STATE DEPARTMENTS OF TRANSPORTATION'S VISION TOWARD DIGITAL TWINS: INVESTIGATION OF ROADSIDE ASSET DATA MANAGEMENT CURRENT PRACTICES AND FUTURE REQUIREMENTS

Ashtarout Ammar[1 *], Hala Nassereddine[1], Gabriel Dadi[1]

[1]Department of Civil Engineering University of Kentucky, Lexington, Kentucky, USA (ashtarout.ammar@uky.edu; hala.nassereddine@uky.edu; gabe.dadi@uky.edu)

KEYWORDS: Digital Twins, Departments of Transportation, Transportation Asset Management, Roadside Assets, Data Management, Data Collection.

ABSTRACT:

Transportation Asset Management (TAM) is a data-driven decision-making process to maintain and extend the serviceability of transportation assets throughout their lifecycle. TAM is an extensive data process that requires accurate and high-quality information for better decision-making. A significant challenge faced by state Departments of Transportation (DOTs) is the need to allocate their limited funds to optimize their assets' performance. The criticality of this challenge increases when state DOTs need to manage a wide variety of assets distributed along with a vast network. To address this challenge, a new paradigm of digitizing the management of the built environment is emerging and is perceived to highly depend on the integration of several technologies namely on Digital Twins. Digital Twins, by definition, are the connection between the physical and the digital aspects of an asset, thus, aligning with the overarching objective of asset management of leveraging the use of the asset information (i.e., digital aspect of the asset) to improve the asset's performance throughout its lifecycle (i.e., physical aspect of the asset). At the core of implementing Digital Twins is having the right data collected for use throughout the lifecycle of the asset. Thus, realizing the potentials of Digital Twins in supporting state DOTs to manage their transportation assets and the anticipated benefits, this paper investigated the current practices of state DOTs in digitizing the Data Collection for Roadside Asset Systems by developing and distributing a web-based survey. Five major Data Collection variables and seven Roadside Asset Systems were considered. Furthermore, this paper presents a case study from a leading DOT in digitizing the management of the built environment to further understand the requirements of implementing Digital Twins to support transportation asset data management.

1. INTRODUCTION

1.1 Transportation Asset Data Management

The American Association of State Highway Transportation Officials (AASHTO) defines Transportation Asset Management (TAM) as "a strategic, systematic process of operating, maintaining, upgrading, and expanding physical assets effectively throughout their lifecycle. It focuses on business and engineering practices for resource allocation and utilization, with the objective of better decision making based upon quality information and well-defined objectives." (AASHTO TAM Guide, 2021).

In the United States (US), state Departments of Transportation (DOTs) are responsible for managing and maintaining transportation assets. State DOTs usually prioritize the management of high-value and high-visible assets such as pavements and bridges. However, transportation systems extend beyond pavement and bridges to include a wide variety of ancillary assets (or auxiliary assets) such as lighting structures, roadway signs, pavement markings, guardrails, and technology hardware equipment, among others (AASHTO TAM Guide, 2021). Typically, state DOTs develop Transportation Asset Management Plans (TAMPs) to act as the key source for information about the assets, the management strategies, long-term expenses forecasts, and business management processes. TAMPs are considered necessary management tools to connect all related business stakeholders and business processes to better understand and commit to improving assets' performance (FHWA, 2019).

Managing highway assets requires different data input, thus, state DOTs collect, store, manage, and analyze vast amounts of data to support TAMs. Consequently, TAM is an extensive data process and requires a data-driven decision-making process to maintain, preserve, and extend the long-term service life of transportation assets (Yuan et al., 2017).

Every year state DOTs conduct hundreds of projects, such as creating new roads, bridges, signs, guardrails, and making changes to existing ones. New constructions require the generation of new sets of data to be included in the current database. On the other hand, reconstruction, rehabilitation, asset demolition, or other major maintenance activities requires revising and updating the database. All changes conducted for assets through the project execution and its whole lifecycle should be collected accurately and promptly to ensure effective TAM and proper operation and maintenance (O&M) (Le et al., 2018). Asset data represents the end product shared and contributed by different stakeholders and becomes a vital component of any infrastructure management practice (Le et al., 2018).

Le et al. (2018) created a "cradle to grave" life of a transportation asset where an asset's life initially starts with a new construction project and then passes through several rehabilitation and renewal projects before its deterioration with time. The authors highlighted that construction activities are necessary within different periods to maintain the asset's condition to a particular desired level of service. As a result of ongoing construction projects, the data flow is positively correlated with the up-to-date revision versions of the asset's current state that is required to track the history of the asset inventory and to evaluate the asset performance. Usually, the asset management team will receive

* Corresponding author

the updated data related to the asset conditions, and will, in turn, continue to update the existing database at the end of the construction project. Therefore, each asset will have several revision versions resulting in massive amounts of data that state DOTs should handle. Thus, this demands proper data handling and data flow between asset management teams and the maintenance teams (Le et al., 2018).

Moreover, the challenges associated with asset data collection and management and resource allocation faced by state DOTs are amplified by the deterioration of the transportation system, limited funds, and the increasing demand for a user-oriented performance system. Thus, in an attempt to surpass these challenges, the Architecture, Engineering, Construction, and Operation (AECO) sector is undergoing a digital transformation toward adopting new approaches which are affecting their ways of managing the business. Thus, a new paradigm of digitizing the management of the built environment is emerging (Saxon et al., 2018). Among the technologies encompassed in this new paradigm is Digital Twinning, where a digital version of the physical asset is created in the form of information or processes and used in software platforms throughout the asset lifecycle. This digital environment allows for the integration between the digital and the physical asset systems (Saxon et al., 2018).

Therefore, DOTs can leverage the use of data to monitor the efficiency of the implemented maintenance approaches, identify performance metrics, and determine possible improvements. Also, specific data related to the asset's conditions can be collected to prioritize repairs, finances, and resource determination (Allen et al., 2019). In general, the cost of asset management systems can be high, however, collecting data still pertains to the most increased expenditure during any system's lifespan (Allen et al., 2019). Therefore, state DOTs must integrate technologies and management systems to change the culture of the organizations toward transforming data into accurate and reliable information that can be translated into proactive and actionable decisions (Ammar et al., 2022).

1.2 Digital Twins for Asset Management

Asset management can be further explained as the task of connecting the fundamental mission of an organization of operating the infrastructure i.e., connecting the digital aspect of the asset and its physical aspect to ensure better asset operation and maintenance and support decision making (Garramone et al., 2020). The concept of asset management of connecting the physical and digital aspects of assets intersects with the concept of Digital Twins for infrastructure. The Center for Digital Built Britain (CDBB) defines Digital Twins as "a realistic digital representation of assets, processes or systems in the built or natural environment. What distinguishes a digital twin from any other digital model or replica is its connection to its physical twin" (Bolton et al., 2018). Moreover some practitioners define it as "a static representation of the physical asset at the design phase, later construction telemetry, Geographic Information System (GIS) and sensor data might be integrated to make a digital twin a virtual simulation of a physical asset" (Broo and Schooling, 2021), or that Digital Twins are "a realistic digital representation of physical assets, processes, and systems", it should represent something "real" and should be connected bi-directionally to the asset, process or system of assets (Callcut et al., 2021).

Digital Twins represent an innovative solution to manage assets throughout their lifecycle by collecting and integrating asset data to improve the design, construction, operation, and maintenance of infrastructure assets resulting in the sustainable development of infrastructure systems (Broo and Schooling, 2021; Chen et al., 2021).

The potential of Digital Twins in leveraging the value of asset data and making use of information to support decision-making attracted the attention of several researchers, organizations, and practitioners. Chen et al. (2021) conducted a comprehensive review of the implementation of Digital Twins in asset management and most of the selected cases included the O&M of complex construction projects e.g. university campus buildings. Moreover, Callcut et al. (2021) reported on some identified use cases of Digital Twins in the transportation system including the sectors of railroads, highway and autonomous vehicles, and bridges. Furthermore, Highway England expected that by the end of 2035 live as-built Digital Twins will become standard for the construction and maintenance of highways (Highways England, 2020).

The concept of Digital Twins is emerging in the construction industry and its benefits are started to be realized especially in the asset management sector. The highway asset system within a state DOT can be classified into three categories: bridges, pavements, and ancillary assets. For instance, for the fiscal year of 2021, Utah DOT was responsible for managing a transportation system with a total value of $51 billion where bridges amounted to $10 billion (19.6%), pavement accounted for $30 billion (58.8%), and ancillary assets totalled $11 billion (21.6%). Several researchers investigated the implementation of Digital Twins for bridge management (Dang et al., 2018; Shim et al., 2019; Kaewunruen et al., 2021; Kang et al. 2021; Zhao et al. 2022) and for pavement management and performance (Steyn 2020; Yu et al., 2020; Fox-Ivey et al., 2021; Steyn and Broekman, 2021). However, and even though ancillary assets pertain to a great monetary value of a highway transportation system and most assets are related to highway safety, to our knowledge, no study has yet investigated the implementation of Digital Twins for the management of ancillary transportation assets in the US. Additionally, studying this asset category is critical because creating a holistic Digital Twins for the highway transportation system all assets must be considered. Therefore, the overarching aim of this study is to fill the gap of knowledge by investigating the status quo of state DOTs in managing their ancillary transportation asset data by focusing on their current practices that are related to data collection, and data management and governance. Three types of ancillary asset systems were considered: Roadside Asset Systems (e.g. sidewalks, bike paths, pavement markings, etc.), Drainage Structure Systems (e.g., drain inlets and outlets, curb and gutter, pipes, etc.), and Electronic Systems (e.g. Information Technology System (ITS) equipment, roadway lighting, traffic control signals, etc.). As this paper is part of this ongoing research effort, the scope of this paper will focus on studying state DOTs' current practices for Data Collection of the Roadside Asset Systems. Furthermore, this study investigates the state DOTs' requirements toward digitizing the management of the Roadside Asset Systems' data collection and their vision toward implementing Digital Twins for transportation systems.

2. METHODOLOGY

This research paper is part of an ongoing effort to investigate innovative solutions by integrating different technologies to support state DOTs in managing their ancillary transportation asset data. The first step toward achieving the objective of this research is to investigate and document the status quo of state DOTs and their current practices in managing ancillary transportation asset data. To fulfill the goal of this paper, a web-based survey was developed and approved by the Office of Research Integrity before distribution. The survey was then distributed to the AASHTO maintenance transportation committee members and the Federal Highway Administration

(FHWA) asset management contact list. A total of 29 complete responses were received from the state DOTs. Complete responses received from state DOTs are represented in Figure 1. Responses were mainly collected from the DOT's asset management division (a total of 20 complete responses) and the operation and maintenance division (a total of 9 complete responses). Responses from the asset management division included responses from the office of asset management administration, asset management engineers, asset performance measures, and asset data management analysts. Responses from the asset operation and maintenance division included responses from the office of asset maintenance management administration and services, highway maintenance engineers, and roadway asset operations.

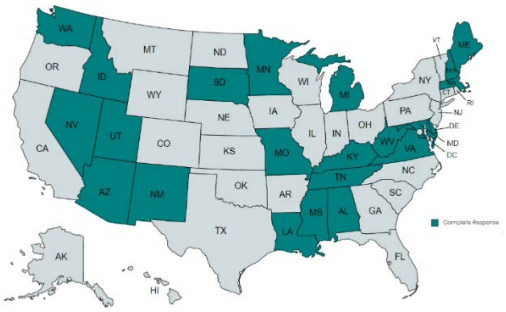

Figure 1. Complete responses from state DOTs.

As the scope of this paper is to document the current practices of state DOTs for managing the data of Roadside Asset Systems and to focus primarily on the practices related to Asset Data Collection, the various assets encompassed under Roadside Asset Systems need to be identified. Roadside Asset Systems mainly include the following assets: sidewalks, roadside assets, fence, turf, brush control, roadside hazard, landscaping, access ramps, bike paths, signs, guardrail, guard rail end treatments, impact attenuator, other barrier systems, and pavement markings (NHI, 2017). The list of these assets was provided in the survey and respondents were asked to select all the assets included in the Roadside Asset System that their agency usually collects data for. The total count/asset is represented in Figure 2.

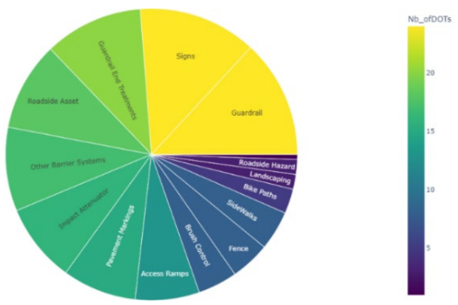

Figure 2. Count of state DOTs that collect data for the assets included in the Roadside Asset System.

For better data analysis, assets with more than 15 data points (i.e., assets that 15 or more state DOTs replied that they collect data for) were selected as the scope for this paper. The selected assets, sorted alphabetically, and the number of state DOTs that collect data for the selected assets are summarized in Table 1.

Roadside Asset Systems	**Nb. of DOTs**
Guardrail (G)	24
Guardrail End Treatment (GET)	20
Impact Attenuator (IA)	16
Other Barrier Systems (OBS)	17
Pavement Markings (PM)	15
Roadside Asset (RA)	18
Signs (S)	24

Table 1. Selected Roadside Asset Systems.

To properly depict how state DOTs manage their transportation asset Data Collection, five major Data Collection categories (i.e., variables) were investigated and each category can be described by three to five subcategories (i.e., variable levels). The first variable covered the format of data collection/asset (i.e., what is the data format in which the asset data is collected) and the associated levels are: 1) Paper-Based, 2) PDF, 3) Smart PDF, 4) 2D Models, and 5) 3D Models. The second variable investigated the required level of detail of the collected data/asset and the associated levels are: 1) Not Available (NA), 2) L4, 3) L3, 4) L2, and 5) L1. The description of the level of details was adopted from the FHWA report "Asset Management Data Collection for Supporting Decision Processes" (Flintsh and Bryant, 2009). It was noted in the survey that L1 represents the most comprehensive level of detail, L2: level of detail sufficient for comprehensive analysis, L3: Sufficient details to conduct elementary methods of maintenance, and L4: basic details with summary statistics of inventory, performance, and utilization. The third variable was concerned with the project phase during which the state DOT usually collects data for each of the selected assets and the associated levels are: 1) during the Design phase, 2) during the Construction phase, 3) during Project Closeout (i.e., as-built), and 4) during the Maintenance phase. The fourth variable focused on the features used by the state DOTs to inventory the asset and its associated levels are: 1) Location (i.e., indexed location of the asset), 2) Dimension, 3) Material Type and Properties, and 4) Asset Condition. The final variable included the techniques used by the state DOTs to collect asset data and the associated levels are: 1) Manual Collection, 2) Automated Collection, and 3) Remote Collection. The description of the levels was adopted from FHWA (2020) where the manual collection can use handheld computers, global positioning system (GPS) units, or pen-and-paper records; automated collection involves a vehicle equipped with technologies to identify and document transportation assets; and remote collection uses photo logs, video logs, and satellite images.

Once data was collected, descriptive and statistical analysis were employed to further investigate state DOT's existing Data Collection practices of Roadside Asset Systems data. Next, a case study was conducted with a selected DOT that is on its journey toward implementing Digital Twins to investigate the requirements of transitioning toward a fully digitized ancillary asset Data Collection and the roadmap to implement Digital Twins for the highway transportation system.

3. SURVEY RESULTS AND DISCUSSION

This section describes the results obtained from the 29 valid responses collected from the web-based survey. The analysis mainly focuses on the current practices of state DOTs that are related to Roadside Asset Systems Data Collection. Major data collection variables, the type of the variable, and the associated levels and codes (for ordinal variables) are summarized in Table 2.

Variables	**Variable Type**	**Variable Level and (Code)**
Data Format	Ordinal	Paper-Based (1), PDF (2), Smart PDF (3), 2D Models (4), 3D Models (5)
Data Level of Detail	Ordinal	NA (1), L4 (2), L3 (3), L2 (4), L1 (5)

Data Collection/ Construction Phase	Categorical	Design, Construction, Project Closeout (as-built), Maintenance
Data Features	Categorical	Location (i.e., indexed location of the asset), Dimension, Material Type and Properties, Asset Condition
Data Collection Techniques	Categorical	Manual Collection, Automated Collection, Remote Collection

Table 2. Roadside Asset System Data Collection major variables and variable levels.

For each variable level, the responses per asset were aggregated and the relative percentages were calculated. The aggregated percentages were then stacked and clustered into the five major variables of Data Collection. The stacked and clustered results are presented in Figure 3.
Two types of data analysis were conducted, descriptive data analysis and statistical data analysis.

3.1 Descriptive Data Analysis

Data Format: for all the selected Roadside Asset Systems 2D Models were the major data format used by the surveyed state DOTs with a usage range varying between 42% and 52% of all the adopted data formats. 2D Models were then followed by Paper-Based with a usage range varying between 17% and 27%, then by PDF data format with a usage range varying between 13.8% and 14.3%. Conversely, 3D Models were used for certain assets with relatively low percentages (equal or slightly higher than the usage of Smart-PDF). It is worth noting that for the roadside asset, no DOT reported the usage of 3D Models. Moreover, Smart-PDF was the least data format used by the surveyed state DOTs.

Data Level of Detail: for guardrail, guardrail end treatment, other barrier systems, pavement markings, and roadside asset the surveyed state DOTs mainly adopt a level of detail L3 (i.e., sufficient details to conduct elementary methods of maintenance) while for signs they mainly adopt a level of detail L2 (i.e., level of detail sufficient for comprehensive analysis). For the impact attenuator, the relative percentages for L2 and L3 were equal. For guardrail, other barrier systems, and signs 20% to 22% of the adopted level of details refer to L1 (i.e., the most comprehensive level of detail). It is also worth noting that some state DOTs reported that they don't have a required level of detail for the selected Roadside Asset Systems, the corresponding relative percentages are represented as NA in Figure 2.

Data Collection/Construction Phase: for all the selected Roadside Asset Systems, it was found that all surveyed state DOTs mainly collect the asset data during the Maintenance phase with a relative percentage ranging between 40% and 74%. The Maintenance phase was followed by Project Closeout (as-built) with relative percentages varying between 21% and 35%. The relative percentages for data collection during the Design of the asset were slightly higher than those during the Construction of the asset except for guardrail end treatments. For other barrier systems and pavement markings, none of the surveyed state DOTs reported that they collect data for those assets during the phase of the Project Closeout (as-built).

Data Collected Features: for all selected Roadside Asset Systems, the data feature that is most collected by the surveyed state DOTs was the Indexed Location with a relative percentage varying between 29% and 33% of the collected data features. The relative percentages for the asset's Dimension and the asset's Material Type and Properties were relatively close with the former slightly higher for most of the selected assets. Moreover, the Asset Condition was the least feature collected by the surveyed DOTs with a corresponding relative percentage varying between 12% and 21% of the collected asset data features.

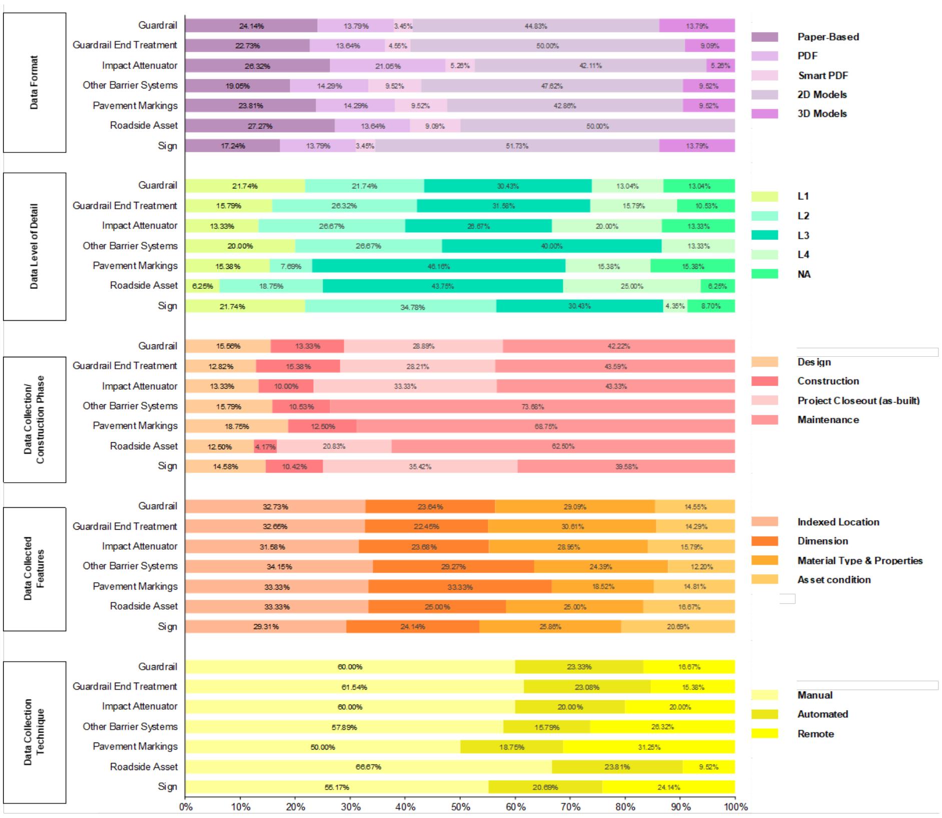

Figure 3. Stacked and clustered results for major Data Collection variables and variable levels.

Data Collection Techniques: it was found that for all selected Roadside Asset Systems the surveyed state DOTs mainly collect asset data manually where the corresponding relative percentages for Manual Data Collection ranged between 50% and 67% of the adopted data collection techniques. For the assets, guardrail, guardrail end treatment, and roadside asset the relative percentages for Automated Data Collection were higher than that for Remote data collection. However, for the assets of other barrier systems, pavement markings, and signs the relative percentages for Remote Data Collection were higher than the corresponding relative percentage for Automated Data Collection, and for impact attenuator the two relative percentages were equal.

3.2 Statistical Data Analysis

To detect the existence of any statistical significance in the data collection practices across all assets and within each asset individually, further statistical analysis was conducted. For Ordinal variables, i.e., Data Format and Data Level of Detail, the plotting of the different variable levels across the seven assets under consideration is represented in Figures 4, and 5 respectively.

Figure 4 shows that for all selected assets, at 95% confidence level, there is no statistical significance between the usage of different Data Formats across all assets, and this was statistically validated by using the Kruskal-Wallis test with a resulting p-value of 0.8748 (greater than 0.05).

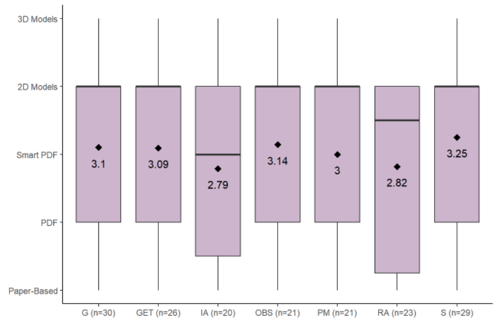

Figure 4. Comparing state DOTs' usage of different Data Formats across the different assets of the Roadside Asset Systems.

Moreover, to investigate the level of Data Format digitization for individual assets, Wilcoxon Test was conducted as shown in Table 3. Assuming that beginning to use Smart PDF is a step toward digitizing asset data format, the Null Hypothesis H_0:mu_0=2 and the Alternative Hypothesis H1=mu_1>2 were tested. As shown in Table 3, all assets resulted in a significant p-value at the 95% confidence level, thus failing to support the null hypothesis and indicating that for all of the seven Roadside Asset Systems, state DOTs are heading toward digitizing the Data Format.

Roadside Assets	P-Value	Sig. at 95% Confidence Level
Guardrail	0.0001676	Significant
Guardrail End Treatment	0.0007308	Significant
Impact Attenuator	0.007246	Significant
Other Barrier Systems	0.0007662	Significant
Pavement Markings	0.001913	Significant
Roadside Asset	0.002524	Significant
Signs	4.966e-05	Significant

Table 3. Results of Wilcoxon Test for Data Format usage across the different assets of the Roadside Asset Systems.

Figure 5 shows that for all assets there is no statistical significance at 95% confidence level, between the usage of different Data Levels of Details across all assets. This was statistically validated by using the Kruskal-Wallis test with a resulting p-value=0.4817 (greater than 0.05).

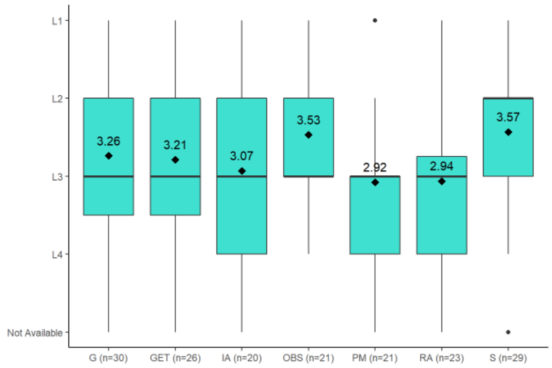

Figure 5. Comparing state DOTs' usage of different Data Levels of Detail across the different assets of the Roadside Asset Systems.

Moreover, to investigate the level of data quality represented by how comprehensive the Data Level of Detail is collected by state DOTs for individual assets, Wilcoxon Test was conducted as shown in Table 4. Assuming that starting to use L3 is a step toward improving the quality of collected data and therefore increasing the ability to conduct a comprehensive analysis, the Null Hypothesis H_0:mu_0=3 and the Alternative Hypothesis H1=mu_1>3 were tested. As shown in Table 4, other barrier systems and signs are the only assets that resulted in a significant p-value at 95% confidence level, and therefore for all the other assets the quality of collected data is still at the level of detail L3 which is described as including sufficient details to conduct elementary methods of maintenance.

Roadside Assets	P-Value	Sig. at 95% Confidence Level
Guardrail	0.1899	Not Significant
Guardrail End Treatment	0.2463	Not Significant
Impact Attenuator	0.4451	Not Significant
Other Barrier Systems	0.0328	Significant
Pavement Markings	0.6026	Not Significant
Roadside Asset	0.6226	Not Significant
Signs	0.02586	Significant

Table 4. Results of Wilcoxon Test for Data Level of Detail usage across the different assets of the Roadside Asset Systems.

For the three remaining categorical variables, i.e., Data Collection/Construction Phase, Data Features, and Data Collection Technique, the results of the analysis are presented in Figures 6, 7, and 8 respectively. Given the nature of these variables, proportion was used as the unit of analysis and the proportion of the different variable levels was calculated. The Standard Error (SE) (equation1) was then obtained to present the variability in estimating the proportions and to compute the 95% Confidence Interval (CI) (equation 2).

$$\text{Standard Error (SE)} = \sqrt{\frac{p(1-p)}{n}} \quad (1)$$

Where:

- p is the sample proportion calculated as $p=\frac{x}{n}$ with x denoting the number of successes out of a sample of size n
- n is the sample size

And the general formula of the CI is:

$$\text{Point of estimate } \pm Z_{\frac{\alpha}{2}} SE \quad (2)$$

Where:

- point of estimate is the sample proportion p
- $Z_{\frac{\alpha}{2}}$ is the z-score

From the standard normal distribution, for 95% CI, α equals 0.05 and therefore $Z_{\frac{\alpha}{2}}$ =1.96 and thus the 95% CI is obtained by calculating p$\pm 1.96SE$.
Once the CIs were computed, they were visually represented using error bars as shown in Figures 6, 7, and 8 to represent the variability of the corresponding data.
Figure 6 indicates that the CIs for the Maintenance phase for the assets of guardrail, guardrail end treatment, impact attenuator, and signs overlap with the phase of Project Closeout (as-built) indicating that state DOTs collect data mainly during the Maintenance phase and the Project Closeout phase. While for the assets of other barrier systems, pavement markings, and roadside assets the CIs for the Maintenance phase doesn't overlap with any of the other phases indicating that state DOTs mainly collect data during the Maintenance phase, highlighting the dominance of this phase for these assets.

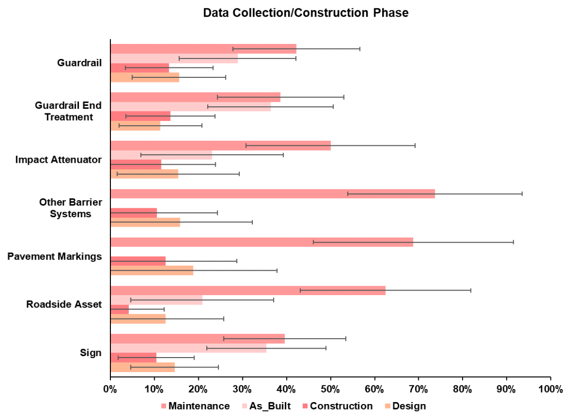

Figure 6. Breakdown of state DOTs' Data Collection across the different Project Phases for the different assets of the Roadside Asset Systems.

Figure 7 indicates that the CIs for all Data Features across all assets overlap, indicating that state DOTs collect data on the asset indexed location, asset dimension, asset material type and properties, and asset condition for all seven assets similarly.

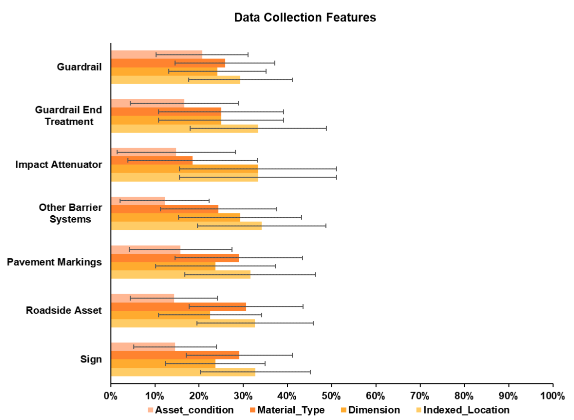

Figure 7. Breakdown of state DOTs' Data Collection Features for the different assets of the Roadside Asset Systems.

Figure 8 indicates that the CIs for Manual asset data collection for all assets except for impact attenuator and other barrier systems don't overlap with other techniques, indicating that for all the selected roadside assets except impact attenuator and other barrier systems, Manual data collection is the dominant data technique adopted by state DOTs.

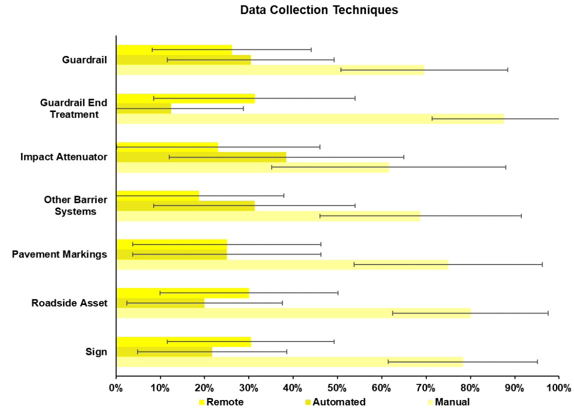

Figure 8. Breakdown of state DOTs Data Collection Techniques used for the different assets of the Roadside Asset Systems.

4. DIGITAL TWIN VISION: A CASE STUDY FROM UTAH DEPARTMENT OF TRANSPORTATION

In addition to understanding existing practices associated with the Data Collection of Roadside Asset Systems, this study aimed to investigate the requirements of Digital Twins in supporting transportation asset data management from the perception of a state DOT that is well rounded with the concept and already established a strategic plan for Digital Twins implementation. To achieve this, an interview was conducted with the state-wide asset data managers of the Utah Department of Transportation (UDOT). UDOT has recently established a vision for Digital Twins as an information management strategy to connect the enterprise asset information to a geospatial model of the individual physical assets. Digital Twins is envisioned to support the documentation of the planned and as-constructed (as-built) updates and therefore to fill the gap in the information. Digital Twins are perceived to enable UDOT to conduct complex and holistic data analyses and support comprehensive decision-making to improve highway safety, optimize mobility, and maintain transportation infrastructure.
The guiding vision for an ideal Digital Twins at UDOT is based on two critical elements that are necessary for the development of Digital Twins. The first element is the automation of the collection of the enterprise component of transactional data, and the second element is the development of a data governance framework that ensures one true source of reliable and relevant information.
UDOT defines Digital Twin as a "digital representation of a physical asset that contains a geometric model of the asset as well as information about the asset such as its properties, functions, evaluative properties, and other analytical context. Digital Twins are scalable, adding geometric and contextual detail as needed to support the business uses". The DOT's journey toward the development of a digital project delivery started by investing in Geospatial Information Systems (GIS) and the use of LiDAR technology to collect mobile LiDAR data and integrate it with GIS-based asset registries. The asset information database is connected to other databases including construction costs, safety, and traffic volumes. With the available digital delivery tools, UDOT is capable of producing Digital Twins of assets, and in the

spring of 2021, the DOT received an Accelerated Innovation Deployment (AID) grant from FHWA to advance its digital construction by digitizing data capturing.
Since the focus of this paper is on Roadside Asset Systems Data Collection, interviewees were asked to share their perspective on the following for each of the five Data Collection Variables:

- Reflecting on the survey findings, what are your thoughts on where state DOTs stand nationally?
- From UDOT's experience, what are the Digital Twin requirements for each variable/asset?

The interviewees' feedback on the different variables is reported next.
Data Format: The resulting average of the selected assets of Roadside Asset Systems on a national level was not surprising (see figure 4). However, at UDOT, the interviewees noted that their agency has retired the use of smart PDF and is using data formats that enable the DOT to extract information that can be pushed to its databases. As such, the DOT is heading toward creating 2D models and 3D models for their Roadside Asset Systems. Moreover, in their efforts to automate data collection, and depending on the used techniques and the collected data attributes, the DOT can use 3D models on certain projects. Currently, the agency is piloting Digital Twins on certain projects, and eventually, all assets will be represented by 3D models. Establishing 3D models for assets should be created by the project designers, however, creating3D models for the whole state is very expensive compared to the earned value, thus before reaching a holistic 3D model of the asset system intermediate steps should be taken by selecting certain assets to be represented in 3D models and the rest in 2D models. In the long term, UDOT is heading toward representing all assets in 3D models as a requirement for creating Digital Twins for the transportation asset system.
Data Level of Details: The average for the selected assets of Roadside Asset Systems on a national level is reasonable and the results align with the current state at UDOT where the data level of details is between L3 and L2. However, the interviewees noted that the resulting level of detail is highly correlated with the method of data collection and with the classification system that UDOT adopts for its asset management. For instance, UDOT has detailed information for signs and barrier systems compared to other selected assets and this aligns with the relatively high resulting averages (see figure 5). However, UDOT is pushing toward reaching a level of detail L1 for the assets that it classifies as Tier 1 i.e., assets that require accurate and sophisticated data collection to conduct predictive modeling and risk analysis, especially for those that are related to traffic and safety such as signal devices. Moreover, for other assets that are classified as Tier 2 and don't require sophisticated data collection but need information about the asset condition that is necessary to conduct a risk assessment based on the asset failure, a level of detail between L3 and L2 is sufficient. Such assets include signs, barriers, and pavement markings. Finally, for assets that are classified as Tier 3, such as bike lanes, the needed information is used to conduct general condition analysis and can be represented with a level of detail L3.
Data Collection/Construction Phase: It is reasonable that state DOTs collect asset data mainly during the maintenance phase since asset management is based on the data collected by the maintenance crews. Currently, state DOTs are using the data collected during maintenance to create Digital Twins, however, creating a Digital Twin for an asset should start at the construction/as-built phases. While certain types of data (e.g. standards) can be extracted from the design phase, populating the full database with the design information was found to be ineffective. At UDOT the main focus is to pull data from the construction phase and the as-built asset condition to create the asset Digital Twin. Moreover, the collected as-built asset data will allow for the generation of databases, and the asset data collected during the maintenance phase will continue to flow into the databases throughout the remaining stages of the asset lifecycle (i.e., rehabilitation and renovation, demolition, and removal) and will keep the asset's Digital Twin updated.
Data Collection Features: To create a Digital Twin of an asset, it is crucial to define the attributes and features that should be collected per asset. For instance, the indexed location is a data feature that should be collected for all assets. For certain assets, such as pavement markings, dimension is a feature that is important more than other asset features, while the material type is critical for certain assets such as impact attenuator. The major gap in the collected data features is in the asset condition and this pertains to the lack of standards and the absence of a clear classification system that describes different asset conditions (e.g. what good, fair, and poor entails/asset). Additionally, for some assets, the condition is not important since the asset doesn't deteriorate before it becomes obsolete, and the decisions for those assets are based on the material type and age of the system instead of the asset condition. As such, it is essential to define for each asset the data attributes and features that should be collected and integrated into the asset's Digital Twin that would enable running cost-benefit analysis and risk analysis and generating deterioration curves to support the asset management and decision-making.
Data Collection Techniques: The data collection technique is highly correlated with all of the above-mentioned variables and varies from one asset to another. For instance, for pavement markings and signs, most state DOTs are using vehicles that can collect retrospective reflectivity by adopting automated techniques. As such, for some assets, automated techniques can be adopted and are efficient; however, for other assets, manual data collection will continue to be considered, especially for assets where better quality and accurate information is required. Moreover, automated data collection is usually scheduled every two years, and depending on the used technology, the collected data might not include all the necessary information such as the asset type. To ensure accurate and reliable information that will keep the asset's Digital Twin updated, the future vision is to adopt a combination of manual and automated data collection, especially between the cycles of the scheduled automated data collection. With the available techniques, the collected level of detail is close to L3 and in the meantime, the only data collection method that will result in an L1 level of data quality is through manual data collection to support conducting comprehensive data analysis.

5. CONCLUSION AND FUTURE RESEARCH

As the benefits of the new paradigm of digitizing the built environment are becoming tangible and innovative technologies such as Digital Twins are emerging, state DOTs in the US can benefit from the flowing wave of digital transformation to surpass the challenges they face in managing their transportation asset data. This paper investigated the status quo and current practices of state DOTs that are related to digitizing the Data Collection of Roadside Asset Systems. Five major Data Collection variables: 1) Data Format, 2) Data Level of Detail, 3) Data Collection/Construction Phase, 4) Data Features, and 5) Data Collection Technique for seven assets of the Roadside Asset System (guardrail, guardrail end treatment, impact attenuator, other barrier systems, pavement markings, roadside assets, and signs) were investigated. It was found that for the selected assets, and on a national level, state DOTs are using 2D Models as the adopted data format, the level of detail of the collected data is between L3 (i.e., sufficient details to conduct elementary methods of maintenance) and L2 (i.e., level of detail sufficient for comprehensive analysis), data is mainly being collected during the Maintenance phase, the asset's Indexed Location is the data feature primarily collected for all assets and the Asset Condition is the least data feature collected, and Manual data collection is the most used technique to collect asset data.

This paper also showcased the requirements for state DOTs to implement Digital Twins for asset Data Collection by offering a case study that discussed the Digital Twin vision of UDOT, a department of transportation that is considered one of the leading DOTs in digitization and whose vision is based on two elements for implementing Digital Twins for transportation asset data management: data collection and data governance.

This paper contributes to the existing body of knowledge by investigating current asset data collection practices and discussing the requirements to prepare data collection practices for implementing Digital Twins. It is important to up-skill the levels of knowledge for transportation agencies in the US and other countries by providing detailed and comprehensive information on the data requirements necessary for creating and adopting Digital Twins.

It is important to note that defining the data attributes that should be collected and ensuring integrability with the generated databases is critical and can be achieved by establishing effective and reliable data governance. A successful Digital Twin cannot be achieved by generating a plethora of databases that are overwhelming to understand and process but rather by creating databases of high-quality information that can be easily accessed and integrated with the Digital Twins and can result in a beneficial return on investment. Future research will investigate the "ideal" environment that fulfills the vision of Digital Twins and how to transform Digital Twins from concept to action, where transportation agencies can harness all the benefits and capabilities of Digital Twins to optimize the performance of existing transportation systems, and design for a resilient and sustainable infrastructure of tomorrow.

6. REFERENCES

Allen, B., Ram, P., Koonce, J., Burns, S., Zimmerman, K. A., and Mugabe, K. 2019. Handbook for Including Ancillary Assets in Transportation Asset Management Programs(No. FHWA-HIF-19-068*)*. Federal Highway Administration. Office of Research, Development, and Technology., United States, 97.

American Association of State Highway Transportation Officials (AASHTO), Transportation Asset Management (TAM) Guide. (2021). URL https://www.tamguide.com/ (accessed 4/1/22)

Ammar, A., Dadi, G., & Nassereddine, H. 2022. Transportation Asset Data Management: BIM as a Holistic Data Management Approach. In: *Construction Research Congress 2022.* Presented at the Construction Research Congress 2022, American Society of Civil Engineers, Arlington, Virginia, (pp. 208-217).

Bolton, A., Butler, L., Dabson, I., Enzer, M., Evans, M., Fenemore, T., Harradence, F., Keaney, E., Kemp, A., Luck, A., Pawsey, N., Saville, S., Schooling, J., Sharp, M., Smith, T., Tennison, J., Whyte, J., Wilson, A., and Makri, C. 2018. Gemini Principles. Apollo - University of Cambridge Repository.

Broo, D. G., and Schooling, J. 2021. Digital twins in infrastructure: definitions, current practices, challenges and strategies. *Int. J. Construct. Manage.*, 1–10.

Callcut, M., Cerceau Agliozzo, J.-P., Varga, L., and McMillan, L. 2021. Digital Twins in Civil Infrastructure Systems. *Sustainability*, 13(20), 11549.

Chen, L., Xie, X., Lu, Q., Parlikad, A. K., Pitt, M., and Yang, J. 2021. Gemini Principles-Based Digital Twin Maturity Model for Asset Management. *Sustainability*, 13(15), 8224.

Dang, N., Kang, H., Lon, S., and Shim, C. 2018. 3D Digital Twin Models for Bridge Maintenance. *Proceedings of the 10th International Conference on Short and Medium Span Bridge*, Quebec City, QC, Canada. Federal Highway Administration (FHWA). Asset Management. 2019. URL https://www.fhwa.dot.gov/asset/plans.cfm (accessed 3/1/22)

Federal Highway Administration (FHWA). Roadside Safety Hardware Identification Methods. 2020. URL https://safety.fhwa.dot.gov/roadway_dept/countermeasures/reduce_crash_severity/id_methods/ch1.cfm (accessed 5/1/22)

Flintsch, G. W., & Bryant, J. J. 2009. Asset management data collection for supporting decision processes. Federal Highway Administration.

Fox-Ivey, R., Laurent, J., and Petitclers, B. 2021. Using 3D Pavement Surveys to Create a Digital Twin of Your Runway or Highway. *Airfield and Highway Pavements*, 180–192.

Garramone, M., Moretti, N., Scaioni, M., Ellul, C., Re Cecconi, F., and Dejaco, M. C. 2020. BIM and GIS Integration for Infrastructure Asset Management: A Bibliometric Analysis. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,*, VI-4/W1-2020, 77–84.

Highways England. 2020. Connected and Autonomous Plant Roadmap to 2035. National Highways, Birmingham.

Kaewunruen, S., Sresakoolchai, J., Ma, W., and Phil-Ebosie, O. 2021. Digital Twin Aided Vulnerability Assessment and Risk-Based Maintenance Planning of Bridge Infrastructures Exposed to Extreme Conditions. *Sustainability*, 13(4), 2051.

Kang, J.-S., Chung, K., and Hong, E. J. 2021. Multimedia knowledge-based bridge health monitoring using digital twin. *Multimedia Tools and Applications*, 80(26–27), 34609–34624.

Le, T., Asce, A. M., Le, C., Jeong, H. D., Asce, A. M., and Associate, P. R. (2018). Lifecycle Data Modeling to Support Transferring Project-Oriented Data to Asset-Oriented Systems in Transportation Projects. *J. Manage. Eng.*, 34(4), 13.

National Highway Institute (NHI). 2017. Enhanced Maintenance Management Systems: Principles and Practices for Enhanced Maintenance Management Systems. National Highway Institute Course 134112. Participant Workbook. FHWA-NHI-10-102. Federal Highway Administration, Washington, DC.

Saxon, R., Robinson, K., and Winfield, M. 2018. Going Digital-A guide for construction, clients, building owners and their advise*rs*. UK BIM Alliance, UK.

Shim, C.-S., Dang, N.-S., Lon, S., and Jeon, C.-H. 2019. Development of a bridge maintenance system for prestressed concrete bridges using 3D digital twin model. *Struct. Infrastruct. Eng.*, 15(10), 1319–1332.

Steyn, W. J., and Broekman, A. 2021. Development of a Digital Twin of a Local Road Network: A Case Study. *J. Test. Eval.*, 51(1).

Steyn, W. J. van der M. 2020. Selected implications of a hyper-connected world on pavement engineering. *Int. J. Pavement Res. Technol.*, 13(6), 673–678.

Yu, G., Zhang, S., Hu, M., and Wang, Y. K. 2020. Prediction of Highway Tunnel Pavement Performance Based on Digital Twin and Multiple Time Series Stacking. *Adv. Civ. Eng.*, (J.-R. Lin, ed.), 2020, 1–21.

Yuan, C., McClure, T., Cai, H., and Dunston, P. S. 2017. Life-Cycle Approach to Collecting, Managing, and Sharing Transportation Infrastructure Asset Data. *J. Constr. Eng. Manage.*, 143(6), 04017001.

Zhao, H., Tan, C., OBrien, E. J., Zhang, B., Uddin, N., and Guo, H. 2022. Developing Digital Twins to Characterize Bridge Behavior Using Measurements Taken under Random Traffic. *J. Bridge Eng.*, 27(1), 04021101.