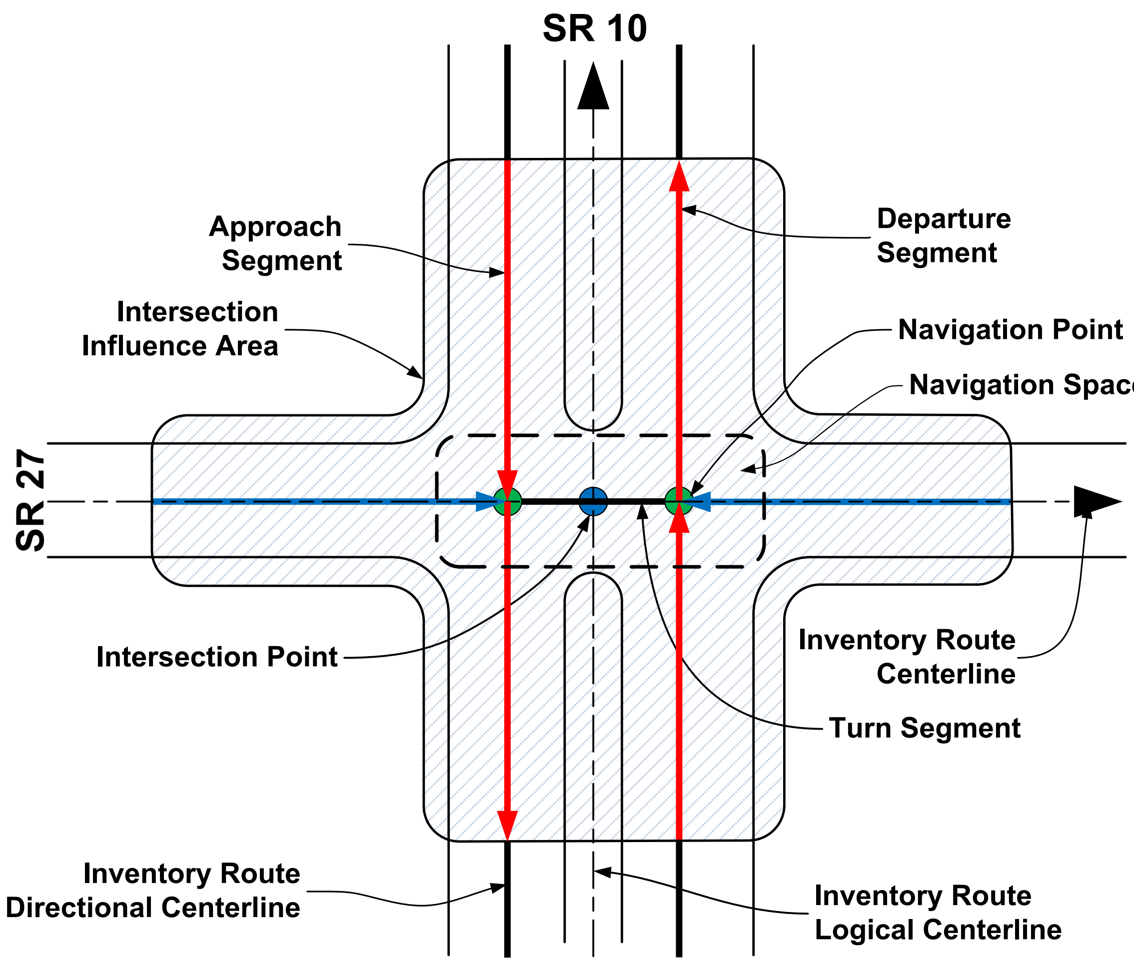


Figure 15. Intersection components.

[72] Some States have adopted a business practice of including the ramp extension in the LRM. That option was not endorsed for the NRBM due to the ambiguity inherent in determining the path of a ramp extension and the more common business practice of treating the pavement described by the ramp extension as part of the limited-access highway and not the ramp. Stopping the ramp inventory route at the physical gore avoids double-counting the centerline mileage for both the ramp and the limited-access highway.

Chapter 4. Fundamentals of the NRBM Specification

Chapter 1 set the framework for the National Roadway Base Map (NRBM) data specification. Chapter 3 established its functional requirements through a series of definitions and business rules. This chapter describes its fundamental elements, including business data and cartographic features, as a means of showing how the proposed NRBM data specification as presented in Chapter 5 was derived.

The NRBM specification provides a structure for how data should be created, edited, and published. The specification itself is both a *what* and a *how*. The *what* is a set of data expressed as a technical list of object class and attribute descriptions. The key to generating the data described in the specification is the *how*, which is formed by clear definitions, business rules, consistent application of those business rules, and a workflow guide.

Most importantly, the how content is exemplified by what the implementing States have to do: develop enterprise data governance, which leads to enterprise data management. This is because the only way to generate the specification data in the prescribed manner is to pull the entire agency together to function as a single, asset-focused data machine called *civil infrastructure management*. Data governance is covered in Chapter 2. This chapter focuses on the technical foundation of the data structure.

The proposed NRBM specification is designed to allow the States to provide a national transportation dataset in a coordinated but decentralized manner. Portions of the dataset may originate with local, regional, and Tribal governmental entities for incorporation into the State submission. The intent is to produce a map with a national extent that is suitable for project-level analysis. The fundamental NRBM deliverables would include the following:

- A segment-based cartographic product depicting all public and private roadways with minimal business data attached.
- A segment-based topological network.
- Business data with LRM position information.
- A route-based cartographic product suitable for mapping the business data.
- LRM objects representing travel restrictions that allow the topological network to support the routing of specific vehicles by considering weight, width, and size limits.
- Intersection elements represented as intersection points and approach segments.
- Interchange elements with interchange point cartography and ramp extensions.
- Terminus elements with point geometry.

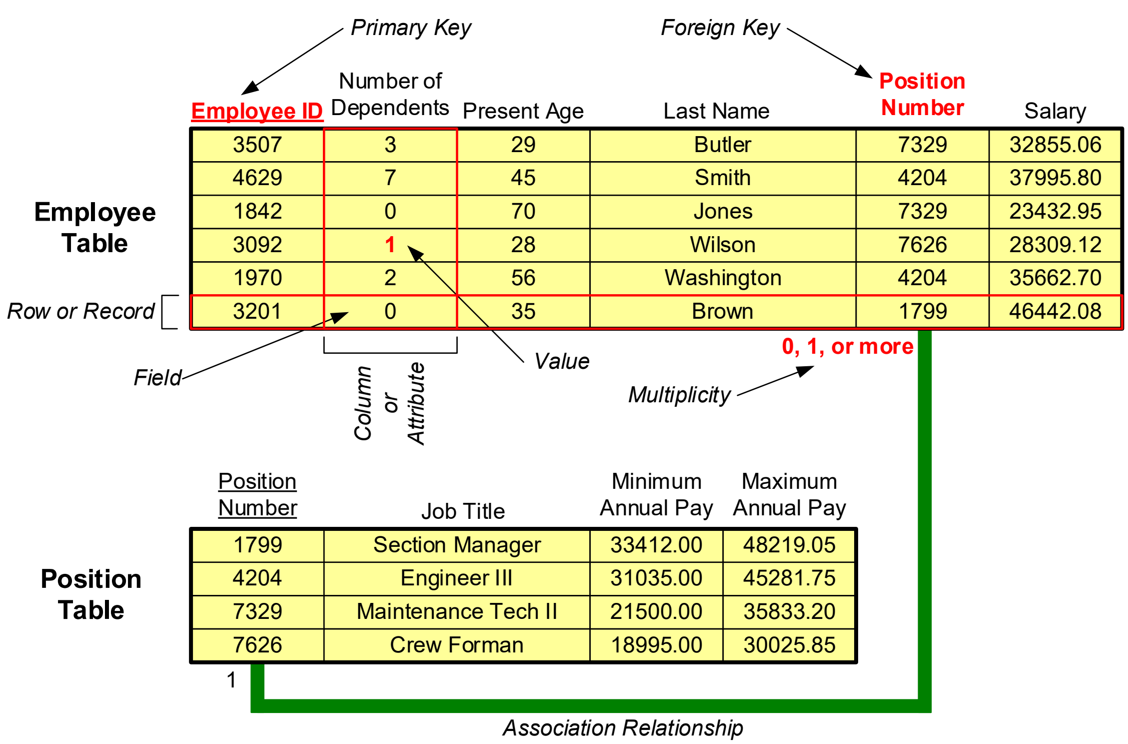
In addition to these core components provided at the national extent scale, intersections, interchanges, and other elements of the transportation should be available at larger scales for a more focused analytical capability, such as required by MIRE 2.0, to support a variety of safety and operational analyses. Large-scale intersection components include navigation points and turn segments contained within the navigation space. MIRE fundamental data elements (FDEs) are attached to intersections, approach segments, and interchanges and may be attached to navigation points. Photographs, diagrams, and other visual aids are also supported as supplements to the tabular business data.

The core dataset components provide the national coverage. A proposed set of optional components accommodate roadway elements with independent geometry and additional business data, such as:

- Roadway cross-sections
- Noise walls
- Guardrails
- Bridges and tunnels[73]
- Pavement sections
- Projects
- Curves

Base Concepts of Temporal Databases

This chapter begins with the core structure and moves to progressively more complex components of the total specification. The basic concepts are presented as a series of logical data models. The type used here is called an entity-relationship diagram (ERD). Entities are shown as boxes and relationships as lines. Each instance of an entity is an object—think of it like a row in a table. The relationships may include labels but will always include numbers at each end where they connect to entities. These numbers represent the multiplicity of the relationship. Picture this as the number of rows in this table that can be tied to a single row in the other table. Figure 16 shows an example of two tables and their relationship. It also serves to illustrate the various terms used in relational database design concerning tables.



Employee ID	Number of Dependents	Present Age	Last Name	Position Number	Salary
3507	3	29	Butler	7329	32855.06
4629	7	45	Smith	4204	37995.80
1842	0	70	Jones	7329	23432.95
3092	1	28	Wilson	7626	28309.12
1970	2	56	Washington	4204	35662.70
3201	0	35	Brown	1799	46442.08

Position Number	Job Title	Minimum Annual Pay	Maximum Annual Pay
1799	Section Manager	33412.00	48219.05
4204	Engineer III	31035.00	45281.75
7329	Maintenance Tech II	21500.00	35833.20
7626	Crew Forman	18995.00	30025.85

Figure 16. Example tables showing database terminology.

[73] The NBI treats tunnels as bridges, so there is really no distinction in terms of content.

Relationships are generally characterized as one-to-one, one-to-many, and many-to-many. Only one-to-one and one-to-many relationships will be used in the data models. Since they cannot be directly implemented in software given today's technology, any high-level, many-to-many relationships will be resolved to a pair of one-to-many relationships and an attributed relationship entity. (See Figure 17.) A single number or a range of numbers may be shown for the multiplicity value. Any range that begins with a zero means it is optional as to whether any of the entities at that end of the relationship exist. These concepts will become more apparent as the models are presented. Figure 17 shows a pair of one-to-many relationships being created to resolve a many-to-many relationship.

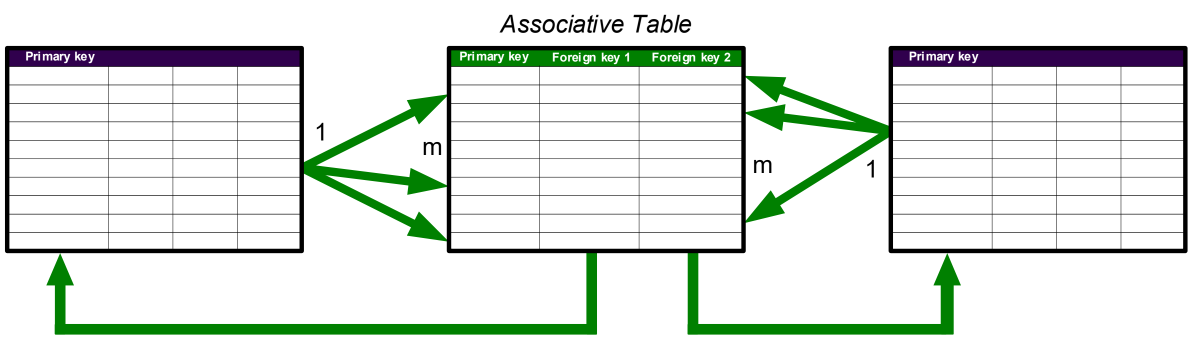


Figure 17. A many-to-many relationship converted to a pair of one-to-many relationships.

The most fundamental difference between the NRBM specification and most State roadway inventories is its intrinsic support for temporal views of the database. A key ARNOLD-derived NRBM requirement is support for change management. For example, the ARNOLD Reference Manual says that the core system should "employ a data structure that tracks inventory projects and roadway/route changes so that questions regarding data changes can be answered."[74] Part of this direction deals with the temporal extent of the NRBM database, such as the need to include "planned, unbuilt facilities, as well as abandoned or destroyed roadways, in the dataset."[75] Retaining previous states of the LRS database—what the ARNOLD Reference Manual calls "geoarchiving"—is another aspect of the temporal support mandate.[76] These various requirements were generally accepted during the AEGIST development process by peer exchange participants and have begun to be implemented by the States. A number of business rules presented in Chapter 3 dealt with temporal aspects of the database design and related business practices.

The NRBM has two temporal vectors, one for tracking when edits occur and one for changes in the real world. The editing vector supports asynchronous, or transactional, data transfers. This means the user (recipient of the dataset) can periodically download only those new records written since the last update was done, rather than require the user to download a complete file. The advantage of the transactional update is the reduction in workload imposed on the user, particularly when the user is integrating the NRBM data into their own structure. This transactional update feature is the result of the specification's use of a consistent set of temporal fields. These fields are part of every table.

[74] *Op. cit*, p. 9.
[75] *Op. cit*, p. 20.
[76] *Op. cit*, p. 79.

Rather than repeating these standard temporal fields in every table description, inheritance is used as a way of constructing an abbreviated version of the various standard tables.[77] As shown in Figure 18, inheritance allows omission of the repeating attributes that were inherited from the parent abstract *Object* class. (By industry practice, the name of an abstract class is shown in italics.) So, instead of showing the full table description on the right, the shorter version in the middle is used, with the understanding that **`RouteSegment`** inherits all the fields from its parent abstract class, *Object*. The open arrow connecting *Object* and **`RouteSegment`** points from the child class to the parent class. In the following discussion of the many tables proposed in the NRBM specification, the Object class and the inheritance relationship are omitted. Note that *every table* will have those additional fields to supply temporal and editing record information.

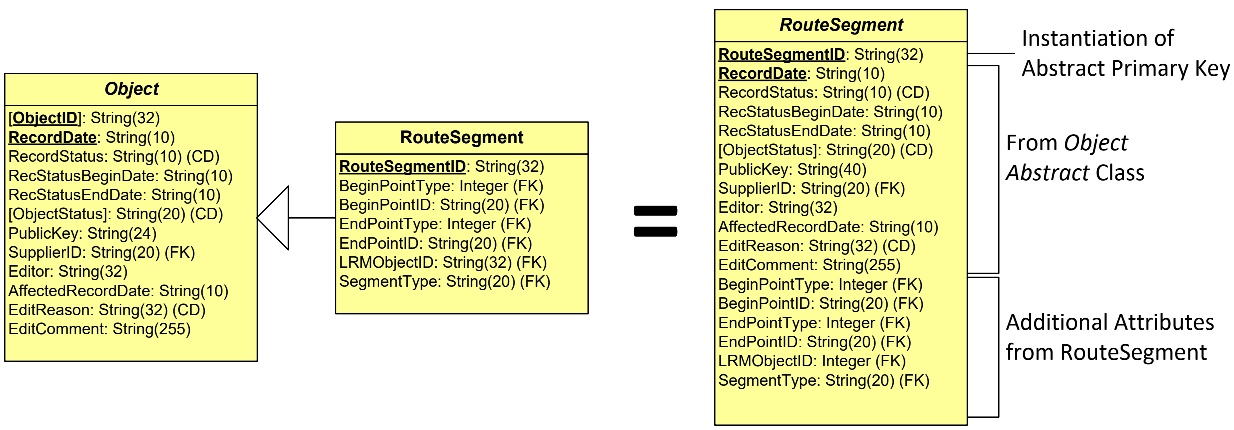


Figure 18. The RouteSegment class inherits the standard temporal and editing fields from the abstract Object class.

The ARNOLD Reference Manual additionally recognizes the need for adjustments to the raw data within each State and provides a taxonomy of changes that may be made due to internal and external factors.[78] Changes in this context include the following:

- Spatial change.
- Attribute change.
- Spatial and attribute change.
- New object.
- Delete object.

To provide the required temporal functions, the entire database must be temporal. This means nothing can ever be deleted. Historical data is just as valuable as current data when one of the NRBM objectives is to be able to show the state of the transportation system at any point in time: past, present, and future. The old roadway alignment needs to be retained as does its attributes. If a road is widened, the old pavement width and number of lanes values must still

[77] This discussion is not intended as a general introduction to object-oriented design or data modeling. Its purpose is to provide a basic introduction to these concepts in order to allow the novice reader to better understand the graphical notation used in the AEGIST Guidebook.

[78] *Op. cit*, pp. 31-32.

be in the database. In the same way, users need to be able to show future states of the transportation system in order to see the results of programmed projects. Even look-up tables for values, such as pavement condition ratings, must be temporal. Users will need to know that the pavement quality rating method changed in 2021, for example, and that the coded values stored in early records have a different meaning than those contained in later records.[79] Both the old and new methods need to be in the database. Also, given the large size of the NRBM dataset, users need to be able to update their copy by downloading just the records that changed.

The *Object* class template used to create all the object classes contained in the NRBM dataset is shown on the left side of Figure 19. (Object templates are also called stereotypes.) The template is kept simple to facilitate data exchanges without relying on specific software capabilities.

Object
[**ObjectID**]: String(32) **RecordDate**: String(10) RecordStatus: String(10) (CD) RecStatusBeginDate: String(10) RecStatusEndDate: String(10) [ObjectStatus]: String(20) (CD) PublicKey: String(24) SupplierID: String(20) (FK) Editor: String(32) AffectedRecordDate: String(10) EditReason: String(32) (CD) EditComment: String(255)

Figure 19. The NRBM specification object class template (stereotype) with its temporal fields.

Before getting into the details of the *Object* abstract class, first is a summary of the symbols, abbreviations, and descriptive terms used in the data modeling process. The name of the class is shown at the top above a horizontal line. Below that are the attributes of the class. In a relational database implementation, these are the names of columns. The data models are presented as classes that are equivalent to relational database management system (RDBMS) tables, so there are no methods to list under the attributes, as there could be for an object-oriented presentation. The AEGIST Guidebook uses mnemonic names for the attributes, which means the name is supposed to remind the user of its meaning. To be compatible with most RDBMS, the Guidebook has adopted the practice of omitting spaces and starting each "word" in the attribute's name with a capital letter. So, **`RecordDate`** can be readily understood to mean "record date"; i.e., the date the record was created. It also adopted the convention of using a monospaced font to indicate a term is the name of a class attribute, so that it cannot be confused with an equivalent English word.

Separated from the attribute name by a colon is the type of data this attribute contains. The two choices shown are Integer, which means a whole number, and String(x), with (x) indicating a number. A String data type is one that allows a series of alphanumeric characters, which includes all the capital and lowercase letters, numbers, and certain symbols—basically everything present on a keyboard. The

[79] Normally, Active records will all be converted to the same coded value domain, but earlier Retired records will likely persist in the database and will need to be accommodated by being able to recall the previous domain.

maximum number of characters is stated in the parentheses. Thus, "String(20)" means the field can contain up to 20 alphanumeric characters.

String values are used for most attributes to provide the greatest flexibility possible in accommodating the form of source data. Remember that this data model represents a publication specification. The actual source data may have any form, and the simplest publication process is to accept whatever the source offers. Date attributes are a good example. All RDBMS products on the market today provide internal date/time data types. The NRBM specification cannot use them because it has to remain vendor neutral. As a result, the specification uses a simple `MM/DD/YYYY` format, which requires a string data type field. An example value would be "05/10/2019." This string field can be subdivided or converted by recipients to a form supported by their particular RDBMS platform.

There are other abbreviations shown in parentheses. "CD" means the field must contain a choice from a coded domain of values. Functional class is an example of an attribute that has specific coded values that can be used. Street type, which includes ST, CT, AVE, DR, PKWY, etc., is another. These are all presented to the editor as a pick list of choices. Normally it is best to avoid having an editor enter data by typing, to avoid misspellings and other mistakes. It also helps to classify the data when users can tightly control the choices. When "(CD)" is in the class, expect to find a domain description somewhere. (There is an NRBM specification for how these are structured.) Not shown is the related "NR" abbreviation, which stands for numeric range. For example, number of lanes is an attribute that could be shown to have a numeric range constraint.

The first listed field is **`ObjectID`**, which is the globally unique identifier (GUID) of that object instance. **`ObjectID`** is really a placeholder for the real attribute name that will be used when the template is instantiated, which is why it is shown in brackets. The actual field name when this template is applied will be changed to replace '**`Object`**' with the actual object name, such as **`RouteSegmentID`**.

The temporal elements of the class template come in three groups that track two temporal vectors: the period of validity for the record and the period **`RecordDate`** (the date the record was created). Values are in the format of a fixed 10-character field: `MM/DD/YYYY`—two characters for the month, a forward slash, two characters for the day, a forward slash, and four characters for the year. The number of characters shown for each part is mandatory; i.e., June is '06'. This is not a native date format for storing the time a record is created in a computer database. It is expected that the date the record is written will be part of an automated process using the specific date/time capabilities of the native editing environment. Most data sources will want to expand this field in their internal editing databases to include time, but such capabilities are implementation-specific, as they vary depending on the database management software used. This part of the specification is designed for data exchange.

`ObjectID` and **`RecordDate`** are both shown in bold, underlined type to indicate that together they form the complex primary key for the class. It is not enough to just use the values in **`ObjectID`** to uniquely identify each row in the table. This is because a user may have many records with the same value as new entries replace old ones. A temporal database of the type proposed for the NRBM never deletes any record because one of the objectives is to be able to recover past states of the roadway system. As a result, **`RecordDate`** is included in the primary key to separate the various versions in time. Users have to know not only which object, but when.

The next three fields work together to provide the first part of the temporal capabilities of this template. No information is ever deleted from the database. Instead, the status of a record is changed to reflect that it is no longer a valid description of the object. **`RecordStatus`** provides this information and has three possible values: Active, Retired, and Replaced. An active record is currently valid. **`RecStatusBeginDate`** will always have a value when an active record accurately reflects the object's attributes. The starting date can be any time point in the future or the past. An active record will normally not have an ending date, which would be stored in the **RecStatusEndDate** field.[80]

Retired records reflect a past state of the object and will have both starting and ending dates. If a user seeks to learn what the transportation system looked like at some point in the past, he or she is likely to find a combination of active and retired records. A Replaced record was never valid; it was a mistake. It will not have a value for either the beginning or ending date of validity. **RecStatusBeginDate** and **RecStatusEndDate** have the same date format as **RecordDate**; i.e., `MM/DD/YYYY`.

Figure 20 shows how the **RecordStatus** field operates to allow users to retain history as the data are edited. **RecStatusBeginDate** and **RecStatusEndDate** describe the applicable period for the given **RecordStatus** value. The initial state shows a table with six records. A change is made for Entity 5 in Updated State 1. The current record is retired by changing the value of **RecordStatus** to 'Retired' and adding a **RecStatusBeginDate** value. The value in **RecordDate** is the date this edit was made. A new active record with the correct **RecStatusEndDate** is added to the table. Later, it is discovered that the entry for Entity 7 was wrong. This record must be replaced with one that has the correct information. To do so, the value of **RecordStatus** is changed to 'Replaced' and the value for **RecStatusEndDate** is removed. A new active record with the correct information is added. A replaced record will never be discovered when looking for the state of the system on a particular date, but it remains in the database as a historical record.

This design provides four primary capabilities. First, it allows the state of the database to be recovered for any point in time by looking at **RecStatusBeginDate** and **RecStatusEndDate**. The user would perform a search for the records that apply on a specified date by finding records that have a **RecStatusBeginDate** value on or before the specified date and a **RecStatusEndDate** value that is on or after the specified date. A null value for **RecStatusEndDate** means that **RecordStatus** is 'Active' for the date specified.

Second, if someone wants to update their copy of the database, they only need to download all the records with a **RecordDate** later than the date of the last update process. Once those records are downloaded, the user would first replace all records in their NRBM copy that match those with a **RecordStatus** value of 'Retired' or 'Replaced', and then add all new 'Active' records. A match is determined by finding the most recent record in the target class with the same value for **ObjectID**.[81]

Third, this process allows asynchronous edits, where changes can be made to a Retired record in order to more completely preserve the history of that object. An editor would call up the state of the record as it existed on the date the edit should have been made, make the historical modifications, and save a new Retired record. The dataset recipient would get these revised records in their update download, even if they applied to a previous update cycle.

[80] There are instances when there could be multiple Active records with different validity periods. For example, there could be a future construction project where a series of realignments occur while a roadway undergoes a major reconstruction. These Active records would have future periods of validity, so the user would receive the version that was expected to apply on the date of interest. In general, these are the exception and almost always would apply to future conditions.

[81] The presence of asynchronous edits in a transactional update dataset may make this process a bit more complicated, as the latest update could contain changes to a record that was not the last one previously downloaded from the source. To address this complication, it is necessary to include an **AffectedRecordDate** attribute in the *Object* stereotype so that the entire primary key of the changed record can be matched to the same record in the recipient's database.

Fourth, the design allows an almost endless "Undo" process, where edits can be rolled back by recovering the information contained in the 'Retired' or 'Replaced' record and removing the 'Active' record.

Initial State 0

EntityID	RecordDate	RecordStatus	RecStatus BeginDate	RecStatus EndDate	EntityStatus	[Additional Attributes]
1	02/17/2017	Active	11/06/1985		In Use	
2	02/17/2017	Active	11/06/1985		In Use	
3	02/17/2017	Active	11/06/1985		In Use	
4	02/17/2017	Active	11/06/1985		In Use	
5	02/17/2017	Active	11/06/1985		In Use	
6	02/17/2017	Active	11/06/1985		In Use	

Updated State 1

EntityID	RecordDate	RecordStatus	RecStatus BeginDate	RecStatus EndDate	EntityStatus	[Additional Attributes]
1	02/17/2017	Active	11/06/1985		In Use	
2	02/17/2017	Active	11/06/1985		In Use	
3	02/17/2017	Active	11/06/1985		In Use	
4	02/17/2017	Active	11/06/1985		In Use	
5	**02/18/2020**	**Retired**	11/06/1985	**01/15/2020**	In Use	
6	02/17/2017	Active	11/06/1985		In Use	
7	05/21/2018	Active	03/28/2018		In Use	
5	**02/18/2020**	**Active**	**01/16/2020**		**Closed**	

Updated State 2

EntityID	RecordDate	RecordStatus	RecStatus BeginDate	RecStatus EndDate	EntityStatus	[Additional Attributes]
1	02/17/2017	Active	11/06/1985		In Use	
2	02/17/2017	Active	11/06/1985		In Use	
3	02/17/2017	Active	11/06/1985		In Use	
4	02/17/2017	Active	11/06/1985		In Use	
5	02/18/2020	Retired	11/06/1985	01/15/2020	In Use	
6	02/17/2017	Active	11/06/1985		In Use	
7	**04/27/2020**	**Replaced**			In Use	
5	02/18/2020	Active	01/16/2020		Closed	
8	06/19/2018	Active	05/13/2018		In Use	
9	06/19/2018	Active	09/02/2019		In Use	
7	**04/27/2020**	**Active**	**03/28/2018**		**In Use**	

Figure 20. How temporal editing works.

`ObjectStatus` is similar to the record status field but refers to the entity in the real world rather than to the record in the database. As with **`ObjectID`**, the 'Object' portion of the field name will be replaced by the name of the object class. Each class may have its own domain of possible values. For example, buildings may have the status values of 'Proposed', 'Under Construction', 'Occupied', 'Abandoned', 'Closed', and 'Demolished'. A roadway may have the object status domain of 'Proposed', 'In Design', 'Under Construction', 'Open to Traffic', 'Closed to Traffic', 'Abandoned', and 'Removed'. Multiple classes may use the same coded domain. This attribute works with **`RecordStatus`** and the two validity date attributes to display past, present, and future transportation system views.

Example: A new bypass route around a small town is to be constructed in 2023. Design will occur in fiscal years 2021–2022. The bypass will take 3 years to build and will open to traffic in 2026. All of these events can be in the NRBM at one time using different record validity ranges and object status codes:

RecordStatus	RecStatus BeginDate	RecStatus EndDate	ObjectStatus
Active	10/01/2021	09/31/2022	In Design
Active	10/01/2022	09/30/2025	Under Construction
Active	10/01/2025	[no entry]	Open to Traffic

Centerline symbology could be used to indicate that the roadway was in design or under construction, with its full appearance as a facility open to traffic starting on October 1, 2025. Of course, any portions of the original route would need to also have coincident changes stored in the database.

The next field is **`PublicKey`**, which is the GUID constructed from concatenating State FIPS code, County/Parish FIPS code, a source or agency identifier assigned by the State, and the agency's identifier for the object.[82] The public key conforms to the format shown in Chapter 3:

- 2 characters for the State FIPS code.
- 3 characters for the County/Parish FIPS code.
- 3 characters for the supplying agency's identifier assigned by the State.
- 32 characters for the supplying agency's unique object identifier.

All of these components are integers, but each can begin with a zero, so they have to be converted to string values. To make it simpler to split the field into its four parts, the supplying agency's object identifier should be right justified with zero fill to get to 32 characters in length.[83] A geographic identifier is provided by the first two fields. For example, the public key starting with '12097' would mean the object is in Orange County Florida ('12' for Florida and '097' for Orange County).

Since the object template has a complex primary key consisting minimally of the attributes **`ObjectID`** and **`RecordDate`**, **`PublicKey`** is not sufficient to actually identify a single record when multiple versions exist, each with its own **`RecordDate`** value, but it could be used with **`RecordDate`** to find the same records. Using just the **`PublicKey`** value will return all the records for the identified object in that class.

`SupplierID` is the identifier for the source of the data contained in the record. It is the Agency identifier portion of the value stored in **`PublicKey`**. This field is primarily useful for States that use external sources of data, such as local governments or contractors, but it can also identify the workgroup within the State DOT that provided the information. Storing this attribute allows the source agency to extract its own information and the State DOT to go back to the original source to confirm the data or notify the source of any detected data quality issues.

[82] **ObjectID** is a computer-assigned value that avoids the possible erroneous entry of **PublicKey** values by a data source. Industry practice strictly avoids using a human-entered value as part of the real primary key.

[83] The 32-character length is set by the requirements of a GUID, which consists of 32 hexadecimal digits.

The next four attributes work together to store a record of the editing process; i.e., who created the record and why. **`Editor`** tells us who created the record. This field could be automatically populated by the editing software. **`EditReason`** provides a general classification of the reason for the edit. The entries in this field are controlled by a coded-value domain. **`EditComment`** allows the editor to provide more information about this particular action; this is an open text field.

Included in the group of editing-related attributes is **`AffectedRecordDate`**. This attribute stores the previous **`RecordDate`** value for the edited record. When recipients obtain a revised record during a transactional update process, they have to match the existing records with the new records. All classes that inherit from the Object abstract class use a complex primary key consisting of **`ObjectID`** and **`RecordDate`**. To know unambiguously which record was changed in the latest set of edits, the recipient has to find the old one. Normally, this is just the latest record with a matching **`ObjectID`** value, but in the case of asynchronous edits, where a more recent update can affect a record that was not in the latest set of edits, this is not a foolproof method. In other words, the recipient cannot assume that it is the latest record in the old copy of the database that was changed. To ensure the recipient knows which one was changed, the editing process needs to save the original **`RecordDate`** as the value in the **`AffectedRecordDate`** field.

The NRBM design directly supports the editing process, both in terms of making changes and in distributing them to users. It is not merely a publication specification, although it can serve only that role. The power of the design comes from a State's ability to natively provide the NRBM dataset when it adopts the transactional aspects of the design. Some degree of migration to the NRBM specification as an internal structure is necessary in order to provide the required temporal data. A State that does not have an internal data editing environment that provides temporal aspects of the NRBM will not be able to publish the data in accordance with the specification. This is the case with most States. To the extent that software vendors support the NRBM's temporal data structure, the NRBM specification offers a guide for how States can structure their roadway data editing environment. The data can continue to be supplied to internal users in the way it always has been.

The migration path to publishing the temporal NRBM dataset can be incremental. The minimum data to supply are the two fields that form the complex primary key (i.e., **`ObjectID`** and **`RecordDate`**) in the required format. A value for **`RecordStatus`** is also required, but they will all initially be Active. The value for **`RecordStatusBeginDate`** can be the same as **`RecordDate`**. Since every record is Active, there is no need to put anything in **`RecordStatusEndDate`**. The value for many object classes can also default to a standard value, such as 'Open to Traffic' for inventory routes. There is no requirement to enter a value for **`PublicKey`**. **`SupplierID`** is the value chosen for the State DOT. The three edit description fields can all be omitted. Therefore, a State publishing an NRBM temporal dataset will need to do the following:

1. Convert or create a primary key that consists of a unique integer identifier for each record and the date the record was created/published (as `MM/DD/YYYY`).
2. Enter the value 'Active' for **`RecordStatus`**.
3. Make the value of **`RecordStatusBeginDate`** the same as the record date.
4. Supply a standard value for **`ObjectStatus`** that is appropriate to the object class.

This dataset will not allow recipients to perform transactional updates, nor will it provide temporal data, but it does allow data to be published in accordance with the minimal requirements of the NRBM. It's one small step for a State; one giant leap for a national transportation dataset.

Route Segment, Intersection, and Terminus

At its fundamental level, the roadway system is a collection of route segments, intersections, and route termini. A route segment is an atomic linear cartographic feature with attributes. It extends from one intersection to another, from a terminus to an intersection or from an intersection to a terminus. Intersections and termini are represented as points.

Each route segment is a line between a terminus and an intersection or between two intersections. A divided roadway must have two route segments, one for each direction of travel. Mid-block transitions from divided to undivided status are marked by a continuity intersection.

Figure 21 illustrates the basic roadway segmentation approach. Several of the segments are represented by directional centerlines. Intersection 2476 (the leftmost blue dot) is a continuity intersection where the roadway splits into directional centerlines; the others are "real" intersections where roadways connect. All three will need a GUID and a type indicator. Each route segment will also need the following attributes:

- Route segment type.
- Identifier of the beginning intersection or terminus.
- Identifier of the ending intersection or terminus.
- Road name(s).
- Functional classification.
- Road status (Open to Traffic, Closed to Traffic, Planned, Removed, Restricted Use, Seasonal, etc.).

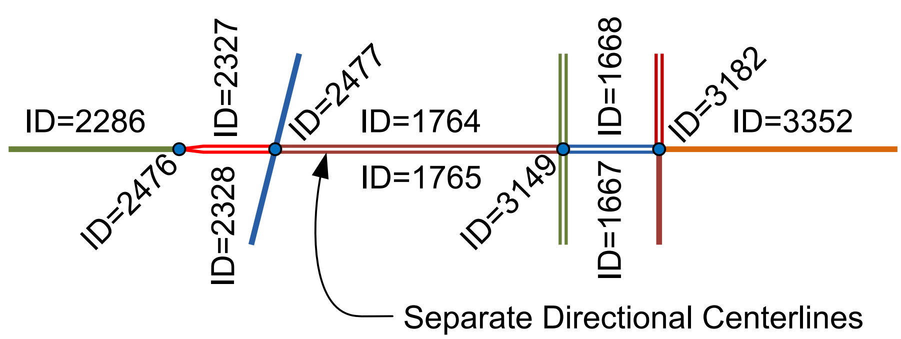


Figure 21. Roadway segmentation must occur at intersections in order to produce a topologically complete network for pathfinding.

Figure 22 presents the core data model showing the entities of Route Segment, Terminus, and Intersection. This figure uses the conventions of an entity relationship diagram (ERD). Each entity is represented by a rectangle. The blue background color indicates these are cartographic features, which combine vector geometry with attributes. All data model entities in this example are called object classes. A class is a template for creating instances of that object, like rows in a table. The relationships are depicted using lines connecting the rectangles. Some relationships state their explicit role. Here, 'Begin' and 'End' are examples of a role.

As shown on the left side of the figure, the fundamental entity of the database design is the intersection. Route segments are a way to manage connections (relationships) between intersections.

This concept can be flipped to become segment-centric rather than intersection-centric by saying the fundamental entity is the route segment and intersections are a way to manage the connections (relationships) between segments. Whichever view is preferred, route segment termini that are not intersections need to be added to provide topological objects at the end of a route segment that does not coincide with an intersection; i.e., a dead end or cul de sac. The result is the Terminus class. Of course, the database design includes many more classes than these three, but the foundation is very simple.

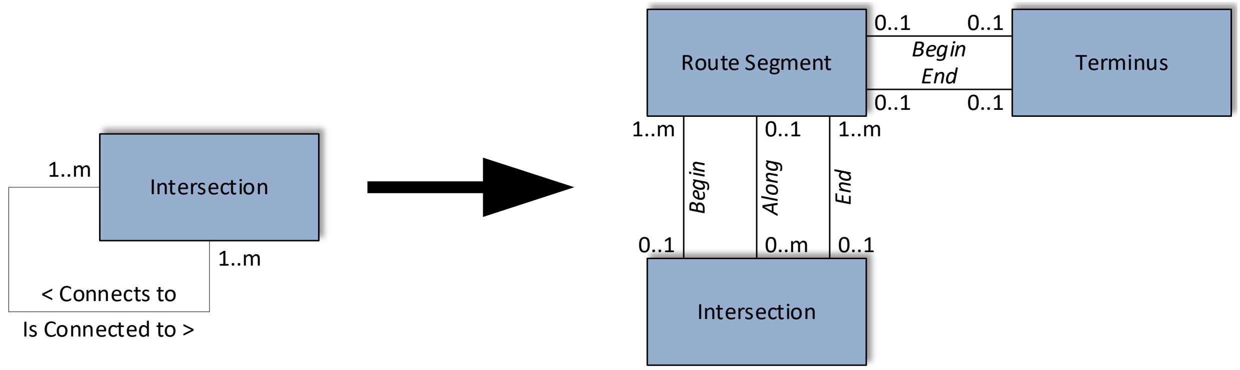


Figure 22. The core NRBM data model.

There are multiplicities shown at the end of each relationship line expressing the number of instances that may or must exist. For example, the 'Begin' Line connecting Route Segment and Intersection has 1..m at the Route Segment end and 0..1 at the Intersection end.[84] Reading from the bottom up, it says that one intersection may terminate 0, 1, or more route segments. Reading down, it says that a Route Segment may begin at an intersection. These are the characteristics of a one-to-many relationship; i.e., one intersection may be related to many route segments. This relationship is why a Route Segment needs to know its terminal objects, but those terminal objects cannot know what route segment(s) they terminate. The relationship of Route Segment to Terminus is similar, but has a key difference: a terminus may only be related to a single route segment.

The complete way to read the data model is to say that a route segment is a piece of a linear transportation facility that extends from one intersection to another, from a terminus to an intersection, or from an intersection to a terminus. These are abstract creations used to describe the real world. Intersections and terminus features serve to terminate a route segment, but have different attributes and behaviors, so they are shown as different entities in the data model. Each route segment must know the type and identifier of the two objects that define its extent. The multiplicity for the relationships starts at 0 to reflect the fact that an intersection or a terminus may have only a single role; i.e., to begin or end a route segment. The model does not preclude a single intersection from terminating both ends of a route segment, as might be the case for a self-intersecting route.

[84] This discussion and the described data model avoids the complexity of recognizing the process of creating route segments, where intersection and terminus objects may exist first in order to create the route segments that connect them. If the terminal objects must exist first so that their identifiers may be stored in route segment attributes, then the initial condition must be that an intersection has no mandatory relationship with any route segment. This would make the multiplicity on the Route Segment end of the relationship become 0..m. Therefore, while this condition may technically be true for a given State, it fails to properly convey the intent of these relationships to the reader.

Since the relationships between route segments and their terminating objects are labeled as ‘Begin’ and ‘End’, the order for the intersections and a direction for the route segment are implicitly imposed. A single intersection may begin and/or end one or more route segments. A terminus may only begin or end a single route segment. Existing rules in the ARNOLD Reference Manual state there must be two route segments created for divided roadways, one for each direction of travel. Both directional route segments will terminate at the same intersections, but the ordering of the termini will be reversed as appropriate for the direction of travel.

In this context, ‘Begin’ and ‘End’ and the direction they create serve two purposes. The first is for supporting vector graphics, the type of geometry used in the NRBM. Vector graphics, such as the roadway centerline features to be created to represent route segments on a map, must be given a direction. The second purpose is to support linear referencing. Route segments do not have linear referencing. Their attributes apply to the entire segment. However, route segments will be assembled to form inventory routes, which do have linear referencing. If users are to be able to locate transportation elements along a route segment as part of an inventory route, then they must know the point of origin from which the offset distance is applied.

Route segments know their terminal objects, and this information can be passed along to segment centerlines. However, that is not enough information to support pathfinding, which depends on intersections to serve as decision points where the pathfinding algorithm can select the next step in the path. The intersections in this basic data model do not have that information. To provide connectivity, a topological network needs to be constructed.

The topological equivalent of a route segment is an edge. The topological equivalent of an intersection or terminus is a node. Edge and Node are dimensionless objects. A single node may begin or end many edges. The result must be that each edge begins and ends at a node.

Within the editing environment, the rudimentary topology provided by the basic route segment data model is sufficient to ensure that cartographic rules are enforced; i.e., that route segments terminate at intersection and terminus points, with no gaps or overlaps. This is provided by a sort of “topology on the fly,” where cartography rules are enforced. For route segments, one might utilize the function that says a route segment endpoint must be covered by (i.e., be coincident with) an intersection point or a terminal point.

There are actually two forms of topology in the NRBM. There is, of course, the cartography-derived connectivity provided by the coincidence of route segment termini at intersections. However, the topological structure of edges and nodes as a geometric network is not an actual deliverable. Most GIS software platforms combine the segment centerline geometry and the edge topology to create a geometric network using a “build topology” function that utilizes the geometry of lines and points to create edges and nodes. In this way, edges and nodes get information about the route segments and intersections they represent from the geometric features. The process also readily identifies any “free” route segments, intersections, or termini, which would be errors. Geometric networks are beyond the scope of this Guidebook. Complete pathfinding applications need more information on travel restrictions, such as one-way traffic and left-turn prohibitions, which are attributes of the nodes that regulate movement within the node between edges. Accommodating these attributes in turn table structures is not addressed in the NRBM specification and is also beyond the scope of this Guidebook.

The other kind of topology provided by the NRBM is the data supplied in the attributes of a route segment; i.e., the beginning and ending points. Some GIS software can use this information to build topological networks without having to build a geometric network consisting of cartographic objects.

Topology does not care about geometry, per se, only the cost of travel along an edge. Length is one way to define cost, but the NRBM seeks only a routable network. Cost is no issue.

The initial definition of a route segment said they were cartographic features with attributes. This means that any implementation of the basic data model shown in Figure 23 will combine the Route Segment and Segment Centerline entities into a single Route Segment feature class. Practically, this means the supplier will draw a route segment centerline feature and then attach attributes to it. Similarly, Intersection and Intersection Point will become an Intersection feature class and Terminus and Terminus Point will become a Terminus feature class. These points may be generated through an automated process that places an intersection at all locations where two or more route segments terminate and a terminus at all locations where only a single route segment terminates. The final set of objects to be produced will be these three feature classes.

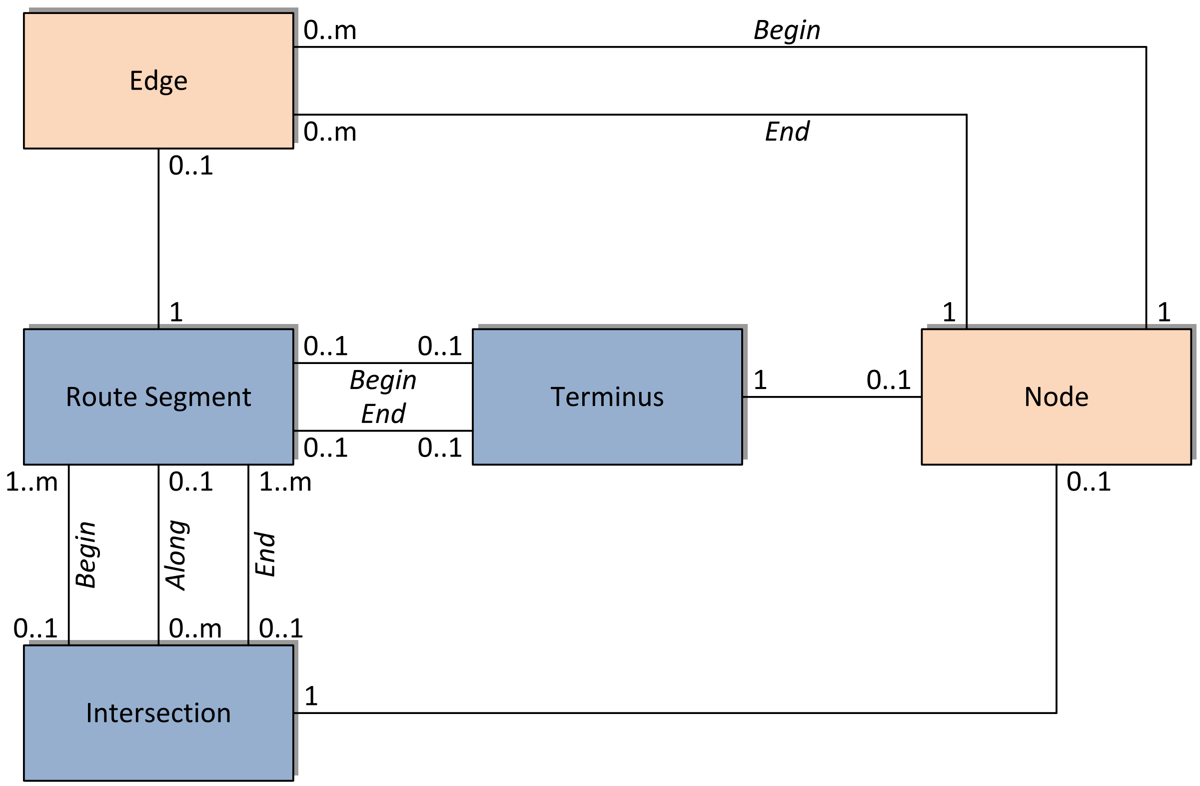


Figure 23. Adding geometry and topology to the core model.

Points are dimensionless objects in the vector graphics world created by GIS. No matter how much the scale increases, the point remains exactly the same size. Lines similarly retain the same width regardless of scale. What can change, though, is the shape of the line. In the case of a route segment feature, this change is expressed by the scale at which vertices are determined. A large-scale route centerline will need many more vertices than a smaller scale centerline because larger scales provide greater detail.

ARNOLD sets a scale requirement for the NRBM dataset of 1:5,000 (1” = 417’). The baseline scale of 1:5,000 does not mean the data are compiled at that scale, nor does it mean that maps will be produced at that scale. The scale can express the level of positional accuracy required to meet national map accuracy standards and set the foundation for making decisions about the LRM requirements, level of abstraction, etc. Centerline geometry is commonly compiled using aerial photography in the form of orthophotographs at scales approaching 1:600 (1" = 50'). This implies that the number of vertices provided will approach the level generated by this much larger scale.

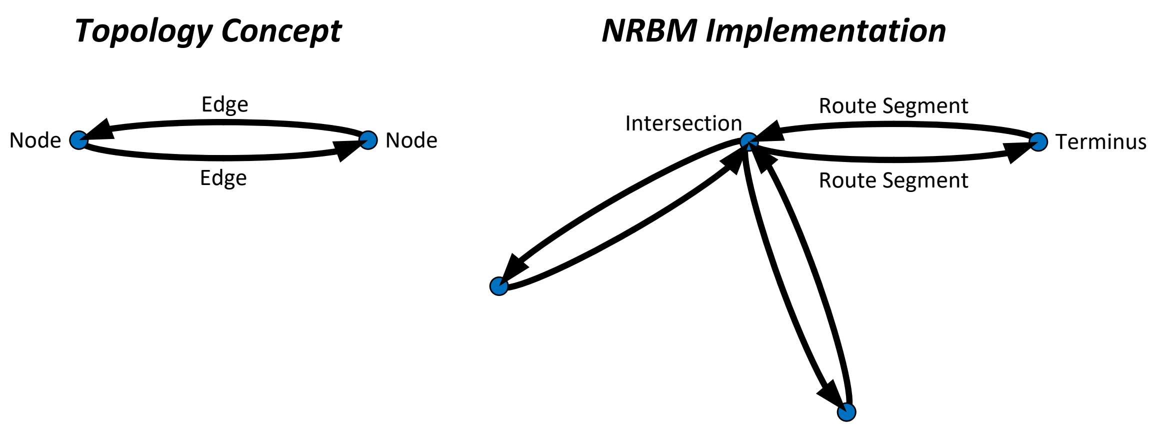


Figure 24. The topology concept and its implementation in the NRBM.

The problem with trying to meet national map accuracy standards, which prescribe that 90 percent of all points need to be within 14 feet of their true position on the ground at a scale of 1:5,000,[85] is that these standards are based on the ability to know what the true ground position is. The true ground position of an abstract roadway centerline or intersection point cannot be known because these are not actually recoverable locations on the Earth. Road centerlines are abstract objects that represent a facility that can exceed 100 feet in width. A centerline is a fairly poor representation of something that could be 2 inches wide at a scale of 1"= 50'. Plus, LRMs establish 1D datums, which means the only position that matters is where the point is relative to the origin of the centerline. There is no "horizontal LRM position." The 1:5,000 scale specification deals more with defining the appropriate level of abstraction.

Figure 25 shows the implications of the 1:5,000-scale specification. A highway with dual carriageway representation is shown with the separation of carriageway features at a scale of 34 feet—a dimension determined by 12-foot lane widths, two lanes in each direction, and a central median 10 feet wide (10/2 + 24/2 = 5 + 12 = 17-foot offset from the median centerline). At this scale, the two parallel carriageway line features are 0.08-inch apart (a little more than 1/16-inch). The top portion of the illustration shows the carriageways using a fairly typical 2-point line width. The bottom portion shows the centerlines with a scale width of 24 feet and a space between them of 10 scale feet to represent the median. Note that this is the largest possible display scale for the centerline abstraction. Cartographic specifications proposed herein are determined by applying this largest possible display scale.

[85] This refers to the traditional, scale-dependent national map accuracy standard, which says that the minimum accuracy of a dataset should be 90% of the points being within 1/30-inch of their true horizontal position at a publication scale of 1:20,000 or larger. This standard was established in 1941 by the U.S. Bureau of the Budget to set accuracy standards for all federally produced maps. The standard was revised in 1947. Based on this standard, the permissible error is ±14 feet on the ground at a scale of 1:5,000.

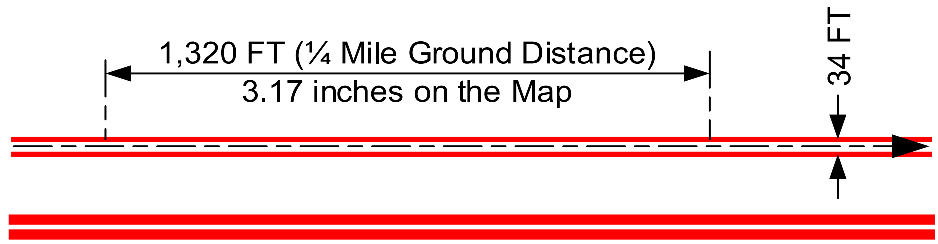


Figure 25. ARNOLD scale requirements.

Figure 26 illustrates what a typical street might look like on a 1:5,000-scale map. The divided roadway centerlines are the same distance apart as in Figure 25. The three closely spaced cross-streets are 300 feet apart; the intersection point symbols are 1/8-inch in diameter. Intersection mapping is one of the main design issues for the NRBM specification for MIRE 2.0 data, but, as this example shows, internal details of the intersections are completely invisible at the required mapping scale.

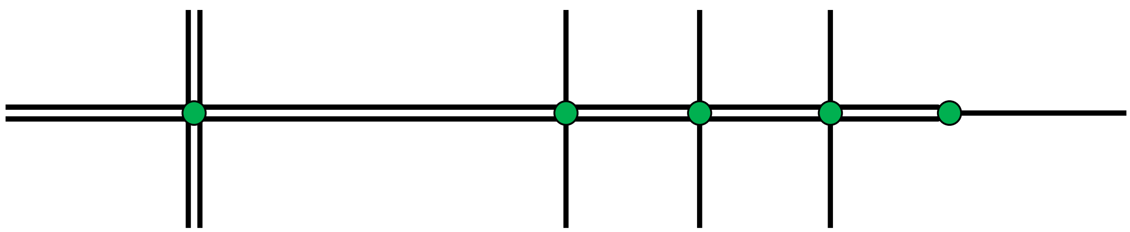


Figure 26. Example street system.

The route segments, intersections, and termini specified earlier are for constructing linear road networks. In addition to terminating a route segment, an intersection can also be used to attach basic data specified in MIRE 2.0, and, with the appropriate software, create approach segments for additional MIRE 2.0 data elements. A point intersection feature is adequate for these system-wide tasks.

The NRBM specification mandates that divided roadways be represented using directional centerlines, one for each direction of travel. It also requires there be a single intersection point for each intersection and a topological connection between intersections and segments. The only way to meet these requirements is for divided roadway segment centerlines to bend from their physical path to turn toward the intersection point and begin or end at that intersection point. Figure 27 shows one way to bend segment centerlines to meet at a common intersection point. This structure has been referred to as the “star” or “spider” intersection. Very few States have implemented this approach.

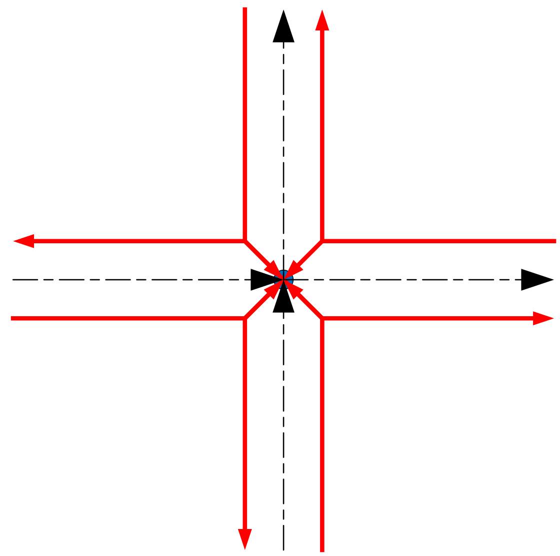


Figure 27. Segment centerlines at intersections in the NRBM specification.

Most States use a "hashtag" approach to mapping intersections, as shown in Figure 28, because this most closely represents the driven path and the way LRM-based data are collected in the field. There are four cartographic intersections (blue dots) in Panel A where route centerlines cross. However, these are not coincident with the intersection point or the LRM-based positions of the intersection on the route centerlines; i.e., they are not actual intersections. (They could be navigation points, but these objects are not part of the NRBM's national extent.) Nor are they at the right location to represent the LRM position of the intersection on that centerline, as shown in Panel B. Route segments cannot begin and end at the cartographic intersections shown in Panel A. Doing so would create four tiny route segments equal to the width of the intersection.

The four implied intersection points (blue dots) stored in the LRM-based roadway inventory reflect the intersection's measure value on that route, but do not actually coincide with the intersection point. However, from a map making perspective, route segments could begin and end at these locations, if topological connectivity to the intersection point was not required.

Since the native object is the inventory route and the route segments will be derived from them, it is expected that dynamic segmentation (dynseg) will be used to "clip" segment centerlines stored as LRM-defined linear objects from inventory route centerlines. Those route segment objects would terminate at the points shown by blue dots in Panel B. It is apparent, though, that subdividing these inventory route centerlines at the LRM-based intersection locations will not produce topologically complete route segment centerlines for divided roadways. None of the resulting segments will connect to the "real" intersection point. To meet the topology requirements, the route segment features must terminate at an intersection point (or a terminus object).

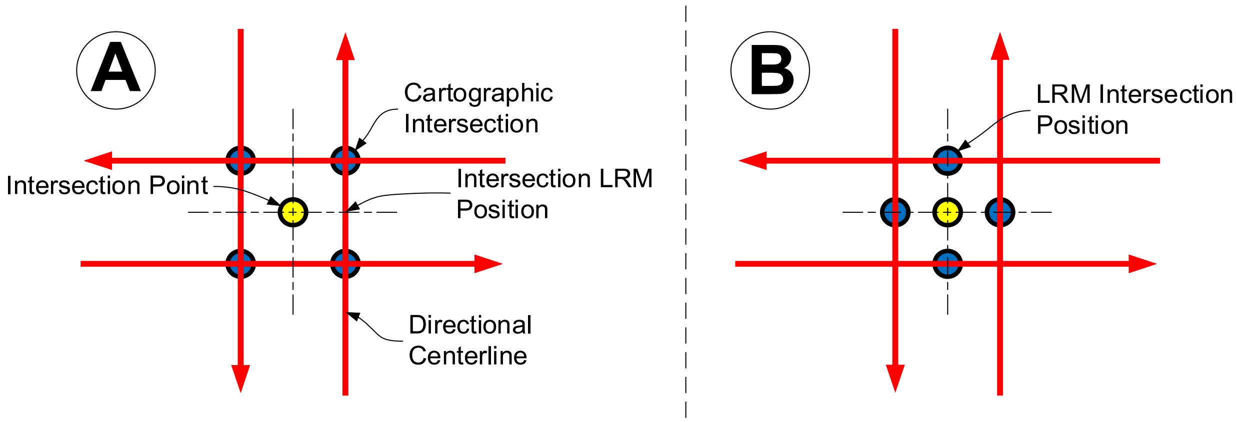


Figure 28. Hashtag intersections, where route centerlines do not meet at the central intersection.

Fortunately, common GIS software functionality can provide a solution. The basic approach is shown in Figure 29. The initial state is shown in Panel 1. Only inventory route centerlines exist. Route segments are defined as linear LRM objects whose endpoints are defined by the LRM-based position of the terminating elements (intersection and terminus objects). Panel 2 shows the use of dynseg to create centerline segments corresponding with the LRM extent of each route segment. Panel 3 shows the result of using GIS software functions to create lines that go from the end of each route segment to the intersection point. These short topological connections are then added to the route segment centerline to produce the final result: a topologically complete segment centerline. This solution meets the many applicable business rule requirements without placing any extraneous requirements on inventory routes or their representative centerline features.

Of course, the NRBM specification requiring that segment centerlines artificially bend to meet at a single intersection point is not a perfect solution. The artificial appearance that is visible at the scales used in Figure 27 and Figure 29 will not be visible at NRBM mapping scales. The intersection points used in Figure 26, for example, completely mask all internal intersection geometry. Large-scale mapping of intersections should use a completely different approach, as shown in a later section of this chapter.

In addition to the controversy surrounding the choice of a spider or hashtag intersection structure, there was considerable debate regarding the treatment of roundabouts. The final decision was to treat them the same as other intersection types. The internal structure of an intersection is beyond the capabilities of an LRM-based roadway inventory, which deals with larger geographic extents and smaller mapping scales. At the typical map scale for which the NRBM dataset will be used, it is not possible to see the internal structure of an intersection, even for a fairly large roundabout. The NRBM roadway inventory is an abstraction of the real world. Attempting to accurately describe the internal structure of intersections violates the central tenants of the LRM abstraction.

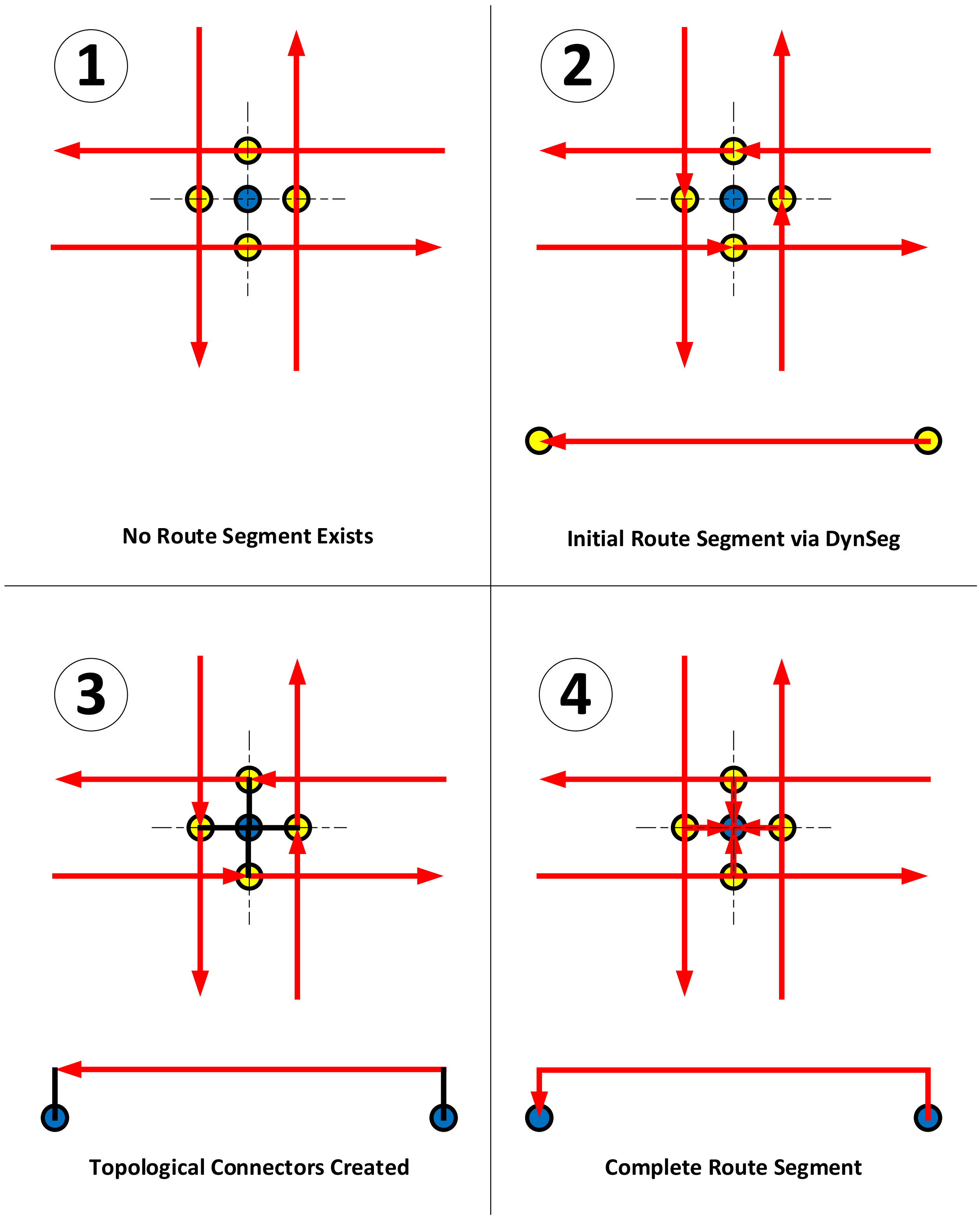


Figure 29. Deriving route segments from hashtag centerlines.

There is one route segment type that needs additional attention. Traditionally, highway inventories enforced a business rule that said the end of an interchange ramp was located at the physical gore. This is the point where the ramp pavement merges with the limited-access highway pavement. The rule was based on the conclusion that the portion of the ramp beyond the physical gore was another lane on the limited-access highway. Extending the ramp all the way to a visually pleasing point of intersection with the directional centerline of the limited-access highway would result in a double-count of the centerline miles for the ramp. This rule worked well until GIS became commonly available and users expected to see the ramp connect to the limited-access highway. ARNOLD actually adopted the requirement for such cartographic connections and for the general requirement that centerlines match travel pathways.

The NRBM specification enforces the ARNOLD requirement, as shown in Figure 30 and Figure 31. Both illustrations show that the route segment for the ramp with LRM measures still ends at the physical gore. The measure value for the intersection of the ramp on the limited-access highway is given the LRM position of the physical gore. There is a ramp junction type of intersection located at this point. To complete the path to the merge point, which is further downstream on the limited-access highway's directional centerline, the NRBM specification requires that a ramp extension centerline be created. This centerline will not become part of an inventory route, but it will be available for mapping and MIRE 2.0 analyses.

This visual solution still leaves us with the issue of how to topologically connect the ramp's inventory route, which stops at the measure value representing the physical gore, to the intersection point located on the limited-access highway's directional centerline. The answer is again provided by the topological connector. The route segment centerline will be extended from the ramp junction intersection to the intersection point on the directional centerline of the limited-access highway using the topological connector.

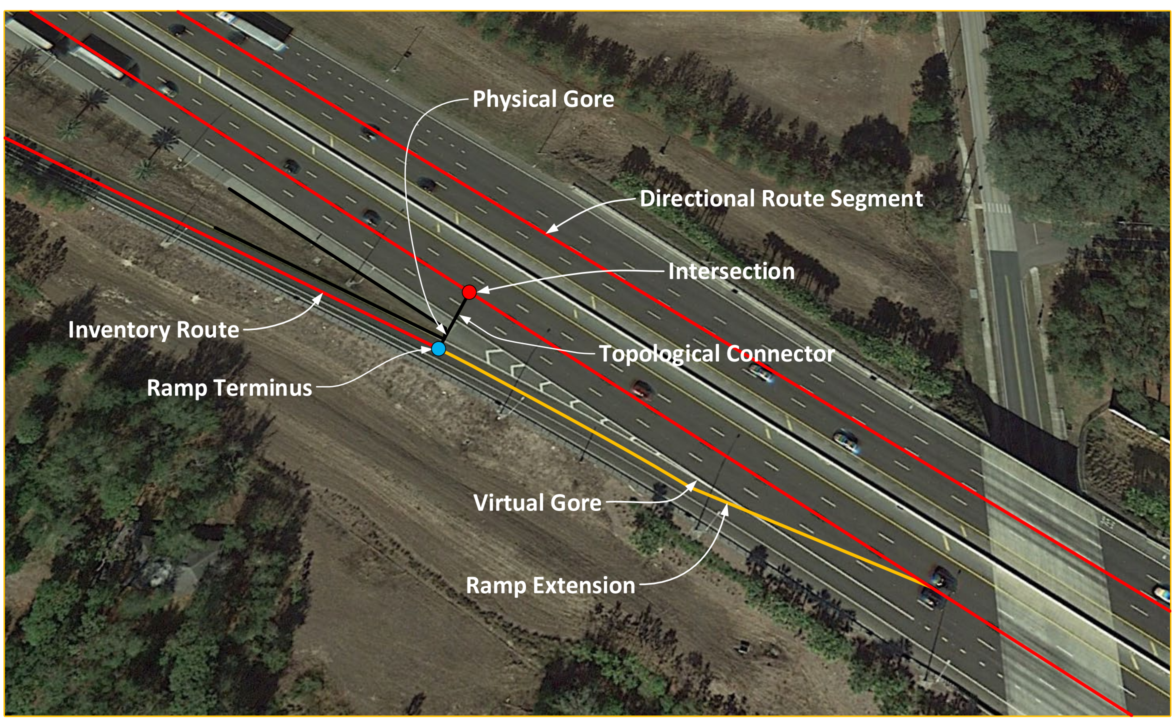


Figure 30. Interchange ramp—acceleration/merge route segment example.
Map data ©2019 Google® Earth™

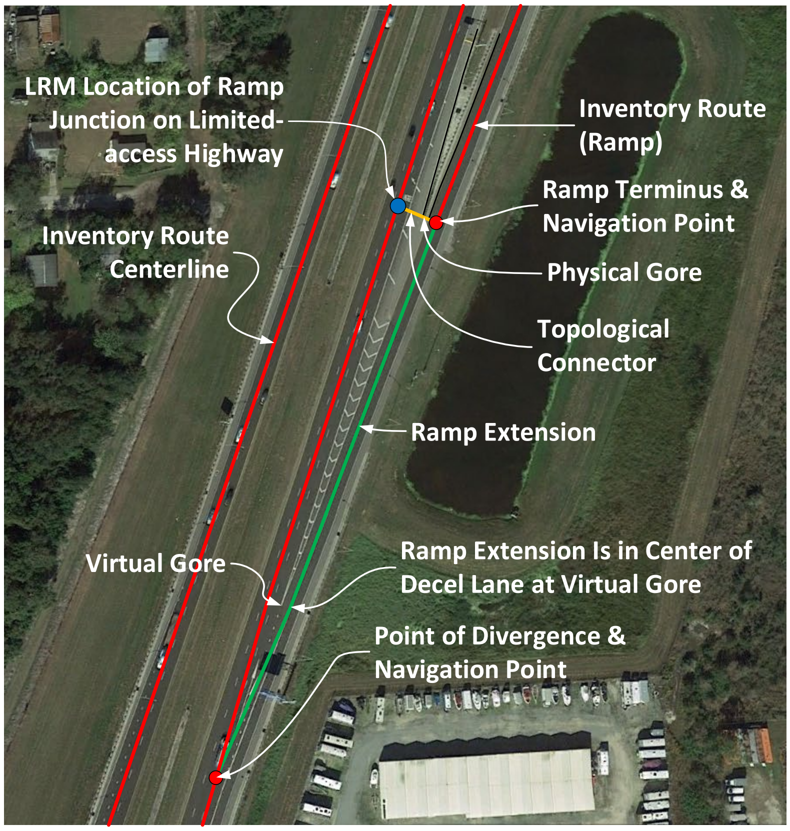


Figure 31. Interchange ramp—deceleration/diverge route segment example.
Map data ©2019 Google® Earth™

As these two examples show, the general recommendation is for the ramp extension to follow the center of the marked lane from the physical gore to the virtual gore, at which point it turns slightly to intersect the limited-access highway's directional centerline. There is no specific requirement for the shape of this portion of the extension's centerline. Pavement width may be a guide and consistency is a desirable characteristic, but these are abstractions of the real world that have been included for topological completeness and large-scale map appearance. The same approach should be taken when a ramp turns into a new lane or terminates an existing lane on the limited-access highway.

Figure 32 shows a full interchange with its many route segments, ramp extensions, ramp termini, and intersections. It also shows the interchange point.[86] Such point features differ from intersections in that an interchange point exists simply to attach to interchange attributes and to provide a small-scale indication of the interchange's location on a map of the transportation system. The interchange feature is not involved in the network topology and does not represent a node in the network or a termination point of a route segment. There is no specific rule regarding its geographic location.

[86] Interchanges may also be defined by a bounding polygon for specific analytical processes, but the NRBM does not include them.

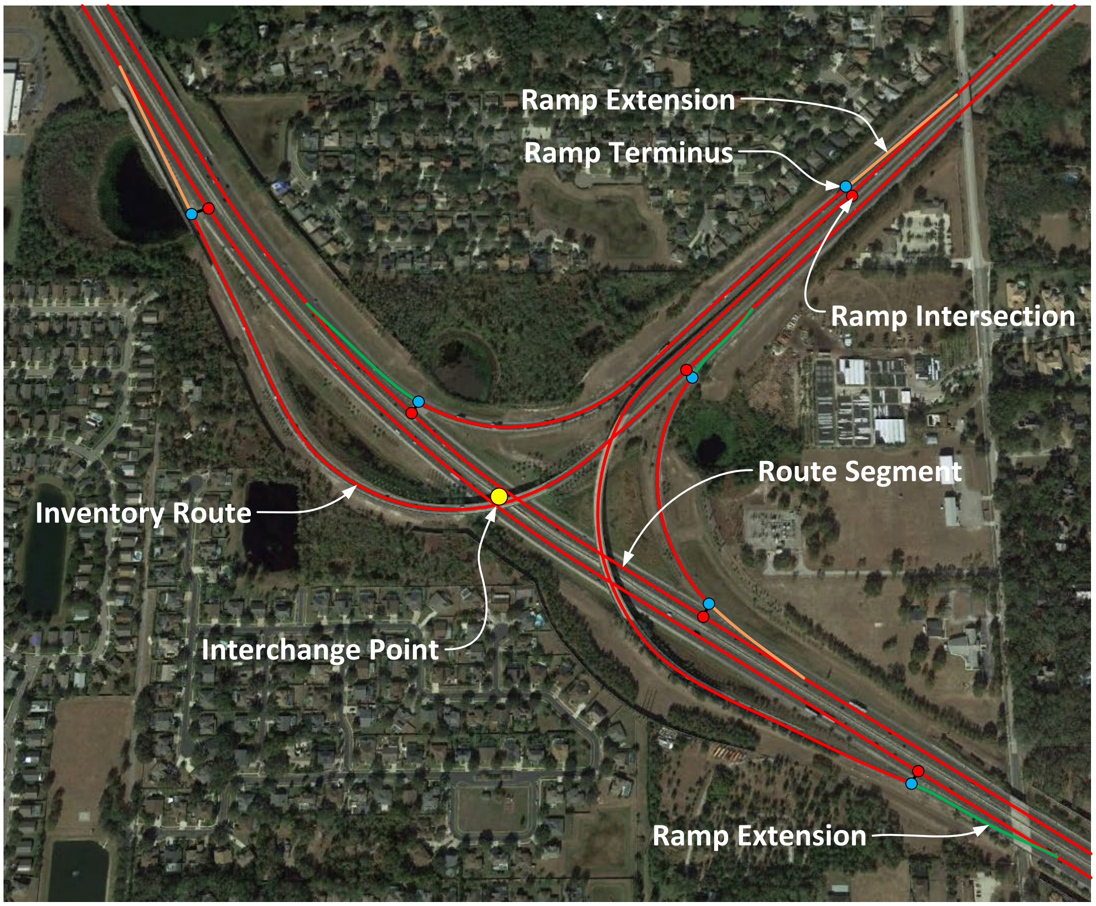


Figure 32. A complete interchange showing ramp extensions, route segments, intersections, and the interchange point.
Map data ©2019 Google® Earth™

Although the interchange point is not topological and does not have a location on all the involved routes, it should be given an LRM position on the limited-access highway. Doing so will permit transforming the measure values along the highway into a mile marker form that will better serve some users. The Interstate Highway System and other limited-access highways have a side-of-the-road linear referencing system using signed mileposts. Many States assign interchange (exit) numbers based on the mile marker system. Some have gone so far as to install fractional milepost markers as an aid to motorists who report traffic crashes and other incidents using cellphones. Police officers may also refer to the distance from a mile marker in reporting the location of traffic crashes.

Often, the point where a roadway transitions from divided to undivided or vice versa is at a mid-block location. To represent this transition graphically, the ARNOLD Reference Manual directs the States to shape their carriageways to follow the flow of traffic.[87] The NRBM adopts this convention. This modeling of medians differs from that provided by HPMS- and LRM-based inventory systems, which have set the point of divergence or convergence at the ends of the physical median. ARNOLD suggests using a cartographic taper around the median obstruction for a better map appearance that more realistically conveys the path of vehicles. This practice results in longer centerlines than implied by the related

[87] Op. cit, p. 55.

median LRM measures. In other words, to be compliant with the ARNOLD guidance, the directional centerlines will be longer than the median that creates the divided roadway.

The distance measurement for the divided portion of the roadway goes from the physical length of the dividing median to the effective length of the division between directional traffic flows. The actual length of the median has not changed, of course, but the longer distance between divergence and convergence points reflects actual traffic flow lines. This Guidebook recommends that the median length, stored as a linear LRM object and reported in HPMS, be left as it is; i.e., the length of the physical separation. The median will be stored as an element with an LRM position on both directional centerlines.

Figure 33 illustrates how to provide consistency between the starting and ending points of the directional centerlines and the LRM-derived length of the median. At the top is a drawing depicting the current HPMS practice. The median length is the length of the physical divider. Because an LRM is 1D, there are no connectors from the end of the undivided roadway to the two parts of the divided roadway.

For the NRBM specification to be consistent with guidance in the ARNOLD Reference Manual, instead of starting and ending exactly at the face of the median barrier, the beginning and ending points will be determined by the driven pathways. Figure 33 shows that these points of transition are classified as continuity intersections. The length of the median within the context of LRM measures would be extended to match the distance from navigation point to navigation point. This approach to divided roadway transitions is consistent with the standard for segment centerlines at intersections.

A continuity intersection is one where a decision point exists, but only one choice is actually available to the driver or pedestrian. Another example is where two route segments connect at a curve. A match point intersection, which can be viewed as a subtype of the continuity intersection, will be created at data supplier boundaries. A match point provides the common place of connection that is used by all data suppliers. States will assert the match points within their boundaries. Absent a State-to-State agreement as to match point locations, the USDOT will declare their location at State boundaries.

It is possible to create topological connectors for routing continuity at points of median divergence and convergence, as was done for route segments at intersections. That approach was not selected because they serve different purposes. Route segments with topological connectors are not expected to be used for mapping, which should use the route segments clipped from inventory route centerlines. The topological connectors added to route segments to provide pathfinding continuity actually detract from cartographic continuity. The modified approach and transition geometry showing the divergence and convergence of the directional centerlines with the bidirectional centerlines at each end of the divided segment is required for mapping continuity at the 1:5,000 scale.

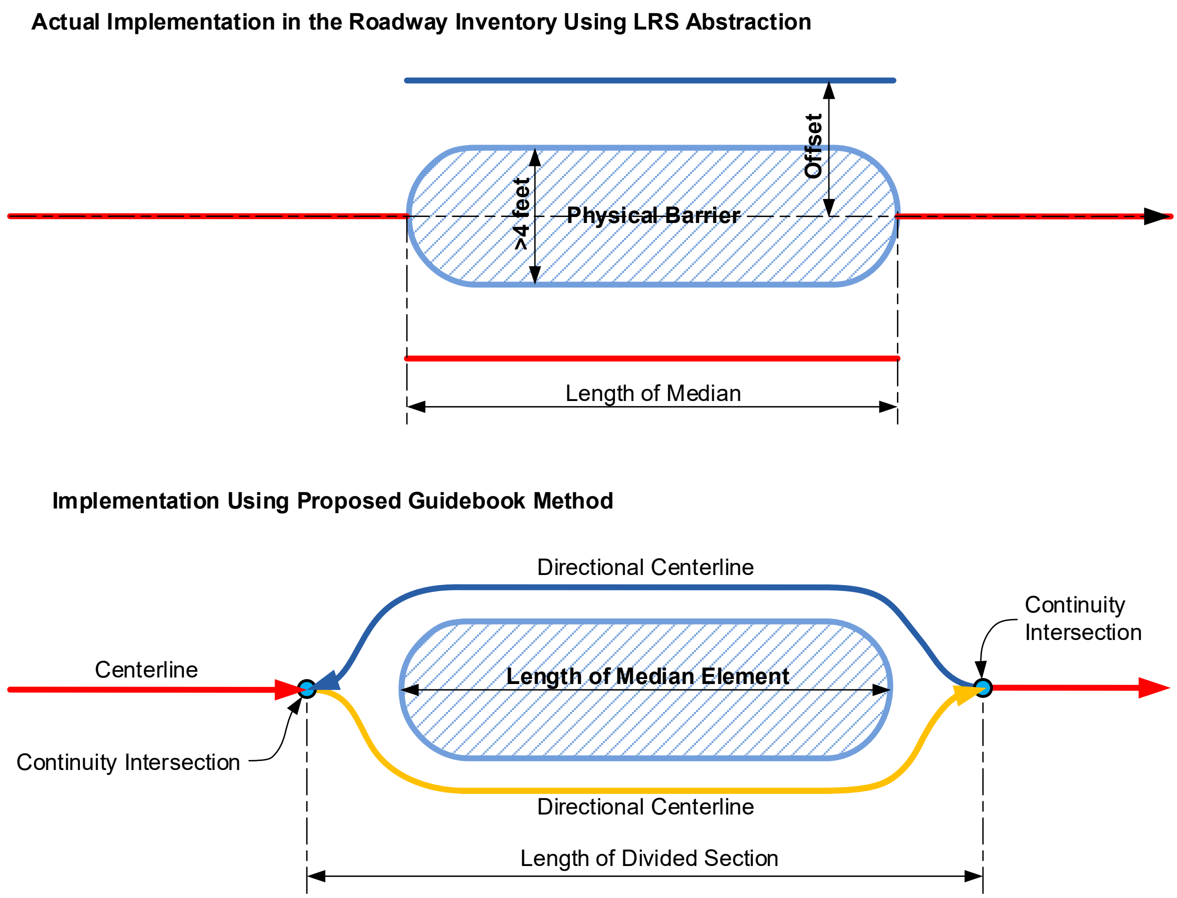


Figure 33. Directional centerlines need to connect to terminal intersections for topological continuity.

Not all intersection types break route segments. Figure 34 shows the use of at-grade, roadway access, and median cut intersection types along a divided roadway. The at-grade Tee intersection near the center of the photograph breaks the route segment for this side of the roadway, but has no impact on the paired directional centerline. The roadway access and median cut intersections have no effect on either directional centerline, but can serve as potential reference LRM objects for attaching MIRE data, turning movement counts, and crash records. Simply looking at an aerial photograph may not be sufficient to determine the proper intersection type.

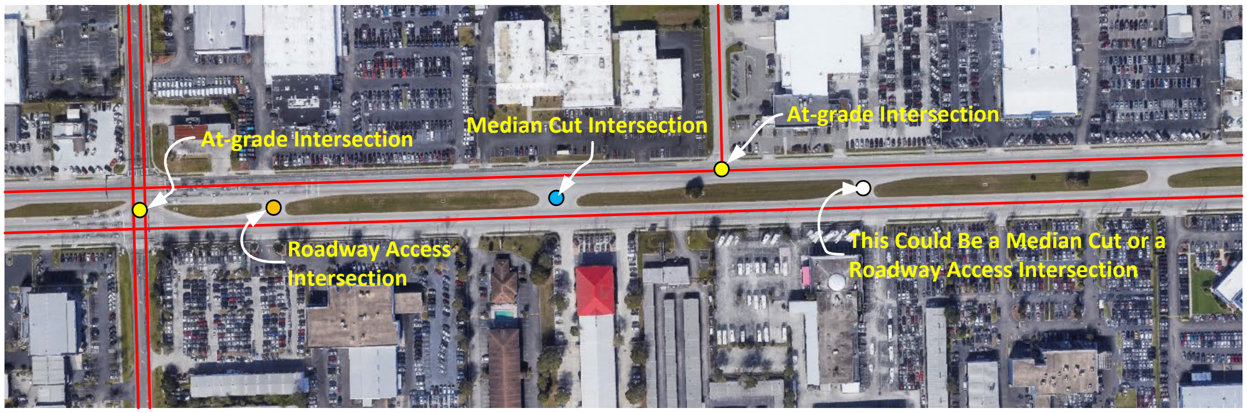


Figure 34. Typical use of median cut and roadway access intersection types.
Map data ©2019 Google® Earth™

Terminus types also have an influence on route segment geometry. Figure 35 shows the two variations of cul de sac. On the left is a dead end terminus that begins a route segment. For these types of terminus, which will have a radius of zero feet, the feature is placed at the physical end of the roadway. In contrast, the cul de sac on the right side of the figure will have a positive radius, so the terminus feature is positioned at the center of the cul de sac. By their nature, cul de sacs typically end short route segments, so the difference in length will not be noticeable at NRBM map scales. In fact, mapping efforts may require that inventory routes shorter than some minimum length be omitted from the map.

There is one last point to note regarding the basic data model. In addition to the 'Begin' and 'End' relationships between Intersection and Route Segment, there is an 'Along' relationship. This is provided to allow a single route segment to span multiple intersections; i.e., to create "long" route segments. This capability has been provided to allow each supplier to moderate route segment lengths. For example, urban areas often include very short block lengths. At the 1:5,000 scale, maintaining route segments of less than one inch in map length may impose greater workloads than justified, particularly if they all have the same attribute values. Intersection objects would still exist along the route segments so that shorter, intersection-to-intersection route segment might be generated when they are needed.

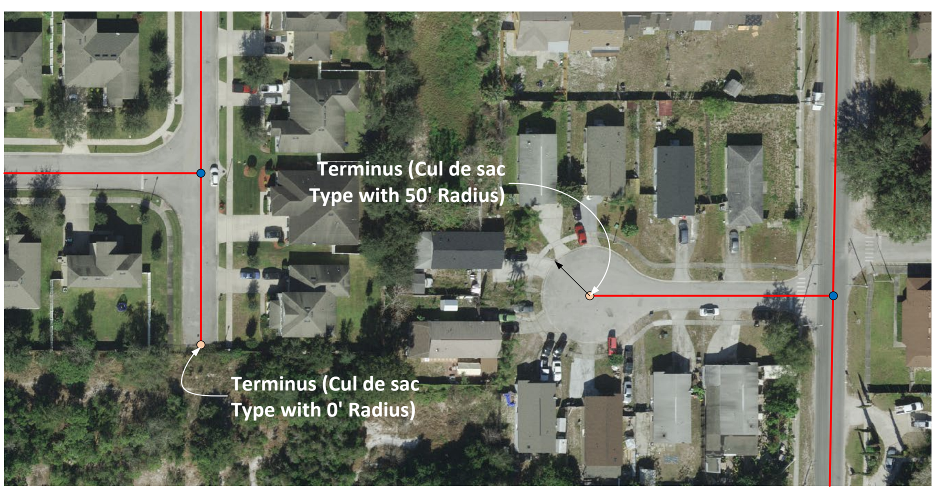


Figure 35. The impact of cul de sac radius on route segment geometry.
Map data ©2019 Google® Earth™

Inventory Routes and Linear Referencing

Conceptually, route segments are combined to form inventory routes, which are the objects that have LRM values. An inventory route is a continuous, ordered sequence of whole route segments. If a segment has two directional centerlines, these will produce two directional inventory routes. Figure 36 shows the data model that supports the process of combining route segments to produce inventory routes. These are the 1D datum objects that apply the LRM to the transportation system.

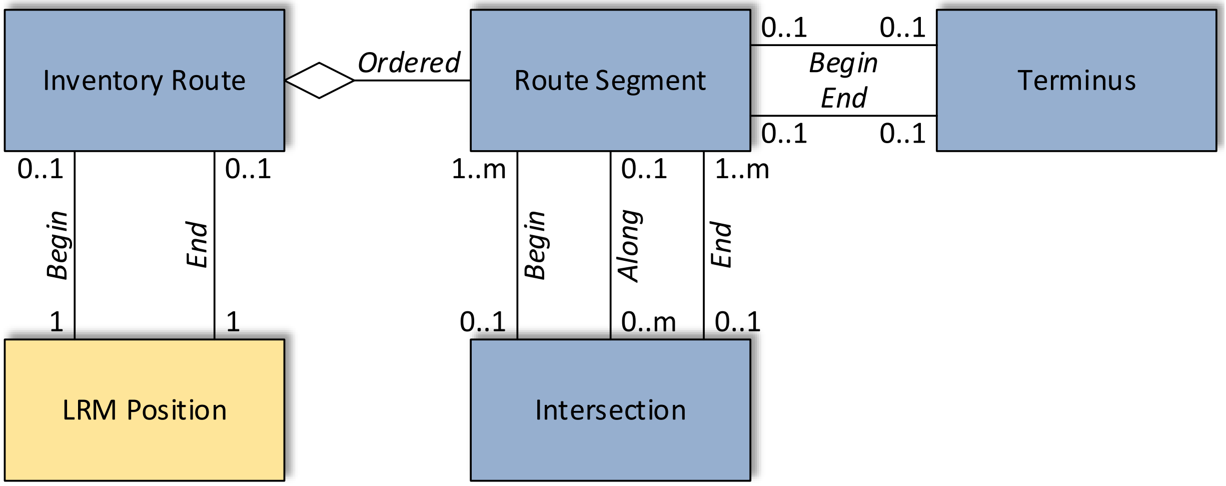


Figure 36. Adding inventory routes and linear referencing to the basic model.

Figure 36 introduces a new relationship type, the aggregation, which is shown by the diamond on the left end of the line. This means that an inventory route is a collection of route segments and the annotation says the collection is ordered; i.e., there is a specific sequence of route segments that form an inventory route. The blue background for Inventory Route means that it is also a feature class. (It has to be because it is a collection of features.) Below Inventory Route is LRM Position. The cardinality relationships shown indicate that one LRM position will begin an inventory route and that another will end the inventory route. The tan background for LRM Position means that it is not a feature class; i.e., it does not include geometry. An LRM position provides a measure value for the beginning and end of an inventory route.

Although the NRBM concept says that route segments are combined to form inventory routes, the reality is that the reverse is the actual production sequence: inventory routes are subdivided to construct route segments. This is because the States are the source of NRBM data. State DOTs normally have inventory routes that span fairly long distances, such as traversing entire counties. The route segments shown in the preceding data models are unlikely to exist natively at most State DOTs. Route segments are generally found in local government transportation databases. Therefore, while the preceding fundamental data models express the core of the NRBM database and the implementation guidance provided in this document, they are an unlikely place to begin the process of actual implementation. They were, however, a good place to begin the process of expressing the NRBM specification.

An LRM position assigns a measure value to a location on an inventory route in order to describe that location in a manner consistent with the LRM. LRM positions in the NRBM dataset are assigned to objects located at positions along an inventory route according to their real-world distance, in miles, from the inventory route's origin. This distance is called a measure and is stated in units of 0.001, which is abbreviated by the Greek letter *mu* (M). Initially, an LRM position must be provided at each end of an inventory route to create a linear datum. One LRM position is the origin and all positions along the inventory route are defined in terms of their distance from the origin. The two defining LRM positions will correspond with the location of an intersection or terminus since an inventory route consists of one or more whole route segments, and route segments must begin and end at an intersection or terminus. In NRBM, an inventory route begins at 0.000 and ends at a measure equal to the LRM length of the route. Those two values constrain the range of measure values that may exist along the inventory route.

Figure 37 adds to the LRM capability of the data model. The focus in this figure is on inventory routes and the LRM, which essentially says that an LRM position is used to place LRM objects on an inventory route. An LRM object is a characteristic, element, or event—anything that can be tied to a location along a route. LRM positions can now be located along an inventory route. Each LRM object must be located by at least one LRM position. Each LRM position may locate one or more LRM objects.

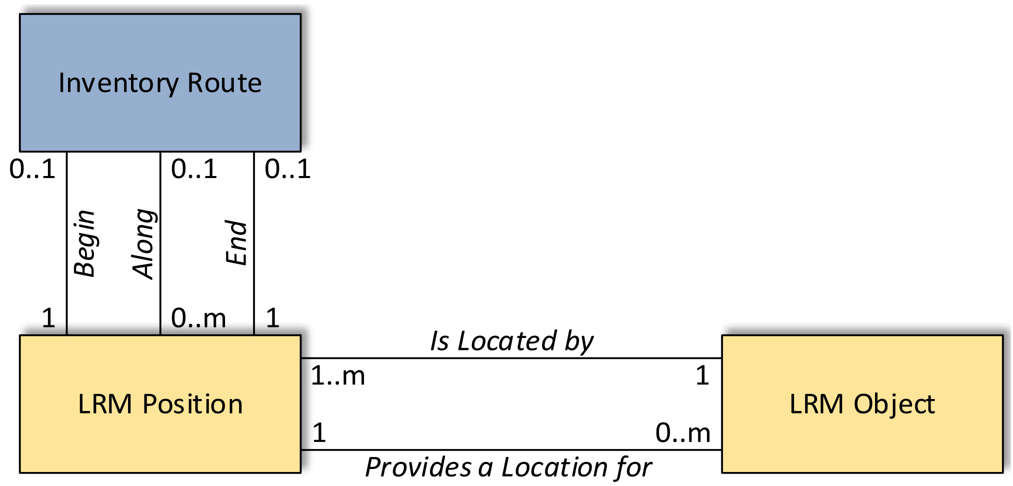


Figure 37. Adding more LRM positions and what they locate on the inventory route.

The data model in Figure 37 is simplified by providing only a single implied LRM datum. The NRBM model supports multiple LRM datums, so the model needs to be expanded by adding a Linear Datum class. This addition is shown in Figure 38. Also needed are defined positions within geographic datums. Figure 39 adds two more classes to accomplish that task.

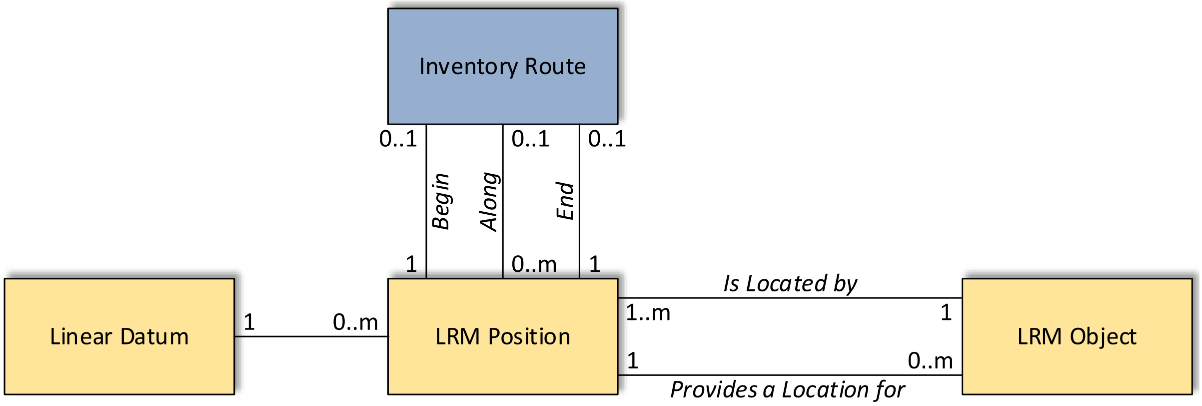


Figure 38. Adding other LRM datums to the model.

Note in Figure 39 that both ends of the relationship between LRM Position and Geographic Position have a multiplicity that starts with zero. This means that the relationship is optional. LRM positions can exist without equivalent geographic positions, and some geographic positions will not be linked to LRM positions. They will not describe a location along a roadway, at least not a location defined by an LRM position. There may, however, be a spatial coincidence of a point along a route centerline. As with linear datums, the model provides for multiple geographic datums to exist in the dataset.

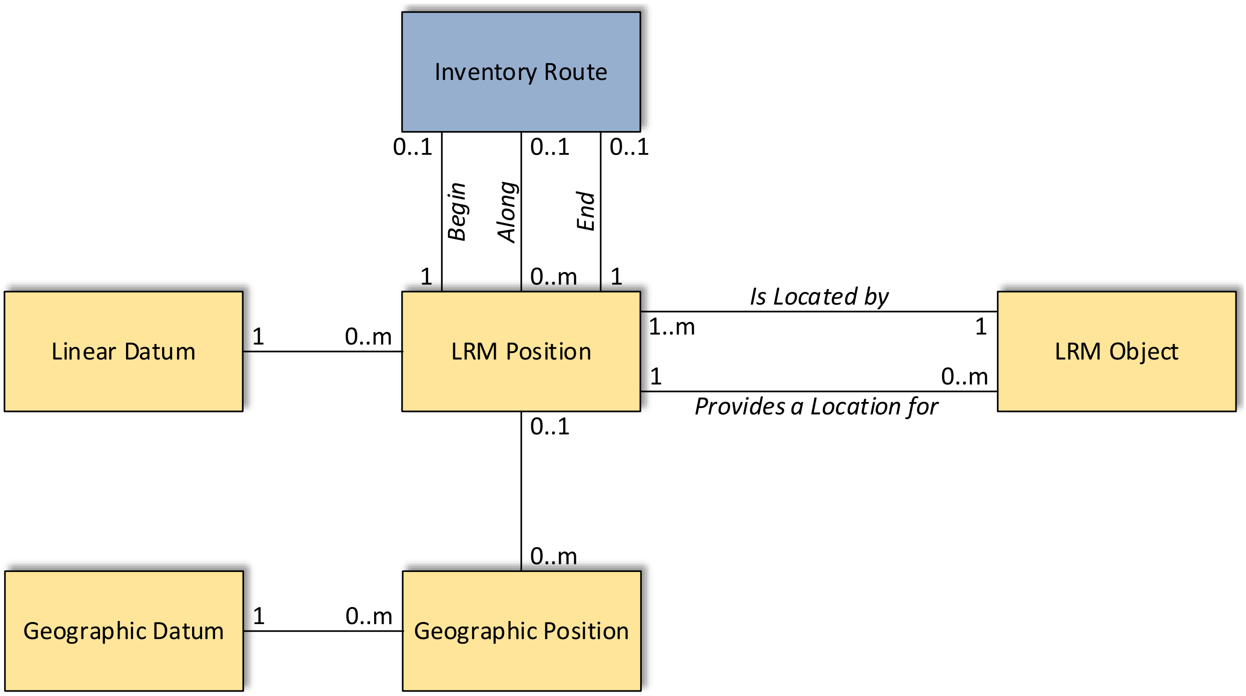


Figure 39. Adding support for geographic locations to the model.

Route segments and intersections—and many other roadway characteristics, elements, and events—will have LRM positions assigned to them. They will be stored as LRM objects. Dynamic segmentation (dynseg) can be used to convert the linear route segment LRM objects into route segment centerlines. This model extension is shown in Figure 40. The beginning and ending point features can then be added as attributes of the route segments by finding the intersection and/or terminus LRM object that has the same LRM position value as each end of the route segment.

A paradox is presented by the data model in Figure 40. What started as a simple model that showed inventory routes are an ordered sequence of route segments is now a model where route segments are derived from inventory routes. This is the classic chicken and egg problem: which comes first? In this case, it is the chicken. Inventory routes come first, with route segments being derived from them. It is still conceptually correct to say that an inventory route is composed of one or more route segments, however, that is not the preferred method for constructing route segments from existing State roadway inventories.

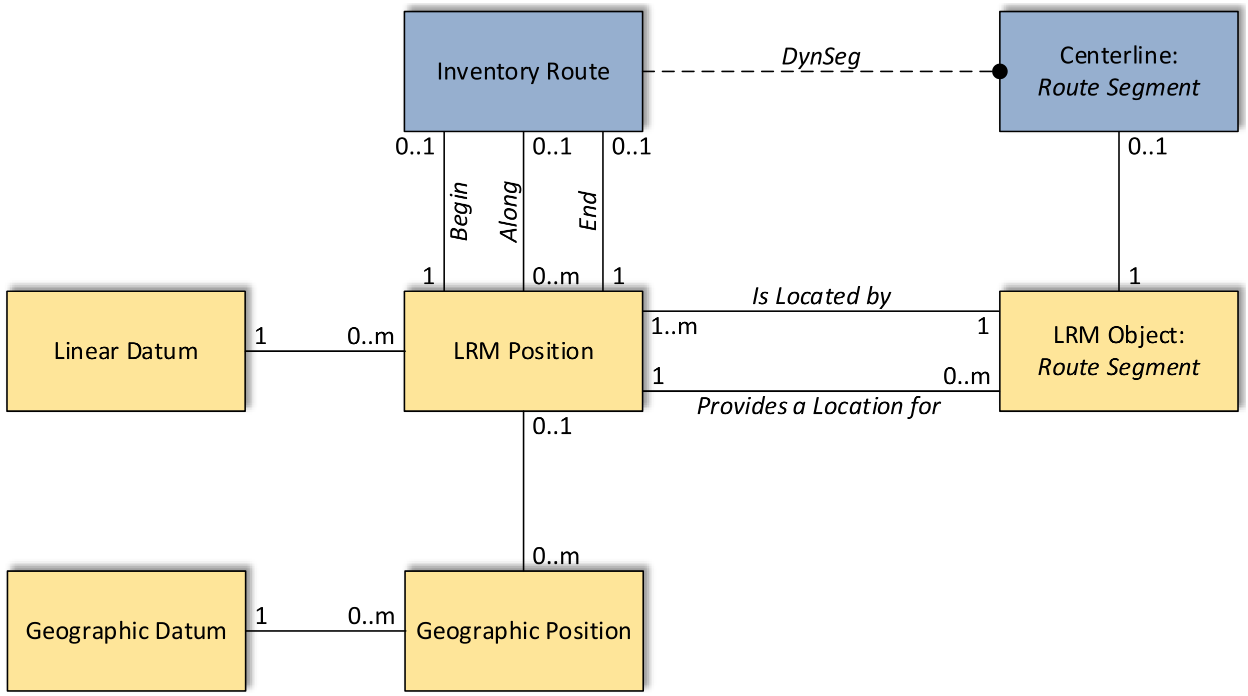


Figure 40. Route segments become LRM objects.

There is another problem: Inventory Route cannot be a feature class. An inventory route is defined as an ordered sequence of one or more route segments, and route segments are no longer cartographic features. In addition, what provides the ordering of route segments? LRM positions do. In fact, an LRM position defines the extent of each inventory route. An inventory route may be cartographically displayed using a centerline feature, but it has to first exist as a non-geometry class, as shown in Figure 41. A Centerline class has been added with a one-to-many relationship with Inventory Route, which means that one inventory route may be represented by any number of centerline features. The multiplicity of centerline features is an expression of both retaining historical versions that have been retired and multiple scale-specific versions for making maps at many scales.

The process has now gone full circle from a data model that had everything as a feature class to one that has feature classes included only as representations of classes that are not cartographic in nature. The Intersections and Terminus classes have become LRM objects. Like route segments, intersections and termini are assigned LRM positions that can create point objects along inventory routes using dynseg.

Unfortunately, the data model shown in Figure 41 is wrong. It is missing a key ingredient. The NRBM centerline cartography editing process starts with a simple line that describes the path of an inventory route. It has no measure values. Dynseg depends on the centerline having measure values assigned to each vertex. How are measure values added to a centerline feature?

The answer is provided by calibration points, which are the cartographic expression of an LRM position. They are used minimally at the ends of the inventory route centerline to make it LRM aware by assigning a measure value to the start and end measure. These calibration points are the LRM position values for the applicable intersection and/or termini that begin and end the inventory route. The calibration process can use these values to interpolate all the measure values on vertices located between the beginning and ending points. Additional calibration points may be used to ensure that LRM positions are properly mapped along the centerline.

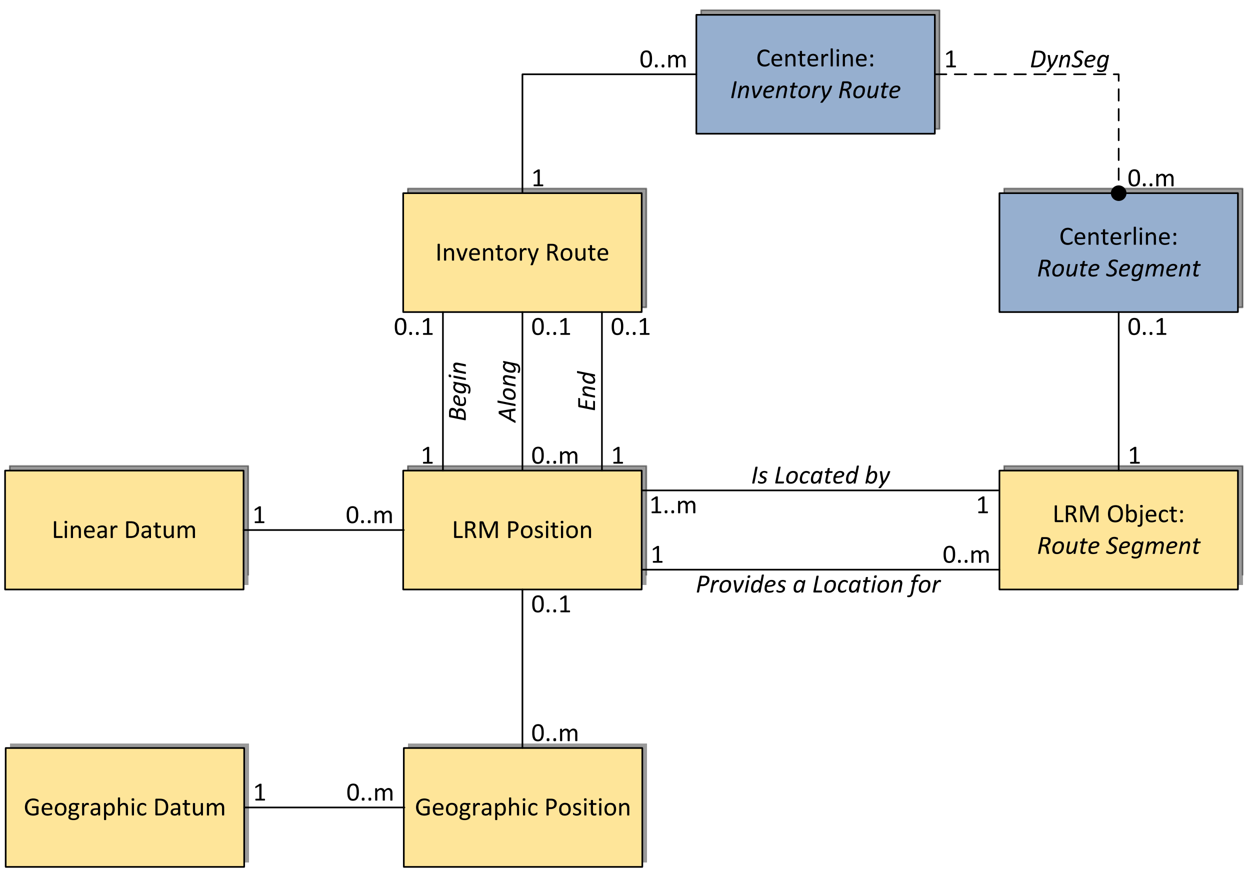


Figure 41. The refined LRM data model.

The basic requirements for a calibration point are that it has a defined LRM position and is readily identifiable in the field and on aerial photography that may be used to create the centerline feature. As shown in the data model in Figure 42, LRM position information is acquired from the LRM object that is present at the location of a calibration point. This means the calibration point knows both its LRM location and the type of LRM object it represents. Potentially, every intersection point, bridge end, and railroad grade crossing could be used as a calibration point. Since centerline editing will involve the addition, removal, or modification of a whole segment centerline, converting all intersection points to serve as calibration points is strongly recommended. (Terminus LRM objects will always be a calibration point, as they represent the end of an inventory route.)

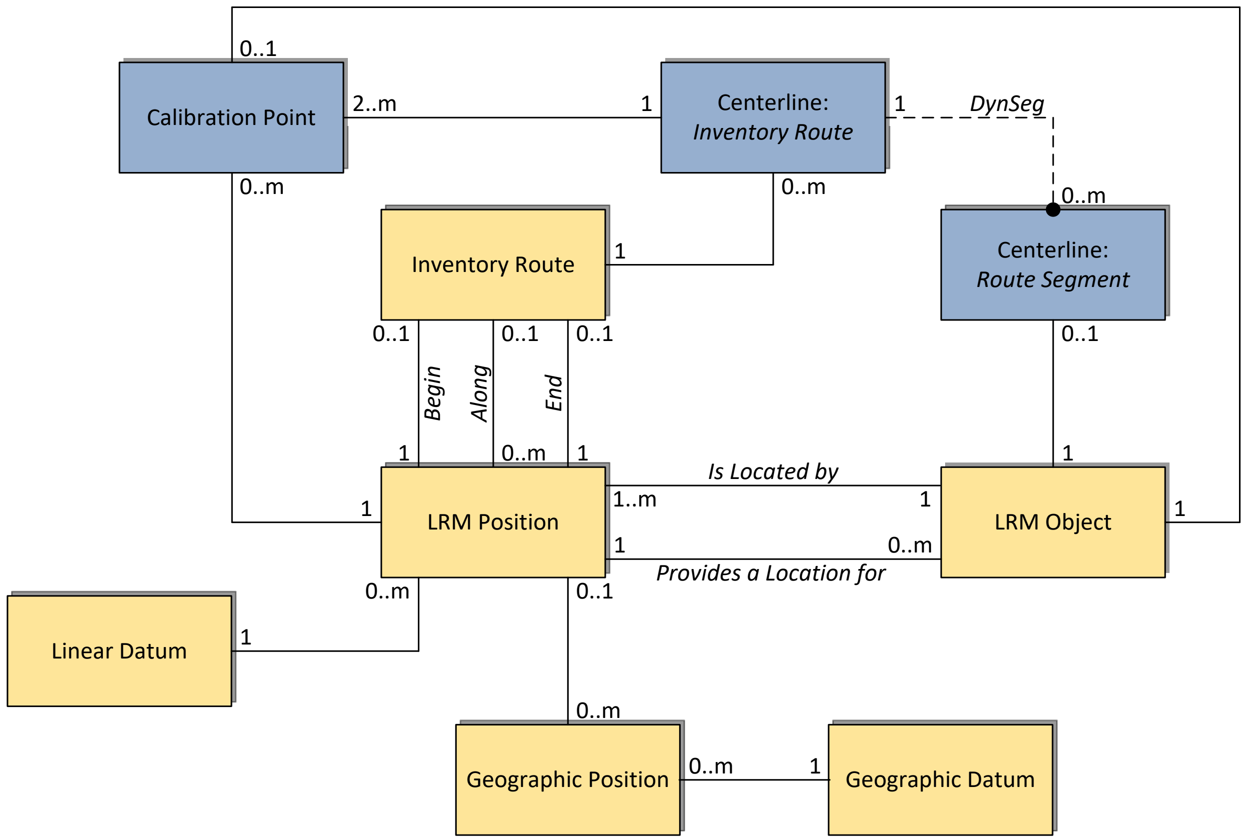


Figure 42. The complete LRM data model.

Prior to calibration, inventory route centerlines contain only the route identifier and have no *m* values applied to their vertices. The calibration process adds the *m* values. This design approach permits a single centerline to have multiple LRM datums applied to it and allows a single route to be illustrated by more than one centerline feature. It also simplifies the editing process, which can ignore measure values, and precludes the negative impacts of centerline editing on measure values stored in the business data tables. Centerline editing is completely separate from business data editing.

A fundamental business rule stated in Chapter 3 is that cartographic objects cannot alter LRM positions. Measure values are only applied to and not derived from the inventory route centerline. Regardless of the shape and length of a centerline segment, LRM objects will be placed in the "correct" interpolated location and will not affect the location of objects on downstream segments. An LRM object should not "move" (i.e., experience a change in its measure value) when the centerline cartography is edited.

Except for the inventory routes, everything with an assigned LRM position is an LRM object.[88] The model shows that an LRM object exists independently of its position description provided by the related LRM Position records. The alert reader will have noticed that the relationship role of "is located by" has a

[88] Technically, inventory routes are LRM objects in that they are assigned an LRM position to define the starting and ending measure values for the route, but they have different applications and behaviors, so they are in a separate class.

cardinality of one-to-many: one LRM object may be related to one or more LRM position records. This capability is useful for things that involve more than one inventory route, such as an intersection or bridge. Depending on the application or scale of representation, some roadway elements, like bridges, may need both point and linear position descriptions.

Figure 43 shows how one LRM position identifier can actually be assigned to several records, each giving a unique position description. In this example, an at-grade intersection has been given an **`LRMPositionID`** value of 3295. This intersection involves Inventory Routes 42 and 96. The database includes two LRM datums: State milepoint and local address block range. The complex primary key for **`LRMPosition`** includes the standard attributes of **`LRMPositionID`** and **`RecordDate`**, but it also includes **`InventoryRouteID`** and **`DatumID`**. Having two inventory routes and two LRM datums means that **`LRMPositionID`** 3295 has four position records describing one location. The intersection can be discovered following either inventory route (or its traversal derivative) using either LRM datum.

LRMPositionID	LRMType	InventoryRouteID	FromMeasure	ToMeasure
4723	StateMilepoint	67	239	526
3295	StateMilepoint	42	6372	
3295	StateMilepoint	96	2756	
3295	LocalAddress	42	2300	
3295	LocalAddress	96	1900	
6429	StateMilepoint	64	766	6352
2084	StateMilepoint	65	392	

Figure 43. One LRMPosition record may include many position descriptions for one location.

Although many State DOTs have adopted the practice of equating a signed route to an inventory route, it is important to stress that an inventory route has no direct relationship with a signed route in the NRBM specification. The route number or street name assigned to part or all of an inventory route—or to parts of several inventory routes—are attributes of those inventory routes; i.e., they are linear LRM objects. Data can still be published using signed routes as the means of expression, but data editing should be based on inventory routes that are independent of—and thus unaffected by changes in—the path of a signed route. A State may have traced the path of signed routes to initially create its inventory routes, but the connection stops there. Changes in signed routes are reflected in edits to business tables, not inventory routes or their centerline features.

Local governments are likely to use and maintain segment-based roadway databases. Attributes are denormalized, which means each route segment gets a complete set of attributes—even if the next route segment has exactly the same values. It is a fairly straightforward process to combine route segments to construct inventory routes for the imposition of linear referencing. LRM measures may be imposed on the resulting inventory routes based on map length or through other automated means. Once the linear datum has been established, the next step is to dissolve linear attributes to create LRM objects. Terminal intersections may also need to be created for each segment. LRM objects corresponding to route segments and intersections will serve as the basic mechanism for exchanging data between State DOTs and Local agencies. All LRM objects derived from route segment databases will begin and end at intersections due to the nature of the source data.

Thus far, the process has been to show that route segments are created and then combined to form inventory routes. This sequence may be reversed by subdividing an inventory route into its component route segments. For State DOTs and other agencies with linear referencing infrastructures, inventory route centerlines may be relatively long, spanning the entire extent of a facility across a county or more, and route segments may not exist as cartographic features. The task then becomes creating route segments, segment centerlines, terminal points, and intersection points. By creating linear LRM objects that correspond to the extent of each included route segment, dynseg may be used to create the route segment centerlines, terminal points, and intersection points. Intersection points may also be created by selecting the calibration points that represent intersections, when they exist.

Route segments, intersections, and termini are all stored as LRM objects, which means they have an LRM position along an inventory route. Dynseg can be used to “clip” segment centerlines from inventory route centerlines using the LRM position information stored for each LRM object. Only some of the entities need to exist as a starting point. For example, if only intersections are present in the LRM inventory database, then they can be used to subdivide inventory route centerlines into route segments, with termini being placed at all remaining centerline ends.

Once these objects are created, the editing protocols take over. The temporal editing process applies equally to feature classes and database tables. Editing route centerlines presents a number of problems. One of the more difficult problems presented by traditional LRM-based data systems has been the lack of clear demarcation of spatial edits, such as the result of improvements in centerline data sources or to reflect actual changes in roadway alignment. The adoption of a segmented centerline structure makes the route segment feature a logical spatial extent for geoarchiving. Edits to a centerline should produce an action to retire the entire segment centerline and store a new one after the needed modifications are completed. Edits to an intersection will likely require that all affected segments be retired, in addition to the intersection elements that were changed, when intersection edits alter the location of navigation points.

A new inventory route centerline feature will need to be constructed after all route segment changes have been made. This means the overall sequence is to create route segments by subdividing inventory route features at intersections, edit those route segment features that need to be modified, and then reassemble the inventory route centerline features. To ensure that measure values found in business tables describing LRM objects are not altered by geometric feature editing, the AEGIST Guidebook proposes that the editing process for geometric features and business data be separated. In addition, route segment centerlines should be stripped of measure values prior to editing them. Each segment centerline impacted by editing must be retired and a new active record created (unless the real world object ceases to exist). Once the editing is completed, the inventory route centerline should be reassembled and the calibration process reapplied.

Figure 44 illustrates the editing process involving route segments as the basis for revising geometric features and the objects they represent. The drawing depicts the changes that happened to a State highway crossing a river using Bridge No. 1735. This was a narrow bridge, so the State DOT constructed a new bridge, No. 2763, which had to be built on a different alignment to allow traffic to continue to use the old bridge during construction. After the new bridge and its approach sections were opened to traffic, the signed route for the State highway was placed on the new alignment and the western approach to the old bridge was removed when the bridge was demolished. The eastern approach was retained to provide access to property, was reclassified as a local road, and maintenance responsibility was transferred to the local government.

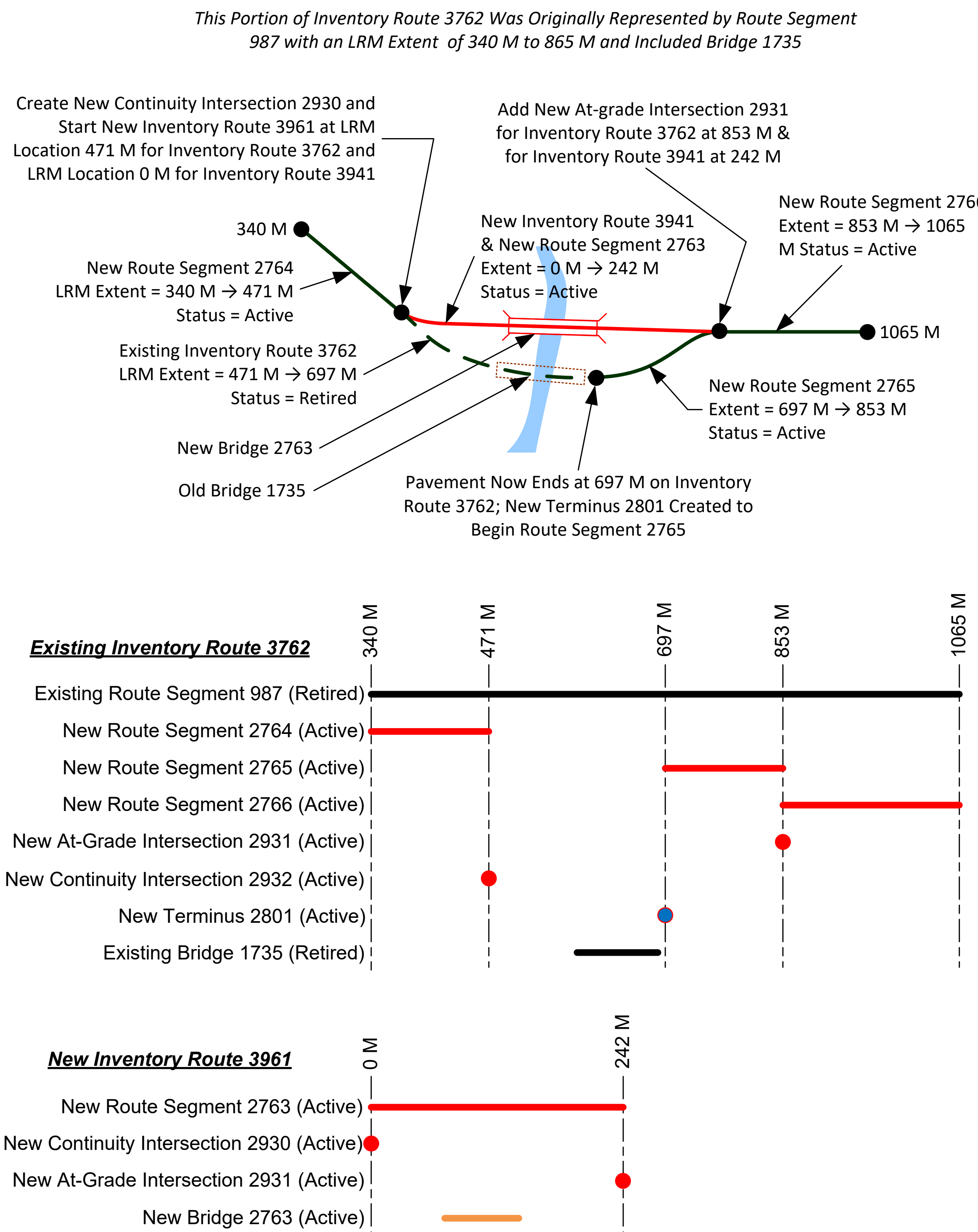


Figure 44. How temporal editing works for inventory routes and route segments.

Many States would move the signed route to the new alignment, calculate the measure values for the new alignment, and recalculate all the measure values until the end of the inventory route in order to accommodate the change in length. That amounts to a lot of work. In addition, changing measure values for LRM objects farther down the route can cause confusion for users. It also creates chaos for a temporal database, which will now store two measure values locating the same object on the same route.

Rather than revise the path of the inventory route to reflect the new path for the roadway, the advice is to create a new inventory route that matches the new part of the roadway. The bottom portion of the figure shows the inventory route, route segment, intersection, and terminus records that must be modified or created to record these events in the NRBM.

The process starts by creating Inventory Route 3941 with a beginning measure value of 0 and an ending measure value of 242. This inventory route could have been created as soon as the alignment was set by the design team, at which time it would have an **`ObjectStatus`** of 'Planned for Construction'. A single route segment, No. 2763, would span the length of this new inventory route. A new continuity intersection will be placed at Measure 0 on new Inventory Route 3941 and at Measure 471 for existing Inventory Route 3762. A new at-grade intersection will be placed at Measure 853 on Inventory Route 3762 and at Measure 242 on Inventory Route 3941. An LRM Object record should also be created for the new bridge. Assuming these changes were made when the project was proposed, all of these objects would be given an **`ObjectStatus`** of 'Planned for Construction'. Later, the **`ObjectStatus`** value may change to 'Under Construction', followed by 'Open to Traffic'.

Once the new alignment is open to traffic, Route Segment 987 needs to be retired, which includes this portion of Inventory Route 3762. **`EditReason`** will be 'Realignment'. **`EditComment`** will include a note saying that a portion of the route was removed from service. Next, Terminus 2801 is added at Measure 697 on Inventory Route 3462, and the route segment that goes from this terminus to the at-grade intersection placed at Measure 853 on Inventory Route 3762 (also described as Measure 242 on Inventory Route 3941) is constructed. Lastly, the new route segment that goes from this intersection to the end of the original route segment at Measure 1065 is constructed. Once that is done, Inventory Route 3762, which now has a discontinuous centerline is reconstructed. The relationship of Inventory Route to Centerline is now one-to-many for this route. Depending on how quickly the old alignment and bridge are demolished, a new route segment for the portion of Inventory Route 3762 that will be demolished may need to be created and given an appropriate **`ObjectStatus`**, such as 'Under Demolition'. The figure does not show this action.

For the business data contained in the LRM objects, the old Signed Route characteristic record for the portion of the route that was abandoned as a State highway is retired and a new Signed Route record for the new inventory route is created. Then, other LRM Object records can be created for the many characteristics and elements that exist on the new roadway. Similar objects that were on the old alignment are retired or modified. It would be helpful if the editing software supported a cascade of retirements for all LRM objects on an abandoned section of roadway. No edits are needed for the LRM objects that are outside the limits of the section affected by the realignment. Downstream objects will not change their LRM positions. No LRM object should "move" solely because another portion of the inventory route (or signed route) was realigned.

It needs to be restated that inventory routes are not signed routes or named streets. Assigned route numbers and names are attributes of inventory route sections; i.e., they are linear LRM objects. The same section of an inventory route can have any number of signed route designations. Constructing a signed route is merely a matter of selecting all inventory route sections that have the desired signed route designation. A route segment that is part of both SR 10 and SR 38 needs to be included in the path of each signed route. This requirement eliminates dealing with dominant and subordinate route designations in the editing environment. However, such practices may continue to exist in the published data, such as by publishing the inventory by following dominant routes and transforming inventory route measures to conform to the path of the dominant route. Inventory routes, named streets, and signed routes should consist of one or more whole route segments. The database will show there is a single inventory route segment with multiple signed route designations; data are not duplicated for both signed routes. To more readily accommodate multiple route designations for a single route segment, multiple route name fields are provided for the Route Segment object. LRM-aware inventory routes use a different approach.

Figure 45 shows a variety of ways in which the States presently deal with signed route overlaps when they form inventory routes. Commonly, the numerically lowest signed route number controls dominance; i.e., the dominant route has the data, which is why there is a solid red line illustrating the path of SR 10 on the right side of the figure. In order to avoid double counting and LRM object duplication, the overlapped portion of the two signed routes is either omitted completely from SR 38 or an exception linear LRM object is created to tell the user that all the roadway data is tied to SR 10 for the distance of the overlap.

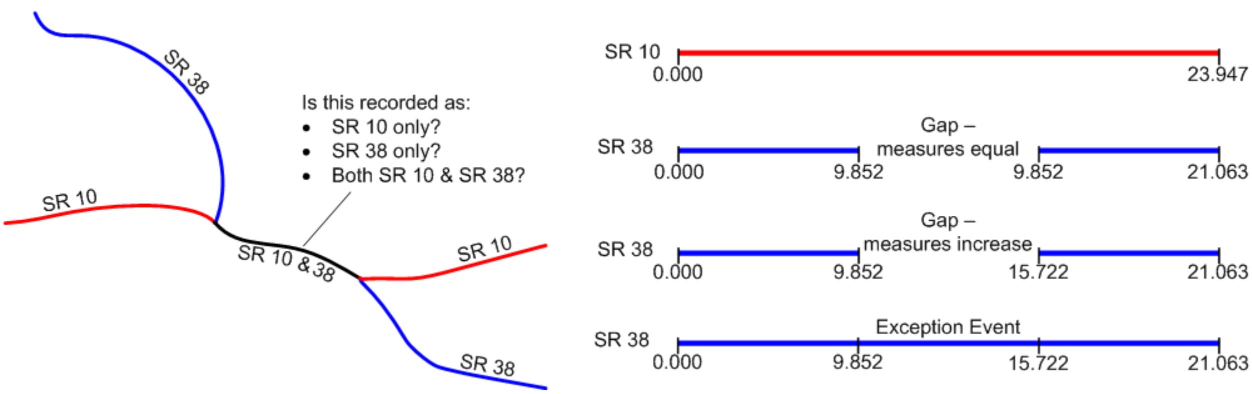


Figure 45. Various State approaches to signed route overlaps.

Figure 46 shows how the NRBM addresses the issue of overlapping signed routes. The figure shows the same roadways as in Figure 45. Route Segment 1640 is the one with the overlapping signed route numbers (10 and 38). It is part of an inventory route (blue line) that includes Route Segments 746 and 5392, both of which are only signed for SR 10. Route Segments 3287 and 17522 are signed only for SR 38. Since a single route segment cannot be part of two inventory routes in order to eliminate double counting, the two segments that are signed only for SR 38 are part of different inventory routes. The inventory route containing Route Segment 1640 will have two LRM Object records storing SR Name values of '10' and '38'. To find the path of a signed route, the user selects all the linear LRM objects that contain that route number. Signed routes will transition from one inventory route to another only at intersections.

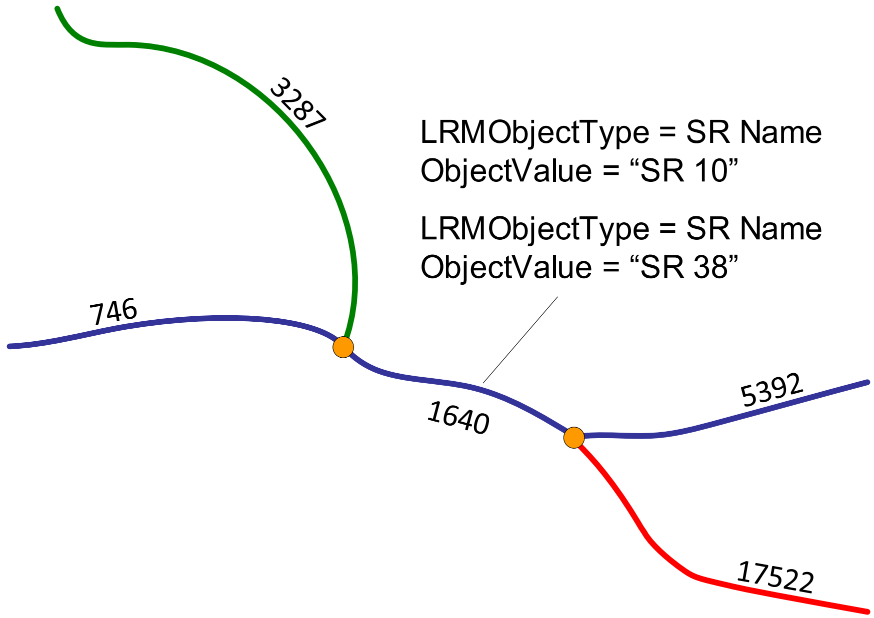


Figure 46. How the NRBM treats signed route overlaps.

A signed route is a traversal, which is an ordered sequence or inventory route sections that create a path through the roadway system. There are persistent traversals, like signed routes, and transitory traversals, such as the path one might ask a Web mapping service to provide to get from the office to a new restaurant for lunch. If an agency’s users expect to see data delivered using signed routes, it can transform the inventory route structure to produce the signed route structure following the traversal defined by each signed route number. It will also be able to transform the measure values to match what they should be for the signed route traversal. Measure values can be provided for the entire length of every signed route or in whatever manner is expected by users.

The next section shows how to link the LRM database to the rich resources offered by external business data suppliers by introducing the concept of LRM objects that are physical elements of a roadway, such as guardrail, noise walls, traffic signals, and bridges.

Elements

A roadway element is any physical part of a roadway stored in the NRBM as an inventory route. As Figure 47 shows, the typical roadway environment has a large variety of potential elements for inclusion in the roadway inventory database. Examples in this image include pavement markings, signs, bridges, guardrail, right-of-way parcels, drainage ditches, retention ponds, toll plaza, traffic signals, and fencing. Each element type probably has a State DOT workgroup devoted to its design, installation, and maintenance. There is also likely to be a specialized database with information about the elements they work with. Many of these workgroups will need geographic location data in addition to business data in order to put them in the proper 2D or 3D location.

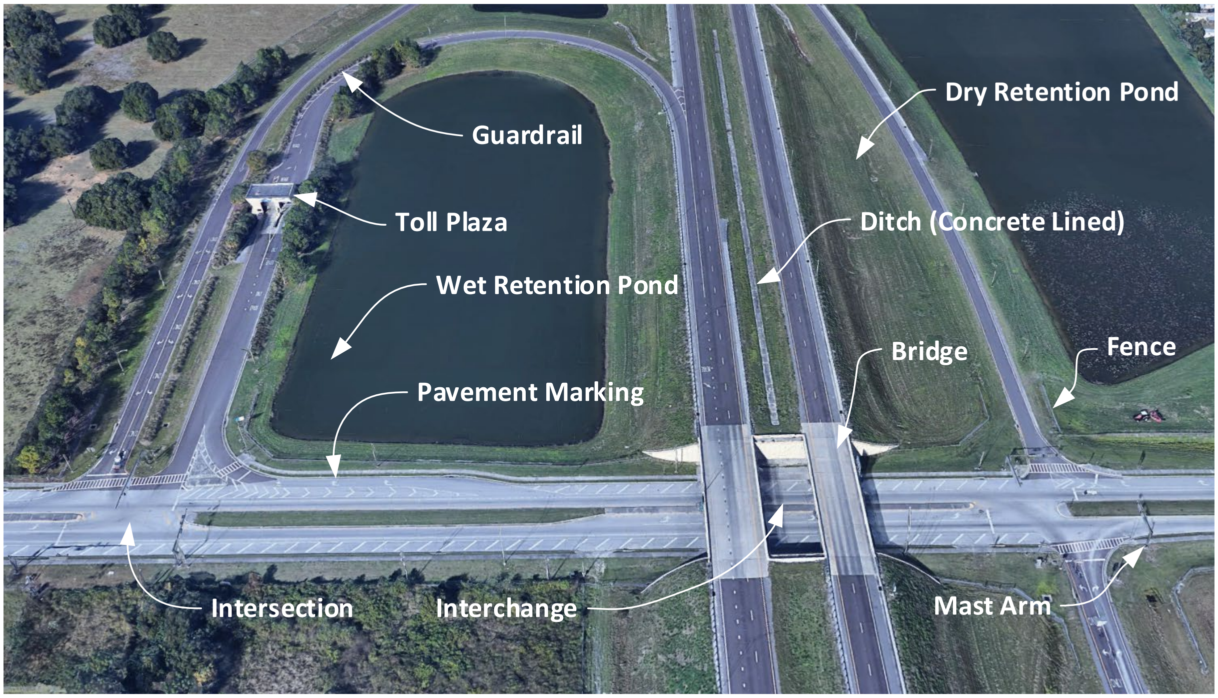


Figure 47. Example roadway elements.
Map data ©2019 Google® Earth™

An LRM position will say where they are located along the roadway (1D location). An LRM object will connect the LRM position to the element record by storing the element identifier; e.g., **`ObjectType`** = 'BRIDGE' and **`ObjectValue`** = [Bridge_ID]. If there were multiple representations or levels of abstraction stored for bridges (point, line, and polygon), then there may be a need to create multiple LRM objects, such as 'BRIDGEPT' (point LRM object) and 'BRIDGELINE' (linear LRM object).

Tying all these elements to inventory routes through LRM objects allows a user to virtually travel down the inventory route to "discover" elements of interest in the order of increasing measure values. Relational joins can be done to create a table or feature class that combines the LRM data with the business data stored in a separate, element-specific database using the stored element identifier as a foreign key for the search.

An inventory route creates the linear datum for LRM positions to be established and applied to LRM objects, which can have an abstract representation as a point (one 'At' measure) or a line (two measures ordered as 'From' and 'To'). The default assumption is that LRM objects are cartographically expressed through dynseg as points along a centerline or as a clipped portion of the centerline. These relatively simplistic abstractions work well at typical map scales. Some State DOTs have enhanced this simplistic approach by providing for lateral offsets from the centerline, either through capabilities provided by GIS software or by explicitly storing offset information in *LRM Object*. The NRBM specification supports these additional position attributes. Large-scale mapping may require something more.

To provide that "more," the data model supports independent elements that exist within the transportation system. An element is a discrete physical component, such as a bridge, guardrail, or sign, which needs to have multiple attributes. *LRM Object* acts like an attributed relationship to say that this element is at this location along an inventory route. It also provides the LRM gateway to the *Element* record that provides additional business data. That business data can contain one or more alternative map geometries, such as providing a polygon to show the physical extent of a bridge or a line to represent the path of a noise wall.

Some elements may need to describe multiple components. For example, an asset inventory that includes traffic signals will need to separately track the controller, conflict monitor, mast arms, span wire poles, vehicle detectors, traffic signal heads, and all the other pieces of a traffic signal. These components do not require a pictorial aspect. All components have the same LRM position as the element they comprise, so no additional map geometry beyond that of the element is supported by the specification.

Figure 48 shows the inventory route data model extended to include elements and their components. One element may have zero, one, or more components. Both elements and components may have map geometry independent of the inventory route centerline. The data model also includes support for a static image that shows the object or component from a different perspective. For example, a sign assembly may need to track each of the signs it supports. A sign face image may be used to clarify each sign's legend, or to show the arrangement of signs within the assembly. This static image could also be a photograph. However, Illustration can also be a link to much more complex means of describing an element or component. It can also be a link to a CAD file, a BIM application, or a videolog segment.

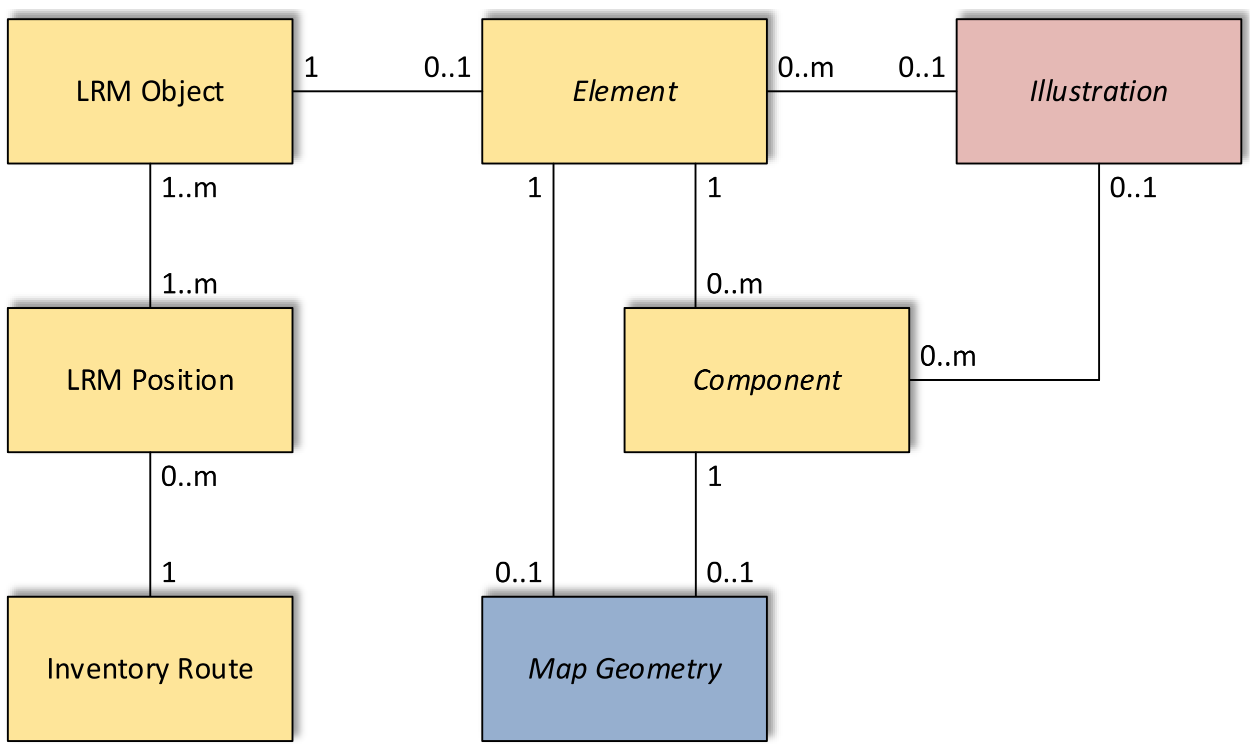


Figure 48. Supporting elements, which have connections to outside databases.

Element, *Component*, *Illustration*, and *Map Geometry* are abstract classes. These are all placeholders for the real business tables maintained by the various workgroups. For example, *Element* could become the Bridge table maintained by the Structures Office. Element and Component work together to provide a more complete view of complex objects. *Element* could become Sign Assembly (with a photograph) and *Component* could become Sign (with an image from the Manual on Uniform Traffic Control Devices) and Mounting. Noise walls and guardrails would likely need their own geometric representations for mapping since the road centerline is a bad substitute in most cases. Regardless of what the Element extension to the model is instantiated to support, such as to include a number of element-specific components, the intent is to allow users to discover the existence of the element and its component objects at the proper location along the inventory route.

There are dozens of potential elements that could be included, ranging from sign assemblies to culverts. The NRBM specification mandates inclusion of only the following roadway elements, each of which has a list of essential and optional attributes:

- Intersections and their component approaches.
- Restrictions.
- Interchanges, with their component ramp route segments, ramp extensions, and intersections (on limited-access highways only).

Intersections and interchanges are treated as independent elements in the LRM-based roadway inventory. Mapping criteria for specific intersection configurations are not established by NRBM's map scale specifications; they are set by MIRE 2.0 and other operational and safety applications. MIRE 2.0 includes 18 attributes for at-grade intersections and 38 attributes for intersection approaches. Although these attributes are substantially scale-independent, they imply a much larger working scale for intersections in order to provide an abstract space to place the required elements. A point feature will not get the job done. The scale to be effectively applied at intersections is much larger than what applies to the segments between intersections, something on the order of 1:1,200 (1″ = 100′) or more. But, again, this is not the real world; it is an abstract world to be created for attaching and displaying business data. Even at the largest NRBM-supported scale of 1:5,000, this abstract world is effectively invisible. LRM-based data do not properly locate all the elements that comprise an intersection at an operational scale. To the roadway inventory using LRM position descriptions, an intersection is a point. MIRE 2.0 requires that this point become a polygon.

Figure 49 shows the two intersection treatments provided by the NRBM. In the upper left is the LRM-based version, where intersections are a single point. To the right is an enlarged view of the middle portion of the LRM map. This scale allows the internal components of an intersection to begin to appear. This is not just an enlarged view of the LRM map; it is an entirely different cartographic product. A still larger view of the left intersection appears on the far left of the figure.

The NRBM specification provides the element database connection in the LRM-based roadway inventory but it does not mandate the inclusion of intersection diagrams and other detailed pictorial data. The construction of intersection approaches is left to the user. The function of intersections, within the context of this Guidebook, is to anchor route segments and to support various MIRE 2.0 intersection attributes, some of which apply to intersection approaches. Intersection geometry must also support such operational needs as crash diagramming and turning movement analyses. This level of detail is highly localized and is supported, but not mandated, by the NRBM specification.

An intersection is expected to be a point feature and to be stored as a point LRM object. A few States have included the component objects in their LRM-based roadway inventory. To accommodate them, the NRBM specification allows an intersection to be stored as a linear LRM object. The length of the intersection would extend from the first navigation point to the last that is encountered by that inventory route. These States will still need to supply the single intersection point and terminate all route segments at that point, but there are GIS processes that can be developed to automate that task.

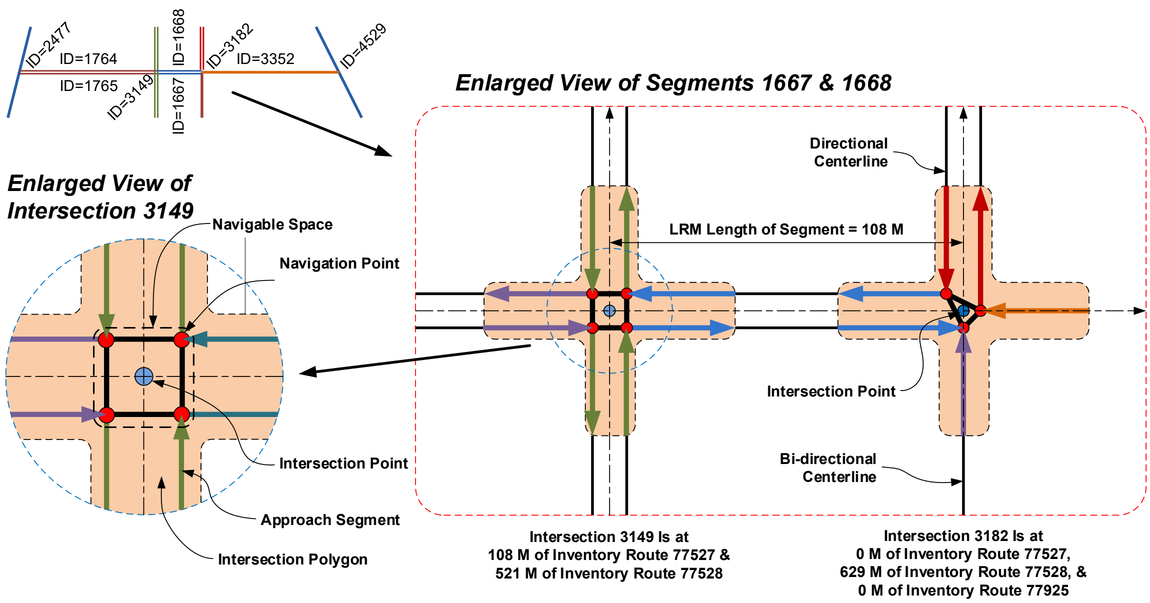


Figure 49. Peering into the basic components of intersections.

This is also how roundabouts should be treated. As shown in Figure 50, some States have constructed roundabouts as a collection on intersections where every route enters the facility, sometimes also creating a separate inventory route for the internal circulator component and bifurcating routes at the entrance islands. Because of scale and abstraction considerations, the NRBM specification treats these all as internal components of the roundabout. The roundabout itself is represented as a single intersection point and a linear LRM object on each involved inventory route. All route segments will terminate at the intersection point.

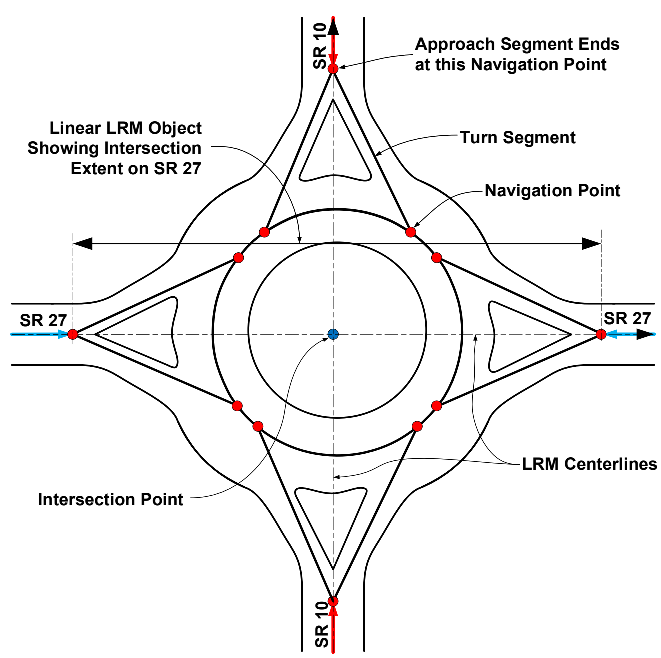


Figure 50. Convert a roundabout into a simple intersection located by linear LRM objects.

Figure 51 shows an intersection where its representation by a linear LRM object will not be the same for both directional inventory routes. In one direction, there is a long, channelized turn lane that creates a turn segment and navigation point. There is no such turn lane in the opposite direction for this Tee intersection. If a State is treating intersections as linear objects because they have included turn segments and navigation points in the basic inventory route structure, then intersection lengths will need to reflect the actual geometry of the facility.

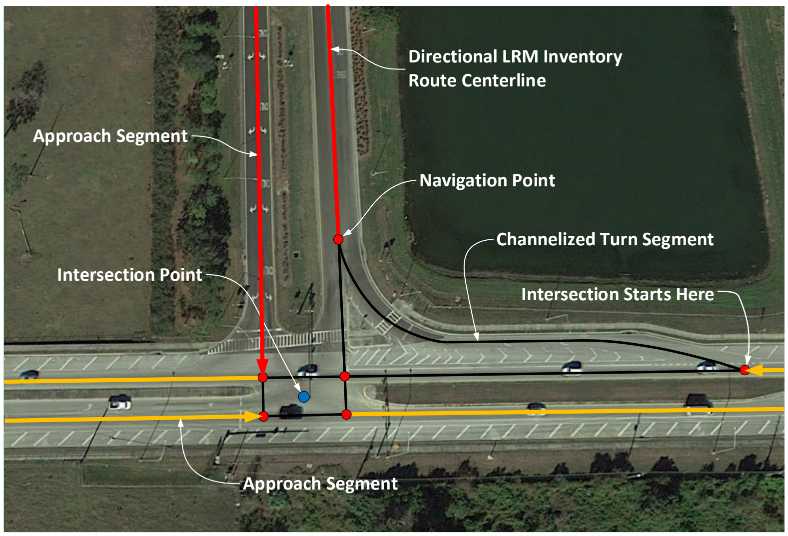


Figure 51. Example intersection with different lengths on the paired directional inventory routes.
Map data ©2019 Google® Earth™

Figure 52 offers another example of a Tee intersection that presents a different configuration problem. The Tee intersection shown involves only one side of the divided highway. The LRM object representing the location of the intersection will only be on this inventory route. The directional median turn lane imposes a navigation point on an inventory route without actually breaking the underlying route segments because the intersection is on the other direction of travel. This intersection is a point location on the bottom inventory route and would not appear on the other directional inventory route.

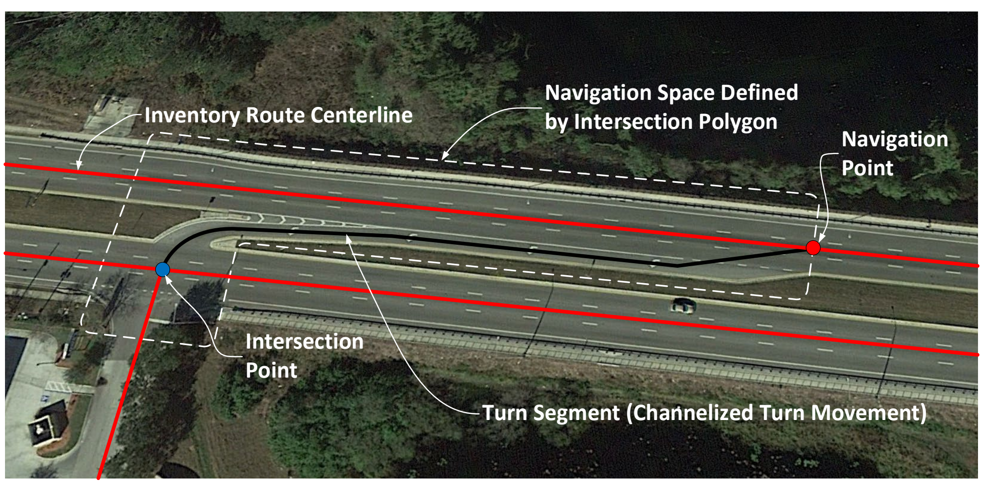


Figure 52. A navigation point is on an inventory route that is not involved in the LRM intersection.
Map data ©2019 Google® Earth™

While on the topic of mapping difficult intersections, there are a number of problematic configurations. Many of these deal with intersections where there is a transition from divided to undivided roadway geometry. Figure 53 shows a few examples and how the fundamental structures of navigation points and turn lanes differ from the way they should be represented by inventory route centerlines. Essentially, these examples attempt to use the intersection point to serve double duty. It is both an intersection on two or more inventory routes and a point of directional centerline transition for divided roadways.

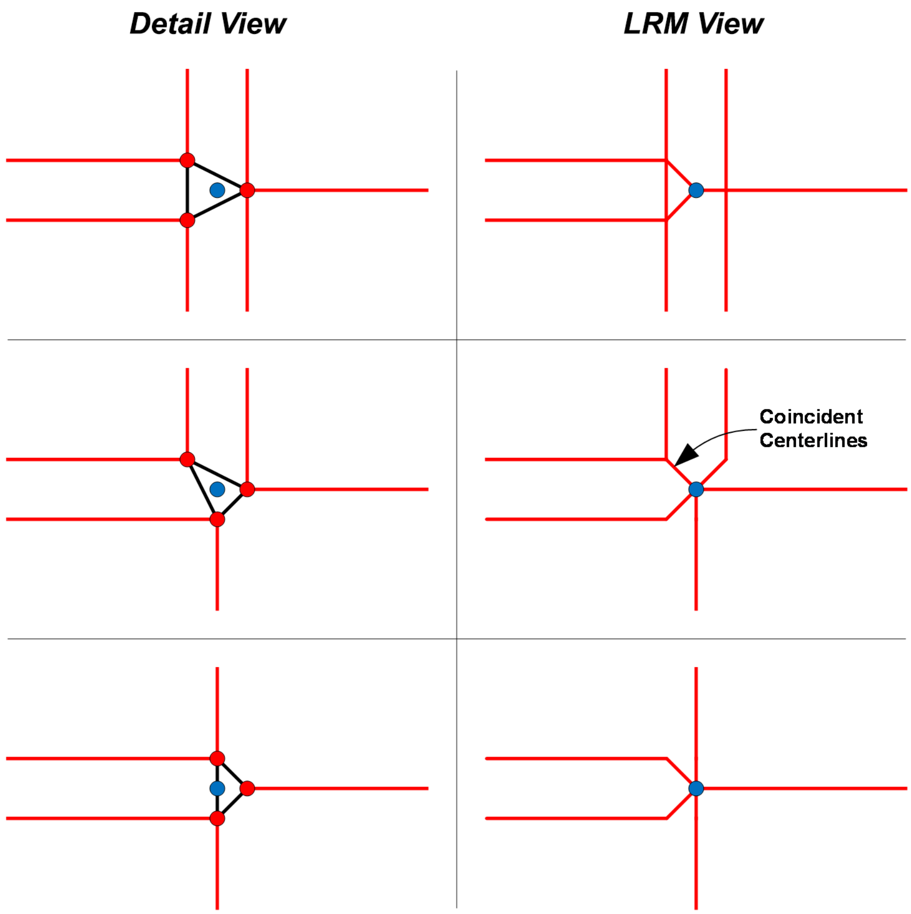


Figure 53. Dealing with problematic intersections where medians begin and end.

The detail view on the left side of Figure 53 shows how navigation points (red dots) and turn segments (black lines) would be constructed for a large-scale view of the internal components of the intersection. The right side shows possible inventory route centerline configurations that connect at the intersection point (blue dot). Intersections of divided roadways will rarely have a turn segment that goes through the intersection point. Only the intersection at the bottom of the figure has a turn segment that goes through the intersection point. The differences between the detail and LRM views of the intersection are not significant at the base scale of 1:5,000 and would be invisible at most mapping scales. The impact on dynseg operations for all but the shortest route segments would be within the stated accuracy specification.

As noted above, intersections and interchanges are roadway elements. Figure 54 presents the element extension model for intersections and interchanges. The Interchange and Intersection classes are instantiations of the abstract *Element* class shown earlier. Ramp, Approach, Turn Segment, and

Navigation Point are all instantiations of the *Component* abstract class. Interchange Point and Intersection Point are instantiations of *Map Geometry*, although Intersection Point could be created through dynseg from the intersection LRM object. (Interchange points do not necessarily fall on a roadway centerline, so they really are independent geometric objects.) Intersection Diagram is a version of *Illustration*. The classes in the bottom part of the figure are not mandatory classes in the basic NRBM dataset. They are primarily expected to exist for only part of the roadway system in a State. Turn Segment, Navigation Point, and Approach are all covered in the Definitions section of Chapter 3.

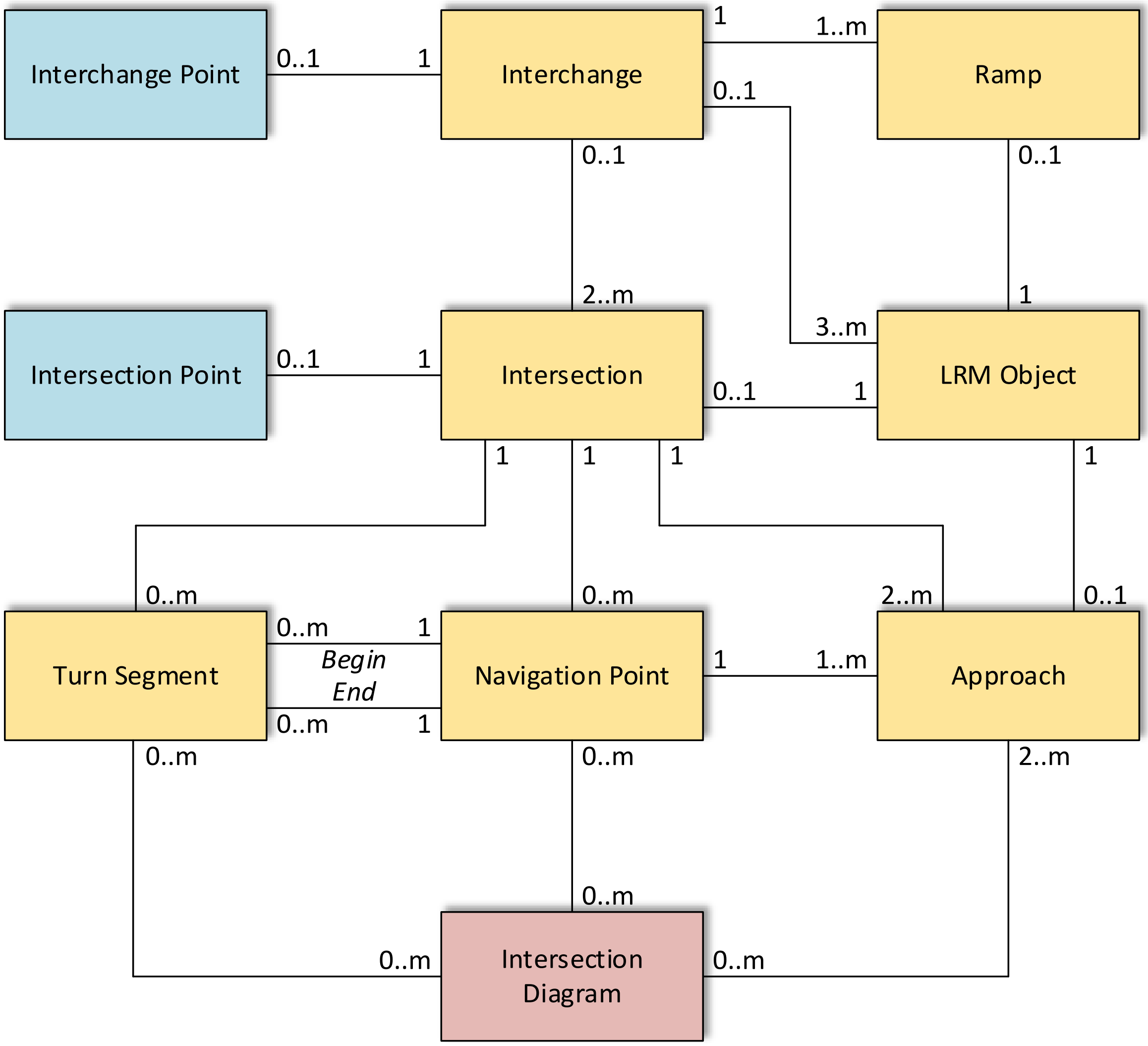


Figure 54. Instantiating the data model's Element Extension for interchanges and intersections.

Approaches may be LRM objects since they are formed from a portion of an inventory route,[89] but turn segments and navigation points are not visible in the LRM world.

[89] At the basic 1:5,000 scale, approaches are LRM objects that terminate at the intersection point, as recorded by an LRM position. At larger scales, approaches are intersection components that terminate at the first navigation point, if there is more than one.

The NRBM specification mandates two other element types: medians and restrictions. A median record is required whenever a roadway is represented by directional centerlines. The median is located on both directional centerlines by using an LRM object with an assigned LRM position. Figure 55 shows a section of a divided roadway where the median type changes from 'Curbed' to '2-way Left-turn Lane'. The blue line in the figure is used merely to indicate where the median is located. A separate median geometry is not required by the NRBM specification; however, it may prove to be useful for safety analyses and other applications.

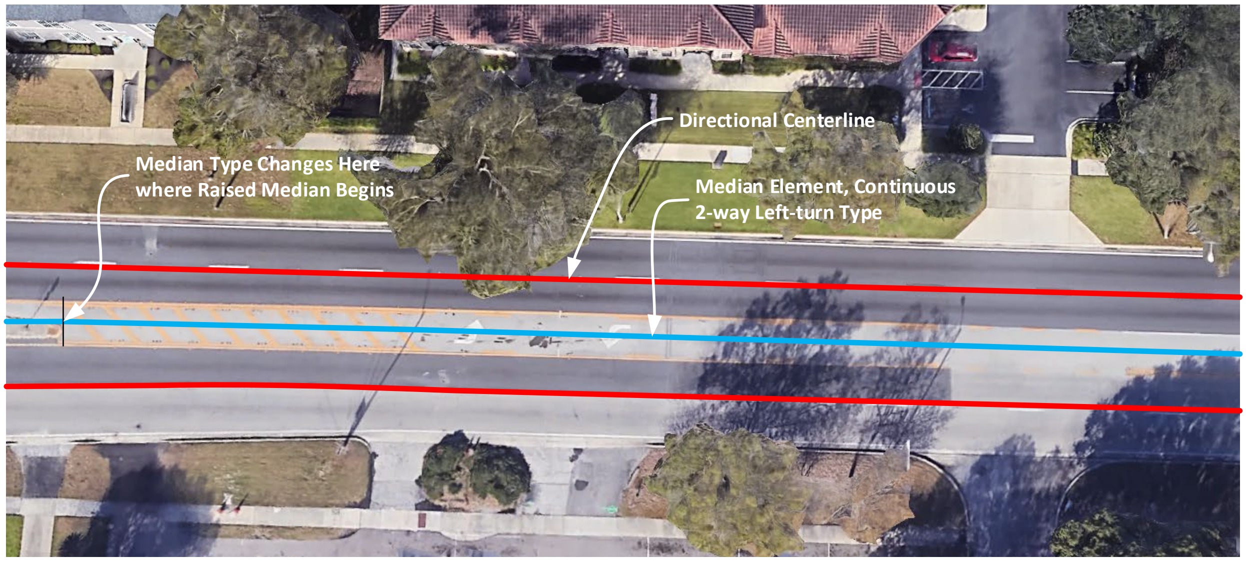


Figure 55. Median elements. The illustration geometry is not required for a median element.
Map data ©2019 Google® Earth™

A median element can actually be a fairly complex object and include many components, like concrete barriers and inside shoulders, which can have width, type, and other attributes. A median element may also be part of a cross-section element that provides a complete lateral view of the roadway. At least one GIS software vendor has demonstrated the ability to pull information from multiple data sources to construct a cross-section element.

A restriction element is anything that may impede travel along a route segment. Restriction elements are used by routing applications to determine whether a given route segment is available for the vehicle of interest to traverse. Bridges, tunnels, toll plazas, gates, and time-of-day regulations are all possible restriction types depending on what type of vehicle is being routed and the time it seeks to travel. In its initial deployment, the NRBM includes only physical impediments to travel posed by structures that may have weight, height, and width limitations. The expected use of the restriction elements is to remove affected route segments from the network prior to running the routing algorithm. For example, if the application is to move a 100,000-pound load that is 16 feet wide, the first step would be to remove all route segments (edges) that contain a bridge that cannot support the weight and any structures or other travel impediments that cannot provide the horizontal clearance required.

Connecting the States

The NRBM specification describes a national database being produced by the States and other entities. This process requires the specification to address the need to connect adjacent datasets at State boundaries by using match points that coincide with those of adjacent jurisdictions.[90] The ARNOLD Reference Manual notes that establishing match points at State boundaries "will be critical in facilitating the edge-matching of data between neighboring States, and ultimately stitching together a nationwide roadway dataset."[91] Although offering no specific solution for how such match points would be established, the ARNOLD Reference Manual states the core requirement for match points is that they be placed "at unambiguous, verifiable locations that bordering jurisdictions can agree upon," perhaps through agreement "between GIS managers for the adjacent jurisdictions, and ideally would be endorsed by their respective senior executives."[92]

Match points must also be established within each State in order to combine data from sub-State sources, such as local governments. In the absence of such a process of voluntary agreement, the ARNOLD Reference Manual suggests that the States "create these points on evident criteria, and provide them to Local Government Agencies (LGAs) for review."[93] Noting that most State DOTs do not have the resources to mediate disputes between local governments, the Manual states "the local jurisdictions will need to work this out between themselves." Match points must be established whenever multiple datasets are combined in a single geographic space.

The NRBM proposes a technical solution rather than an organizational one. The proposed NRBM method for establishing match points is for the State DOT to declare match points within its territory and along the outer edges of its NRBM dataset coverage. The FHWA would then either declare boundary match points for the States to use or construct zero-length topological connectors to connect adjacent States to each other. (In this context, "zero-length" means the LRM length of the route segment is zero.) The States may use a similar process for resolving disconnects within a dataset caused by differences between the centerline termini provided by their local, regional, Tribal, and other data supply partners.

The NRBM specification is fairly straightforward and involves only a few data entities, but the logic behind the design is often complex and subtle. Chapters 1–4 have laid the foundation for the specification. It is now time to present it in Chapter 5.

---

[90] See, for example, *Op. cit*, pp. 86-87.
[91] Ibid.
[92] Ibid.
[93] Ibid.

Chapter 5. The NRBM Specification

The National Roadway Base Map (NRBM) is a national data specification that facilitates exchange of information on roadways throughout the United States in a uniform manner. It provides information in two data structures, one using discrete route segments with a few key attributes and another with extensive data tied to linear referencing. Such information includes roadway characteristics, like the number of lanes or functional classification, roadway elements, like bridges and intersections, and events that relate to roadways, like construction projects and crashes.

The NRBM dataset also provides a basic topological network for routing applications. The data included in the NRBM dataset supports mapping roadway conditions; conducting safety analyses, including use of Model Inventory of Roadway Elements (MIRE) data; exploring the data contained in the Highway Performance Monitoring System (HPMS), National Bridge Inventory (NBI), and Statewide Transportation Improvement Program (STIP); and many other functions. The data stored in an NRBM dataset is temporal and can provide a past or future view of the roadway system.

The NRBM, though, is more than a data specification. It is also a guide for describing best practices in data management and database design for internal use by the States as they migrate to more advanced information management infrastructures. This part of the NRBM is founded on the principles of civil infrastructure management (CIM), which provides a framework to organize the complete life cycle of an asset and to support effective enterprise data management. For CIM to be effective, however, the State must engage its internal and external stakeholders in enterprise data governance. Data management provides the implementation of policies and procedures adopted by the data governance system.

CIM looks at an asset as moving along a timeline, with each station along that timeline belonging to a specified workgroup. As the asset moves through its life cycle, it also moves from one workgroup to the next. For each workgroup to do its job well and efficiently, data about what the previous workgroup did and the state of the asset at the time of hand-off needs to be conveyed. In this way, CIM data management is about information access. Task 1 is to make sure the data needed by the following workgroup is defined and understood by both workgroups. Task 2 is to actually compile the information in the required format and deliver it on time. This is where data management comes in. All of these asset management workgroups should be part of the enterprise governance system, where major policy decisions are made. The governance system says who will do what and when. The data management system takes this direction and determines how it can best be done.

As presently envisioned, the NRBM would be constructed by the States in cooperation with Federal, local, regional, and Tribal partners in accordance with the proposed specification. The NRBM's foundation was developed by the Applications of Enterprise GIS for Transportation (AEGIST) project, which involved extensive interaction with practitioners throughout the country and at all levels of government. This group of practitioners set three basic requirements for the NRBM:

1. Any data to be mapped at a national extent must have a uniform definition, collection method, maintenance process, and delivery mechanism.
2. The data source must be authoritative. The suppliers of NRBM data must be the original sources of the information.
3. All users should get the same answer when asking the same question using data from the original source and the NRBM.

The practitioner group also determined there are two groups of users to be addressed. First, the editors who compile, maintain, and publish the data. This group has a transactional focus. Second, the analytical users of the data who utilize the information in various business units. This group has a cross-cutting need to combine data from multiple sources, often to answer complex questions. The operational guidelines provided in the AEGIST Guidebook serve the editors. This target audience is provided with database designs and operational information that will simplify the editing process and enhance the information being provided to users.

The database design is transactional in nature. It offers the simple route segments needed for basic mapping of the roadway system, along with a topological network that supports rudimentary routing applications, such as discovering whether it is possible for an over-dimension vehicle to get from Point A to Point B. These are derived from a transactional database centered on linear referencing and its ability to provide data discovery and access to a rich trove of information maintained by all the workgroups. It meets the three core requirements for the NRBM by supplying clear and complete definitions, business rules, and authoritative sources of information. It also serves as a set of functional requirements for software vendors, whose efforts to develop a complete GIS-based roadway inventory management solution have been hampered by the disparate methods and rules employed by the States today.

By far, the biggest task toward actually producing a national dataset in accordance with the NRBM specification is the need to first establish enterprise data governance, and then to reform practices and modify data structures in the States. Because of this necessary first task, the development of this Guidebook is being followed by a pooled-fund study that will financially support multiple States as they attempt to adopt and deploy the recommendations contained herein.[94]

During this follow-on phase of the AEGIST project, the specification contained in this chapter and the business rules and definitions on which it is based may be modified, deleted, or enhanced by additions. While multiple peer exchanges were held and a Technical Advisory Group participated in and reviewed the contents of this Guidebook, there is no way to know for certain whether the project team got everything right until it is put to the test.

During the TPF is also when Chapter 6 of the AEGIST Guidebook will be written to provide information to the States on how to implement the specification. This chapter will offer step-by-step direction to the States on how to adopt the NRBM practices and produce the data as a reliable product, based on the experiences of the pooled-fund study. Information of this type that is available now is contained in Chapter 2 (governance) and Chapter 4 (NRBM fundamentals).

[94] This phase of the AEGIST project is Transportation Pooled Fund Program number TPF-5(431). For more information, go to https://www.pooledfund.org/Details/Study/654. At the time of this writing, five States had committed to participating in the program.

The Basic Requirements

This section contains the fundamental requirements set by the NRBM specification with regard to accuracy, precision, resolution, and extent.

Geographic Extent: The 50 States, the District of Columbia, and the Commonwealth of Puerto Rico.

Topical Extent: All public and private roads located within the geographic extent.

Cartographic Scale:

a. 1:5,000 for the national dataset. This is the scale of abstraction for the entire dataset.

b. 1:1,200 for intersection diagrams. This is the scale of abstraction for intersections relative to the production of approach segments, navigation points, and turn segments; i.e., the internal basic components of an intersection that are not included in the linear referencing system.[95] There is no specific requirement for intersection diagrams used for crash analyses or similar applications to be drawn to any scale.

Linear Referencing Method (LRM) Position (Measure) Resolution: 0.001 mile, which is a unit of measure symbolized by the Greek capital letter *mu* (M). Measure values are integer values; no decimal places.

Accuracy:

a. The total length of an inventory route, as stated in measure units, must be within 5% of its true value.
b. The total length of a route segment must be within 5% of its true value.
c. LRM objects must be located along the defining inventory route within 158 feet (30 M) of their true position in urban areas and within 264 feet (50 M) of their true position in rural areas.

Precision: At least 95% of the LRM objects must meet the accuracy specification.

The NRBM database design supports the following data operations:

Spatial Changes –

a. Modification of a geographic object (feature).
b. Modification of a position description (LRM Position).
c. Additions; new objects are created and are assigned a starting date of validity.
d. Retirements; previous data is retained for historical views of the system.
e. Replacements; previous data is removed from access but is retained to preserve the historic record of database edits.

Attribute Changes –

a. Changes in a stored value.
b. Changes in object type within a class.

[95] Intersection approaches (a.k.a. approach segments) may be defined in the LRS with termination at the single intersection point. These approaches will change in dimension and termination when the display scale is enlarged to reveal the intersection's internal components. At these larger scales, approaches will terminate at the outermost navigation point.

c. Additions; new objects are created and are assigned a starting date of validity.
d. Retirements; previous data is retained for historical views of the system.
e. Replacements; previous data is removed from access but is retained to preserve the historic record of database edits.

All but one object class contained in the specification data model are based on an abstract *Object* class stereotype described in Chapter 4. This template provides the standard attributes to support temporal aspects of the database design. This stereotype is included in the data dictionary, which is the subject of the next section. Its functional requirements are as follows:

- Provide a nationally unique identifier for each object and (optionally) a public key that may be more commonly used to identify the object.
- Indicate the status of the record and the real-world object it describes.
- Provide two temporal vectors for the data, one tracking changes to the database and another stating the period of validity for any object in the database; i.e., when the information in the record accurately reflects the real-world condition.
- Identify the supplier of the data.
- Classify changes to the database in terms of the reason for making the change and who did it.

The fields included in the abstract *Object* class and inherited by almost every other class[96] are:

`ObjectID` – The stereotype name of the object identifier as a 32-character globally unique identifier (GUID). Although the GUID is a number, it is expressed in Base-16 and must be of the string data type in order to include the non-decimal numerals of A through E used in hexadecimal notation. The 'Object' portion of the attribute name will usually be the name of the class.

`RecordDate` – The date the record was created; expressed in the form of MM/DD/YYYY.

`AffectedRecordDate` – This attribute stores the previous RecordDate value for the edited record.

`RecordStatus` – The status of the record; the domain is '-1' for Replaced, '0' for Retired, and '+1' for Active.

`RecStatusBeginDate` – The date the record status is initially true; expressed in the form of MM/DD/YYYY.

`RecStatusEndDate` – The date when the record status ceased to be true; expressed in the form of MM/DD/YYYY.

`ObjectStatus` – The status of the object in the real world; the domain of possible values is specific to each object class. The 'Object' portion of the attribute name is replaced with the name of the object class.

`PublicKey` – A candidate primary key that is more recognizable to the user than a GUID. In the case of an inventory route, for example, the public key could be a so-called "intelligent identifier," such as may be constructed by a sequence of component parts that can be derived from attributes of the object.

---

[96] The only initial mandatory class that does not inherit from the *Object* stereotype is Usage, which manages the presence of an attribute within a class. The Photo class also does not inherit from the *Object* stereotype because these records are found by using the file name of the image.

`SupplierID` – The identifier of the data supplier for this object. The value is assigned by the State DOT. It may be useful for specific workgroups within the State to receive their own supplier identifier.

`Editor` – The name or other identifier of the person who created the record. This could be a UserID value created within the State's computer network.

`EditReason` – The reason the record was created; the domain is '1' for New Object, '2' for Object Status Changed, '3' for Object Field Value Changed, '4' for Error Detected in Record', '5' for Roadway Realigned, and '6' for Roadway Eliminated. This domain is subject to change during the pooled-fund program. The domain provided for this attribute is intended to permit general statistics to be compiled for the nature and number of edits conducted during a given time period.

`EditComment` – A large text field to provide more information about the reason for the record's creation.

Data Dictionary

A major difference between the NRBM and other transportation data specification proposals is its self-describing characteristic. The NRBM specification requires a complete data dictionary to be available for all users as part of the dataset. A data dictionary describes all the dataset components. It lists each object and every field it contains. It also provides the information needed to edit and publish the data. The intent is both to tell users what each piece of data means and where it came from and to provide information that could be used by software vendors to automate the data entry process.

Figure 56 shows the logical data model for the NRBM data dictionary. It consists of four tables: *Object*, Usage, Field,[97] and CodedValue. *Object* is the same class stereotype described in Chapter 4. It serves as a placeholder for all the tables and feature classes included in the NRBM dataset. The Field class stores information about all the attributes (columns) contained in the included classes. Usage is actually an attributed relationship. It resolves the many-to-many relationship between Field and all the classes based on the *Object* class template. (An attribute can be in many objects, and each object can contain many attributes.) The NRBM data dictionary will be supplied by FHWA. Any agency adopting the NRBM for its internal use will need to expand the data dictionary to include additional tables/classes and attributes they include.

[97] In Chapter 3, 'Field' is defined as the intersection of a row and column; i.e., a single piece of data. In the data dictionary, what is placed in the Field class are actually attributes; i.e., columns in a table. 'Field' was used herein because 'Attribute' was thought to be too confusing a term given its broad, general meaning and usage in this Guidebook.

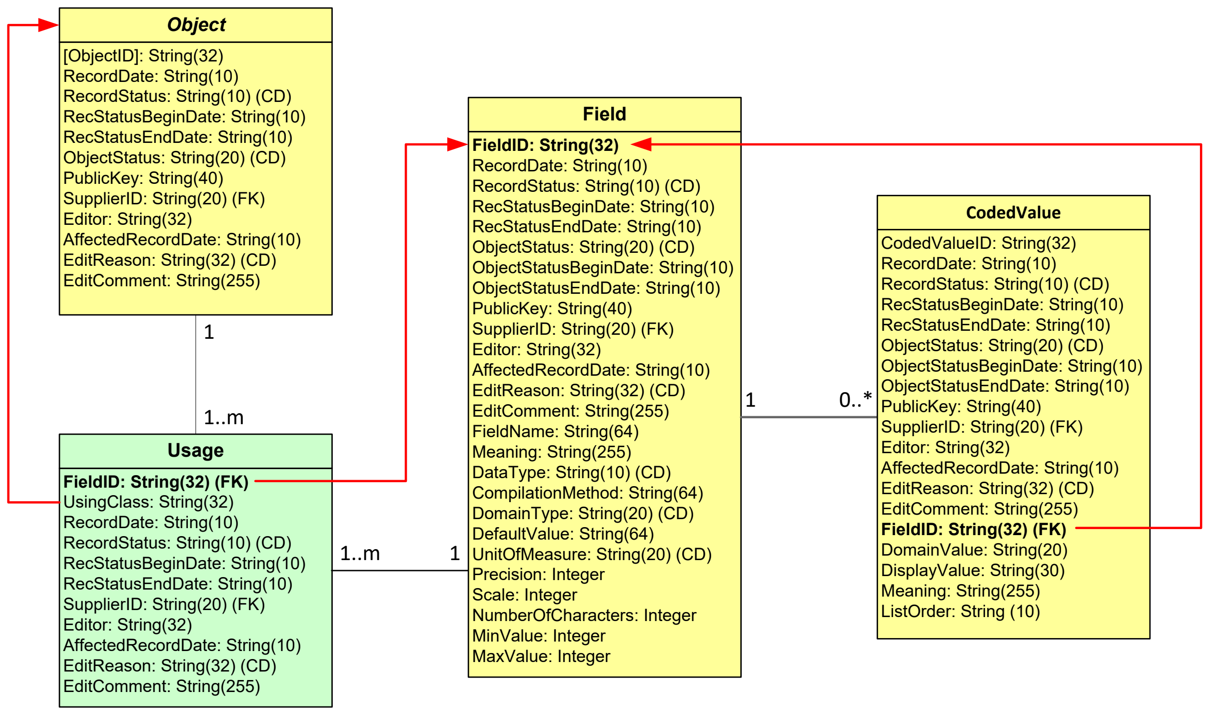


Figure 56. The NRBM data dictionary logical data model.

The Field class describes every attribute contained in all the NRBM dataset classes. An attribute may appear in multiple tables/classes, but it must mean exactly the same thing in every instance. Any needed change in meaning, format, or domain must result in a different attribute being created and defined in the data dictionary.

The information in the Field class can be used to format displayed data and regulate data editing and publication. Some fields contain a coded value domain; i.e., a set of defined choices. A good example is functional classification. All of the valid choices will be contained in the CodedValue table. It would be a fairly straightforward task for a software vendor to construct a single data entry routine that used the information in the Field class to automatically customize a data entry screen for an attribute, offer a pick list of valid choices using the information in the CodedValue class for that attribute, and provide quality control checks on the entered information.

All the classes shown are temporal, a point being emphasized here by including the fields inherited from the *Object* abstract class that is the parent of the Usage, Field, and CodedValue classes. Attribute descriptions can change over time, classes can be added to the dataset, attributes can be added or removed from object classes, and coded value domains may be modified. NRBM has to accommodate change in itself as much as in the transportation system.

Usage is an attributed relationship class. As such, it has a couple of differences from other classes. First, there is no **`UsageID`** field because the combination of **`FieldID`** and **`UsingClass`** is unique and, with **`RecordDate`**, forms the complex primary key for this class. Second, the relationship between the *Object* abstract class and Usage is not like the others shown in the data model. The cardinality is correctly shown; i.e., "Each *Object* class must be described by at least one Usage record" and "Each Usage record must describe the presence of an attribute in one *Object* class." However, these are not relationships implemented by using foreign keys because the relationship is expressed by storing the name of the using object class.

Usage actually has a one-to-many relationship with every other class in the database, because every other class inherits from the *Object* abstract class stereotype. The Usage class has many of the elements of the *Object* class, but it does not inherit from that class; it is a standalone class. Looking closely at the attributes in the Usage class, one can see that several of the *Object* class attributes are not there, such as **`ObjectID`**, **`ObjectStatus`**, and **`PublicKey`**. Strict adherence to the inheritance process would require all the *Object* class attributes to be included in the Usage class. Since the data dictionary includes a description of every NRBM attribute and the data contents of every object class, the attributes of Usage, Field, and CodedValue are also described in the data dictionary.

Figure 56 has the usual cardinality indicators and adds red lines with arrows showing the use of foreign keys to link the classes together. The cardinality indicates a one-to-many relationship between Field and both Usage and CodedValue. These two "child" classes include **`FieldID`** as a foreign key to link the applicable records to the correct Field class object. (The "many" side of the relationship always stores the foreign key attribute.) The relationships are read as, "Each Usage record must relate to one Field record," "Each Field record must relate to at least one Usage record,"[98] "Each CodedValue record must relate to one Field record," and "Each Field record may relate to one or more CodedValue records."

Route Segment, Intersection, and Terminus

The route segment form of the NRBM data publication specification includes three classes. RouteSegment provides the information about segments of the roadway system. Intersection provides information about places where route segments connect. Terminus has information about locations where route segments terminate without a connection to another route segment. The business rules for this deliverable dataset are:

1. A route segment is a discrete portion of a roadway.
2. A route segment must begin or end at an intersection.
3. A route segment may begin or end at a terminus.
4. Except for the Roadway Access and Median Cut intersection types, all intersections "break" a route segment into two separate segments.
5. An intersection, except for non-breaking types, must connect at least two route segments.
6. A terminus may begin or end only a single route segment.
7. Route segments that form a county/parish or State boundary should be subdivided at the boundary. It is optional to break a route segment at other boundaries, such as city limits.
8. A route segment that includes part of a boundary greater than 500 feet in length shall be subdivided so that a single route segment forms the boundary.
9. The minimum route segment length is 500 feet, except in dense urban areas, where the minimum route segment is the length of one block. However, where necessary to properly convey attribute values that differ significantly, a shorter route segment may be created.

---

[98] Technically speaking, the sequence of building the database actually requires that the multiplicity of Field to Usage to be 0..m instead of 1..m. This is because the Field record must exist before it can be linked to a class that includes the described attribute in the Usage class. Similarly, the class instantiating the *Object* abstract class must also exist prior to the creation of the related Usage class records.

10. Directional route segments shall be created for any divided roadway section longer than 400 feet. A divided roadway is established by the presence of a median, as that term is defined in Chapter 3.
11. Channelized turn lanes and other parts of the roadway system located within the navigation space of an intersection do not form route segments.
12. All route segments, except those representing ramp extensions and topological connections, should become part of an inventory route.

Figure 57 shows the basic data model for the route segment deliverable. The blue background color of the three class objects indicates they are feature classes. It includes the route segment class and the two classes that represent objects that can exist at the end of each route segment. This model adopts the convention of not including the inherited *Object* stereotype class attributes.

The model shown in Figure 57 is a logical data model. This means that it shows the various attributes and their specifications in a general form. The conceptual data models in Chapter 3 just showed the classes without attributes. The logical data model presents the internal tabular contents of the class in a general way. A logical data model is as far as the NRBM specification needs to go because this is all it cares about. Many attributes are specified as string data types as a sort of universal translator option. However, if the specification data structure is used as a native guide for how a State should structure its own internal databases, then it is necessary to go one step further to produce a physical data model. Such a data model takes into account the specific needs and capabilities of the State's relational or object-oriented database management system. For example, instead of treating date as a 10-character string, as the NRBM data exchange specification does, a physical data model would choose one of the many date/time data types provided by the State's database management system. The States that participate in the pooled-fund program will need to construct physical data models tailored to their specific database management environments.

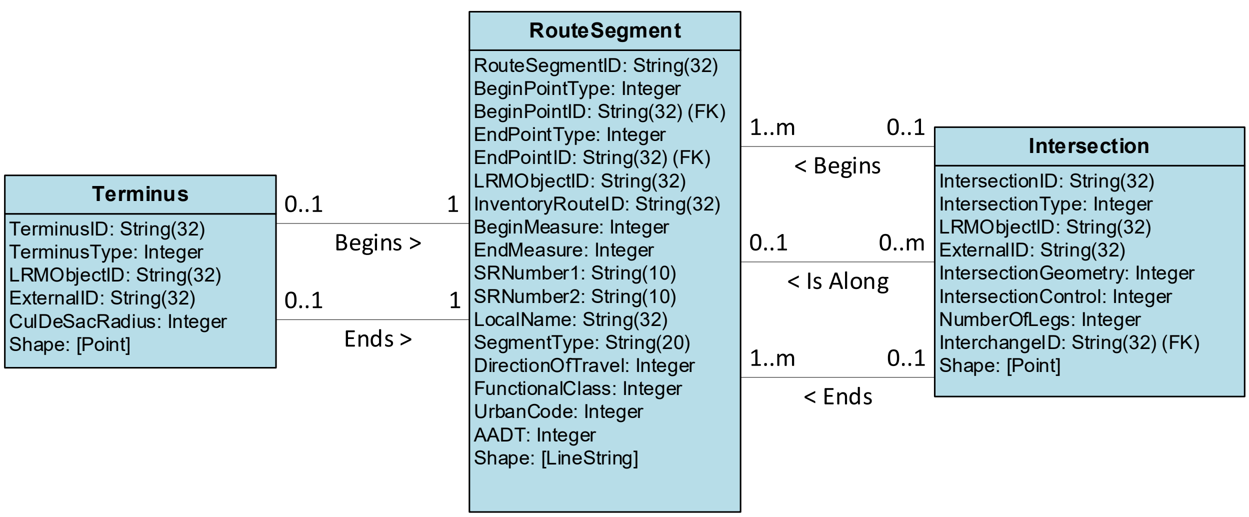


Figure 57. The Route Segment data model containing the three required feature classes.

The Route Segment data model was initially postulated in the Transportation for the Nation (TFTN) Strategic Plan and expanded on in the ARNOLD Reference Manual to include the Intersection feature class. The NRBM specification adds the Terminus class and expands the relationship between RouteSegment and Intersection to include an 'Along' option. This third relationship role allows a supplier to avoid breaking route segments at every intersection, as two types—Roadway Access and Median Cut—do not impose a break in route segments.

There may be instances when it is desirable to break a long route segment into more manageable pieces or there is a midblock change in one of the included attributes. In such instances, a continuity intersection can be added to allow route segments to be subdivided at useful locations.

In addition to the standard *Object* class template attributes, the RouteSegment class includes the attributes of the following:

`BeginPointType` – The type of object that begins the route segment; the domain is ‘0’ for Terminus and ‘1’ for Intersection.

`BeginPointID` – The identifier of the beginning object. This field is a foreign key containing either a **`TerminusID`** or an **`IntersectionID`** value, as determined by the value of **`BeginPointType`**.

`EndPointType` – The type of object that ends the route segment; the domain is ‘0’ for Terminus and ‘1’ for Intersection.

`EndPointID` – The identifier of the ending object. This field is a foreign key containing either a **`TerminusID`** or an **`IntersectionID`** value, as determined by the value of **`EndPointType`**.

`LRMObjectID` – The identifier of the LRMObject record that ties the route segment to an LRM position.

`InventoryRouteID` – The identifier of the inventory route of which this route segment is a part.

`BeginMeasure` – The LRM measure where the route segment begins.

`EndMeasure` – The LTM measure where the route segment ends.

`SRNumber1` – A place to record the primary signed route name. The primacy order is Interstate, US Route, State Highway, County Road, and City Street.

`SRNumber2` – A place to record a secondary signed route name. The primacy order is Interstate, US Route, State Highway, County Road, and City Street.

`LocalName` – The locally assigned name for the roadway of which this route segment is a part.

`SegmentType` – The type of route segment; the domain is ‘Roadway’, ‘Ramp Extension’, ‘Trail’, ‘Ferry Link’, and ‘Topological Connector’. A route segment of the Ramp Extension type will not become part of an inventory route, as it is provided merely to improve the mapped appearance of the area where the ramp joins or departs from the limited-access highway.

`DirectionOfTravel` – An indicator as to whether the direction of travel is the same as the direction of the route segment, as set by the beginning and ending objects; the domain is ‘1’ for Same as Direction of Route Segment, ‘0’ for Bi-directional Travel, and ‘-1’ for Opposite to the Direction of Route Segment. A value of ‘-1’ should be relatively rare.

`FunctionalClass` – The functional classification of the route segment; domain is defined by HPMS Functional System element: ‘1’ for Interstate, ‘2’ for Principal Arterial-Other Freeways and Expressways, ‘3’ for Principal Arterial-Other, ‘4’ for Minor Arterial, ‘5’ for Major Collector, ‘6’ for Minor Collector, and ‘7’ for Local.

UrbanCode – The Census urban area code; domain is the value assigned by the Census for the urban area (up to five digits, no leading zeros), '99998' for small urban area, and '99999' for rural area.

AADT – Annual average daily traffic, as described in the HPMS Field Manual.

Shape – The Open Geospatial Consortium (OGC) LineString geometric object representing the path of the route segment; this geometry should be provided in accordance with the HPMS Field Manual, except that it does not include measure values on the vertices and the scale of abstraction, accuracy, precision, and resolution.

The Intersection class contains the following attributes:

IntersectionType – The type of intersection depicted; the domain is 'At-grade', 'Continuity', 'Median Cut', 'Multimodal Interchange Point', 'Railroad Grade Crossing', 'Ramp Junction', 'Roadway Access', 'State Boundary Match Point', and 'Supplier Match Point'.

LRMObjectID – The identifier of the LRMObject record that ties the route segment to an LRM position.

ExternalID – The identifier used by a related external dataset.

IntersectionGeometry – The type of intersection configuration; from MIRE Element 116 Intersection/Junction Geometry, which offers 14 potential values. This field should be populated for all at-grade and median cut intersections and for applicable continuity and roadway access intersections.

IntersectionControl – The type of traffic control provided at the intersection; from MIRE Element 121 Intersection/Junction Traffic Control, which provides 12 possible choices.

NumberOfLegs – The number of intersection approaches; from MIRE Element 115 Intersection/Junction Number of Legs. A value between '2' and '8' is to be provided for all at-grade intersections.

InterchangeID – The identifier of the interchange of which the intersection is a part. This value will be absent for most intersections.

Shape – The OGC Point geometric object representing the location of the intersection.

The Terminus class includes the following attributes:

TerminusType – The type of terminus; the domain is 'Cul de Sac', 'Map Limit', and 'Ramp Merge/Diverge Point'.

LRMObjectID – The identifier of the LRMObject record that ties the route segment to an LRM position.

ExternalID – The identifier used by a related external dataset.

CulDeSacRadius – The effective radius of the cul de sac, in integer feet. This value will be absent for all terminus types except cul de sac. A dead end street terminus will be classified as a cul de sac with a radius of zero.

Shape – The OGC Point geometric object representing the location of the terminus.

Some consideration was given to including additional attributes for cul de sacs in the Terminus class, such as an indicator of whether there is a central island.

It is important to note that the multiplicity shown in the logical model is intended to express the relationship of active records. In practice, as retired and replaced records are created, there can be many potential records with the same identifier but different record dates; however, there should normally only be a single active record included in any singularity relationship.

Route segments are extremely limited with regard to the amount of information conveyed. A number of optional attributes have been proposed, but none are included in the NRBM specification at this time. One such example is street address range. These were omitted from the specification in recognition of the fact that the States are not the authoritative sources for address information. Local governments, and particularly 911 agencies, are the authoritative sources of address data. It is better to get such information from these sources than from a State-sourced dataset that could be significantly out of date. In addition, there is no common standard presentation of address data, nor has there been a decision as to whether address point or address range data is more appropriate.

The primary user group for the route segment part of the NRBM dataset is local government. It is also useful for groups that do not need a lot of roadway data, they just need a roadway layer for displaying other thematic information. However, by including the inventory route and LRM object identifiers in the route segment class, all users have the ability to drill down into the LRM-based data in the other part of the published NRBM dataset to extract any attribute of interest. Roadway elements with their own geometric representations can also be added to the route segment dataset.

The route segment class includes the beginning and ending measures assigned to the LRM object that it represents along the related inventory route. This information allows the route segment to become its own inventory route for the location of characteristics, elements, and events that exist along the segment. This extension of the route segment model could be a first step toward deploying full linear referencing system (LRS) data structures by a local government or regional agency. At the very least, it provides the LRM location information that can allow the route segment data user to explore available data for this route segment in the State's roadway inventory.

LRS Database Specification

The information for which the States are the authoritative source is roadway inventory data. These data are natively contained in a linear referencing database. Figure 58 shows the NRBM specification for a linear referencing dataset to be supplied by the States. The classes in this figure omit the template's standard fields in order to better emphasize the added fields and to fit the available space.

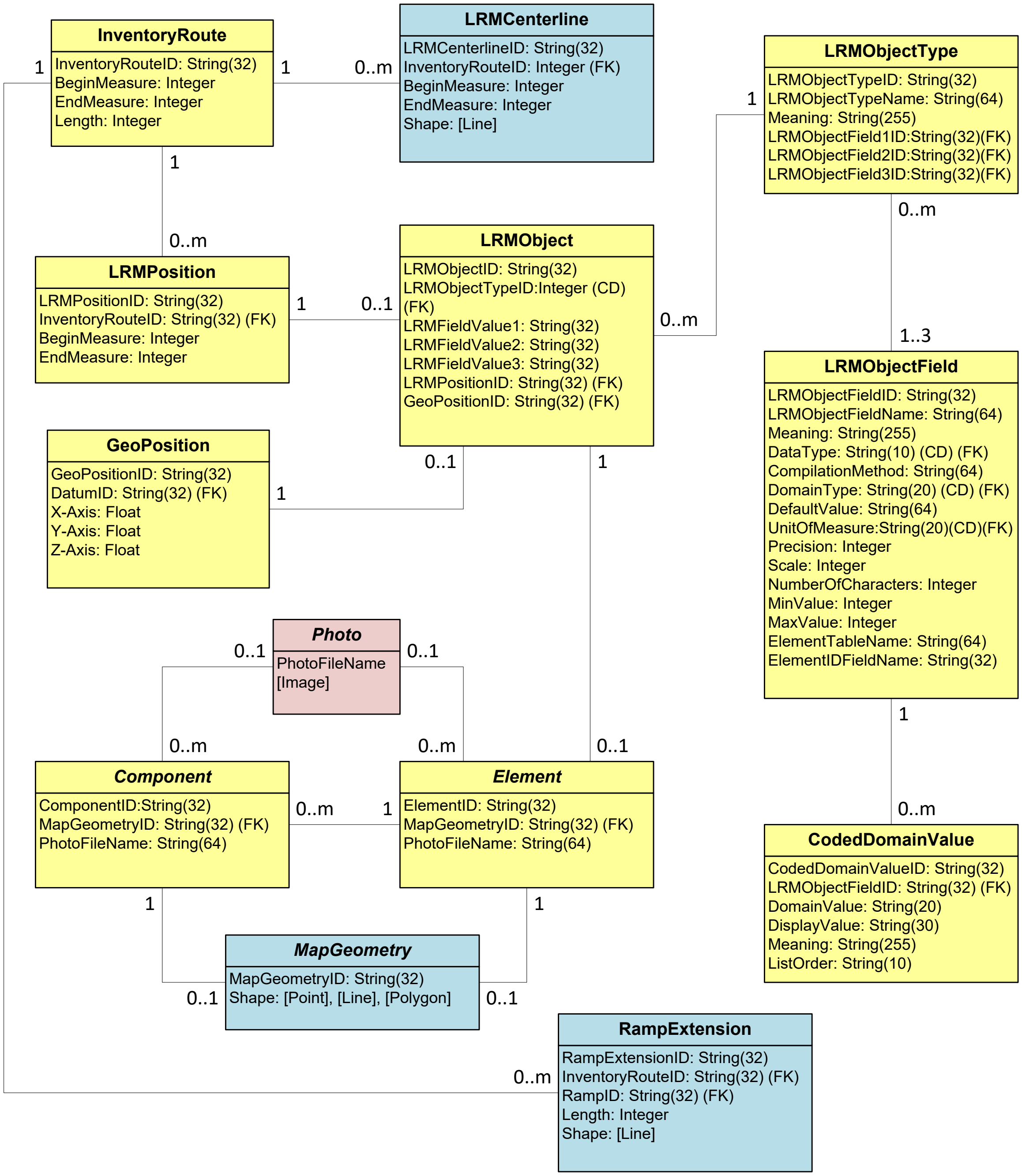


Figure 58. The NRBM linear referencing logical data model.

InventoryRoute is the LRM object that organizes everything else. LRM positions are defined along its length. The logical data model shows the data exchange version of InventoryRoute, where there is only a single NRBM LRM based on monotonic measure values. As a result, InventoryRoute contains the beginning and ending NRBM measure values that set the outer bounds for all LRM positions stated along the route. In an implementation of the data model for the native editing environment in a State, the beginning and ending measure attributes would be replaced by a single **LRMPositionID** attribute. This LRMPosition record would provide the beginning and ending measure values for all the LRMs used by the agency, not just the one used to publish NRBM data.

At the top of the model are the InventoryRoute and LRMCenterline classes. In the initial logical data model, these were represented by a single InventoryRoute feature class with the geometry to display each inventory route on a map. Later, InventoryRoute was separated from its cartographic representation because there could be multiple versions of a centerline to show the path of an inventory route at various scales. The LRMCenterline feature class now provides the roadway geometry for mapping the LRM-based roadway system.

The separation of InventoryRoute into a table and a feature class was also influenced by the fact that centerlines experience changes during editing processes. When initially created, each centerline feature is a single continuous line corresponding to one or more route segment centerline features. However, when part of a centerline is abandoned or otherwise removed from the dataset, the centerline feature becomes discontinuous. This is not a problem for the atomic route segments, as they can readily come and go from the active records in the dataset. The inventory route could be split into two continuous sections and one or both assigned a new identifier, but doing so would require revising all the LRM position records to reflect the new route identifier and, most likely, the new measure value. It is more efficient to manage these situations with a one-to-many relationship: one inventory route may be mapped using one or more LRM centerline features. (The 0 multiplicity is due to the fact that the InventoryRoute instance has to exist before it can be connected to the LRMCenterline feature(s) that describe it on the map.) If a user only wants to use the roadway layer as a transportation base map layer, LRMCenterline will work.

In the bridge replacement example given in Chapter 4, drawings showed that a new bridge was constructed on a new roadway alignment. Once the bridge was opened to traffic, part of the original route was abandoned, part was removed from the State Highway System, and the old bridge was demolished along with part of the old roadway. This resulted in Inventory Route 34762 having a discontinuous active-status centerline. The old continuous centerline needed to be retained for history. The new two-part centerline had to be constructed to retain the centerline's connection to the parent inventory route and to avoid changing any downstream measures, leading to the one-to-many relationship.

One alternative means of implementing this new relationship would be to simply create multipart centerline features. Unfortunately, it is too difficult to have a vendor-neutral NRBM specification and multipart centerlines, as each vendor approaches the problem differently, and exporting the result is problematic. Fortunately, it is simple to put the inventory route identifier and the beginning and ending measure values into each route centerline segment. These values can then be used to identify which centerline should be used for the dynamic segmentation (dynseg) operation to display LRM-based data.

These changes in the path of the inventory route are not a problem for the included RouteSegment features because whole route segments are replaced through the editing process. Route segments always have a single, unbroken centerline. The referenced bridge replacement example retired one whole route segment and replaced it with three new ones. (It also added a new route segment for the new roadway alignment.) The user clicks anywhere on the roadway centerline and the applicable route segment is highlighted and its attributes displayed.

Inventory routes do not work this way. A user cannot click on an inventory route and see all its attributes. Instead, the user needs to first select the desired attributes from a long list of available LRM object types, and then use dynseg to provide the discrete segments corresponding to homogenous attribute values. In other words, the route segment equivalent features change depending on the LRM objects selected.

At the bottom of the figure is a RampExtension feature class, which is a requirement inherited from the ARNOLD Reference Manual. This class provides the geometry needed to display ramp extensions, which are not part of an inventory route because they actually reflect the distance where the departure or entrance portion of the ramp has become part of the limited-access highway. The inventory route terminates at the physical gore. At larger scales, the ramp extension geometry serves to indicate the path a vehicle might take through this transition area. A one-to-many relationship has been established with a foreign key to link the ramp extension to the inventory route it graphically extends.

The overall design of the LRM-based dataset is one of deconstruction. It is normalized like a transactional database. This is usually not a good design for a data publication specification, where components are generally put together (denormalized) into various final forms. But LRM-based data are inherently difficult to combine into neat packets like route segments. In fact, the whole concept of LRM data structures is that the data cannot be readily denormalized into discrete pieces that will serve everyone's needs. Therefore, the NRBM carries that deconstructed form to the publication specification, where users can put the pieces together however they desire to segment the data where it makes sense to them. The LRM-based component of the NRBM dataset provides the bricks agencies can use to build whatever structure is needed.

The model does show that each LRMPosition instance points to a single inventory route. Naturally, many LRM positions may be tied to a given inventory route, so this is a one-to-many relationship. LRMPosition also has a one-to-many relationship with LRMObject. In other words, each LRM object must be assigned one or more LRM positions. This means an LRM object can be located on one or more inventory routes—it can be in two places at once—at least as the LRM sees it. Most States make the route identifier and measure value attributes of the LRM object, which means each object can get only a single LRM position description. The result is that intersections must be listed at least twice in the database, once for each involved inventory route. Any change to that intersection will require at least two LRMObject edits. If, instead, an LRMPosition record is used to show where the intersection is located on each involved route, then a single intersection object can be discovered regardless of which inventory route is being traversed. No data are duplicated.

Another difference from many State roadway inventory databases is the use of a single LRMObject class for all characteristics, elements, and events. LRMObject stores only two things, in addition to its identifier and that of the LRMPosition records that put it in the right location: **`ObjectTypeID`** and **`ObjectValue`**. **`ObjectTypeID`** points to the related record in the ObjectType class. This class has a record for each type of object (each event type) that explains the information in **`ObjectValue`**. This information is needed to understand the data in **`ObjectValue`**. It can be useful for automating the data entry process, as it contains everything a data entry routine will need to control what the user enters.

At the bottom of the list of attributes for LRMObjectField is **`ElementTableName`**. This field is used when the LRM object is a roadway element. It stores the name of the class where information about that element may be found. A few of these classes are described below. When the LRM object is an element, the data in LRMObject's **`ObjectValue`** field will be the identifier of the record in the table listed in ObjectType's **`ElementTableName`** field.

DomainValue works with ObjectType to store all the lookup table choices for LRM objects that need that information. The design allows the user to see a longer version of the choices than what might actually be saved as the value in LRMObject. It also allows GIS software to support tool tips by allowing the data supplier to provide the meaning of each lookup table choice.

This middle part of the data model best illustrates the type of database design required to support full temporality. Splitting LRM object data into multiple classes allows each to evolve independently. Reshape a centerline and none of the business data needs to change. Modify the domain values for an LRM object type to add new options and no existing data values need to be changed.

As with the prior logical data model, the attributes inherited from the *Object* stereotype are not shown. The additional InventoryRoute class attributes are as follows:

> **`BeginMeasure`** – The beginning measure of the inventory route. In accordance with the NRBM specification, this value should normally be zero. It may become non-zero when the inventory route is shortened from its point of origin.
>
> **`EndMeasure`** – The ending measure value for the inventory route as a monotonic value equal to the beginning measure value plus the stated length.
>
> **`Length`** – The traveled distance of the inventory route in the real world stated in units of measure equal to 0.001 mile.

The LRMCenterline class attributes are as follows:

> **`InventoryRouteID`** – The identifier of the inventory route for which this centerline feature represents all or part.
>
> **`BeginMeasure`** – The beginning measure of the inventory route. In accordance with the NRBM specification, this value should normally be zero. It may become non-zero when the inventory route is shortened from its point of origin.
>
> **`EndMeasure`** – The ending measure value for the inventory route as a monotonic value equal to the beginning measure value plus the stated length.
>
> **`Shape`** – The OGC LineString geometric object representing the path of the route segment; this geometry should be provided in accordance with the HPMS Field Manual except that the scale of abstraction, accuracy, precision, and resolution defined above apply.

The LRMPosition class contains the attributes that place LRM objects along an inventory route:

> **`InventoryRouteID`** – The identifier of the inventory route for which this centerline feature represents all or part.
>
> **`BeginMeasure`** – The beginning measure of the inventory route. In accordance with the NRBM specification, this value should normally be zero. It may become non-zero when the inventory route is shortened from its point of origin.
>
> **`EndMeasure`** – The ending measure value for the inventory route as a monotonic value equal to the beginning measure value plus the stated length. This field will be null for point LRM objects, where **`BeginMeasure`** will serve as an “at” measure.

The basic attributes in all three classes are the same: the identifier of the inventory route along with the beginning and ending measures. This might suggest that these three classes can be condensed into a single class. The reason they are separate is these attributes can carry different values. The total length of an inventory route may not be covered by a single centerline feature, and there are more LRM positions on an inventory route than just the beginning and ending values.

If the editor does not explicitly change the location of an LRM object, then it should not move. This means LRM objects upstream and downstream of any editing process must stay where they were; i.e., retain the same LRM position and geographic location. Linear referencing mandates a historic connection across all changes to ensure downstream locations are not impacted by editing processes. An LRM object on Inventory Route 46233 at Measure 3497 needs to always be at that LRM location in 1D space, regardless of any changes to the shape of the related centerline feature on the map. It must also stay in the same geographic location in 2D space, unless the centerline segment at that location is modified.

LRMPosition is actually a way of managing the relationship between LRM objects and the inventory routes on which they may exist. Most objects, of course, will be on only a single route, but intersections will be on at least two. A bridge may also be referenced to the route going over it and the one going under it. In such cases, there will be only one **`LRMPositionID`** value, but there will be two records for the location, one for each applicable **`InventoryRouteID`** value. As a result, it is necessary to use both **`LRMPositionID`** and **`InventoryRouteID`** to uniquely identify each LRMPosition instance. That is enough information to be able to order all the LRM objects on a given inventory route by beginning milepost measure and “drive down the roadway” to provide information in the highway log format.

The NRBM supports only a single LRM for published data. As a physical implementation for a State’s roadway database design, though, the LRMPosition class would also need to include a linear datum field to support the use of multiple LRMs. For example, the dataset may be published using signed routes as the basis for measure values, or limited-access roadways may have their data also published using mileposts (or KM posts) as the basis of the LRM.

LRMPosition does not include the means for indicating a location offset from the roadway centerline or some other defined object. This is because the scale of abstraction and expected mapping scales do not provide a means of distinguishing an object offset a few feet from the centerline. However, when implemented in a State agency, such information may prove to be very useful. Indeed, some States have already modified their LRM position descriptions to accommodate elements offset from the roadway. Such an addition would typically involve adding attributes to state the distance of offset, the referent element to start the distance offset measurement, and the direction of offset from the referent object.

The GeoPosition class provides the means to store Cartesian coordinates describing the location of an object. This class can be readily converted to a point feature class. Many elements, such as signs, are surveyed using handheld GPS units. These elements can be mapped independently of the roadways by using GeoPosition objects. The following are the included attributes:

> **`DatumID`** – The identifier of the applicable geographic coordinate datum; the NRBM supports WGS84.
>
> **`X-Axis`** – The coordinate for the X axis of the datum.
>
> **`Y-Axis`** – The coordinate for the Y axis of the datum.
>
> **`Z-Axis`** – The vertical coordinate for the Z axis of the datum. This value is optional.

LRMObject is the class where all the data are stored. An LRM object is what has traditionally been called an LRM event, which includes characteristics, elements, and “real” events (things that happen). The following are the provided attributes:

> **`LRMObjectTypeID`** – The identifier of the applicable LRM object type; the domain of possible values is currently open. The primary object types will be established during the pooled-fund program.
>
> **`LRMFieldValue1`**, **`LRMFieldValue2`**, and **`LRMFieldValue3`** – These three attributes all represent a descriptive aspect of the indicated object type. Some object types, like functional class,

only need one of these attributes. Three attributes are provided in order to store related data without having to separately maintain related objects that would have the same spatial extent. One example is inside shoulder, where storing both a type of shoulder (paved, grass, etc.) and its width may be needed. Another is an intersection, where the fields may include the intersection type, a foreign key to an external Intersection table in a MIRE dataset, and a value for the type of traffic control used.

`LRMPositionID` – The foreign key pointer to the applicable LRMPosition record(s) for this object. This field is mandatory for every object.

`GeoPositionID` – The foreign key pointer to the applicable GeoPosition record(s) for this object. This field is optional for any LRM object.

To construct a complete, standalone LRM object table, a user should conduct a relational join between the LRMObject and LRMPosition classes while selecting a single LRM object type. This result could then use the contents of the related LRMObjectType record to put informative column headings on the three **`LRMFieldValuex`** attributes. Information about the units of measure can be found in the up to three related LRMObjectField records. This table could then be applied against the members of the LRMCenterline feature class through dynseg to create a map theme for one or more LRM object types.

An object for calibration points is not shown in this model since they are used for adding measure values (the *m* coordinate) to the centerline's vertices. This process is part of the data creation and publication processes, so it has already happened at the State level before the dataset is added to the national NRBM dataset. Calibration points could be a defined LRM object type and have related LRMPosition and GeoPosition records in order to place them on the map and connect them to the proper centerline for imposing measure values during the calibration process.

The next three classes describe the supported LRM object types. The LRMObjectType class provides the name and explanation for a data type, along with the foreign key identifiers for the detailed information about each of the three available attributes that can be provided for a given object type. Those descriptions are contained in the LRMObjectField class, which says how to understand the values given in the LRMObject class. When one of those three fields stores the value provided by a coded domain of choices, those choices are stored in the DomainValue class.

The LRMObjectType class provides the name and meaning of each **`LRMFieldValuex`** attribute in the LRMObject class for the specified object type. Since the LRMObject class provides space for up to three descriptive attributes for each object type, the LRMObjectType class needs three attributes to point to detailed specifications for each of these fields. The following are the attributes provided:

`LRMObjectTypeName` – The name of the object type; the domain is to be determined during the pooled-fund program. Some obvious choices include the characteristics of speed limit, functional classification, and number of lanes, along with the elements of signs, intersections, interchanges, guardrail, noise walls, and bridges. LRM object types should be declared and specified by FHWA following a national vetting process.

`Meaning` – The text description or explanation of the name.

`LRMField1ID` – The identifier of the record in the LRMObjectField class that provides information about the first LRM field value attribute in the related LRMObject record.

`LRMField2ID` – The identifier of the record in the LRMObjectField class that provides information about the second LRM field value attribute in the related LRMObject record.

`LRMField3ID` – The identifier of the record in the LRMObjectField class that provides information about the third LRM field value attribute in the related LRMObject record.

The expected LRM object types to be defined during the pooled-fund program include:

- Functional classification (functional system and urban area code).
- Annual average daily traffic (AADT) (volume, year, and an indicator as to whether it is based on a field count).
- Number of lanes (total number of lanes and number of thru lanes).
- Traveled surface (type and width).
- Pavement condition (International Roughness Index, percent cracking, year of observation).
- Speed limit (maximum and minimum speed limits).
- Jurisdiction (maintenance jurisdiction using the Federal Information Processing Standard [FIPS] code).
- Federal-aid system designation.

The LRMObjectField class provides the specification of a single LRM field value in the LRMObject class record for the object. There will be one LRMObjectField record for each valid LRM field in the object type. The included attributes are as follows:

`LRMObjectFieldName` – This is the name of the object type field being described.

`Meaning` – The text description or explanation of the name.

`DataType` – The type of data stored in this field; the domain is 'GUID' for a GUID foreign key value; 'S' for a string of alphanumeric characters, 'I' for an integer number, 'F" for a floating-point number (contains digits after the decimal place), and 'D' for date. This domain may need to be refined during the pooled-fund program. For internal use at a specific agency, the domain may need to be modified to include the data types supported by the database management system.

`CompilationMethod` – The manner in which the applicable data value is determined; the initial domain is 'Office Research', 'Field Measurement', and 'Other'. This domain will need to be refined and expanded during the pooled-fund program.

`DomainType` – The type of value domain imposed on the values entered; the domain is 'NR' for a numeric range (minimum and maximum allowable values are provided in the MinValue and MaxValue attributes), 'CD' for coded domain of values, and 'NONE' for no domain control.

`DefaultValue` – An initial value that may be displayed for acceptance or modification by an editor; an entry is not required.

`UnitOfMeasure` – The measurement unit, if any, for the attribute's possible values; the initial domain is 'Feet', 'Meter', 'Mile', 'Square Foot', 'Square Yard', and 'Square Meter'. This domain may be modified during the pooled-fund program. The field is valid only for numeric data types.

`Precision` – The total number of digits permitted in the numeric value; this field will be omitted for non-numeric data types.

`Scale` – The number of decimal places provided by a floating point number; this field will be omitted for other data types.

`NumberOfCharacters` – The total number of alphanumeric characters permitted for a valid entry; this field is applicable only to string fields and will be omitted for other data types.

MinValue – The minimum permitted value for a numeric entry controlled by a continuous range of numbers; this field is valid only for numeric data types.

MaxValue – The maximum permitted value for a numeric entry controlled by a continuous range of numbers; this field is valid only for numeric data types.

ElementTableName – The name of an external relational database table, GIS theme, or object class that contains information about the LRM object type when the object is an element with an external dataset; this field is valid only for an element with an external data table/class.

ElementIDFieldName – The name of the field in the external element table or class that contains the primary key for finding additional information regarding the LRM object. One of the three LRMObject field value attributes will contain the actual identifier. **ElementTableName** provides the name of the external table or class, **ElementIDFieldName** says which attribute to search, and **LRMFieldValue*x*** provides the actual value for which to search. This field is valid only for an element with an external data table/class.

The CodedDomainValue class works with the LRMObjectField class to provide the actual values needed to populate a lookup list of choices from which the editor may choose, or to provide a recipient of the NRBM dataset to understand the data being provided. This is the class that explains such things as what a Pavement Type of 5 means, or what Functional Class 4 means. There will be one CodedValueDomain class record for each valid choice for a given object type. The included attributes are:

LRMObjectFieldID – The identifier of the LRMObjectField class record to which this DomainValue record relates; this is a foreign key.

DomainValue – This attribute stores one of the valid coded value domain choices. This is the value actually stored in the database.

DisplayValue – The coded domain value that is displayed to the user. For example, a pavement condition value stored in **DomainValue** may be 2. The **DisplayValue** field might contain the word 'Fair' as the value shown to the user.

Meaning – The text description or explanation of the name (domain value).

ListOrder – An alphanumeric attribute that allows ordering of the coded value domain entries. For example, rather than listing pavement condition values in alphabetical order, they may be placed in the order from 'New' to 'Poor'. New values may be inserted in a pick list by controlling the value placed in this field. For example, a new entry could be inserted between 3 and 4 by making it choice 3F.

The final four classes relate to roadway elements. One element may consist of multiple components. Both elements and components may have their own cartographic feature representation for mapping and/or a photograph. Note that the relationship shown in the figure for the Photo class is not really present in a foreign key structure. The model shows a Photo class being included to list all the available photo file names. The file names are shown as being foreign keys in the Element and Component classes, but they would not comply with the globally unique identifier (GUID) requirement. The Photo class could be implemented simply as a collection of files without an organizing Photo class being created. All four element classes are abstract; they are templates for creating a specific class for each type of element. All the indicated attributes in these four abstract classes will be listed and described, as these templates are working examples that need to be refined for specific element types during the pooled-fund program. However, three elements—intersections, interchanges, and restrictions—are mandatory inclusions in the dataset and must be described.

Figure 59 shows the prescribed element classes related to intersections, interchanges, and travel restrictions. The tabular Intersection class is joined with the IntersectionPoint feature class to produce the required Intersection feature class in the published route segment dataset. Ramp is a type of inventory route; it is expanded here to meet the requirements of MIRE and HPMS. The RampExtension feature class is related to the Ramp class record it extends. An interchange is composed of one or more ramps and one or more intersections, and is represented by an intersection point. The multiplicity of the relationships between intersections and interchanges to their geometric feature class represents the possibility that the intersection point may be placed at different locations depending on the map scale.

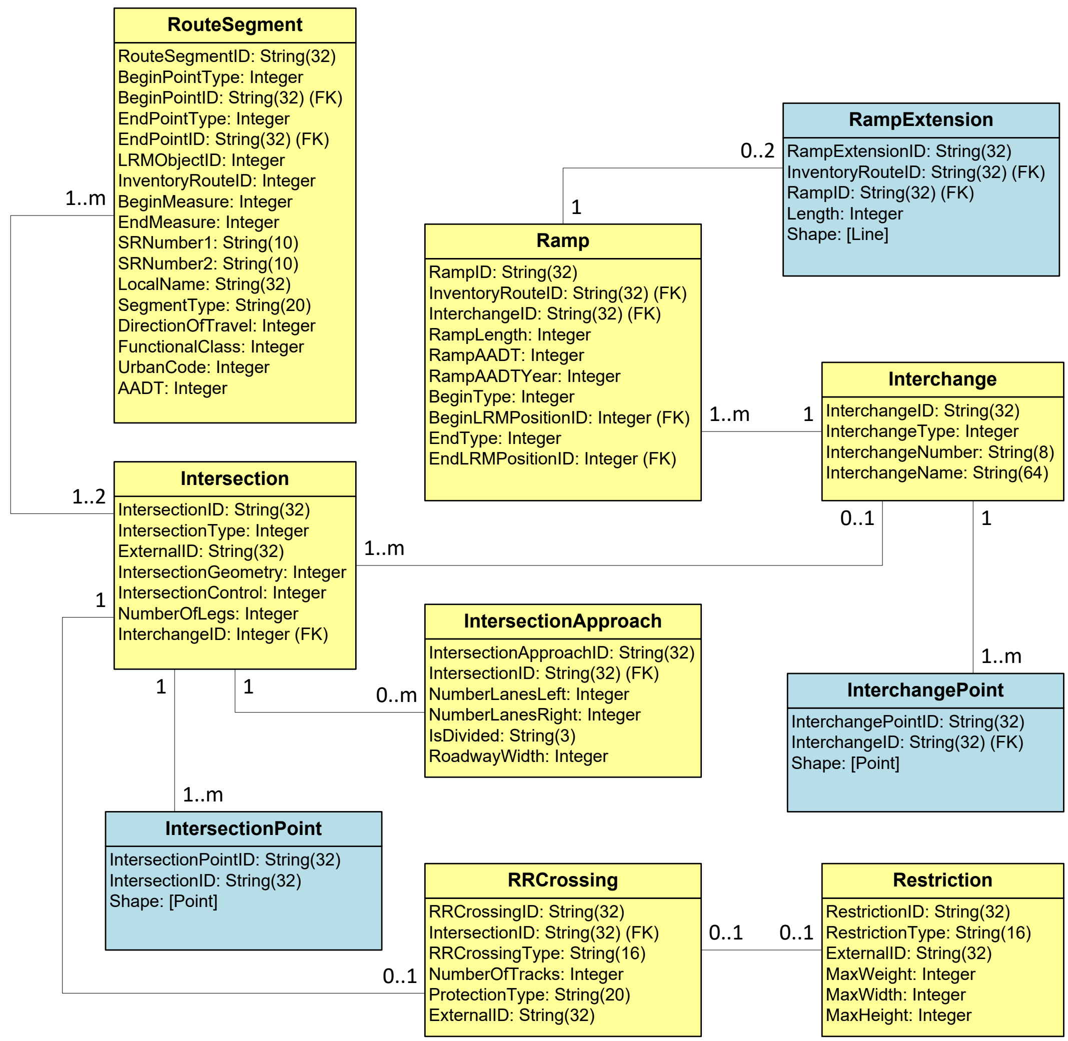


Figure 59. The prescribed NRBM element classes.

The classes included here along with their relationships are roughly consistent with the concept of elements and their components shown in the general LRM data model of Figure 58. Intersections are components of interchanges but are more commonly not associated with an interchange. Railroad crossings are a type of intersection or a travel restriction, not a component of them. Other restriction types include bridges, tunnels, and narrow roadways. Intersection approaches are a component of an intersection and are included to meet the MIRE need for these objects.

The RouteSegment class is shown here without its related point feature, which can be added using the dynseg function provided by the GIS platform. The logical data model for this class includes all the attributes of the published version; however, it could internally be implemented as the output of a process that pulls its various data elements from a variety of other LRM object types. If constructed in that manner, the only attributes required in the element RouteSegment class would be those not available elsewhere. The assembly of published route segments will be a topic developed during the pooled-fund program.

The LRM business rules are contained in Chapter 3. The most important ones are:

1. An inventory route shall consist of one or more whole route segments.
2. Each inventory route shall be assigned a GUID. Any desired intelligent primary key or identifier may be placed in the object's **`PublicKey`** field.
3. Measure values provided on an LRMCenterline feature must be monotonic for the published NRBM dataset.
4. Signed and named route designations are linear characteristics (LRM objects) of all or part of an inventory route. Rerouting of the signed or named route shall not modify the underlying inventory routes or measure values of other LRM objects unless the action directly modifies the extent of the involved inventory routes and other objects.
5. A route segment without topology may be created by using the dynseg function of a GIS platform by applying the route segment element class to the LRMCenterline class. The members of this feature class may then have topological extensions created by the GIS platform to form the topological route segment features required by the NRBM specification.
6. When the geometry of an inventory route is edited, one or more of the affected route segment features must be retired and new route segment features created, as needed. Once the editing process has concluded, the LRMCenterline feature that contains the modified/added route segments shall be recompiled. If the result is for an inventory route to require multiple, discontinuous LRMCenterline features to map its full length, then each such feature shall be related to the same inventory route but contain the measure values appropriate for the extent that remains.

One element that is not shown in the logical data model is the Terminus class. This published class is included in the route segment component of the NRBM dataset. It has few attributes and can be implemented by merely using an LRM object type. The related terminus point can be generated by applying the dynseg function. Thus, the Terminus feature class is a published entity and not one maintained in the active database. In contrast, the Intersection class has more attributes and requires its implementation as an element class with a separate cartographic feature class.

The Ramp, Interchange, RRCrossing, and Restriction classes are models for what might be contained in the NRBM specification once there is more experience with what is needed during the pooled-fund program. The attributes included in this logical data model are some, but not all, of the elements specified in MIRE. A published Interchange feature class could be readily added to the route segment component of the NRBM dataset.

The LRM component of the NRBM is a gateway to all the data available to describe the structure and performance of the roadway system. The many data sources inside and outside the States can be readily incorporated into the LRM structure. Chapter 4 raised the question of whether to include an explicit Median feature class. While such an inclusion would provide many benefits, its implementation may be difficult in the early version of the NRBM given that such features would need to be manually compiled by almost every State. Other expected element classes to be defined during the pooled-fund program include the following:

- Median (element), along with potential components of Barrier, Inside Shoulder, and other parts of a divided highway median
- Traffic Section
- Sign Assembly (element) and Sign (component)
- Bridge
- Tunnel
- Noise Wall
- Guardrail
- Pavement Section
- Cross-section
- Project
- Curve

Remaining Work

This AEGIST Guidebook is just the next step in developing a national transportation dataset that can serve the needs of the States, Federal agencies, local governments, Tribes, regional agencies, and other users. Many key decisions have been made regarding the structure of the national dataset for roadways and how it can be constructed by the States. The guidebook serves as a living document. As future work is incorporated, the specification will be revisited.

There are other potential elements of a complete NRBM to be developed during the pooled-fund program. One example is the implementation of a data error detection feedback loop, where users of the NRBM dataset can convey problems they may have found in the data. Such a feature should be provided as part of the overall governance structure in each State. Another possible extension is to provide guidance to a local government on how to migrate to using linear referencing or to a State for adoption of the NRBM database design as an internal data structure. It would also be useful to develop a work process flowchart for how a new workgroup can be added to the governance structure and data set. Detailed instruction for how to publish the various components of the NRBM from a different data structure will be a necessary part of each State's participation in the pooled-fund program.

It was noted earlier in this document that the field of the BIM method of designing and constructing transportation facilities was fast moving. Readers of this Guidebook are expected to have an interest in integration of CIM and GIS, particularly with regard to transferring information from the BIM platform to the GIS platform to reflect a change in status, such as "open to traffic."

The key to building the NRBM is enterprise data governance and its implementation of data management through CIM. This prerequisite to success is not expressed in the publication specification or data models presented here. An initial guide to developing enterprise data governance was provided in Chapter 2. More specific guidelines will be developed in each State that participates in the pooled-fund program. Software vendors are providing the tools to solve technical aspects of the task. It remains for the participants in the process to devise and implement the organizational aspects. In particular, as States implement the AEGIST guidelines and FHWA attempts to create a merged dataset of all States' data, it is expected that the specifications will change to accommodate new features, changes to existing definitions, and to add flexibility as needed.