TECHNICAL MEMORANDUM

Summary of Feedback Gathered on Strengths and Challenges of Mobility Data Platforms

Task 1: Evaluate Mobility Datasets

SUMA
Support for Urban Mobility Analysis

A technical memorandum to

Support for Urban Mobility Analyses (SUMA)
FHWA Pooled Fund Study

Authors:
Gargi Singh
Michael Gerke
Kartikeya Jha
Vijayaraghavan Sivaraman
Ed Hard

August 2022

Submitted by the

Texas A&M Transportation Institute

Disclaimer

The views expressed in this document do not represent the views of the Texas A&M Transportation Institute and are based on the feedback gathered through interactions with the participating agencies. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration (FHWA).

Executive Summary

The increasing availability and sample penetration of passive sources of mobility data drives the need to develop a documentation of the current state of services of platforms analyzing and providing such data. Extensive knowledge of data sources, types, and their merits and constraints is needed to match the data to their correct use cases. A series of surveys and online interactions with several users of various mobility data platforms and tools were organized under this task as part of the Support for Urban Mobility Analyses (SUMA) project. The objective of this task was to help the SUMA participating agencies identify the merits and challenges of such platforms to better align their project requirements with what the data platforms deliver.

For this study, nine discussion sessions with eight participating agencies highlighted various strengths and challenges in each platform. The platforms discussed in this document are used by the participating agencies. To develop a better sense of the strengths and shortcomings of each of these platforms, the discussion sessions covered details about the platform attributes such as ease of use, data availability, preferred application, and tech support. Table 1 summarizes participant feedback from the discussion sessions.

In general, users prefer platforms for their accessibility and data management. However, this also comes with a trade-off: not all platforms provide clear documentation. This limitation might make some platforms more black box than others. Therefore, many users use the platforms primarily for their data download feature since they give users control of the analysis.

Several users remarked that the nature of the project dictates which platform is used, and agencies assess the suitability of platforms for specific applications on a case-by-case basis. The choice of platform is also governed by the application of the project or the analysis tools available on specific platforms. The level of technical support offered by the platforms varies considerably and is also factored in by agencies considering the different platforms.

This technical memorandum summarizes the user experience across platforms shared with the research team through discussions. The takeaways from these interactions are expected to help agencies make an informed choice for their future needs. While most of the challenges discussed in this memo can be addressed or have one or more solutions, there are a handful that are either by design or inherent to the data generation process. For instance, the lack of features (example: data downloader) or inadequate technical support can be resolved over time. However, other challenges such as sparse data coverage across regions and(or) at finer geographic resolutions could either be due to data anonymization, privacy agreements in place, lack of representation or due to poor coverage of technologies utilized in collection of such data. This document aims to develop awareness among users the strengths and challenges associated with such data sources and platforms, some of which might not have any resolution.

Table 1: Summary of Strength and Challenges of Platforms

PLATFORMS	STRENGTHS	CHALLENGES
StreetLight Data	- **Data options:** Multiple passive data technology (GPS and LBS) and data types (auto, transit, and demographic data). - **Transparent**: Methodology available through white papers. - **Good Technical support** **- Popular applications:** Origin-Destination (O-D) Studies, AADT	**- Lack of historical data in some areas** **- Low sample size among some data types** (e.g., Transit data) - **No data downloader feature** but data can be retained, if requested.
Moonshadow	**- Allows ingestion of big data:** Users' custom data or data from vendor through the platform - **Real-time stream** of mobility data from several data sources - **Finest resolution** compared to other data platforms, as the data is raw and not aggregated unless plotted on maps or charts. **- Popular applications:** O-D matrices**,** select link analysis	- **Platform can be slow:** Observed when using real-time data
Iteris ClearGuide	- **Real-time Data**: Additionally, historical data is available - **Data Downloader feature** for bulk data download requests. - **Good User Interface** and produces separate analysis reports for each route **- Popular applications:** Congestion contour plots – daily and monthly	- **Mass data downloader** does not download all the attributes generated on the platform or produce aggregated data for download - **Shapefile changes often** without advance notification or cross-references, makes it challenging to relate data to maps as roadway **- Volume data unavailable**
Regional Integrated Transportation Information System (RITIS)	- **Data Downloader feature:** Aggregated data available to download - **Stable and mature platform**: Flexible in defining the analysis zones - **Flexible pricing structure** **- Popular applications:** Congestion scan tool, trend maps, Map-21 widget, and data downloader	- **Road segments:** Not ideal for longitudinal analysis as historical data does not match current roadway segments. - **Vehicle weight class information** is unavailable in the probe data sample - **Mismatch between downloaded and plot data**
INRIX Roadway Analytics Tool	- **Maintains the same roadway segments** for current and historical data; therefore, a preferred tool for before and after studies. - **Shorter road segments** allow localized analyses **- Quick response user interface** because it produces raw data and saves time on aggregating data - **Popular applications:** Performance charts, bottleneck ranking	- **Instance of platform crashing:** Platform works better when using smaller road segments

Table of Contents

Introduction

Using web-based platforms to deliver emerging big passive data for mobility analysis has become commonplace among transportation professionals in recent years. These web-based platforms provide a seamless and easily accessible medium for end-users to retrieve big data in an easily digestible format. However, using these platforms also comes with the challenge of making the right decision to match an agency's mobility analysis needs with the services offered by such platforms. Thus, there is a need to gain situational awareness pertaining to platforms and their data sources, use cases, merits, and constraints.

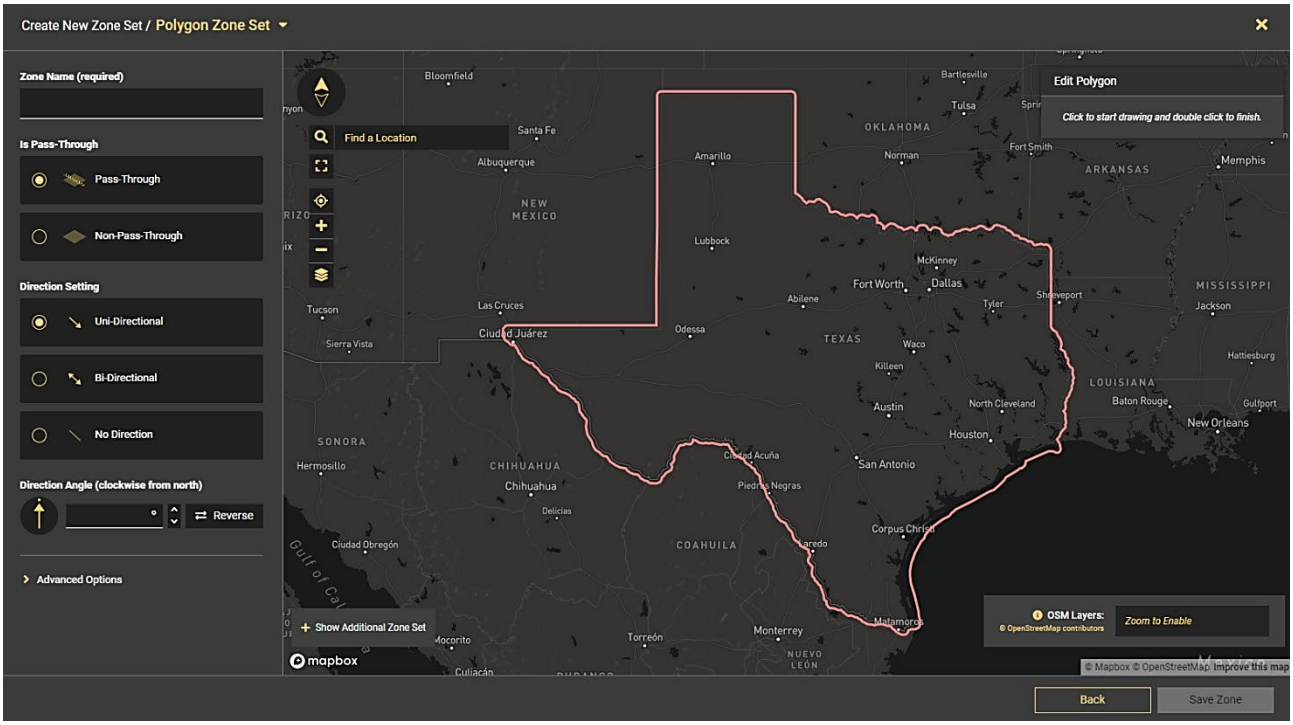
Under this Support for Urban Mobility Analyses (SUMA) task, the Texas A&M Transportation Institute (TTI) research team organized surveys and online interactions with several users of various mobility data platforms and tools. The objective was to help the SUMA participating agencies identify platforms aligned with their application needs by sharing experiences of mobility data platforms with users from other participating agencies. Nine interviews with eight participating agencies were conducted, followed by a webinar delivered virtually in January 2022 to present findings from user assessments of mobility data platforms. The webinar also provided an opportunity for all participants to discuss these data platforms collectively and share their experiences.

The process of gathering feedback from these interactions and surveys focused on various platforms and tools used by participating agencies, as described in this technical memorandum. To develop a better sense of the strengths and challenges of each platform, the discussion sessions covered details about the platform offerings, such as ease of use, data availability, preferred application, and tech support. The engagement, interaction, and feedback from sponsors throughout the user assessment process was an enriching experience. This technical memorandum describes the user experiences and observations shared by the participating agencies for different mobility data platforms.

For this study, the users' experience was of interest but not the identity of the individuals. For that reason, the feedback received from the discussion groups has been kept confidential to protect the privacy of the platform users. This confidentiality encouraged users not to be reticent in sharing their experiences.

StreetLight Data

The StreetLight Data platform (Figure 1) provides mobility data such as annual average daily traffic (AADT), origin-destination (O-D) and route information, intersection turning movement counts, vehicle miles traveled (VMT), and vehicle hours of delay as well as some inferred demographic data (i.e., income, race, education, and family status information at the neighborhood level) to users.[1] The platform also provides analytical applications through its tool catalog, such as select link analysis (i.e., trip volume to or from given O-D zones along with trip directionality and traveler attributes), top routes analysis, and data on inferred trip purpose, trip speed, travel time, and length. Data sources for StreetLight include INRIX (navigation-GPS data supplier) and Cuebiq[2,3] (location-based services [LBS] data supplier).



(Source: StreetLight Data, https://www.streetlightdata.com/)

Figure 1: StreetLight Data Platform Dashboard Interface

[1] StreetLight Data. Essential Metrics for Everyday Traffic Analysis. StreetLight Insight. https://www.streetlightdata.com/transportation-metrics.

[2] Cuebiq, Inc. StreetLight Data Partners with Cuebiq to Help Transportation, Retail and Logistics Companies Use Location-Based Data. September 20, 2016. https://www.cuebiq.com/press/streetlight-data-partners-cuebiq-help-transportation-retail-logistics-companies-use-location-based-data/.

[3] StreetLight Data. *StreetLight InSight: Our Methodology and Data Sources*. October 2018. https://www.streetlightdata.com/wp-content/uploads/StreetLight-Data_Methodology-and-Data-Sources_181008.pdf.

Strengths

This platform is reported to have good spatial coverage, and its sampling rate (i.e., penetration) was satisfactory for most observed areas and data types. The platform also offers and uses multiple data sources as part of its analytical platform and offers GPS, LBS, auto, transit, and demographic data. This allows several types of analyses to be performed on the platform, such as park-and-ride studies, model validation, and select zone analysis, as was shared by some of the respondents.

Another key strength of StreetLight Data is that the methodology for calculating different mobility measures is made publicly available through white papers and journal articles. This allows the user to obtain more insight into the process rather than it being a just a black box. The platform also has some of the best technical support. Therefore, it is not surprising that there was consensus across respondents that they found the platform to be user friendly and it allowed them to accomplish most of their intended tasks.

Challenges

Despite StreetLight Data's popularity, several challenges were highlighted by users across participating agencies. Users pointed out that it would be beneficial to know what smartphone applications (in the case of LBS data) the data is being sourced from and their share in the overall data. This information could help the users understand the data and take steps to reduce the bias if possible. In this context, a user noted, "It seems a lot of data is coming from food delivery apps or ride-sharing apps."

Even though the platform is recognized for diverse data sources and types, users reported a lack of historical data for some areas. Historical data tend to become sparse while performing a longitudinal analysis. Additionally, across all data types, transit data seem to suffer from sample size issues.

Users expressed interest in the inclusion of a data downloader feature that could enable users to download unprocessed data to either have more control over the analysis or perform an analysis that is unavailable on the platform. Users also expressed interest in a feature that would allow users to export graphics from the platform.

Licensing can be expensive for this platform, particularly for users that rely on it for project-specific studies. There are also constraints on zonal analyses based on the license regarding the number of zones that can be analyzed using a given license type. Zones cannot be modified once uploaded or created. Users can create new zones, but those can be limited based on the license type. For the unlimited access license, it can be difficult for consultants to pass on charges to clients compared to pay-by-project licensing.

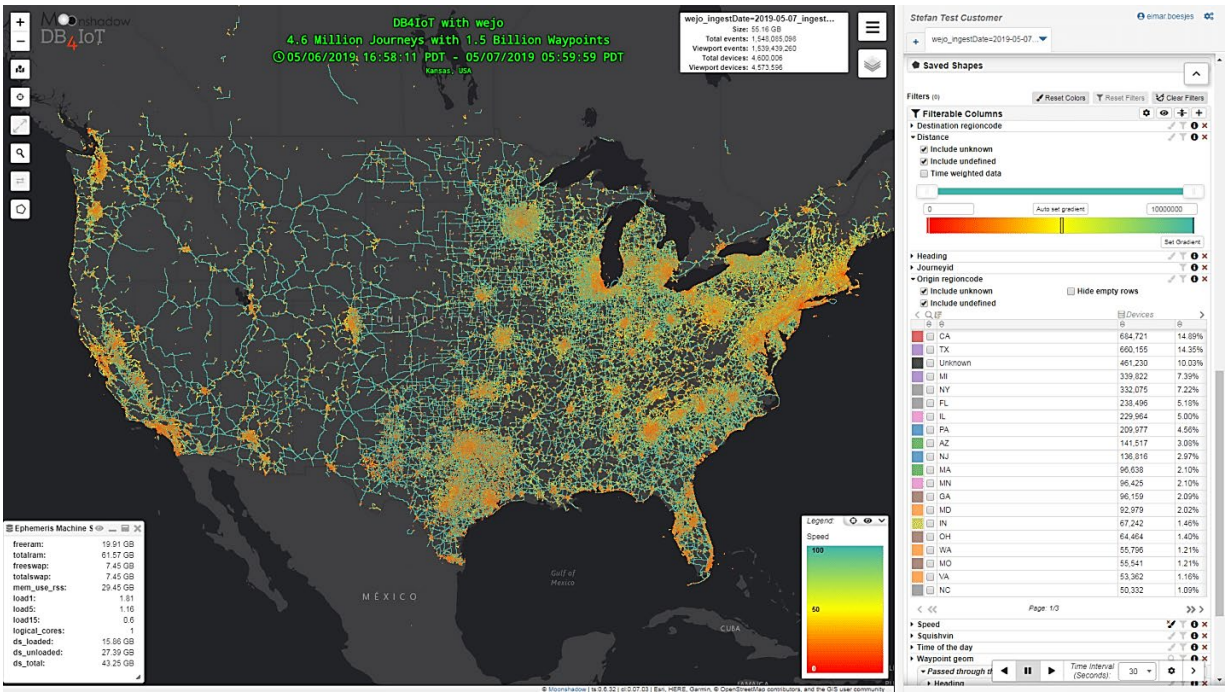
The biggest strength of StreetLight Data is the availability of various types of data and the technical support.

Moonshadow

Moonshadow facilitates geospatial visualization by allowing users to easily recognize key spatial relationships and correlations within large data sets. Moonshadow's database for the Internet of Things is a time-series database engine and mobility analytics platform purpose-built for the Internet of Moving Things and connected vehicle (CV) data analytics. Moonshadow combines data from many multimodal sources and packages them into easy-to-use interactive mobility data analytics and visualizations, which can be used to include maps, charts, graphs, and data in the users' own dashboards or applications.[4]

Standard sources for Moonshadow include major suppliers of CV data including Here, INRIX, Michelin, Otonomo, and Wejo. Users also have the ability to visualize their own data using Moonshadow. The platform allows users to combine CV data with other spatial data such as highway network, geography, weather, traffic incidents, traffic counts, etc.

Apart from being a visualization platform, Moonshadow also has some analytical functionalities such as the ability to generate O-D matrices, select link analysis, traffic distribution, and scenario planning. Figure 2 shows a screenshot of Wejo data loaded on the Moonshadow platform.



(Source : DB4IoT, https://www.db4iot.com/)

Figure 2: Moonshadow Platform Dashboard Interface

[4] Moonshadow. Powering Analytics of Big Data for the "Internet of Moving Things." http://moonshadowmobile.com/.

Strengths

Moonshadow is distinct from other platforms evaluated as part of this task because Moonshadow:

- Is primarily a visualization tool;
- Hosts a vast and continuous real-time stream of mobility data from several data sources; and,
- Allows ingestion of custom big data.

These capabilities allow Moonshadow users to visualize their geospatial data and/or use the platform's available data.

The data hosted in this platform are at their finest resolution, which helps to understand sampling issues, unlike other platforms. According to one user of this platform, "The quality of data is good, as was observed from the results of some validation exercise done, compared to continuous count station data." The users also mentioned that the technical support is very responsive and helpful.

As with other platforms discussed in this technical memorandum, Moonshadow allows users to make selections and download project data. Moonshadow also generates maps or charts using both real-time and historical data.

This platform lends itself to being more useful for traffic operations analysis compared to being used for regional transportation planning.

Challenges

The challenge with Moonshadow, as could be expected, is that it occasionally tends to be slow while running visualization queries, particularly when using real-time data. This can deter frequent use of this platform. Lastly, the platform subscription is costly, making it difficult for many organizations to afford the initial cost.

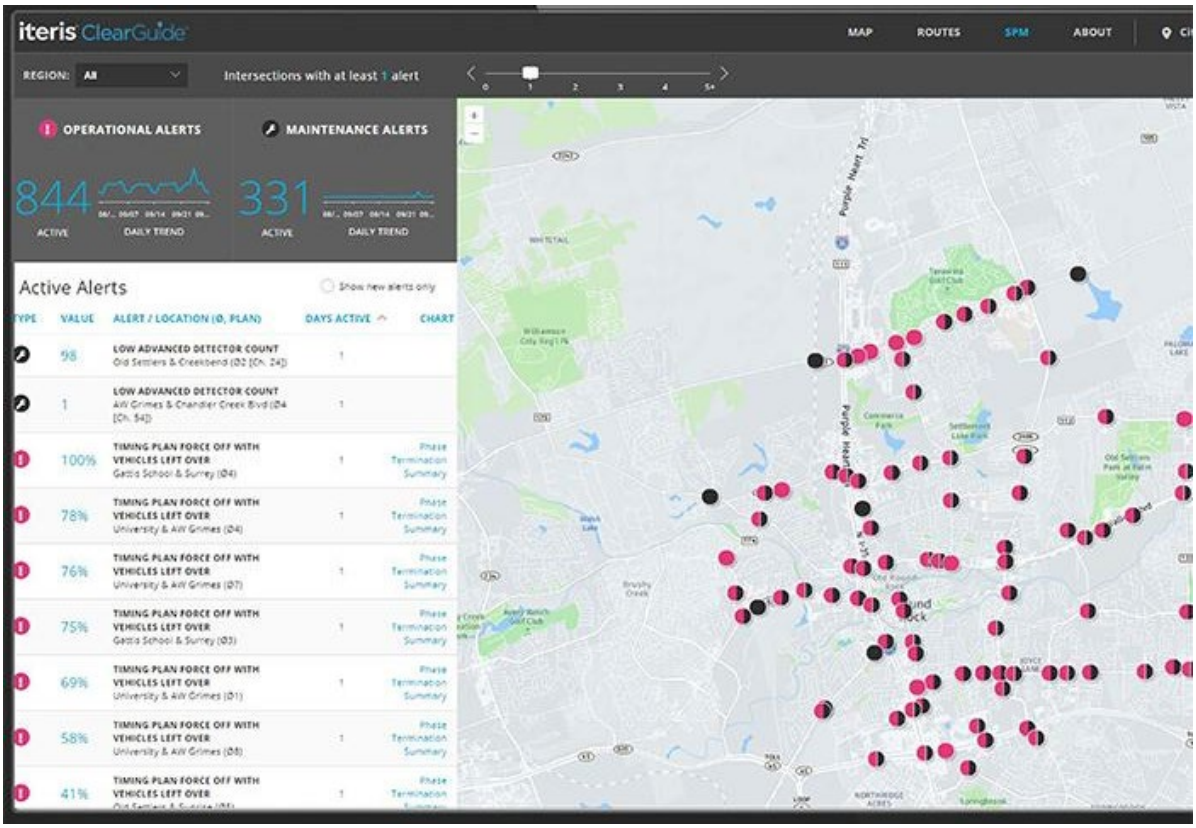
Moonshadow allows users to either visualize their own geospatial data and/or use what is available through their platform.

Iteris ClearGuide

Iteris ClearGuide is a web-based platform that allows users to analyze transportation big data (it primarily hosts data from INRIX and Here). ClearGuide has visualization and some analytical reporting tools. Its visualization options allow users to visualize both real-time and historical mobility data, which could help identify problem areas before traffic congestion worsens.[5] The analytical reporting tools as part of its tool catalog include:

- Signal Performance Measures that can be used to find and resolve operational or maintenance concerns (see Figure 3);
- Arterial Performance Measures that provide metrics on arterial travel times and reliability, and could be used in prioritizing retiming projects, identifying congestion hotspots, and identifying highway traffic impacts surrounding arterials; and,
- Highway Performance Measures that present insights into congestion including how events such as weather, construction, incidents, and special events impact highway traffic.

The platform also includes a mass data downloader feature.



(Source: Iteris, https://www.iteris.com/products/performance-analytics/clearguide)

Figure 3: Iteris ClearGuide Platform Dashboard Interface

[5] Iteris. ClearGuide—Real-Time Contextual Mobility Intelligence. https://www.iteris.com/products/performance-analytics/clearguide.

Strengths

There was consensus among the users that Iteris ClearGuide is the platform of choice for traffic operations and analysis. The platform has good visualization tools (e.g., its ability to produce time-series and congestion contour plots for daily and monthly speed data). The visualizations and downloaded data are available for real-time and archived (historical) data. Iteris ClearGuide also has a mass data downloader feature, which is helpful for bulk data download requests. It is one of the most used features on the platform since it allows retaining downloaded data after the subscription expires.

The user interface (UI) of Iteris ClearGuide is well designed, with options to create separate analysis reports for each route. The interactive UI allows for visualization of multiple-day congestion patterns in calendar format, which is very useful for specific segment assessment and analysis of potential bottlenecks.

Challenges

This platform is a new and evolving product and comes with a few challenges. The mass data downloader, although the strength of this platform, does not generate all the attributes available on the platform. Moreover, the platform does not aggregate the data as part of the mass data downloader process, making it challenging to manage the massive unprocessed granular data after download.

Sometimes, the downloaded data do not match the shapefile, making it difficult to relate the data spatially. This possibly results from the roadway shapefile being updated without advance notification or cross-references. This inconsistency makes it challenging for users to match segment attributes and analysis over time.

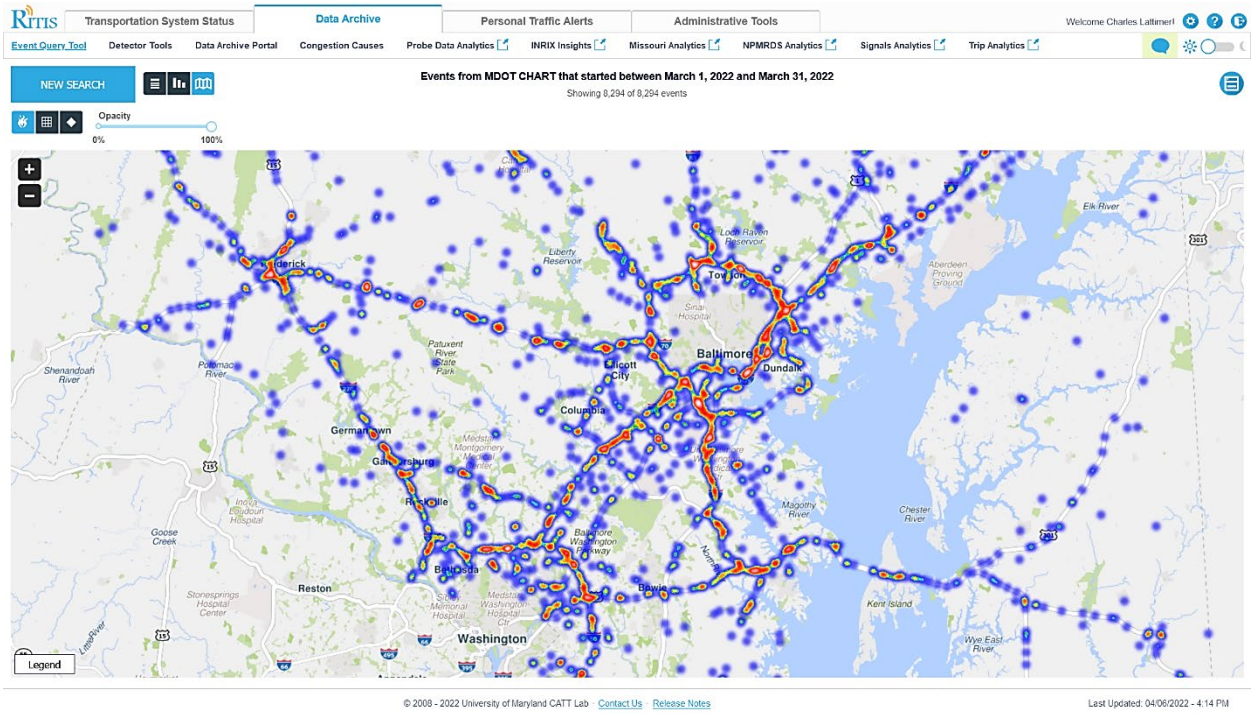
In terms of availability, only speed data are available on the platform. Vehicle volume data, which are of interest to the users, are currently unavailable on the platform. Further, Iteris ClearGuide focuses more on specific routes, making the platform suitable for operations, but lacks features to undertake area-wide analyses, making it undesirable for planning applications.

Users reported the technical support was unsatisfactory because it is not as responsive as other platforms. Further, there is inadequate documentation on the calculations behind the numbers produced by the analytical features of the platform.

Iteris ClearGuide appears to be more popular with professionals working in the operations sector than the planning sector.

Regional Integrated Transportation Information System

The Regional Integrated Transportation Information System (RITIS) platform allows users to query big data from several different sources, including INRIX, HERE, and TomTom. The platform features a wide variety of tools to analyze data hosted on this platform. Some of the prominent tools are trip analytics, segment analysis, and a massive data downloader. There are three main components of the RITIS platform—real-time data feeds, real-time situational awareness tools, and archived data analysis tools.[6] Figure 4. shows various analysis options available on RITIS.



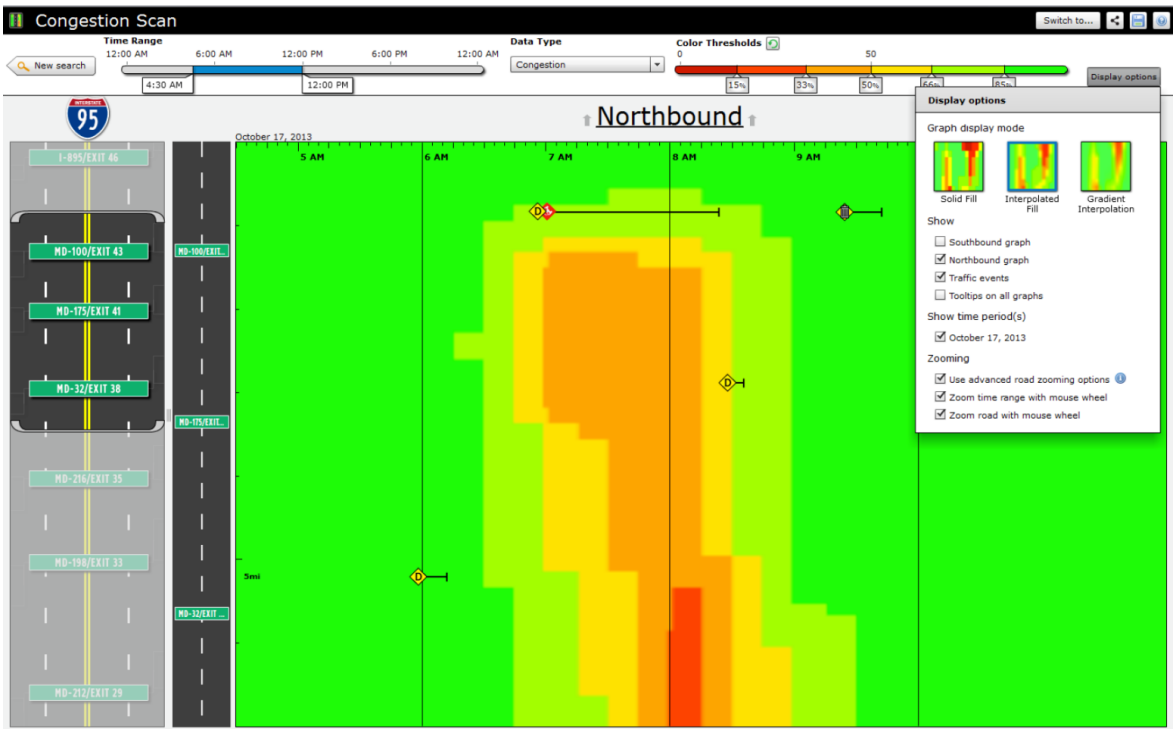
(Source: RITIS, https://ritis.org/intro)

Figure 4: Sample Screenshot of the RITIS Map Showing Duration, Timing, and Heatmaps (Locations) of Incidents

Strengths

The Massive Data Downloader tool is this platform's most frequently used tool. The tool aggregates travel time and speed data by various time intervals (5-, 10-, and 15-minute) and are available to download. Other popular applications of this platform are Trend Maps and the MAP-21 widget feature. In addition to these tools, the congestion scan tool helps communicate results to the public because the results are easily read and well visualized, as shown in Figure 5.

[6] RITIS. RITIS Tool Catalog. https://ritis.org/tools.



(Source: RITIS, https://ritis.org/tools)

Figure 5: Congestion Scan Tool

RITIS is flexible in defining the analysis zones. A user remarked, "The RITIS platform is also more 'mature' or flexible in selecting Traffic Management Centers (TMCs), drawing polygons to define areas of analysis, etc." The platform is also more flexible in pricing structure than some of its peers since adding or removing features or data sources is intuitive in RITIS.

Challenges

In general, users desired additional attributes to the existing data catalog:

1. The roadway network only includes the National Highway System (NHS).
2. Some users expressed that they wished RITIS had trip (O-D) data since that would have been helpful for modeling purposes.
3. Vehicle weight class information is unavailable in the probe data sample, preventing users from separating personal and freight traffic. The availability of freight data could be helpful in planning and operational purposes.

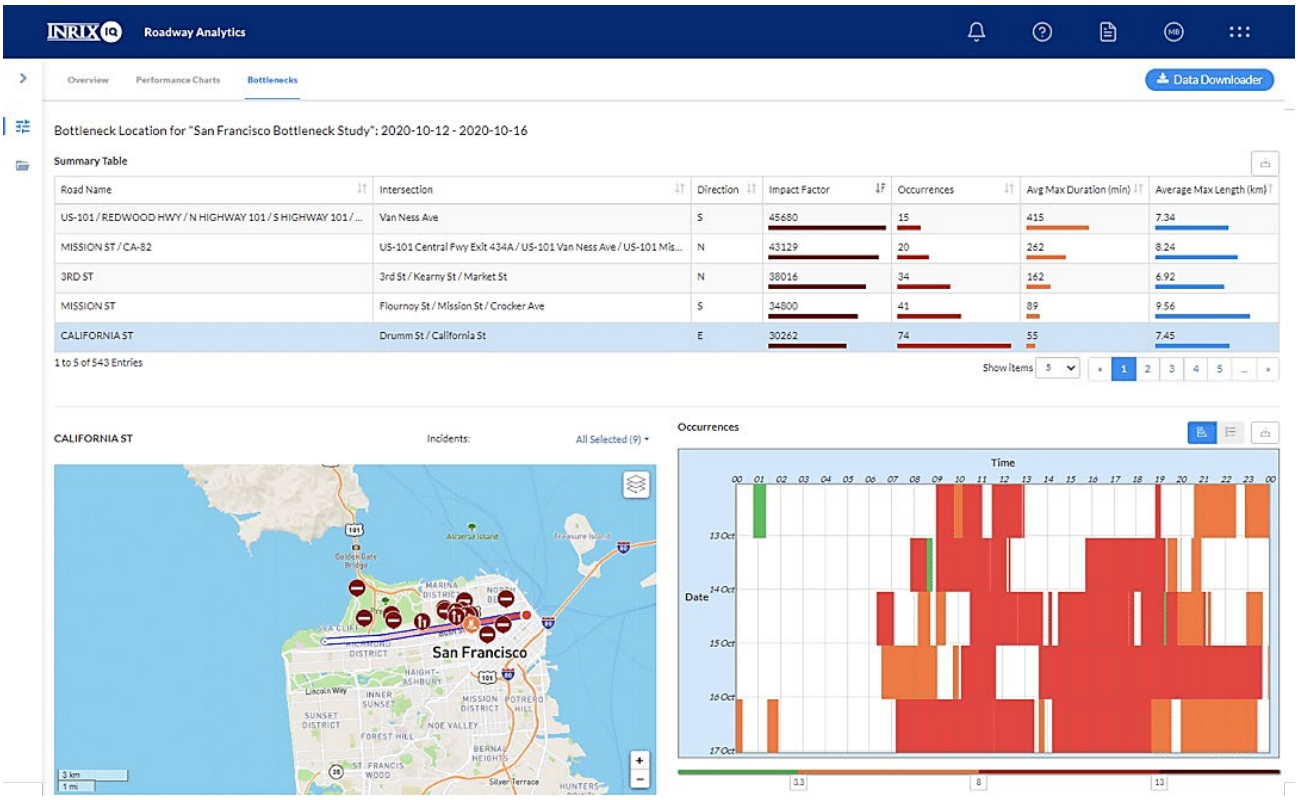
Road segments also seemed to be a challenge when working on the platform. The data vendor of RITIS (currently INRIX) updates its roadway segments for the covered NHS network periodically; however, the roadway segments linked to downloaded data are limited to the road segments for which the data were collected. In other words, the road segments are not ideal for longitudinal analysis as historic data does not match current roadway segments. In addition, while using the Probe Data Analytics tool, for some

state agencies, the TMC segments are much different than the states' segments, and there is a need for additional conflation. Moreover, users feel that the road coverage is wanting. Users would prefer to have either no or a higher limit on the application programming interface (API) calls since they limit the number of segments a user can visualize at a time on the map. Additionally, a user reported that they observed that performance speed plots and downloaded data for the same plots can convey different results if the user looks at multiple days.

RITIS is one of the most popular platforms among users, and its data downloader feature is sought after.

INRIX Roadway Analytics Tool

The INRIX Roadway Analytics platform uses anonymous data from CVs to deliver on-demand traffic data, analytics, and visualizations.[7] As Figure 6 shows, the tool catalog includes Bottleneck Ranking (i.e., identify, prioritize, and visualize locations with recurring congestion throughout the entire roadway network), Performance Charts (i.e., line and bar graphs for before and after inquiries, including comparison studies), Congestion Scan (i.e., identify and pinpoint problem areas along any travel corridor to better target improvement efforts), and User Cost Delay (i.e., combine speed, volume, and economic data to determine how much money is lost by delay on the roadway). More recently, a few users have started using the INRIX Roadway Analytics platform data for signal analytics as well**.**



(Source : INRIX, https://inrix.com/products/roadway-analytics/)

Figure 6: INRIX Roadway Analytics Platform Dashboard Interface

[7] INRIX. INRIX Roadway Analytics. https://inrix.com/products/roadway-analytics/.

Strengths

The INRIX platform is a preferred tool for before-and-after studies because it maintains the same road segments for historical data, letting users conduct longitudinal analyses. Moreover, compared to RITIS, the road segments on this platform are shorter and have greater coverage, which helps in conducting localized analyses, for example, at roundabouts.

The platform provides good visualization tools and has a quick response. The processing time is faster for data requests because it outputs the raw data and thus saves time on processing or aggregating data.

Challenges

The INRIX user interface can be challenging at times. For instance, defining routes during corridor analysis can be difficult. A user noted, "The platform crashed several times and was almost unusable." Users ideally might want to set smaller lengths (e.g., one-tenth of a mile) for signalized intersections because long segments do not work well for the signalized intersections in urban areas.

INRIX currently offers a product trial for the platform; however, the 1–2-week timeline for the trial is too short to complete an analysis. Users would like a more extended trial period to assess the useability of the platform.

The INRIX platform is preferred for analyses that require consistent roadway segments and wider road coverage.

Conclusions

In general, data platforms are preferred by users because they find those platforms easier to work with in comparison to purchasing raw data from vendors. At the same time, the users admitted that there is a trade-off in that the trip data provide more user control on how the data are analyzed and visualized and avoided the black box analysis process. Therefore, some use the platforms primarily for their data download feature. However, some platforms are better than others in publicly making their methodology and documentation available. This avoids a black box perception, helps users understand what they are getting through the data on these platforms, and allows them to adjust outputs if needed.

Figure 7 summarizes the findings from the discussion sessions and surveys. The column headers in the figure were a common theme in all the discussion sessions. *Data retention* refers to the ability to retain the outputs after the analysis, whereas *data downloader* refers to the ability to download unprocessed data from the platform. Not all platforms with a data retention feature offer a data downloader. Moreover, it was observed that different platforms have specific use cases and suitability for planning and operational applications among the agencies. For example, agencies prefer to use StreetLight for planning purposes; Iteris ClearGuide is highly sought after for operations. Similarly, RITIS is used extensively for Federal Performance Measure Rule 3 reporting and planning and traffic operations applications.

	Data Source	Data Retention	Data Downloader	Ease of Use		Historical Data
				Planning	Operations	
StreetLight	INRIX and Cuebiq*	☑	☒	4.3/5.0	3.7/5.0	☑
Moonshadow	INRIX, WEJO, MICHELIN, OTONOMO or Your Own	☒	☒	3.0/5.0	3.0/5.0	☑
Iteris ClearGuide	HERE, INRIX, WEJO	☑	☑	3.5/5.0	4.5/5.0	☑
RITIS	INRIX	☑	☑	4.3/5.0	4.5/5.0	☑
INRIX Roadway Analytics	INRIX	☑	☑	3.7/5.0	4.3/5.0	☑

(Source: StreetLight, https://www.streetlightdata.com/wp-content/uploads/StreetLight-Data_Methodology-and-Data-Sources_181008.pdf)

Figure 7: Summary of User Feedback on Different Mobility Data Platforms

Several users remarked that the nature of the project might dictate which tool/platform is better among the ones discussed in this technical memorandum. This choice may depend on the type of roadway segmentation needed for the project, the suitability of the graphical user interface (GUI), or the features of the platforms. For example, as shown in Figure 8, if the scope of analysis is at a granular spatial level or a before-and-after comparison is needed at a particular location, the user might use a platform with the same segmentation for the before-and-after periods to avoid any external confounding factors. This choice may depend on the roadway segmentation needed for the project, the GUI's suitability, or the platforms' features. Agencies assess the suitability of platforms for specific applications on a case-by-case basis.

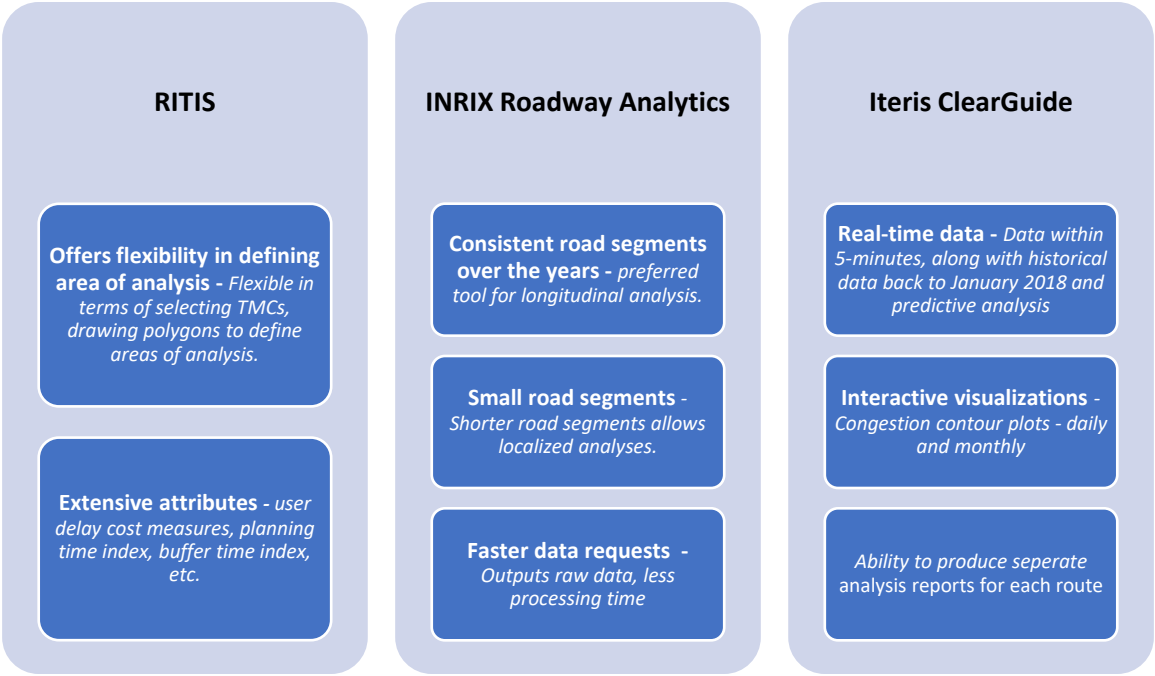


Figure 8: Strengths of Platforms used in Transportation Operations

The age of the platform can also be a governing factor in deciding which platform is to be used. While the platforms developed a few years ago have a more extensive toolbox that the new platforms might lack, newly built platforms are more open to feedback and include new analyses or features. For instance, a few users found the RITIS platform a bit slower than other platforms. Though Iteris and RITIS have similar functionality, many users from the operations side prefer using Iteris for its easy-to-use GUI and additional attributes such as weather data.

Finally, the technical support offered by the platforms is critical to the users. The level of technical support that a platform offers is factored in by agencies while making a subscription decision.

While most of the challenges discussed in this memo can be addressed or have one or more solutions, there are a handful that are either by design or inherent to the data generation process. For instance, the lack of features (example: data downloader) or inadequate technical support can be resolved over time. However, other challenges such as sparse data coverage across regions and(or) at finer geographic resolutions could either be due to data anonymization, privacy agreements in place, lack of representation or due to poor coverage of technologies utilized in collection of such data. This document aims to develop awareness among users the strengths and challenges associated with such data sources and platforms, some of which might not have any resolution.