

GIS and Data Governance Peer Exchange Summary Report

Purpose

The Federal Highway Administration's (FHWA's) GIS in Transportation program sponsored a peer exchange to bring together practitioners of Geographic Information System (GIS) applications to discuss and learn about the implementation of data governance and data management activities and policies across State Departments of Transportation (DOTs). The Arkansas Department of Transportation (ARDOT) agreed to host the peer exchange in Little Rock, AR, and present their efforts to implement data governance and data management policies. Participants included staff from ARDOT and representatives from the Connecticut Department of Transportation (CTDOT), Tennessee Department of Transportation (TDOT), Iowa Department of Transportation (IDOT), Ohio Department of Transportation (ODOT), and Michigan Department of Transportation (MDOT).

The peer exchange was held over the course of two days and included a series of presentations from each participant and several roundtable discussions. The exchange concluded with a discussion of next steps FHWA can take to help State DOTs reach their data governance and data management goals, and final remarks that summarized recurring themes in the event's discussions.¹ This report includes a summary of the peer exchange, including a background, and summaries of each presentation, themes, roundtable discussions, and next steps for the participants.

Background

The peer exchange started with a presentation of FHWA's GIS in Transportation case study on Data Governance & Data Management.² The FHWA completed the case study in 2017 and featured interviews from four State DOTs. According to a survey of State DOT GIS practitioners,³ data governance was a common topic of interest in 2017. The motivation for this case study was to provide a common, modern, and well-known definition of data governance and data management and see how agencies currently define and implement these concepts. The case study found that agencies generally had a common understanding and definition of these concepts, but no formal definitions. Approaches to data governance and data management varied widely in their scope and strategy.

Day 1: State DOT Presentations

ARDOT Presentation

The peer presentations began with ARDOT. The host State DOT presented on how they have approached improving their evaluation, research, dissemination, cleaning, using, and development of data. In the past, ARDOT formed an Enterprise Data Committee to identify key data fields for standardization, implement the use of ArcGIS for server and ArcGIS Online technologies, and

¹ See Appendix A for a list of resources and tools mentioned during the peer exchange, and the peer exchange agenda.

² https://www.gis.fhwa.dot.gov/documents/GIS_Data_Governance_and_Data_Management_Case_Studies.htm

³ GIS Transportation (GIS-T) 2017 State DOT Survey

suggest/identify ideas for an organized data warehousing system. While these ideas/methodologies have a strong foothold, the committee did not have the authority to enforce standardization or govern the uses of the new technologies.

In 2016, ARDOT started working on a Safety and Mobility Data Business Plan. The purpose of this plan is to ensure that ARDOT can meet current and future needs with respect to organizing, collecting, managing, analyzing, and reporting safety and mobility data. The plan included a Department-wide data survey, which had over 233 staff responses that indicated the following:

- Where data are stored.
- How/why data are collected.
- Data items that are used most frequently.
- Confidence in data quality.
- If documentation (data catalogs/dictionaries) on various data items exists.
- How the data are shared.
- Types of geographic information tied to the data.
- Need for a structure to manage data.

The plan identified data gaps around ARDOT and included a follow-up implementation plan to move ARDOT forward toward an established data governance setup.

In 2017, ARDOT established a Data Governance Committee, which includes six senior-level managers and an executive-level manager from across the agency. The data business plan will be a key tool for this group to review documented issues and guide the development of strategies for moving forward. The agency plans on expanding the Data Governance Committee to add more members and working groups.

Ohio DOT Presentation

ODOT has taken a dual approach to data governance, with both short- and long-term activities planned. The agency formed a committee to establish data governance policies and standards. Their long-term vision is that data governance will become second nature as it is so engrained in the day-to-day operations of the agency.

Effective asset management has helped ODOT leverage data governance practices into operations. For example, ODOT formalized its workflows for data collection agency-wide into the following process: a user submits a request for data collection to a council who then decides who the data owner is, what the data standards are, who collects the data, and the overall plan for collection. The agency is making sure that data collection serves multiple purposes—that a data set is owned, properly maintained, and meets agency quality standards. This process helps ODOT continually manage data collection requests and data set creation.

ODOT is currently developing a job description and staffing plan for hiring a chief data officer. Going forward, they will work toward establishing an official data governance framework and strategy, creating data standards, and working with their information technology (IT) department to put a forward-thinking plan in place.

ODOT has had success in justifying data governance activities and showing a return on investment. This includes analyzing the approximate full-time employees (FTEs) needed per year for performing Quality Analysis/Quality Control (QA/QC) on poor-quality data and the organization cost for doing this work.

Iowa DOT Presentation

IDOT strongly recommends creating a Chief Analytics Officer position charged with determining what data to analyze, what data are important, and to begin the data governance process. Starting the process from the analytics side provides an authoritative basis for justifying why data governance issues must be solved, and in what way they need to be solved to effectively analyze and use the data.

There have also been some challenges and issues IDOT faced during their data governing processes. These issues include:

- Difficulties finding, accessing, and understanding data sets.
- Lack of open data for peer review.
- Quality of data available to provide in emergency and critical situations.

Metadata is an often-neglected aspect of data governance, according to IDOT. They stressed that there needs to be a way to better manage metadata. A possible solution is to hire someone with a background in library science who understands metadata management, filing, and card catalogues. IDOT cited Minnesota DOT's (MNDOT's) metadata catalogue as a successful example of this approach since a library science professional designed their catalogue. Demonstrating how this approach also allows them to collaborate with other organizations, IDOT compiles road conditions from MNDOT and integrates and feeds them back to weather stations for consumer use—a process that is possible because they have identical data standards.

Connecticut DOT Presentation

When developing data management standards, CTDOT created the Transportation Enterprise Data (TED) Warehouse, which stores many different types of data. TED started as a collaboration between the planning and engineering departments at CTDOT, using FHWA's Safety Planning Model as a template. TED is an overarching initiative, employing a new philosophy on how data are used to make enterprise decisions, collaborate, and ultimately to maintain authoritative and multi-use data. CTDOT's Data Governance Program helped expand data integration efforts, enhance data products, and spur collaboration across the agency.

CTDOT also created an Executive Oversight Team and Data Governance Council to assist in identifying data sets for formatting, as well as set and enhance data standardization. These groups meet three to four times per year and are accompanied by a lower level staff group doing the same work at a grassroots level. There are also working groups for different functions that feed back into the formal committees. These groups work to exchange and integrate data from other State agencies and metropolitan planning organizations (MPOs).

When discussing lessons learned, CTDOT highlighted the following:

- There are no shortcuts—enterprise-wide data standardization is a large undertaking.
- Involving all partners early on is crucial, as is cross-department collaboration.
- Executive level support is key to success.

- Communication and feedback from TED users and stakeholders are very important in the process of continuous improvement.

Michigan DOT Presentation

For data management and governance, MDOT houses 30 databases through Oracle and uses the agile application development method. GIS data account for 10 to 20 percent of the total data that MDOT owns—they recognize these data are an important part but not the largest part.

Agile development gets things out quickly, according to MDOT, but it does not allow the proper development of data governance questions. MDOT mentioned a need to add steps in the project management office that examine data architecture and data governance questions. Overall, data dictionaries and business glossaries are both needed as well.

MDOT built an intelligent transportation steering committee with the following sub-groups: a Data Governance Council with IT agency liaisons; a Data Governor and Data Stewards; and Communities of Interest for staff. The agency also enlisted a consultant to help implement a plan going forward for data governance. MDOT learned that building a data repository should only be undertaken if its composition and uses are well defined; these repositories often get built and then torn down because their functions do not align with goals (or vice versa).

Finally, MDOT stressed the importance of cross-departmental collaboration between IT and GIS, as well as also recommending a chief data officer position. Such collaborations are key to limiting roadblocks in the data sharing and managing process.

Tennessee DOT Presentation

TDOT has been less successful in their efforts to foster and facilitate data governance. According to TDOT, their business applications are not driving the agency in that direction. These applications are mapped to an organizational chart to learn how they are related, who is working on them, and what can be improved. Users at TDOT already have an established data management method, in which others within the organization understand to approach them for their data needs. Metadata, however, is inconsistent and makes finding the correct data very difficult. The agency completed some preliminary work with data scientists to begin the data governance process, but TDOT believes it must find some way to facilitate the needs of regional offices and create buy-in across the agency for this effort to be successful.

Among its challenges, the agency has lost a large amount of institutional knowledge due to people retiring, and new staff only staying at the agency for an average of three years. Currently, no legacy manual exists that shows new staff how to perform their expected tasks, which TDOT believes is necessary to improve data governance.

There are different efforts underway at TDOT aimed at improved data management and data governance. GIS leadership, for example, developed a list of tentative members for a preliminary Data Advisory Committee, along with a list of potential data categories to be put in front of this committee. A data dictionary initiative is also underway for the entire agency, building on preliminary data standards developed by the IT department. Even with these efforts, however, there remain challenges ahead for TDOT in implementing successful GIS data practices. When evaluating its capability maturity in 2018, the agency found that staff do not see GIS as central to their workflows and are unsure who to ask for help

when a GIS question arises. To best address this, IT and GIS groups at TDOT are working together to better establish data processes and operating procedures.

Roundtable Discussion: Data Governance Policy

All peer exchange attendees agreed a unified, comprehensive definition of data governance is very important to generate useful plans for work and encompass the entire set of problems that may need to be addressed. Finding a common philosophy is a major challenge in getting staff to work together and find a path forward. Establishing this philosophy is critical, however, as it leads into executive management buy-in.

Participants agreed that having an executive-level officer is more functional than a group of executives across the agency when developing data governance policy. However, the executive-level officer needs to have dedicated staff for data governance activities as well. Ensuring representation of GIS staff needs can be done through other means as well, such as consultation from a GIS group solely for higher-level decisions. Additionally, there exists need for a working group at the project level. Participants noted there is a need to have a chief data officer reside outside of the IT department, and the need to have adequate staffing in their department. The position also needs to be high level enough to have authority in enforcing data standards and to attract experienced and capable staff. Participants noted the difficulty in competing with private industry pay for the same positions.

Other Observations

- Data catalogues can show how the data can be used, how it can be queried, and who to contact about it.
- Databases should be unable to be duplicated, and centrally located.
- Crafting the job description to get good candidates and fulfill the skill sets and justify the higher salaries is important.
- ODOT's Services Lifecycle is used to organize operations and take inventory of what the departments/agency can do. For instance, ODOT uses the Services for talking to engineers about what they need to define the product they are getting.
- Putting legacy applications into a database viewer to be visualized can be a useful solution.

Day 2: Roundtable Discussion: Technical Tools for Data Governance

Data Access, Collection, and Storage

Controlling access to data is a necessary step toward data governance—IDOT and CTDOT offered to provide their ArcGIS Online (AGOL) user governance documents for States creating a set of user protocols. Approaches to open data portals may include releasing only vetted data sets for staff to work with, or waiting until the data is clean, formatted, and verified to be submitted to a portal. Monitoring data entry is necessary, but it is unclear who should supervise this; this could be done through the system administrator, or a formalized oversight group or process (much like ODOT's approach, which has a highly documented process involving their Technology Council for all data requests).

IDOT reported improving their data storage infrastructure by upgrading AGOL to an M1 and M2 subscription for unlimited storage at determined price—offering cost savings and faster speeds. The

agency also used the credits for trainings and invited Esri to implement services at the agency. IDOT is using Esri's 50-day training pack for IT staff in order to help them understand the GIS system architecture.

Feature Manipulation Engines (FMEs) are simplified software products for extracting, loading, and moving spatial data without prior coding knowledge or skills. IDOT uses an FME to publish all of its data to open data portals, and they can also push the data to AGOL. IDOT's FME runs on all of the agency's websites and it helps to translate data into different formats. This is especially useful in situations where staff are not utilizing a standardized data format. Similarly, ODOT is also using an FME to produce many of its reports.

There are multiple warehousing options for data. The options selected, however, must be compatible with all the tools an agency has invested in. Agencies can have multiple data warehouses or servers but it is necessary to minimize the number of data formats, tools, and architectures to eliminate duplication of function and reduce the complexity of the overall system. Data warehouses must be the sole source for reporting and should be fed data by smaller databases that may or may not talk directly to one another. Users agreed this architecture must be mapped to show how all the systems are related and identify owners for each element. Once a warehouse is set up, it must be updated on a regular schedule and meet the deadline for Highway Performance Monitoring System (HPMS) reporting.

Valuing Data and Making Investments

The level of accuracy and detail required in data should be dictated by the importance of the data to an agency's work, which should subsequently influence the direction of data governance efforts. There are so many technological options that it is important to focus on what will drive business to invest in the right technologies. In one case, MDOT analyzed the interconnections of their databases (through the software "NeoForge"), and mapped the data shared between them. The application "Orient DB" can be used to create the same type of map for free.

Funding

Peer exchange attendees agreed agencies should seek to justify funding of data governance by looking at how much money is routinely spent on cleaning data. ODOT, for instance, analyzed the cost of data errors and cleaning, compared to a hypothetical scenario with data governance policies in place, and used this exercise to demonstrate the value of investing in data governance. Agencies can also look for synergies with other funding priorities, such as safety and emergency response, transit asset management (TAM), or transportation performance management (TPM). FHWA funds for Enterprise GIS can also be used for data governance. As a recommendation, communication materials for executives are highly useful in acquiring funding and ensuring buy-in for data governance projects.

Communications Materials

Internal documents can be designed to serve multiple communications goals such as: executive summaries for communicating to leadership the purposes and benefits of the data governance program, introductory level material to help familiarize staff with data governance, and technical information to assist in completing and operationalizing data governance processes and procedures. MDOT's communication guidebook, for example, contains a business glossary and an executive-level summary of data governance, what the strategy is, and the benefits that will accrue from this work. This document

helps explain the complexity of the interrelated databases and the ways that creating a central repository will enable better data use.

Starting the Data Governance Process

The first step toward data governance should be conducting a capability maturity model (CMM)-style assessment with data owners and users. These staff need to be well versed in GIS, but also open to change and hearing negative feedback. This group should also help to set realistic targets for development, and make sure that as the assessment is performed, the standards for each level of performance and the tangible actions and resources to move the data through those performance levels are defined. It is also beneficial to have a designated “division champion” to demonstrate the value and cost plus time savings of data governance processes. A lead liaison position between top executives and business units can also help drive data governance forward.

When implementing data governance policies, hands-on opportunities for workshops are very beneficial for an agency. It can be difficult to find time in day-to-day jobs for self-directed training and reading. Workshops can also yield insights on what might best work for an organization’s individual needs, providing insight into different strategies and plans.

Roundtable: Data Governance Staffing and Strategy

Putting the appropriate staffing structure in place will define whether or not a data governance strategy is successful. Agencies may need to create a new job series, new positions in their current structure, or new staff with an IT background (the challenges of having mixed departments makes it difficult to supervise GIS work when managers do not have that background). Agencies need formalized staff in positions where they can take charge of data governance and data QA/QC. They must be given authority to make changes in policy and the required resources to be effective. Setting the correct pay scale to attract applicants with the required skillsets is also important.

Some agencies have experienced conflicts between their IT and GIS departments, with the IT department setting up system architectures before the GIS department reviews them. Contracts with Esri can be used for free in-person staff development plans, technical assistance activities, and staff trainings. Esri trainings can help IT staff understand the architecture of the agency’s GIS system. This will facilitate getting IT and GIS staff on the same page and engaged with data governance efforts. A process for grading the collaboration between different departments can then be used to assess gaps.

Concluding Roundtable Discussion: What Can FHWA Do?

Participants reported they would like to see formal guidelines, a guidebook, and/or user group website related to data governance and data management. Participants also reported that having guidance on providing justification for data governance to State DOT leadership and providing direct FHWA communications on the importance of data governance would be extremely helpful. This could include fact sheets that can concisely show the benefits, and communicate the importance of, data governance. Participants also noted they would like to see a webinar on this topic, and guidance on how to form working groups.

Appendix A: Sample List of Resources as Identified By Peers

- A Data Governance Conference is held in San Diego every year (<http://www.debtechint.com/dgiq2018/>)
- Caltrans Data Governance Manual (Used by MDOT)
- OrientDB—free open-source NoSQL multi-model graphing database
- Neo4j—online graphing platform
- John Ladley’s book “Data Governance”—includes data governance ideas and implementation guides
- Tools for monitoring data governance progress and data integrity:
 - The Urban and Regional Information Systems Association (URISA) CMM
 - Graphic Information Systems Technology CMM
 - Business Intelligence (BI) tools
 - Power BI
 - Tableau
 - Hexagon
 - Insight (Esri)
- Feature Manipulation Engine (FME)—simplified software product for extracting, loading, and moving spatial data without prior coding knowledge or skills.
- InfoSphere—manages database so you can enter a data model and manage the connection to business terms. It allows developers to update a data model and track those updates.
- American Association of State Highway and Transportation Officials’ “The 7 Principles of Data”

Appendix B: Peer Exchange Agenda

Monday, September 17, 2018

Arrive to Little Rock

7:00 PM Meet and Greet (likely in/near the hotel) – optional

Tuesday, September 18, 2018

7:00 – 8:00 Breakfast – Hotel

8:00 Leave for ARDOT – Meet in the lobby

8:30 – 9:00 Welcome and Introductions, Overview of Purpose and Need – FHWA

9:00 – 9:30 Presentation of Data Governance and Data Maintenance case study (Volpe)

9:30 – 10:00 Presentation from Arkansas DOT

10:00 – 10:30 Presentation from Ohio DOT

Break

10:45 – 11:15 Presentation from Iowa DOT

11:15 – 11:45 Presentation from Connecticut DOT

11:45 – 12:15 Presentation from Michigan DOT

Lunch

1:30 – 2:00 Presentation from Tennessee DOT

2:00 – 2:30 Presentation from Alaska DOT

Break

2:45 – 3:15 Presentation from Nevada DOT

3:15 – 3:45 Roundtable Discussion: Policy and Data Governance

3:45 – 4:00 Day 1 Wrap-up - FHWA

6:30 Informal Dinner

Wednesday, September 19, 2018

7:00 – 8:00 Continental Breakfast and Check-out – Hotel

8:00 Leave for ARDOT – Meet in the lobby

8:30 – 8:45 Day 1 Recap – FHWA

8:45 – 9:45 Roundtable Discussion: Technical Tools for Data Governance

Break

10:00 – 11:00 Roundtable Discussion: Data Governance Staffing and Strategy

11:00 – 11:30 Roundtable Discussion: What Can FHWA Do? – All Participants

11:30 – 12:00 Day 2 Key Points / Wrap-up / Adjourn – FHWA

12:00 Optional Tour of ARDOT TMC