



Regional Collaborative Maintenance Model For Road Centerline Data

Proposed Model Description

Version 2.5

January 25, 2011

Written by



For

MetroGIS

St. Paul, Minnesota

1. TABLE OF CONTENTS

1. Table of Contents.....	2
2. Introduction.....	3
2.1. Current Centerline Situation	4
2.2. Threats to Sustainable Solution	5
3. Current and Emerging Needs.....	6
4. Collaborative Maintenance Model Description.....	10
4.1. Model Overview.....	10
4.1.1. General Use Cases	12
4.1.2. Roles & Responsibilities	13
4.1.3. Logical Data Workflow	13
4.2. Policy & Practice.....	14
4.2.1. Authoritative Data Ownership.....	15
4.2.2. Authoritative Data Stewardship	16
4.2.3. Authoritative Data Curation	17
4.2.4. Organizational Participation and Outsourcing.....	17
4.3. Typical Use Scenarios	18
4.4. Technical	20
4.4.1. Integrated System of Software Applications	20
4.4.2. System Characteristics	22
5. Glossary	23

2. INTRODUCTION

On September 26, 2011, over 20 Metro area representatives from state agencies, regional organizations, county and city governments, and private companies met at the Metro Counties Government Center in St. Paul. The workshop attendees discussed the shared need for a **public domain**, authoritative street centerline spatial dataset representing the seven-county, Minneapolis-St. Paul metropolitan area, which at some point could expand to include the entire state of Minnesota and the border counties in neighboring states.

During the workshop, present day issues were discussed and small group breakout sessions brainstormed potential systemic solutions that could be implemented in the near term, evolve over time, and remain an effective collaborative program for the next 10 to 15 years. The workshop articulated the threat to a long-term collaborative solution that relies on privately held proprietary data. The group endorsed a vision that calls for transitioning from the current solution based on the licensed, proprietary road-centerline data file from NCompass Technologies (formerly TLG) to a public-domain, collaborative maintenance model comprised of federated systems, shared responsibility across authoritative jurisdictions, and broad regional commitment and participation backed by appropriate legislative policy. Overall, consensus was reached that the current model – MetroGIS managed distribution of licensed third party (NCompass) data – is not sustainable and cannot support either the current or emerging needs for shared road centerline data.

It was agreed that the next step would be for the project consultant to propose this work plan, generally tackling two initiative objectives, to the extent possible, simultaneously:

*1) **Build the technical infrastructure and related policies and practices to create a “hub of communication” for identifying and sharing information about street centerlines.***

*2) **Develop a strategy to transition from the current proprietary street centerline dataset, owned by NCompass Technologies and endorsed by MetroGIS as the Regional Street Centerline Dataset, to a public domain dataset consistent with the desire for a federated, transaction-based, collaborative maintenance model***

The workshop participants will continue to provide input to the collaborative maintenance model and will comprise the initial membership of the Centerline Steering Committee (CSC). The CSC will meet periodically and provide overall guidance and support during the collaborative maintenance model’s planning, implementation, and on-going operational phases to ensure that the two key objectives are achieved.

Separate workshop breakout groups independently focused on different aspects of accomplishing the policy, technical, and procedural aspects of the first objective. During workshop discussions, the idea of a full road-centerline data management system emerged and participants agreed it should be further developed. The consensus model comprises a federated suite of transaction-based applications for creating and maintaining a common set of **unique road segment identifiers** with minimal additional data elements. The full system makes it possible to create updateable derivative data sets and facilitate the collection, update, and distribution of authoritative regional road centerline data from a single well-known source. This “*hub of communication*” will enable peer-to-peer information sharing among data consuming organizations who need to share road related information and associated attributes ranging from first responder dispatch to roadway maintenance and highway funding distribution. Uniquely identified segments, associated with a central data model, support creation and update of multiple derivative data sets. This sustainable solution meets centerline data consumers’ identified business needs for on-going shared data maintenance and peer-to-peer collaboration and communication.

The creation of a technical infrastructure and organizational workflow to support the first objective requires the realization of the second objective to transition to a public domain solution. Much of the work necessary to achieve a public domain solution is described in this document (e.g., the identification of authoritative sources and the aggregation process to compile those sources into a single regional dataset). However, in order to utilize the existing NCompass centerline data, the public domain solution described must also address intellectual property and contractual issues. Resolution of these non-technical, non-organizational issues is critical to the realization of a shared public domain solution.

Much of the material from all groups has been used in the development of this document, a description of the proposed maintenance model and a work plan defining a set of initiatives to plan and implement the model. This brief outline can be used as a framework for decision making regarding the development of an integrated system of software applications, data workflow processes, policy-based organizational commitments, dedicated authoritative data ownership, and well-defined roles for sharing update and distribution responsibilities.

Due to the complexity and variations in requirements, a successful regional solution for centerlines cannot produce a single solution to meet all the needs of all stakeholders. In recognition of this issue, the model explored during the workshop proposes that its primary function is the maintenance and wide distribution of well-known unique road segment identifiers. Other useful (and to some, necessary) information, such as road geometry and address ranges, may come out of this model, but this information should be seen as a by-product of the core function of providing unambiguous identification of road segments. A system of common and well-know road identifiers will enable interoperability and communication between all centerline stakeholders and data consumers. Current and future collaborative initiatives, such as addressing multi-modal routing functions and other potential data needs within roadway right-of-ways, will be supported by the model, but those efforts are beyond the scope of this road centerline collaborative maintenance model.

2.1. CURRENT CENTERLINE SITUATION

In 1998, the Metropolitan Council on behalf of MetroGIS began to distribute NCompass Technologies (formerly The Lawrence Group – TLG) core product, geospatial data containing road centerlines for the Twin-cities metropolitan area. Currently approximately 200 licensed participants including counties, cities, academic institutions, watershed associations, school districts, the US Census, and several US states use this data in a wide variety of geospatial tasks. For example, MetroGIS uses the NCompass road centerline data for a number of analytical and mapping purposes including identifying bus routes and determining the functional class of roadways.

As the manager of the license agreement with NCompass, MetroGIS plays a central role in the administration current of the regional centerline data solution. In this role, MetroGIS manages license agreements and the distribution of the, currently, quarterly updates. Like some other consumers of the regional data, MetroGIS maintains a value-added derivative of the NCompass centerline data for its own business purposes. This derivative is not distributed to other centerline licensees. Apart from negotiating and administering the NCompass contract, MetroGIS has a limited role in a collaborative regional centerline data-management process. It has limited funding, minimal authority for make decisions regarding the regional data management processes, and no mechanism to enforce or require action among its constituents. For instance, MetroGIS cannot mandate that local jurisdictions be responsible for maintenance of geospatial data for local roads. However, MetroGIS can act as a catalyst for change and can perform some coordination and communications functions that may further the development of a collaborative regional solution.

Many licensed organizations represented at the workshop began using the NCompass street centerline data for their internal business needs, but over time, value-added changes (e.g., added topology and splitting segments at key points, such as municipal boundaries) made to the data by license holders have resulted in numerous derivative data sets. Once updated, these data offshoots lose their provenance and their updates – even if authoritative – never make it back to the source NCompass file. In addition to these NCompass permutations, other stakeholders have developed non-licensed centerline files for business needs not possible with the NCompass-licensed data, such as routing. One county currently maintains three versions of road datasets: Right-Of-Way, survey, and centerline. As this complex situation evolves, it presents broad intellectual-property ownership questions and precludes the function of a “single authoritative source” for a streamlined update process.

The workshop group recognized the stakeholder community, including NCompass, has made big investments in existing datasets and attendant processes. They also feel that there is a strong need for an integrated public-domain solution that could be developed from the current datasets; it does not need to be built from scratch. NCompass, represented at the workshop, stated that the company is open to negotiating the terms and particulars of a new model where the current license agreement is converted into a public-domain solution where NCompass relinquishes intellectual property rights to the current road centerline data. Nonetheless, local producers of centerline data raised concern about the readiness of cities and counties to give up the *total* control over the data and incurring risk by using data produced by other entities.

The various aspects of the current centerline situation (third-party data ownership, management of derivative data sets, aggregation of authoritative updates) can be summarized but this question: How do we, Metro area centerline stakeholders, effectively eliminate the need for derivative centerline files and move back to a common maintenance model while accommodating the consumer’s need to support many varied business processes?

2.2. THREATS TO SUSTAINABLE SOLUTION

During the workshop, the barriers that might prevent the realization of shared collaborative model were discussed. This section describes some of the key concerns raised at the workshop that will need to be addressed and as the shared maintenance model evolves.

- 1. Management of Intellectual Property** – Jurisdictions, at all levels, are making significant investments to create and maintain road centerline data. The current centerline file that is in use is owned by a private company that makes investments to maintain it. Some jurisdictions view their centerline data resources as a valuable asset and are reluctant to contribute their data in a regional solution because they anticipate a reduced return on their investment. Other jurisdictions feel that devising a system whereby centerline data updates generated by governmental organizations are used to enhance a private company’s asset is an inappropriate use of public resources. As geospatial technology becomes ubiquitous, demand will increase for highly accurate information about road networks. The most viable and cost-effective sustainable road centerline solution that protects the investments made by government is one where the intellectual property rights to the centerline data are in the public domain.
- 2. Incomplete Participation** - Not all MetroGIS members have the same level of interest or ability in the creation of a shared and collaborative centerline data-maintenance model. Some organizations see their investment in road centerlines as a proprietary resource that should be protected. Some jurisdictions may need to be convinced that the lack of short-term benefits is out-weighted by the long-term regional gains. Others may lack the technical capability or monetary resources necessary to contribute. These

factors complicate or/ or preclude the creation of complete metrowide publically accessible authoritative centerline data. Without 100% participation from all authoritative sources, the value of a regional solution is diminished, but not eliminated.

3. **Agreement on Authoritative Sources** - Later in this document, the concept of legal road authorities and authoritative data stewards is discussed. As these issues are worked and road segment are assigned stewardship jurisdictions, the identification of the intra-jurisdiction workflows may be problematic. One department might manage roadway geometry and another may be responsible for route names. The current practice of multiple permutations of centerline data, each with their own customized update process, masks the need to identify a common workflow through many organizations. The effort to establish an authoritative workflow within a jurisdiction may influence the production of authoritative output from some jurisdictions.
4. **Conflict Resolution** –The current lack of a well-known structure for authoritative updates enables multiple data modification processes and users creating the need to resolve redundant and conflicting updates. There is a need to identify authoritative source *within* a jurisdiction and there must be a process for resolving conflicts *between* jurisdictions. An example of conflict might be a non-trunk state road that traverses a local municipality. Even if the municipality and the state have a clear understanding of authority responsibility, a lack of well-known conflict resolution process could result in low data quality.
5. **Non-tangible Transition Costs** - Collectively, most stakeholders agree that a common solution would be well worth the costs to establish a collaborative model. However, large organizational and systemic transitions are difficult and expensive. A broad education and outreach effort will be necessary to increase awareness, articulate benefits, and achieve a critical mass of participants. The tipping point will be reached when letting go of some local control is perceived to be more valuable to organizational efficiency (i.e., with collaboration procedural, and economic efficiencies can be realized).
6. **System Complexity** – This model is broad in scope and has several overlapping and interconnected aspects: organizational, technical, procedural, political, etc. Common understanding, definitions, and agreements in all these areas will be necessary to design and implement this model. This complexity imposes significant risk to the fruition of this initiative. A prolonged and iterative approach that is flexible and adaptable will help mitigate this risk.
7. **Unfunded Mandates** – If the implementation of this model is seen as an imposition of an unfunded mandate, then it will be difficult to garner the support and cooperation of all necessary participants. In particular, road authorities, who have legislative authority and responsibility to maintain roads, may see the additional tasks and responsibilities for spatial data management as an additional burden. They may need to acquire additional revenues sources or may not fully participate.

3. CURRENT AND EMERGING NEEDS

Authoritative centerline data that accurately represents the current state of the roadways is needed for not only the seven counties that comprise the Metropolitan Council, but is also needed for the counties that adjoin the Metropolitan Council area. These geographic coverage requirements for centerline reference data extend across the St. Croix River to include at least three Wisconsin counties. Looking ahead to a solution that continuously evolves to satisfy requirements for the next 10 to 15 years should consider supporting the entire State of Minnesota with all transportation connections to the four neighboring states and two Canadian Provinces.

This section describes a number of business requirements for shared centerline data. A number of these requirements exist today, but as geospatial technology expands and becomes commonplace, additional emerging requirements highlight the need for a more comprehensive solution than exists today.

- **Support for multiple business use cases** - Many processes at multiple levels of government require centerline data (e.g., emergency response routing and analysis, mapping, regional planning, computer-aided dispatch). The specific use of the data determines the suitability of data formats and value-added capabilities. For instance, centerline data used for vehicle routing does not necessarily need to be spatially accurate, but centerline data used for mapping right-of-ways requires a high degree of cartographic precision. It is unlikely that one data format can satisfy all constituencies. Rather than attempting to determine a single format that meet all needs, a sustainable solution suggests that a central data model contain enough information to enable the creation of multiple derivative data sets that each can meet their specific business need yet retain key information to allow association with the central data model for on-going maintenance and peer-to-peer collaboration.

One example raised at the workshop articulated the conflicting needs of line segment nodes at jurisdictional boundaries. A solution to support vehicle routing should not have a node where a road segment crosses the town line between intersections. However, in order to manage the number of miles of road within each jurisdiction to support equitable funding distribution, identification of the town line location – possibly as a distance along a road segment using LRS– is required.

- **Support for derivative data set maintenance** – Many centerline stakeholders report that either they have, over time, created internal centerline data sets that are derived from data received via the Metropolitan Council NCompass licensed data or they have created their own file based on locally collected data. This local centerline file creation was typically produced to satisfy a particular business need, such as vehicle routing that could not be supported by the licensed data. Usually, these derived datasets cannot be updated to reflect changes in the original baseline data because of formatting issues (i.e., the derived dataset does not retain the original file’s unique identifiers) or due to a lack of integration capability built into the licensed centerline file maintenance or the derivative file’s maintenance process. The result is often duplicated data maintenance or, more significantly, unsynchronized datasets due to updates being applied to one dataset but not the other.

An integrated maintenance model must recognize derivative data sets and provide a coordinated mechanism to generate update transactions, approved by the segment’s road authority that can be applied to the derived dataset.

- **Frequency of updates** – The current NCompass process provides quarterly updates under agreement with the Metropolitan Council. There appears to be demand for more frequent updates, but that has not yet been clearly articulated by the workshop stakeholders. To-date, no workshop attendees participate in that update-cycle buy-up program. The long update cycle has caused some users to update derived copies of the licensed data, a process that essentially disconnects (forks) a local copy from the NCompass source data resulting in multiple authoritative centerline files that are difficult to keep synchronized. Since the workshop, LOGIS is exploring the possibility of providing the funds necessary for buying-up to monthly update cycles (\$9,117 over a five-year period – weekly updates would cost \$18,234). Access to more frequent NCompass updates encourages consumers to use the NCompass data rather than face the risk that derivative files may not be properly updated. The current license agreement stipulates that once the fee is paid, all licensees will have access to the more frequent updates. The LOGIS buy-up would

relieve some of the issues associated with the current proprietary solution, but other issues such as authoritative sources and lack of a public domain solution still need to be addressed.

- **Multiple geographic reference systems** - Various uses dictate the need for multiple geographic reference systems. For instance, the current NCompass licensed data does not provide Linear Reference System (LRS) support, which is an emerging need among municipal centerline consumers.

Data standardization - There is no metrowide data standard for road centerline data. The current centerline data-management methods of deriving local versions of the NCompass licensed data, or an organization's own source file, produce localized formats. These multiple baseline datasets (state, county, NCompass) each have their own identification data element. These individually formatted local datasets inhibit integration of the data into a metrowide resource for regional uses such as emergency response, which follows the National Emergency Number Association (NENA) standard.

- **Centerline attributes** - Because centerline data is used in many business processes, there is demand to associate many types of attributes with a line segment. Not all users need, desire, or have the ability to access to all centerline attributes. A centerline file that spans the entire metro area and contains all of the necessary data elements for the most widely anticipated use cases would be too large and unwieldy for many local and regional business process owners.
- **Road name aliases** - There are many needs for multiple names and route numbers for road segments. Lack of full alias support often necessitates creation of local datasets to solve local needs.
- **Jurisdiction** - Jurisdiction and timing Issues related to inclusion of roads on private property (e.g. mall access roads, apartment complexes), tribal authorities, and federal property produce gaps in coverage, long update cycles, and out-of-date information.
- **Road status** - Use determines inclusion of road segments based on status (e.g., platted, under construction, built). Most centerline files in use do not have status codes or similar filterable attributes. The general practice is to include only those segments that are needed for the intended use of the file being created, resulting in further differentiation and difficulties in sharing data across (or within) organizations.
- **Regional Routing Network** – There is a strong need for a regional routing network for public safety response, Metropolitan Council buses routing management, and many other public and private use cases. The need for accurate and current routing data that includes multiple modes of transportation (vehicle, pedestrian, bicycle, etc.) is growing.

[During the workshop, it was realized that pursuit of a solution to support the network/routing business functions is a separate and distinct task from managing a regional road-segment identification dataset. Each requires separate roles, responsibilities, and custodial commitments. The consensus among workshop stakeholders is that collaboration on routing initiatives will happen outside this scope of the model proposed in this document.]

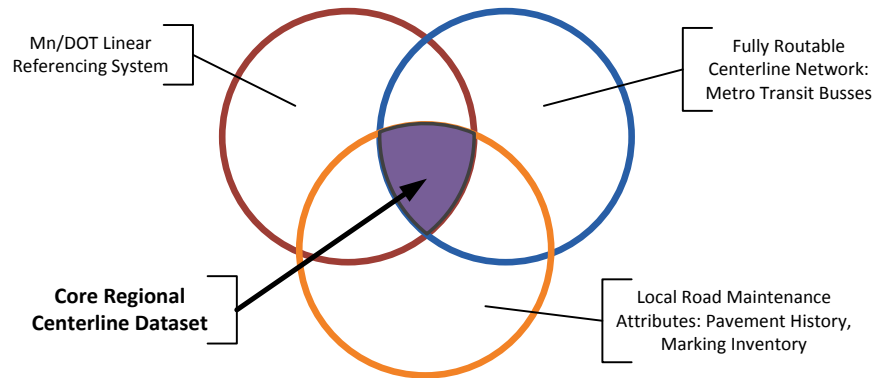
- **Next Generation 911 (NG-911)** – NG-911 programs are more heavily reliant on, and integrated with, geospatial information systems. Though the emergency response routing needs to support 911 belong to the individual dispatching organizations, the NG-911 system relies heavily on road centerline data to geolocate emergency events for call routing to the proper public safety access point (PSAP).
- **Public Access** - Road centerline data should be readily available to the public. The model should enable the NSDI vision with public and private sectors collaborating on maintenance with both having access to

quality authoritative data. Conversion from proprietary licensed data to a public domain solution is required for this to be achieved.

- **Highway funding distribution** – The Mn/DOT State Aid program must determine the number of miles within a jurisdiction.

4. COLLABORATIVE MAINTENANCE MODEL DESCRIPTION

During the stakeholder workshop, three breakout sessions were held. Each group independently focused on different aspects of the challenge of implementing a public-domain shared maintenance model. This section synthesizes breakout session output, along with other workshop and project material to define the required elements for a sustainable collaborative model for maintaining a regional shared centerline system.



The model described in this document focuses on the management of the minimal core dataset. This model will not contain all of the data elements required for direct support of all business functions. However, it must be able integrate with database extensions developed by the stakeholder community.

It should be noted that the maintenance model is not intended to support all stakeholder business functions. Business solutions will be built, by individual stakeholders and stakeholder group, to use the road key identification data model. The main purpose of the collaborative model is to share of a range of road centerline data and to enable communication between stakeholders as well as manage the authoritative update process.

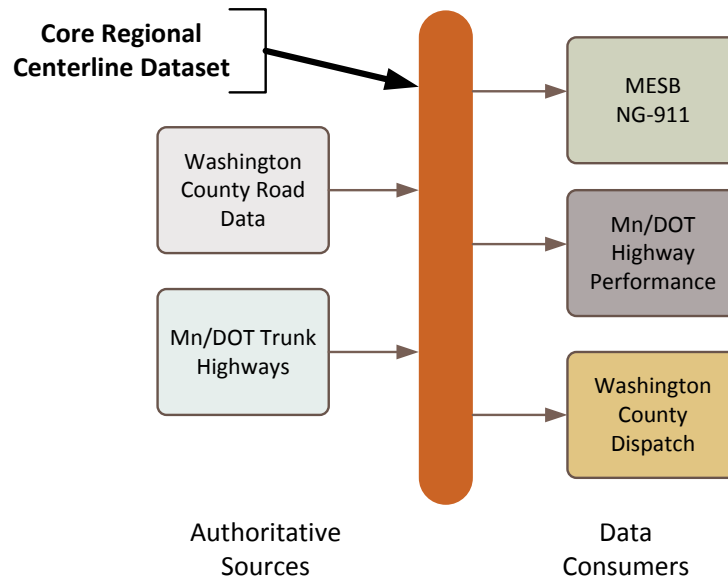
4.1. MODEL OVERVIEW

This section provides an overview of the collaborative data and process model for maintenance of a single source for road centerline data (**Authoritative Source Data**) for the Twin Cities metro area. This federated, regional collaborative model involves assigning data stewards to each road segment, enabling certain data consumers to suggest updates for spatial and attribute updates to existing roads and for the addition of new roads. The model has these key elements:

1. A core set of key data elements are aggregated from authoritative sources and freely shared to centerline stakeholders and other data consumers who can extend the data model with use case specific attributes.
2. Legally defined road authorities delegate data stewardship responsibilities to entities that ultimately share the responsibility for maintaining all of the data elements that define each road segment.
3. The authoritative data provided by each jurisdiction's data steward is aggregated into a centralized authoritative data set managed by a regional data curator.

4. The primary data element for each road segment is a single, permanent unique identifier (**Road Key**).
5. The data curator manages the model guidelines for assignment of unique road keys. For instance, the curator could assign each jurisdiction a 3-character alphanumeric code to be followed by a sequence number unique within the jurisdiction.
6. Each road segment's road key is assigned by the jurisdiction's data steward according to the model guidelines set by the regional data curator.
7. The road key is distributed along with other centralized road attributes to be maintained within all derivative systems and data models.
8. The road key can be used to enable peer-to-peer communication between road centerline stakeholders and other consumers to identify unique road segments and to share extended attribute information
9. Consumers of the authoritative source data can extend the central road-centerline data model to include additional data elements as appropriate to their business use case.
10. Other identifiers, such as road name aliases or additional route numbers, may be assigned to a road segment as necessary, either in the central data model or in the derivative data sets.
11. Certain centerline data consumers may be trained and approved by the data stewards to provide validated updates directly to the authoritative source data. Other general data consumers can suggest updates to the road segment's data steward who will approve or discard the suggested update.

4.1.1.1. GENERAL USE CASES



The model separates creation and maintenance of authoritative source data from distribution and business use of the data. In some cases, the authoritative source may also be the data consumer; however, these two roles should maintain distinct process workflows.

In a typical ***data consumer use case***, one organization might maintain data on traffic accident locations while another organization might maintain data about traffic volumes. Both of these use cases are relatively focused and the data attributes to support them should not be maintained in the central data model. Instead, both sets of data should be maintained in derivative data models that retain the authoritative source data's road key. Users of the derivative data can communicate with each other and correlate their respective data using the authoritative source data's unique road key. Using the road key eliminates the vagaries of road names, route numbers and other data anomalies to increase collaboration across organizations and increases the accuracy of identification of individual road segments.

In a typical ***data updater use case***, the jurisdiction (e.g., County) GIS manager, or another suitably skilled technologist, would be assigned the data stewardship responsibilities for maintaining county road data by the county board, the legal road authority. The steward may train and authorize an individual or role within the county surveyor's office to be a validated updater who can directly provide updates to road names and aliases. Other county departments, such as the public safety operations, may suggest updates to the centerline data, but these must be approved by road data steward, in this example, the county GIS manager.

The ***data aggregation use case*** begins after the data steward approves the changes generated by the data updater use case. The jurisdiction's authoritative data is collected, along with all other jurisdiction's source data by the central data curator who manages the data aggregation process. The curator identifies and resolves data conflicts, oversees the assignment of road keys by the data stewards, and prepares and publishes the authorized source data update for distribution. Potential data curators include Mn/DOT and MetroGIS.

4.1.2. ROLES & RESPONSIBILITIES

The collaborative maintenance model is described throughout this document using the roles identified in the following table. Each role has specific responsibilities and authority. These roles, under the guidance of the Centerline Steering Committee provide the management functions of the collaborative maintenance model described.

Role	Responsibility
Road Authority	Has the legal jurisdiction for maintaining the physical roadway infrastructure and assigning the official road name and route numbers usually a governing body such as the county board or the state Commissioner of Transportation.
Data Steward	The local entity with the authority, delegated by the Road Authority, for final approval of all updates to the authoritative source centerline data. The data steward assigns the unique road key to road segments within their jurisdiction.
Data Curator	The regional role that manages the aggregation process and who is responsible for standards enforcement, resolving data conflicts, and performs edge-matching functions. The data curator manages overall process for assignment of the unique road keys to ensure uniqueness across jurisdiction
Centerline Consumer	Users of the authoritative source data system that extend the central road-centerline data model to include additional data elements as appropriate to their business use case. These users maintain the centerline road key in order to accept future updates to the data, collaborate with other consumers, and provide suggested data updates to the authoritative source data. Updates from centerline consumers require data steward approval.
Validated Updater	Data consumers that are trained and approved by data stewards to provide validated updates directly to the authoritative source data. Updates from validated updaters do not require approval by the data steward.

4.1.3. LOGICAL DATA WORKFLOW

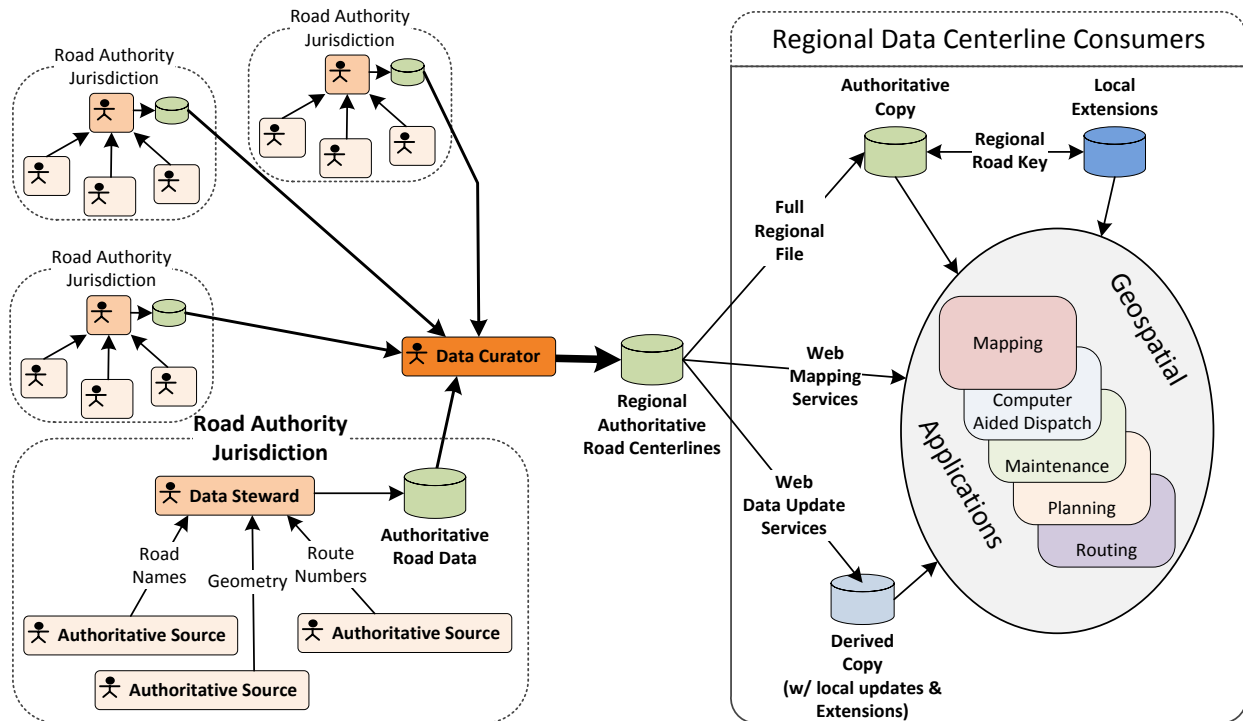
The diagram below shows the logical workflow of authoritative data. Authoritative sources within a road authority's jurisdiction provide updates to a designated data steward within the jurisdiction. The data steward assigns a road key to each segment and transmits the compiled authoritative road data to the regional data curator who aggregates the data from all jurisdictions into a single, authoritative road centerline dataset.

Since the road key will not change over time, this unique and static data element facilitates the on-going updates to each road segment and enables the authoritative regional file to be easily associated with the local extensions developed by data consumers. Consumers can access the regional data in a number of ways.

- A full copy can be downloaded to replace the local version, which may be associated with locally maintained data via the road key.
- Web mapping and feature services could be made available to allow geospatial applications to access the most up-to-date production centerline data.

- The regional data curator could provide web data service that provides updates to consumers who have fully integrated the regional data into their own data structures, but still want to receive transactional updates reflecting changes to the regional data.

Additional workflows, though not directly part of the collaborative data maintenance model, could be built around the road key. Various data consumers could use the road key as a form of communication for absolute identification of a road segment (e.g., when developing snowplow routes between jurisdictions). The road key system also provides for the potential for data consumers to share their value-added local extensions with each other (e.g., traffic volumes counted on state roads could be used for local municipal planning).



This logical workflow allows sources with a road authority's jurisdiction to provide authoritative updates to a single data steward within the jurisdiction. The data steward compiles the authoritative source data and provides it to the regional data curator who then aggregates the data into a single regional authoritative dataset. Various distribution methods can provide data consumers with data updates in a form that meets their particular use case.

4.2. POLICY & PRACTICE

The collaborative data model must have widespread support through practice and policy in order to become *the* source for authoritative data about the Twin Cities Metro Area roadways. Roadway centerline data are consumed by numerous business uses and the data are often augmented with use case specific attributes, sometimes with updated geometry. It is imperative that the model allow corrections, adjustments and additions to centerline data be made as few times as possible, ideally, once the authoritative source and distributed to all consumers. More realistically, data are likely to be updated at the point of use and the collaborative maintenance model must encourage updates to authoritative source. This section describes the policy, organizational, and best practices that are recommended for a successful implementation of a collaborative process model for centerline data.

4.2.1. AUTHORITATIVE DATA OWNERSHIP

The complicated roadway ownership, funding, maintenance agreement, enforcement jurisdictions, concurrent route numbers and authorized roadway uses often confuse who “owns” a road. Each road segment in the authoritative source data must be associated with a single entity that is ultimately responsible for the accuracy and currency of the centerline data. Minnesota statutes define the term “**Road authority**” (below) to identify the governing bodies responsible for public roadways and they identify the road authorities’ tasks and authorizations regarding all aspects of road infrastructure. Though there is no known statute that specifically defines the road authority as the data steward for roadways, it is logical to assume that this data maintenance responsibility could be assumed by road authorities. The assignment of the tasks and responsibilities outlined in this model may be seen as an unfunded mandate by some road authorities.

4.2.1.1. 2011 MN STATUE 160.02 DEFINITIONS, SUBDIVISION 25: ROAD AUTHORITY

<https://www.revisor.mn.gov/statutes/?id=160.02>

"Road authority" means the commissioner, as to trunk highways; the county board, as to county state-aid highways and county highways; the town board, as to town roads; and the governing bodies of cities when the governing bodies or city streets are specifically mentioned.

Realistically road authorities, who have the legal jurisdiction for the maintaining the physical roadway infrastructure and assigning the official road name and route numbers, will likely assign the task of data maintenance to another organizational entity. For this document, the term **data steward** is the entity authorized by the road authority to make the final decision regarding what information is contained within the authoritative source centerline file. It is envisioned that the data stewardship role would generally be performed by a technical organization within the road authority’s jurisdiction. However, the data steward’s duties could be delegated or assigned to another governmental entity or even out-sourced to a third party no-for-profit organization or a for-profit company.

4.2.1.2. NON-PUBLIC ROADS

The Minnesota statute that defines road authority covers trunk highways, county state-aid highways, county highways, and town roads. Several other classifications of roads must be included in the shared centerline file for it to be of full use to all data consumers. Policies will have to be developed to assign road authorities, data stewards, and **validated updaters** for these segments. It is possible that responsibility for these roads be assigned to other entities (e.g. the data steward for the jurisdiction they reside in). The following non-public road types have been identified along with reasons why they should be included in the regional data model.

- Private roads (e.g., access to malls and shopping centers, churches, hospitals, private schools, apartment buildings, mobile home complexes, corporate campuses, etc.)
- Subdivision roads under development, these may eventually be conveyed to other road authorities
- Tribal roads
- Federal roads

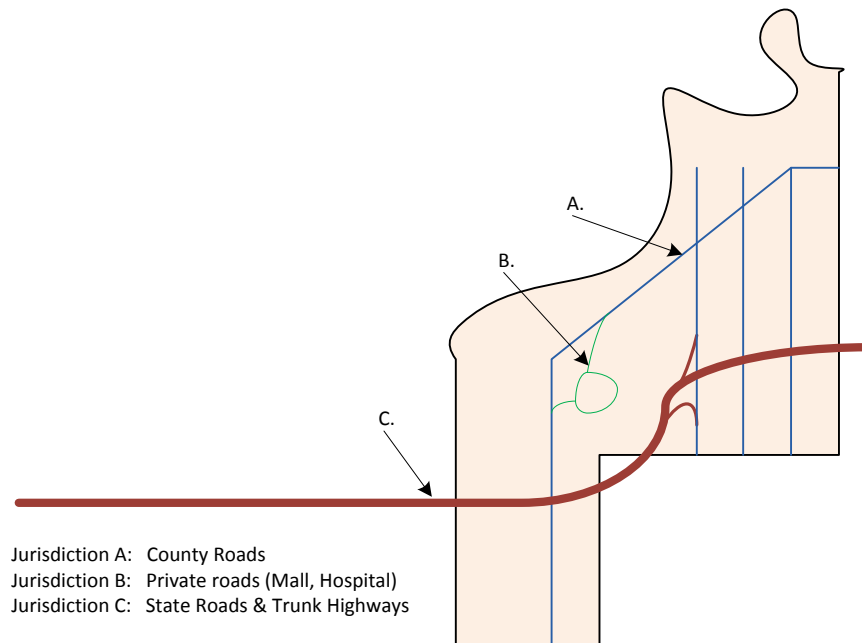
To be effective and ensure the highest level of accuracy and currency, the model should allow selected **centerline consumers** to suggest updates. These updates will come in many forms and from multiple sources; each must be validated and accepted before the authoritative source file is updated and redistributed. Each road segment that

is in the regional authoritative source data must have one, and only one, data owner (i.e., road authority) who is responsible for assigning a data steward to approve update requests to the authoritative source file.

4.2.2. AUTHORITATIVE DATA STEWARDSHIP

Once an appropriate road authority has been identified for road segments, the road authority must identify the data steward (or stewards) to manage the day-to-day processes and systems that ensure centerline data quality and accuracy. Generally, it is anticipated that the data steward will be within the road authority's organization, though it be assigned to another entity, for example in a neighboring jurisdiction or even a private entity. A data steward may then choose to delegate specific tasks or update authority to other entities within, or external to, their organization, but they retain the responsibility for the overall update process and the content of the data. The data steward will maintain overall responsibility for managing the authoritative update process for their jurisdiction and for approving authoritative data source updates. Depending on the overall progress of the system (see "Collaborative Maintenance Mode" below), the steward's update process could comprise local updates to a file that is aggregated with other authoritative data from other jurisdictions, or could use a centralized approach that allows stewards to perform their updates directly to the authoritative source data.

NOTE: A road authority, and therefore a data steward, might not have purview over every road in their geographic area (e.g., trunk highways). This document defines a data steward's **jurisdiction** as the specific road segments of specific classifications within geographic area that have been assigned by the governing road authority (e.g., all state-aid county roads and county roads within a jurisdiction).



Jurisdictions are associated with road authorities, not geographic extents. In this example, three road jurisdictions exist within the same political boundary: State roads, County roads, and Private roads.

Each of these jurisdictions will have an assigned data steward.

The data steward manages the inventory of road segments within their jurisdiction and maintains the lifecycle of a road segment beginning with the assignment of the unique road key continuing through to the road decommissioning and relegation to historical record. The data steward 's oversight tasks include identifying the sources within their jurisdiction who generate authoritative updates and solicits and accepts suggested updates from non-authoritative sources, such as realtors and first responders (e.g., ambulance drivers). In all cases, the data steward retains the responsibility for the quality and accuracy of the data within their jurisdiction.

4.2.3. AUTHORITATIVE DATA CURATION

Each data steward is responsible for managing the source data for their jurisdiction. Their job extends to working with the central **data curator** role to integrate their data with other jurisdictions across the entire region. The curator manages the process for assigning unique road keys and the process for integrating individual jurisdiction's source data into a single authoritative source data aggregation. The updated data may come to the curator as individual real-time or batches transactions, or may arrive as a full jurisdictional source file. Regardless of the method for receiving updates, the curator applies the updates to the authoritative data source. Once the data has passed quality assurance checks and any errors or redundancies have been resolved (by working with the respective data stewards), the curator prepares the data for publication.

In addition to managing the aggregation process, the curator also manages the data distribution process. Because there are many consumers of all levels of need and sophistication, multiple distribution methods are warranted. Initially, a complete regional or jurisdiction level files could be published in a publically accessible data clearinghouse. As the initiative matures, additional services, such as a web mapping service (WMS) or web feature service (WFS) are recommended.

This model is designed for consumers to create derivative data sets that contain, not only the core data elements, but also the necessary data extensions to support their individual business needs. One of the key curator functions is to facilitate the process for making authoritative updates to consumer's derived centerline data files.

4.2.4. ORGANIZATIONAL PARTICIPATION AND OUTSOURCING

As more cities and counties create their own datasets, there is less incentive to submit updates to NCompass for incorporation into the licensed dataset. As a result, NCompass is spending unnecessary and costly time finding and researching authoritative updates. The described model transfers the authority for managing updates as close to the authoritative source as possible, the jurisdiction's data steward. Making updates close to the source increases efficiency and reduces redundant and misapplied updates.

In order for this collaborative model to work effectively, ***all*** authoritative sources must agree to participate in a regional solution. Garnering 100% participation is a large and time-consuming task and may require inducements to encourage involvement:

- **Keep the barrier to entry low** - Make it easy to participate by providing tools, training and outreach
- **Make it a statutory requirement** - Road authorities and other governing bodies will respond to legislation
- **Provide funding** – Unfunded mandates discourage collaboration while funding fosters response
- **Keep it simple** – Authorities should be required to submit updates once; Consumers should have one place to access the data (in the format they require); Derivative dataset owners should have a clear simple, accessible way to incorporate changes to the authoritative source data into their business processes.

- **Minimize change** - Do not ask local authorities to change much about what they are currently doing.
- **Lower data standards** - In areas where the authority does not participate, data may be not up to standards. For example, only main roads could be shown.

Many stakeholders and other authoritative data producers have made significant investments in their current systems and data solutions. Outsourcing authoritative updates to a third-party consultant is one mechanism that can maintain the investment value and significantly reduce the data steward workload. Workflows could be established to enable authoritative update submissions from data stewards to a private vendor (e.g., NCompass) or another centralized organization that could perform portions of the update function belonging to the data steward and portions of the aggregation and distribution functions of the data curator.

Letting go of the control of a local investment (to either the collaborative process in general or an outsourced vendor in particular) produces both risk and stability. A single known process with common workflows, terminology, and best practices increases the overall value of the investments, the quality of the data available, and the general viability of the regional solution. However, the risks to a centralized outsourced solution that should be recognized and mitigated (e.g., loss of control, loss of a vendor, etc.). Any outsources or centralized service should allow for the use of multiple vendors and/or the transfer of data and process from one provider to another by requiring all out-sourced processes to use non-proprietary solutions and reserving intellectual property rights to the processes and methods used.

4.3. TYPICAL USE SCENARIOS

This section provides a brief outline of how various jurisdictions could interact with the collaborative maintenance model. It is unlikely that any two jurisdictions would interact with the model in exactly the same way. The model will be successful and be more quickly implemented if it can be flexible and it is able to support the specific needs of various jurisdictions. These scenarios assume that the collaborative model outlined in this document is operational. Specific organizations, jurisdictions, and requirements examples are not provided in order to generalize how the model could be applied in a variety of situations.

1. Small jurisdiction with no in-house centerline file and little geospatial capability

- A smaller jurisdiction uses geospatial resources for analyzing spatial data and producing maps for planning and other governmental purposes. Its governing board is the designated road authority for some of the roads in the jurisdiction. The board has entered into an agreement with a more affluent neighboring jurisdiction that now performs data stewardship functions. The board has also directed that various departments provide updated information about changes to the jurisdiction's roads to the neighboring steward who performs all data update functions on behalf of the jurisdiction. Periodically, the jurisdiction's sole GIS technician downloads a copy of the jurisdiction's centerline data from the regional data curator's web site. For maps and analysis covering an extent larger than the jurisdiction, the GIS technician uses the web mapping services (WMS) provided by the regional data curator.

2. Midsized jurisdiction with a centerline file derived from NCompass data and some geospatial capability

- This jurisdiction has a small 3-4 person GIS department, which anticipates growing as the jurisdiction grows. A few years back, the GIS department took a copy of the NCompass data for their jurisdiction and began to update it with changes and corrections that were made known to them by various departments. Most people in the jurisdiction know to provide updates to the GIS department, but this process is informal and it is assumed that centerline file is a best guess and is not assumed authoritative. Periodically, a new copy of the NCompass data is downloaded and compared to the jurisdiction's derived

file to discovery updates that NCompass may have collected from another source. It is known that in some areas of the jurisdiction the NCompass data is more accurate than the derived file and vice versa. The derived file with its ad hoc updates is used for most mapping purposes within the jurisdiction.

The jurisdiction has participated in the development of the regional centerline maintenance model and has been consulting with its governing board that is the road authority for the jurisdictions board. The board had delegated data stewardship responsibility to the jurisdiction's GIS manager and directed all county departments to provide updates to road geometry and attributes to the GIS department for processing. Recently, the jurisdiction's board has agreed to perform similar data update functions for the smaller neighboring jurisdiction.

As part of the implementation of the regional collaborative centerline-maintenance model project, the jurisdiction has provided to the regional data curator a copy of the derived file that they have been maintaining. While the data curator merges both the NCompass data and the jurisdiction's derived file into a single authoritative source, the jurisdiction has been receiving training on the new regional centerline data editing application. All parties are anticipating the upcoming process cutover when all centerline data changes will be sent to the GIS department. The new process identifies road segments by the new unique road-key identification field. Some departments are able to adjust quickly to the new workflow; others insist that they need more time and resource to make the necessary conversions. The data steward is able to work with both groups and she will update the regional centerline file using the regional data editing application.

Periodically, a copy of the authoritative centerline data will be downloaded from the regional curator and used for mapping, analysis, and other uses. The jurisdiction anticipates extending the regional data with local value-added extensions and it is working with the Centerline Steering Committee to learn the best practices for using the new unique road-key identification fields.

3. Larger jurisdiction with in-house created centerline and full geospatial capability

- This jurisdiction is large and has many organizations with geospatial technology and expertise. One of these organizations has been selected by the road authority to be the data steward for the jurisdiction. The data steward has worked with peers in various departments across the jurisdiction to identify the authoritative sources for each data element within the regional centerline file. The data steward has worked with the Centerline Steering Committee to create a version of the regional data editing application that can be installed in each department with an authoritative source function. Each department uses their own instance of the data editing application, but the update transactions are collected in a central location and managed by the data steward.

Initial plans call for the centerline file, with updates from all of the jurisdiction's authoritative updates, to be sent periodically to the regional data curator who will update the regional file with the new data received from the jurisdiction. The regional data curator, Centerline Steering Committee, and the data steward are working together to design and implement software and workflows that will convert any update to the jurisdiction's central centerline data file into a transaction that can update the regional centerline file using web services.

4. Data Consuming organization with now authoritative source data

- This regional organization does not have any direct authority for maintaining roads, but has many business critical reasons for needing a highly accurate and reliable centerline file. The organization has been working with the Centerline Steering Committee for years to ensure that the regional solution has the information that they need. The organization has several business functions that the regional solution cannot support. They worked with the committee to ensure that the design of the regional solution can be easily integrated with the organization's systems, which have been extended to provide the functions that the regional data cannot provide. Frequently, a new version of the centerline file is downloaded and the organization's copy is updated. Both the organization, the regional data curator, and the centerline steering committee are working together to develop systems and processes that accept update transactions from the regional centerline file using similar methods developed with the large jurisdiction for applying update transactions to the regional file.

4.4. TECHNICAL

4.4.1. INTEGRATED SYSTEM OF SOFTWARE APPLICATIONS

The shared maintenance model being described is more than a single aggregated data file with attendant distribution systems and processes. The collaborative model should be viewed as a comprehensive suite of federated policy, best practices, organizational structures, as well as the software and server infrastructure that supports the full lifecycle of road centerline data. This section describes some of the software systems required the model. These system components are described at a high level and who uses them (roles), the functions they perform (tasks), and outline how they fit into the fully built out operational model.

Road Key Generator After the initial creation of the authoritative data source, data stewards, who are responsible for creation of new centerline segments, will need a way to assign new road keys that are unique throughout the system. It will be the regional data curator's responsibility to manage the overall process for assigning road keys. However, each data steward will be responsible for assigning a specific road key to each segment. This can be facilitated by the development of a centralized road key assignment application that is managed by the curator and accessed by data stewards.

Authoritative Update Aggregator As the system develops, it is likely that data stewards will create their authoritative data locally and will send that data to the central curator as updated version of a complete file. When the system is complete, there may still be a need/desire among some data stewards to provide authoritative data in batch mode. This component process will be used by the data curator to update the authoritative source data by replacing a single jurisdiction's entire dataset.

Authoritative Update Editor Ultimately, a spatial data editing application, used by all data stewards, validated updaters, and even potentially centerline consumers would provide the most efficient and cost-effective way for managing the authoritative source data. The application could be deployed in a federated manner – along the lines of the MetroGIS address point editor prototype—to allow data stewards to host the application and control access to validated updaters and centerline consumers. It could also be deployed at a central location and managed by the data curator for those data

stewards who have a simple workflow and are unable or choose not to deploy the editor locally.

It is possible that the authoritative update editor could issue web-service based transactions that could be used to update the authoritative source data; however, initial iterations of the authoritative update editor should have the ability to create an export file that contains all the data for the steward's jurisdiction. These extracts should be in a format that is acceptable to the authoritative update aggregator.

This component would need a roles based approval workflow that could be tailored to each jurisdiction. For example, one jurisdiction may allow the planning department to add geometry for a proposed subdivision, but limit status changed (e.g., proposed to routable) to the highway department. General centerline consumers, and other unauthorized roles, could propose changes that would be directed through the workflow for approval or rejection. In the end, the data steward would be responsible for final approval.

Authoritative Update Transaction Processor This system component would be the most complex and would likely be the last to be designed and deployed. Each data steward would choose one method (i.e., update transactions or full export file) for transferring data from the authoritative data editor to the data curator. This component would function as the input queue for update transactions that are produced by the federated authoritative update editors. (Jurisdictions that choose the export file method would use the authoritative update aggregator component.) Approved update transactions would be produced by the locally installed authoritative update editor and would be submitted via web services.

As the overall system matures, this component could accept update transactions from any application that can produce correctly formatted transactions, though initial iterations may wish to limit exposure of this interface to just the authoritative update editor installations.

Distribution To support data consumers of all technical levels the collaborative model should accommodate multiple distribution methods. Some users, such as the public and other read-only use cases would benefit from a Web mapping service (WMS) or Web feature services (WFS). This would eliminate the need for maintenance of a local copy of the data.

Three distribution operations for updating derivative or distributed copies would accommodate the majority of data consumers and business use cases:

1. **Replacement** - Users who use the centerline data for cartographic or relatively simple analysis may desire a simple process that periodically replaces their centerline data with an updated version. These users need a regularly published file available in a Web-accessible repository. Consideration should be made to publish the authoritative data in a complete version as well as versions that are clipped by jurisdiction (or with some reasonable buffer to include full line segments that cross jurisdiction boundaries).

2. **Transaction** – Users who may be maintaining their own extended attributes in a derivative database (or in ancillary files related to a copy of the authoritative source data) may wish to receive individual transactions, or batches or transactions to update their derived data without modification of this user’s extended data elements.
3. **Difference** – Users who do not want to maintain a record of which authoritative source transactions have been applied to their locally derived copy. These users would benefit from a facility that compares their version with the authorized source data to identify the changes that should be applied. The user could decide which updates to apply and which to ignore. This technique would be best suited for an entity that accepts authoritative updates, but also performs (authoritative or non-authoritative) updates to their version of the data.

Consideration should be made to provide a versioned distribution application that is developed centrally, yet resides in the consumer’s environment. This application, common across all installed sites, would use web services to request new update transactions, or the complete dataset, from the authoritative data source servers and use that data to update the local copy of the centerline data, depending on which update method (above) the user has chosen. This application should be designed in a way that would accommodate customer specific centerline file/table layouts. For example, one customer might choose to use all of the authoritative source data elements while another might have any number of additional data elements appended to the centerline attribute table; the update application would modify the centralized data elements and leave extended elements unaffected.

This facilitated distribution process would increase the quality of local copies of the authoritative source data by ensuring consistency and reducing consumer generated variations.

4.4.2. SYSTEM CHARACTERISTICS

Overall, these systems need to have the following characteristics:

- Federated by distributing the processing and authority as close to the point of data creation as possible
- Accommodate users of multiple technical levels and infrastructure capacities
- Achieve the same tasks using multiple methods to accommodate varied use cases: Transaction-based, or difference based.
- Utilize Web services based whenever possible
- Leverage existing software applications and commercial off the shelf software (COTS) solutions

5. GLOSSARY

Term	Definition
Authoritative Source Data	The aggregation of all authoritative data from all data stewardship jurisdictions
Centerline Consumer	Users of the authoritative source data who may provide suggested updates to the data steward
Data Curator	Manager of the authoritative source data that is aggregated from all data stewards
Data Steward	Assigned by the road authority to be responsible for maintaining authoritative updates to a jurisdiction's road segments
Road Authority	Governing body who has legal jurisdiction over a road segment
Road Key	A unique road segment identifier that is assigned to a road segment by a data steward and managed by the data curator
Road Segment	(a.k.a. Centerline) A single section of road between intersections or other nodes that has a unique road key
Stewardship Jurisdiction	(a.k.a. Jurisdiction) The set of roads that are maintained by a data steward – not necessarily geographically bounded (i.e. a county stewardship jurisdiction does not contain state trunk highways)
Validated Updater	Data consumers that are trained and approved by data stewards to provide validated updates directly to the authoritative source data.