

This Implementation Memo is supplemental to *NCHRP Research Report 1071: Application of Big Data Approaches for Traffic Incident Management* (NCHRP Project 03-138). The full report can be found by searching for *NCHRP Research Report 1071* on the National Academies Press website (nap.nationalacademies.org).

The National Cooperative Highway Research Program (NCHRP) is sponsored by the individual state departments of transportation of the American Association of State Highway and Transportation Officials. NCHRP is administered by the Transportation Research Board (TRB), part of the National Academies of Sciences, Engineering, and Medicine, under a cooperative agreement with the Federal Highway Administration (FHWA). Any opinions and conclusions expressed or implied in resulting research products are those of the individuals and organizations who performed the research and are not necessarily those of TRB; the National Academies of Sciences, Engineering, and Medicine; the FHWA; or NCHRP sponsors.

NCHRP Project 03-138, “Application of Big Data Approaches for Traffic Incident Management (TIM)”

Implementation of Research Findings and Products

Technical Memorandum

February 17, 2023

This technical memorandum provides information and ideas associated with implementation of the NCHRP Project 03-138 research findings and products, which include a final research report, summary, and marketing slide deck. More specifically, the research report provides examples and guidelines for transportation agencies to prepare for and implement big data solutions for TIM. These examples/guidelines include:

- Findings from comprehensive assessment of TIM-relevant data sources – Results/findings from an in-depth assessment of 16 TIM-relevant data sources. An overview of each data source is provided, followed by a summary of the challenges and limitations, consequences for use in TIM big data use cases, and associated recommendations. Detailed findings are provided in the appendices in terms of six dimensions of data quality, including timeliness, completeness, accuracy, conformity, consistency, and integrability. These findings should help agencies better understand the quality and limitations of agency-owned and third-party data and where improvements could be made to improve the usefulness and value of the data.
- Four TIM big data case studies – Case studies associated with the four TIM big data use cases/data pipelines developed during this research project. Each case study provides an overview of the TIM big data use case, the datasets leveraged by the research team, a description of the data pipeline, the associated data analysis/products, and lessons learned and recommendations for implementation. These case studies can help demonstrate to agencies the feasibility and value of the big data approach for improving TIM, along with the current limitations of the big data approach for TIM.
- Data pipeline details – Each of the big data pipelines developed in this research project is represented in terms of two different yet equally important processes: 1) a pipeline technical blueprint diagram, or N² diagram, which illustrates the logical data flow through the pipeline and displays the functions performed and items passed between the functions (inputs and outputs of each of the functions) and 2) a pipeline workshop description/diagram, which presents a more practical example of the steps of the workflow, how these steps are performed, and in which systems these steps occur. Also described is the cloud data environments in which the primary functions of the pipelines to collect, analyze, enrich, curate, and store the data were performed. These detailed descriptions could help agencies implement similar data pipelines.
- Example data products – Where practical, the research team developed example data products to visualize the outputs of the data pipelines and demonstrate the practical value of the big data approach. Several example dashboards could help agencies envision how the integration of big data with agency data could improve their ability to detect, verify, and respond to traffic incidents.

- Cost estimates for the running the pipelines – The research team developed estimated average cost ranges for the data pipelines developed for this NCHRP project. These ranges are based on the understanding that the cloud provides flexibility and scalability and can fluctuate month to month based on data storage, cloud processing/queries, etc. Costs are estimated for each use case/data pipeline and are broken down by data storage (e.g., archival storage, real-time curated data storage), data query and analytics (e.g., data ingestion and processing, real-time curated data), and data products (e.g., dashboards). While these costs represent ballpark estimates, they should give agencies an idea of the cost of big data applications, as well as how these costs can fluctuate based on the volume of data stored and the amount of processing being done on the data (e.g., number of users, number of queries, amount of data being processed).
- TIM big data guidelines – Presents 18 guidelines and 80 associated “implementation actions” regarding the development and implementation of TIM big data pipelines. The guidelines build from, enhance, and refine the big data guidelines presented in *NCHRP Research Report 904: Leveraging Big Data to Improve Traffic Incident Management*; pull from big data/modern data management guidelines presented in *NCHRP Research Report 952: Guidebook for Managing Data from Emerging Technologies for Transportation*. The guidelines are presented across six categories: data acquisition and quality; data environment, platform, and architecture; data management; data processing, tools, and mining techniques; data pipeline development and operations costs; and data sharing. These guidelines could help agencies understand the foundational practices for managing big data, which differ greatly from traditional “small” data management practices, including the use of data centralization and cloud/distributed computing and architecture.

This technical memorandum provides recommended actions, potential leading institutions, and challenges to implementation of the NCHRP Project 03-138 research products, as well as ways of measuring the impacts and success of implementation.

Ideas for Putting the Research into Practice

To put the NCHRP Project 03-138 research products into practice, they must be shared and marketed to promote awareness, and agencies need an opportunity to exchange ideas and challenges, as well as hands-on support with implementation. The following are ideas to promote implementation of the research products (Note: These ideas start from the furthest-reaching approach to first promote awareness and then work toward more targeted and hands-on approaches to promote actual implementation of the research products):

- **National webinar(s)** – The research team and panel could conduct a national webinar on the NCHRP Project 03-138 research outputs/products at the end of the project. If TRB cannot host a webinar within this timeframe, a recommended outlet for is the National Operations Center of Excellence (NOCoE). The research team has worked through NOCoE to conduct many webinars. NOCoE promotes and supports webinars as a method of reaching large audiences on topics related to operations and often can attract upwards of 150 participants. For the national webinar, the following specifics are recommended:
 - 3 presentations including the following:

- NCHRP Project 03-138 research team – provide an overview of the research and the resulting research products.
 - 2-3 agency presenters that have developed big data pipelines for TIM (or other operations use cases). (Potential agency presenters include Indiana DOT, Kentucky Transportation Cabinet, Pennsylvania Turnpike.)
 - The webinar should be facilitated by one of the panel members, preferably representing a state transportation agency.
 - Polling questions would be presented to the participants toward the end of the webinar to gauge the participants’ levels of interest in implementing the research products within their agencies. Example questions include:
 - Is your agency implementing any big data pipelines for TIM or other operations use cases?
 - Are you/your agency interested in learning more about the NCHRP Project 03-138 research products?
 - In what ways would you/your agency be interested in learning more about the NCHRP Project 03-138 products (example response choices: participating in a workshop or peer exchange, on-site training, participating in a follow-on implementation project)?
- **Presentations at committee meetings/conferences** – Conference and associated committee meetings are an opportunity to present the outputs/products of NCHRP Project 03-138. Examples include TRB meetings, AASHTO meetings, ITE meetings, and ITS regional conferences. These groups are identified and discussed in more detail in the following section on potential leading institutions.
- **NCHRP Project 08-119, “Data Integration, Sharing, and Management for Transportation Planning and Traffic Operations”** – The objective of this NCHRP project is to develop tools, methods, and guidelines for improving data integration, sharing, and management practices to enable transportation agencies, in collaboration with private-sector and public-sector stakeholders, to make better planning and operations decisions. AEM is also leading this project, and it is nearing completion. As part of this project, the team has developed an online one-stop-shop or “knowledge base” of all things data management, including tools to support agencies. The NCHRP Project 03-138 research products could be featured within the knowledge base for agencies looking for ways to improve TIM through the use of data.
- **Multiagency workshops/peer exchanges** – As part of a potential implementation stage of NCHRP Project 03-138, a multiagency workshop or peer exchange would provide an opportunity to bring together interested agencies (identified as a result of the national webinar polling questions and participants in the workshop for NCHRP Project 03-138) for more in-depth discussions about the research outcomes/products. The research team could organize, develop, and facilitate a 1- to 1.5-day workshop/peer exchange involving representatives from up to 10 agencies representing a range of data capabilities, but all with an interest in leveraging big data to improve TIM.
- **On-site training** – As part of a potential implementation stage of NCHRP Project 03-138, on-site training would provide individual agencies that are highly interested and committed to implementing the research products (as identified as a result of the national webinar polling questions and follow-up conversations, as well as the workshop conducted during the research)

with an opportunity to interact one-on-one with the research team to learn in more detail about developing big data pipelines for TIM. The research team could organize, develop, and deliver a 1.5- to 2-day training workshop at one or two agencies as an initial pilot test of the training. If feedback from the agencies receiving the pilot training is positive, the training could be modified (as per input from the agencies) and offered to other agencies to help expand implementation of the research products.

Potential Leading Institutions

There are three primary institutions that could take a leadership role in promoting the implementation of NCHRP Project 03-138 research outputs/products: the Transportation Research Board, the U.S. Department of Transportation, and the American Association of State Highway and Transportation Officials.

The **Transportation Research Board (TRB)** is a natural fit for marketing and disseminating the resulting research products. Marketing and dissemination of the research products by TRB would likely occur through TRB technical committees and/or an implementation stage of the research (as previously discussed).

The results of the research could be marketed to multiple TRB technical committees, which could help to spread the word and the potential benefits of implementing the research products:

- Operational Resilience Emergency Transportation Operations (ETO) Subcommittee (ACP10-1) – ACP10-1 is a subcommittee of the Regional Transportation Systems Management and Operations Committee (ACP10), which is concerned with regional transportation systems management to maximize transportation system performance in metropolitan areas, including coordinated and integrated decision-making approaches to operations and the harmonization of operations with planning, construction, preservation, and maintenance of transportation facilities. The Operational Resilience ETO Subcommittee was involved with NCHRP Project 17-75 and NCHRP Project 03-138 and is the natural committee to take the lead in promoting the outcomes of the research.
- Committee on Statewide Transportation Data and Information Systems (ABJ20) – The scope of this committee includes research and technology transfer activities pertaining to statewide transportation planning data and information systems for all modes of transportation. A primary concern is the capability of information systems to integrate various transportation-related data sources into a strategic multimodal information database for statewide transportation planning. In order to develop big data pipelines to improve TIM, there is a need to integrate transportation (and partner) data. This could be considered a use case for the ABJ20 committee and one that might be of interest to the members and friends.
- Task Force on Data for Decisions and Performance Measures (A0030T) – This Special Task Force works to understand, share ideas about, and improve the delivery of information to transportation decision makers to support informed, efficient, and data-driven performance measurement and policy, investment, and operations decisions. Its concerns include appropriate metrics, availability and quality of data, strategies for converting data into useful information, and effective ways to deliver that information to decision makers and their constituents. The goal of applying a big data approach to TIM is to develop better metrics and to

inform better decision-making. The availability and quality of data for TIM is a concern, particularly in the context of big data. This project should be of interest to the A0030T committee.

- Committee on National Transportation Data Requirements and Programs (ABJ10) – The scope of this committee includes the development of nationwide and international data on transportation needed to support decision-making and data-related research in all sectors of transportation. All aspects of data development are of interest, including design, collection, analysis, reporting, funding, administration, dissemination, and coordination of statistical programs. Of particular interest is the coordination of transportation statistical programs with non-transportation programs; coordination between national-level and other programs, including international, private, state, and local systems; and the structuring of statistical standards and criteria that guide the development of comprehensive transportation programs.
- Committee on Geographic Information Science and Applications (ABJ60) – The scope of this committee includes all aspects of the spatial, locational, and temporal data used in transportation. The committee is interested in both research into and applications of this information and its associated information systems, commonly referred to as Geographic Information Systems in Transportation (GIS-T). The committee will provide a focal point for and promote coordination of GIS-T activities within the TRB committee structure. Relevant activities include the application of spatial data and spatial sciences across the entire domain of transportation information systems. The NCHRP Project 03-138 research assessed multiple sources of spatial data and relied on spatial data to integrate the disparate datasets for the big data pipelines. Sharing the findings and recommendations with this committee seems prudent.
- Committee on Safety Data, Analysis and Evaluation (ANB20) – This committee is concerned with the study of highway safety. This includes the collection, maintenance and use of crash records and related highway, driver, and vehicle data; the development of theories, analytical techniques, and evaluation methodologies for improving the understanding of highway safety; and the application of these theories, techniques and methods to identify driver, vehicle and/or roadway-based treatments that will enhance the safety of the transportation system. The findings of this research showed limitations in the use of crash data in the context of big data pipelines. Improving the quality of crash data would improve the ability to integrate agency data with other big datasets. Furthermore, the use of related highway, driver, and vehicle data could help to enhance, improve, and expand the TIM big data use cases, should this data be more readily accessible.

The **U.S. Department of Transportation (U.S. DOT)**, through the Federal Highway Administration (FHWA), is an agency with tremendous influence and respect among state agencies. There are multiple paths in which the NCHRP Project 03-138 products could be marketed:

- FHWA's Every Day Counts (EDC) Program is an ideal platform for reaching various engineering and technology professionals. The EDC program was designed specifically to help states implement trusted and proven innovations. Several recent innovations – including the EDC Round 4 *Using Data to Improve TIM* and the EDC Round 5 *Crowdsourcing for Operations* – address the collection, integration, analysis, and use of big data, including big data from emerging technologies. While EDC Round 7 recently started in January 2023 and will continue through December 2024, FHWA will be seeking the submission of ideas for EDC Round 8 from

state and local agencies and industry partners in late 2023/early 2024. These ideas should include proven, market-ready processes or technologies that have the potential to transform the way the highway transportation community does business by enhancing roadway safety, shortening project delivery time, reducing traffic congestion, or integrating automation. As there are proven examples of big data applications of TIM, the recommendations and process outlined in the NCHRP Project 03-138 final products might make a good EDC-8 innovation.

- U.S. DOT’s Safety Data Initiative (SDI) seeks to integrate existing data and new “big data” sources, use advanced data analytics to provide new insights into transportation safety risks, and create data visualizations to help policymakers arrive at safety solutions. The outputs of NCHRP Project 03-138 are very much in line with the goals of the SDI and could be implemented to support SDI efforts already underway or in the works. In addition, the approach of the SDI program to offer demonstration grants could be applied to the implementation of the NCHRP Project 03-138 research products.
- FHWA program areas – With the modern era of computing and data-driven decision-making, every FHWA program could benefit from the promotion of the implementation of modern big data practices. The use of data isn’t going away, and the data will continue to become larger and more complex. Modern big data practices can handle these ongoing changes with greater ease than traditional systems and practices. Below are a few examples of FHWA program areas that could benefit from the application of the modern big data approach to data management:
 - Traffic Incident Management – By nature, TIM involves a wide range of data generated by the various responder disciplines (transportation, law enforcement, fire and rescue, towing, emergency medical services) and an opportunity to further demonstrate the benefits of sharing data across these disciplines to better inform TIM practices, procedures, and policies; reduce incident response and clearance times; reduce incident-related congestion; and improve the safety of responders, incident victims, and the traveling public. Given the TIM focus of NCHRP Project 03-138, the FHWA TIM Program is a natural fit for promoting the outcomes of the research.
 - Work Zone Safety – Work Zone Safety is a use case related to TIM, with events more often planned but impacting traffic and safety in a similar fashion to traffic incidents. Safety in and around work zones could be improved with additional data sources. By combining various sources of data, safety and operations personnel can better understand conditions of the work zone that may prove problematic. Combining GPS-based traffic speed data, smart work zone devices, road weather information systems (RWIS) and other weather data (including crowdsourced weather data), connected vehicle data, Doppler radar, weather forecasts, direction of the sun at the time of crashes, etc. could contribute to a better understanding of work zone impacts and incidents.
 - Road Weather – Road Weather is another use case for big data. For many years, agencies have been trying to better understand road weather and its impact on the public and themselves. Some data that may be combined for a better understanding of road weather include RWIS, Doppler radar imagery, forecasts from various sources, real-time telemetry from fleet vehicles, third-party GPS-based traffic data, crowdsourcing reports, etc. The combination of these data sources provides a very granular and holistic

understanding of road weather events, with any number of data points that can be used to develop a wide range of performance measures.

The **American Association of State Highway and Transportation Officials (AASHTO)** maintains a direct and relevant relationship with state agencies. AASHTO annual meetings, associated regional meetings, and AASHTO committee meetings would all be good venues in which to market and promote the outputs of NCHRP Project 03-138. In particular, the [AASHTO Committee on Data Management and Analytics](#) addresses the collection, procurement, processing, analysis, reporting, and sharing of transportation data. According to its [charter](#), this committee “is dedicated to addressing issues related to knowledge, expertise, resources and tools needed by state DOTs to implement a robust data management and analytics program within their agencies.”

Issues Affecting Implementation and Possible Actions to Address These Issues

“A lot of times, people don’t know what they want until you show it to them.” – Steve Jobs, co-founder of Apple.

There are two primary, high-level issues that will impact implementation of the NCHRP Project 03-138 research products:

- Perception that the research products do not pertain to an agency
- Institutional, organizational, and cultural barriers leading to pushback within an agency or the inability to move in a direction that supports the use of big data

Perception That the Research Products Do Not Pertain to an Agency

One significant misconception by some agencies is that there is no immediate need to get ready for big data. The following is a quote from an agency that demonstrates this perception: “Even though we feel that data is coming, we might understand it at a high level, but we are also balancing needs of 80 other jurisdictions that are not facing a data problem. It isn't a priority for most jurisdictions.”

While a national webinar would provide the opportunity for agency personnel to learn more about the NCHRP Project 03-138 research products, if agencies do not see the need, relevance, or urgency of the topic to them, they are unlikely to participate. What is likely needed for these agencies to see the relevance of the work to them is for them to have more tangible evidence of the benefits of big data and how they could put the results to use. This could come through having access to more in-depth case studies from agencies that have developed big data pipelines, particularly for TIM; participating in a peer exchange with one or more of these agencies; or participating in one-on-one training. Having the opportunity to take part in personal conversations and ask questions to better understand the utilization of many data sources and the speed of distributed computing is normally something that cannot be ignored. As an example, the Kentucky Transportation Cabinet hosted a “Real-Time Data” peer exchange in 2017. In attendance were the Pennsylvania Turnpike, Wyoming Department of Transportation, Iowa Department of Transportation, Florida Department of Transportation, Louisville Metro Government, and the Kentucky Transportation Cabinet. Since that time, many of these agencies have started adopting big data. Having this NCHRP Project 03-138 research report and guidelines could have been useful for those agencies to start their implementation.

Institutional, Organizational, and Cultural Barriers Leading to Pushback Within an Agency or the Inability to Move in a Direction That Supports the Use of Big Data

A lot of agencies talk about “big data,” but many do not fully understand that to use big data requires a paradigm shift in terms of data management. Big data cannot be accomplished using traditional data management approaches in use by most agencies. While these practices have served agencies well for decades with their existing agency data, new and emerging big data sources (e.g, navigation device data, probe vehicle data, connected vehicle data) require new approaches to effectively manage and use for decision-making. Agencies that have internal champions who recognize the need for and benefits of the modern big data management approach are likely to encounter some type of resistance, due to institutional, organizational, and cultural pushback, even in the most technical of organizations. The most common barriers include:

- Traditional business unit and data silos – The implications of big data threaten traditional business unit and data silos that are engrained in most transportation agencies. Examples include:
 - Structure of “Stay in your lane. I’m not going to tell you how to do your job.”
 - Uncomfortable with uncertainty
 - Hesitant to share data across organizations
 - Perception of a loss of control over data that resides in a shared environment
 - Data “vetting” process and the “authoritative” source
 - Data privacy concerns
- IT, procurement, and the inability to properly justify the cost of the cloud – Traditional IT shops often do not have the expertise to implement big data approaches. In addition, procurement is often a slow and cumbersome processes due to lack of familiarity with new services and tools (e.g., pay-as-you-go cost model of cloud services). Examples include:
 - Taking 18 months to get cleared through central IT to upgrade from Tableau desktop to Tableau online
 - Taking 2 years to purchase data (at the point now that they have a contractor purchase the data through a task order rather than purchase it directly)
 - When people are uncomfortable or don't understand it, they do not allow it to be procured
 - Requirement to procure through a regulatory party, which is very political
 - Lengthy legal review
 - IT charges a 100% premium for cloud services, which makes things extremely expensive

Each agency is different, and the barriers will need to be overcome in different ways. Without a top-down directive, however, agency champions should not try to create organization change from the get-go; rather, they should start small while thinking big and work to demonstrate the value of the modern approach through specific use cases, like TIM, thereby slowing and iteratively demonstrating the need for organizational change.

Methods of Identifying and Measuring the Impacts Associated with Implementation of the Findings/Products

Below are ways in which the success of implementation of the research products could be measured:

- Number of website hits, document downloads, and document purchases is a very programmatic method of determining interest in the research products. By recording the number of hits and document downloads using something such as Google Analytics, the results could also be ranked in comparison to other research reports for a better understanding of the overall need by agencies.
- Number of documents that reference/cite the NCHRP research products, following release and publication of the NCHRP Project 03-138 research products. A search on the number of references/citations of the work would indicate – at least from a research perspective – how often the documents were used.
- Through AASHTO and TRB committee meetings, as well as other efforts sponsored by TRB, AASHTO, and U.S. DOT, tracking the following would be a measure of impact or success:
 - Number of states referencing or using the document
 - Number of states implementing TIM big data pipelines
- Number of committees or organizations that discuss the research and guidebook as a point of reference within their group. As previously noted, TRB and AASHTO have multiple committees that could benefit from the information contained in the research products. Surveying the use and adoption of the research products would be one method of measuring success.
- Number of webinar or peer exchange invitations and the attendance of those webinars are additional metrics for understanding how impactful the findings of this research project have been.