Arizona Department of Transportation

Improving Data Quality on the System of Record

Data Management on Arizona's Linear Referencing System

SAFETY DATA CASE STUDY

FHWA-SA-22-006

Federal Highway Administration Office of Safety

Roadway Safety Data Program

http://safety.fhwa.dot.gov/rsdp

U.S. Department of Transportation
Federal Highway Administration

ADOT

Notice

Quality Assurance Statement

Technical Documentation Page

1. Report No. FHWA-SA-22-006	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Improving Data Quality on the System of Record: Data Management on Arizona's Linear Referencing System		5. Report Date August 2022 6. Performing Organization Code
7.Author(s) Ian Hamilton and Lauren Coffey		8. Performing Organization Report No.
9. Performing Organization Name and Address Vanasse Hangen Brustlin, Inc (VHB) 940 Main Campus Drive Raleigh, NC 27606		10. Work Unit No. 11. Contract or Grant No. DTFH61-16-D-00052
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Safety 1200 New Jersey Ave., SE Washington, DC 20590		13. Type of Report and Period Case Study, January 2020-June 2022 14. Sponsoring Agency Code FHWA
15. Supplementary Notes The contract manager for this report was Jerry Roche. Sarah Weissman Pascual provided editorial review and support for technical content.		
16. Abstract This case study highlights a noteworthy example by the Arizona Department of Transportation (ADOT) as the State focuses on the quality of its linear referencing system (LRS). ADOT uses the LRS as the system of record to help manage and integrate datasets between several units and data owners in the agency. ADOT developed a framework for identifying critical issues with its LRS, such as gaps and overlaps in feature records. This process allows ADOT to trace systemic issues and correct them programmatically. This case study also explores how ADOT is advancing the State's LRS to improve safety data integration and support advanced use cases in data management and analysis. Communication with the State's traffic safety team and information technology staff has been critical to collecting and maintaining essential safety data on State highways and local roads. This has helped ADOT build institutional knowledge with respect to key concepts and pursue more advanced safety data applications.		
17. Key Words: Linear referencing system, safety, data, integration	18. Distribution Statement No restrictions.	

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 17	22. Price

Acronyms

Acronym	Description
ADOT	Arizona Department of Transportation
AEGIST	Applications of Enterprise GIS for Transportation
DOT	department of transportation
DQI	Data Quality Index
ETL	extract, transform, and load
FDE	Fundamental Data Elements
FHWA	Federal Highway Administration
FME	Feature Manipulation Engine
GIS	geographic information system
HPMS	Highway Performance Monitoring System
IT	Information Technology
LRS	linear referencing system
MIRE	Model Inventory of Roadway Elements
MPD	Multimodal Planning Division
R&H	Roads and Highways
RCI	Roadway Characteristics Inventory

Table of Contents

Introduction.. 1
Managing Data and the LRS 2
Improving Safety Data 5
Next Steps: Space, Time, and the LRS 6
Challenges 9
Conclusions.. 9
Acknowledgements 10
References 11

List of Figures

Figure 1. Graphic. ADOT's LRS is the common link for many data systems............ 2
Figure 2. Equation. ADOT's section scoring methodology. 3
Figure 3. Graphic. Example of ADOT's LRS data quality dashboard. 5
Figure 4. Graphic. Example LRS spatial topology challenges. 7
Figure 5. Graphic. Developing a spatiotemporal topology. 8

Executive Summary

This case study highlights a noteworthy example by the Arizona Department of Transportation (ADOT) as the State focuses on the quality of its linear referencing system (LRS). ADOT uses the LRS as the system of record to help manage and integrate datasets between several units and data owners in the agency. ADOT developed a framework for identifying critical issues with its LRS, such as gaps and overlaps in feature records. This process allows ADOT to trace systemic issues and correct them programmatically. This case study also explores how ADOT is advancing the State's LRS to improve safety data integration and support advanced use cases in data management and analysis. Communication with the State's traffic safety team and information technology staff has been critical to collecting and maintaining essential safety data on State highways and local roads. This has helped ADOT build institutional knowledge with respect to key concepts and pursue more advanced safety data applications.

Introduction

In 2019, the Federal Highway Administration (FHWA) completed the second U.S. Roadway Safety Data Capabilities Assessment (FHWA, 2019). This nationwide effort documented the safety data processes, policies, and procedures of all 50 States plus Washington D.C. and Puerto Rico. This effort highlighted the current state of practice with respect to all phases of safety data collection, management, integration, and analysis. It also revealed that State departments of transportation (DOTs) were eager to improve their capacity for data management and integration.

The purpose of this case study is to highlight a noteworthy example by the Arizona Department of Transportation (ADOT) focusing on the quality of its linear referencing system (LRS). This case study also explores how ADOT is advancing the State's LRS to be a system of record for integrating data and managing change. An LRS that considers spatial, and potentially temporal, change will allow ADOT to analyze changes to the network over time and analyze safety performance using the roadway conditions at the time the crashes were recorded. This approach can support safety data integration and advanced use cases in data management and analysis.

Purpose and Need

ADOT's Multimodal Planning Division (MPD) has extensive experience with geospatial technology and analysis. The geographic information systems (GIS) team is tasked with managing the LRS, as well as the Roadway Characteristics Inventory (RCI) for federal reporting purposes. As a part of these management responsibilities, the GIS team conducts targeted outreach with business units that manage transportation data. These partnerships help ADOT identify data overlaps and opportunities for integration within the overall GIS program.

Safety data support is a practical example of how ADOT is able to use the LRS as a system of record to help achieve the DOT's goals. At a broad level, the LRS supports transportation data integration in Arizona. As a specific example within safety, ADOT's traffic safety team collects and manages crash data in GIS format, and the MPD GIS team works closely with the traffic safety team to migrate workflows into a geospatial environment to improve data interoperability. This includes management of the federally-mandated Model Inventory of Roadway Elements (MIRE) Fundamental Data Elements (FDEs), as well as support for much of the safety analysis and roadway safety management process. In addition to capturing the spatial dimension of MIRE FDEs on Arizona's roads, ADOT plans to explore the temporal component of these data and refine the State's safety data capabilities.

Target Audience

The target audience for this case study is:

- State DOT GIS Staff
- Information Technology (IT) Staff
- Data Managers, Analysts, and Stewards
- Transportation Professionals in Planning, Engineering, and Safety

Managing Data and the LRS

ADOT began two projects to enhance data capabilities: 1) the RCI to collect MIRE FDEs on the Federal-Aid system, and 2) the Data Review and Visualization Project to support data quality and data management capabilities. These efforts reviewed major transportation-related data systems in the State, including the Highway Performance Monitoring System (HPMS), bridges, pavement, assets, and safety. ADOT recognized that data silos, poor data quality, and a lack of metadata were hindering data sharing and integration. ADOT's approach to data integration involves translating these datasets through the agencywide LRS. This provides a common basemap for all data, improving interoperability between data systems (figure 1).

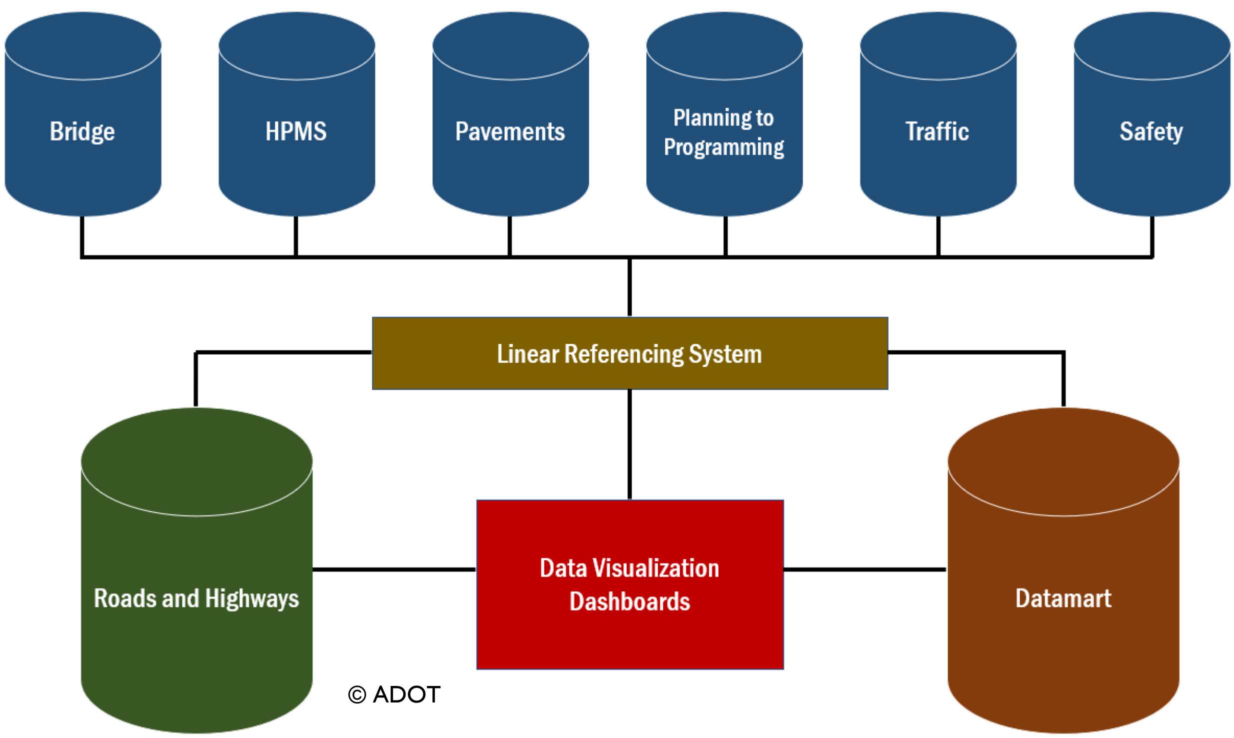

Figure 1. Graphic. ADOT's LRS is the common link for many data systems.

This focus on the LRS as a system of record for interoperability across the agency led to an effort to standardize the LRS and provide consistent workflows and quality checks. ADOT migrated to Esri's Roads & Highways™ (R&H) from its legacy systems in 2014, and the Data Review and Visualization initiative began in 2019. During that time, ADOT moved from an ad-hoc review of GIS data quality to one that uses a suite of Feature Manipulation Engine (FME) quality control checks. These checks form the foundation for data integration in Arizona using ADOT's Data Quality Index (DQI).

The DQI is a framework for evaluating ADOT's LRS and targeting specific issues in the data. The DQI consists of three sections:

- **Source:** This considers the authoritativeness and timeliness of the data.
- **Process:** This measures the sustainability of a process (i.e., replicability or automation).
- **Quality:** This considers confidence that ADOT has in a particular feature.

ADOT adopted a series of checks within each section to compute a total DQI for each data element. These consist of both qualitative and quantitative evaluation criteria. ADOT staff assign qualitative scores manually; the data are either good quality (assigned a value of 1) or poor quality (assigned a value of 0). ADOT automates quantitative scores, and these reflect a measurable component of the data. Qualitative scores are an important component of the overall section score. Figure 2 shows how qualitative scores are multiplied, while quantitative scores are added. Therefore, if a qualitative check fails in the Quality section (i.e., assigned a value of zero) the data element being reviewed will receive a zero for the entire Quality section.

$$Section\ Score = Qual_1 \times Qual_2 \times \ldots (Quant_1 + Quant_2 + \ldots)$$

Figure 2. Equation. ADOT's section scoring methodology.

The total score is computed by adding the three section scores (i.e., Source, Process, and Quality), and the DQI is the total score divided by the total number of individual checks across all three sections.

ADOT also implemented a series of logic checks that help diagnose individual or systemic issues (i.e., "pathologies") in the agency's data, such as:

1. **An event's "From" measure is larger than the "To" measure –** The record's starting position is recorded to a location after its endpoint.
2. **An event's geometric length is different than its spatial length –** The record is longer or shorter than the attributes would indicate.
3. **Multi-part event –** The record is not one single feature, and it is spatially split into multiple locations.

4. **Domain violations –** A record contains a value other than a prescribed set of potential values.
5. **Phantom events –** Records that are not locatable on the network.
6. **Overlapping events –** More than one record overlaps on a particular location on the network.

Using these checks, ADOT reviewed its LRS according to a multi-step process:

1. Evaluate the impact and extent of each issue in Arizona's data.
2. Trace the transactional data recorded in R&H. Tracing helps determine if the issue is an isolated event or a systemic issue and includes:
    a. Time of the transaction.
    b. Version of the software.
    c. Frequency of errors.
    d. Summary of errors by staff member.
3. Replicate the issue to diagnose how the issue occurs.
4. Prevent future possibilities for the issue to occur based on the issue tracing and replication.

This review of institutional processes helped ADOT understand its LRS data as it migrated to a new enterprise system. It also provides a method for evaluating data and promoting quality during edit sessions in real-time. ADOT refers to this as "defensive editing." ADOT adopted a visual dashboard for applying defensive editing to its LRS (figure 3). This interactive diagnosis and analysis application allows ADOT to promote data quality as the LRS and the interoperable data evolve over time.

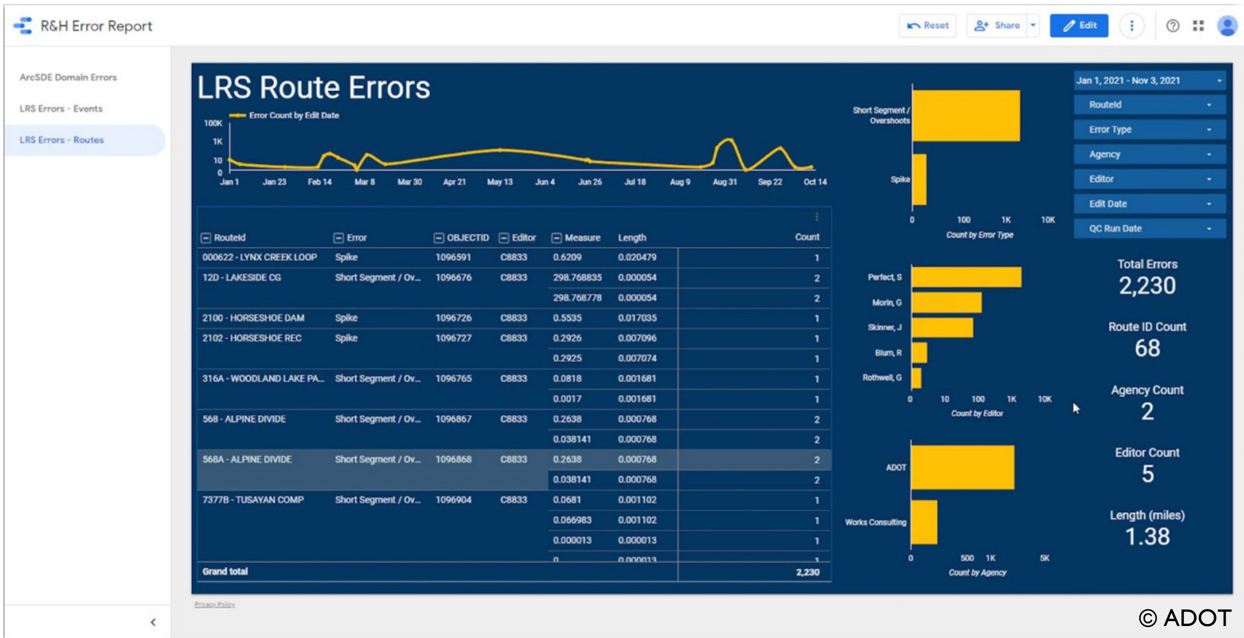

Figure 3. Graphic. Example of ADOT's LRS data quality dashboard.

Improving Safety Data

Through years of cooperative development, the LRS has become the system of record for ADOT to manage the State's transportation data. This approach has greatly supported the State's MIRE FDE inventory. ADOT developed a crosswalk table of existing data inventories and the MIRE FDEs as part of a comprehensive gap analysis. Collaboration with both the traffic safety team and ADOT's IT unit was a critical part of identifying data gaps and refining data schemas to meet the MIRE FDE requirements. These discussions helped establish plans for collecting additional data without causing negative impacts on other business unit reporting.

ADOT's approach viewed roadway and safety data collection holistically. This placed an emphasis on capturing data for all applicable roads in the State to support comprehensive safety analysis and implementation of a new geospatially-enabled safety analysis software. A plan to collect MIRE FDEs for locally-maintained Federal-Aid routes was a key outcome of the gap analysis and the multi-disciplinary discussion. This collection focused on data elements that could be derived from aerial or street-level imagery (e.g., speed limit signs and traffic controls) rather than other types of data requiring resource-intensive efforts (e.g., traffic counts).

As ADOT collected these data, the agency also noted a reciprocal benefit for its LRS as a system of record—route dominance. Route dominance acknowledges that a single physical road may have more than one signed route associated with it (i.e., a U.S. route and a State-signed route, etc.). By establishing route dominance early in the process, ADOT assigned data consistently to appropriate route levels and largely avoided issues of missing and inconsistent data.

ADOT plans to use a suite of tools to help manage the State's data supply chain into the future. This concept refers to the process ADOT employed for documenting the workflows, terminology, and metadata associated with key terms and schemas. The agency used Microsoft OneNote™ and Esri Story Maps™ to engage staff (particularly non-safety or GIS staff) and record institutional knowledge. These systems can potentially help connect the State's transportation data with financial management systems to better understand how the agency spends its resources. As with the LRS, authoritative data and a system of record are critical to effectively collecting data and maintaining it over time.

Next Steps: Space, Time, and the LRS

ADOT is working with FHWA through the [Applications of Enterprise GIS for Transportation (AEGIST)] pooled-fund program to explore innovative enhancements to how the State manages data on the LRS. ADOT is exploring new methods for managing the temporal dimension of its LRS and GIS data in addition to its traditional spatial component. The *Managing Data and the LRS* section provides some insight into how ADOT manages spatial rules, or topology[1], in the State's data. Gaps in the network, overlapping records, and events without a corresponding route location can be screened and addressed (figure 4).

[1] Topology is a GIS-industry term that refers to the spatial relationships of individual data features (i.e., records) and the rules that govern them.

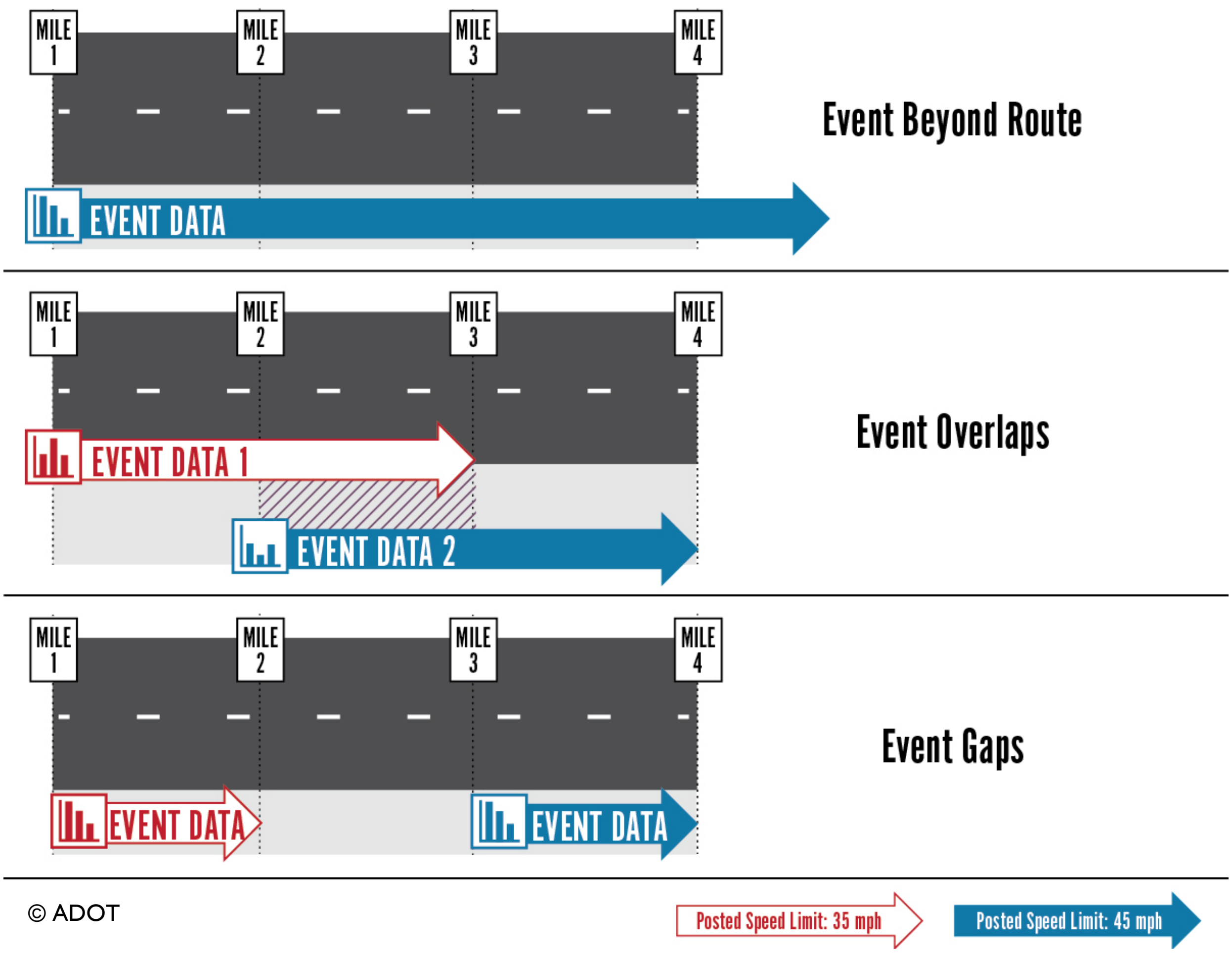

Figure 4. Graphic. Example LRS spatial topology challenges.

Like spatial information, temporal information also has its own topology. With ADOT's migration to R&H, the DOT can track the evolution of data over time according to two axes:

- **Valid time:** The lifespan of an event or geometric event (i.e., the start date of that feature to the date it is retired).
- **Transaction time:** The date database entries are posted (i.e., representing the date the facility is open to the public).

This offers new challenges to data managers, as well as opportunities. ADOT prefers to manage data visually. Examples like R&H provide a framework for visually correcting and managing spatial data, and ADOT can adapt this approach to also address changes in data over time.

For instance, analysts can use these tools to address spatial issues, as well as temporal issues such as:

- Address new and historic issues in spatial topology.
- Archive historic conditions and parts of the network no longer active.
- Manage a composite model of the entire LRS network.

Figure 5 illustrates how LRS visualization and editing tools can help analysts address issues inherent in the data over time. In this example, a portion of Route 2 is retired in 2018, and Route 1 is extended (e.g., reflecting a road name, route designation, or ownership change). However, a gap in the spatial topology emerges as Route 1 is not extended far enough (see example in figure 4). As analysts adjust route extents and retire old records, ADOT could recognize issues with its data over time and proactively address them to match historic data from other business owners.

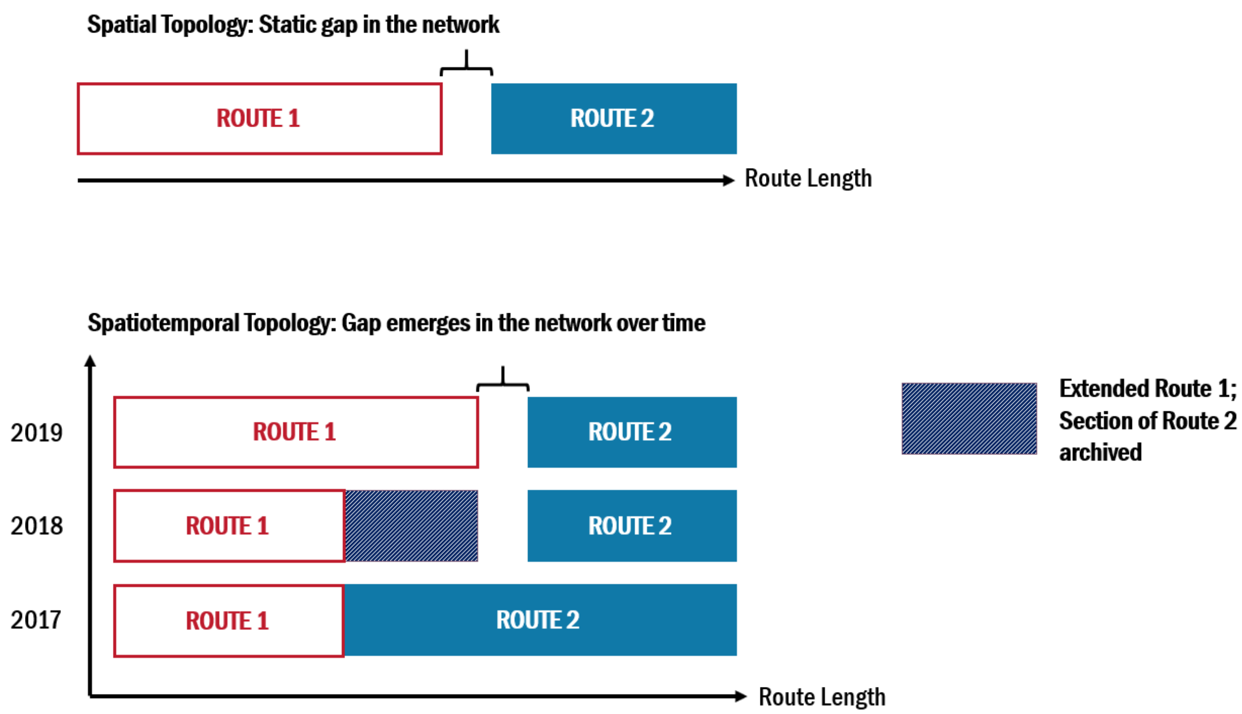

Figure 5. Graphic. Developing a spatiotemporal topology.

Addressing the temporal dimension of ADOT's data should yield benefits for several other business processes that rely on the LRS, including safety. More accurate and reliable data over time can help ADOT link crashes with the accurate conditions of the road network at the time of the crash. Not only could this benefit network screening, but diagnosis and countermeasure selection could be improved with better understanding of the conditions during which crashes occurred. Furthermore, this approach has the potential to support more robust data analytics and use cases, including machine learning for crash analysis and automated methods for feature data extraction (e.g., MIRE FDE data maintenance).

Challenges

ADOT's methodical approach helped address many of the issues inherent with the agency's existing data. However, ADOT noted many challenges as it improved its system of record. As ADOT reviewed many of its existing datasets (e.g., HPMS), it noted that these lacked the specificity in certain cases to meet MIRE requirements. This required ADOT to develop an extract, transform, and load (ETL) process to make these data readily available for safety analysis. Furthermore, datasets that were not readily available prior to extensive data collection, such as intersections and intersection approaches, could be collected according to specifications established during the multidisciplinary collaborative sessions. Finally, even as the agency collected and validated MIRE FDEs for many essential datasets in the State, obtaining timely updates to existing conditions, particularly on off-system roads, remains a persistent challenge.

Conclusions

ADOT's approach to managing data on the State's LRS illustrates how State DOTs can proactively approach data concerns. Authoritative and accessible data are key components of data governance and data integration. By investing in the agency's system of record, ADOT is making it easier to integrate safety data with other datasets. Furthermore, future MIRE datasets, such as the intersection and intersection approach inventories, are more readily integrated into the system due to fewer concerns with the integrity of the LRS and its fidelity to existing conditions.

Collaboration between the MPD GIS team and other business units, such as safety and IT, helps identify data gaps, streamline data collection, and avoid potential conflicts with other business units. These connections will continue to benefit ADOT as the MPD GIS team explores innovative approaches to data management. Formalizing these relationships, roles, and responsibilities are important future considerations for ADOT's data program. Clearly defined business cases for robust analytics and feature extraction methods (e.g., from computer-aided design as-builts and project plans), as well as embracing industry best practices will help mitigate future risk of data issues.

Acknowledgements

FHWA would like to express appreciation to the Arizona Department of Transportation for providing information used in developing this case study. For more information or clarification on the concepts in this case study, practitioners should contact ADOT and carefully consider the applicability of these approaches in their own agency.

References

- Federal Highway Administration. (2019). *Second U.S. Roadway Safety Data Capabilities Assessment*, Report No. FHWA-SA-19-018, Washington, D.C. Available online: https://rosap.ntl.bts.gov/view/dot/55617.
- James Meyer and Bo Guo. (2021, April). *Discover, Assess, Expose, Fix, Immunize: ADOT's Multi-prong Approach to Safe Guard LRS Data Quality* [Presentation].
- Lefler, N., Zhou, Y., Carter, D., McGee, H., Harkey, D., and Council, F. (2017). *Model Inventory of Roadway Elements – MIRE 2.0*, Report No. FHWA-SA-17-048, Washington, D.C. Available online: https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa17048.pdf.
- Patrick Whiteford and Bo Guo. (2021, April). *Anchoring Data in Time & Space: ADOT's Journey in Harnessing LRS Spatiotemporality* [Presentation].
- Patrick Whiteford and James Meyer. (2020, January). *MIRE Data: A GIS Perspective* [Presentation].

Contact

Arizona Department of Transportation

Patrick Whiteford, Geographic Information Systems Group Manager

mpwhiteford@azdot.gov