Integrated Mobility Innovation (IMI) Testing and Deployment of Automated Buses on CTfastrak

Data Management Plan

Report — December 11, 2020

Produced by the Connecticut Department of Transportation
U.S. Department of Transportation
Federal Transit Administration

Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.
The U.S. Government is not endorsing any manufacturers, products, or services cited herein and any trade name that may appear in the work has been included only because it is essential to the contents of the work.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Integrated Mobility Innovation (IMI) Testing and Deployment of Automated Buses on CTfastrak Data Management Plan	December 11, 2020
	6. Performing Organization Code
	1334
7. Author(s)	**8. Performing Organization Report No.**
Lisa Rivers, Eric Dorsey, Peter Calcaterra, Nathaniel Horadam (CTE)	
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Connecticut Department of Transportation 2800 Berlin Turnpike Newington, CT 06131	
	11. Contract or Grant No.
	CT-2020-014-00
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
U.S. Department of Transportation Federal Transit Administration Office of Research, Demonstration and Innovation East Building 1200 New Jersey Avenue, SE Washington, DC 20590	Draft Report
	14. Sponsoring Agency Code
	HOIT-1

15. Supplementary Notes

16. Abstract

The goal of the Connecticut Department of Transportation (CTDOT) Integrated Mobility Innovation Testing and Deployment of Automated Buses on CTfastrak is to provide a unique innovation to: 1) improve Americans with Disabilities Act (ADA) accessibility at platforms through precision docking to eliminate driver error that results in unsafe situations for passengers, 2) increase vehicle efficiency and capacity on the CTfastrak guideway through bus platooning, and 3) reduce the number of accidents resulting in injury or vehicle damage at two intersections along CTfastrak due to cross traffic not stopping at red lights. Project success will be tracked through the following metrics: 1) safe operation of the vehicles with the automated driving system (ADS) engaged, 2) reliable operation, with minimal downtime while in service, 3) consistency in precise docking at stations, 4) effective demonstration of vehicle platooning, 5) reduction in hard stops and collisions at key intersections, and 6) bus operator and rider acceptance. This report documents CTDOT's Data Management Plan (DMP) for this project and provides details regarding how the CTDOT project team will securely and reliably manage the data and provide access to the data to users inside and outside the team's organizations.

17. Key Words	18. Distribution Statement
Bus Rapid Transit, BRT, Automation, Precision Docking, Platooning, Demonstration, Data	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	13	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Contents

1. Project Overview

The following table provides an overview of the Integrated Mobility Innovation (IMI) Testing and Deployment of Automated Buses on CTfastrak project.

Project Title	Project Goals and Objectives	Project Description	Project Lifecycle Phase	Project Performance Measurements
Integrated Mobility Innovation (IMI) Testing and Deployment of Automated Buses on CTfastrak	The objectives of this research project are: 1. Safely and successfully demonstrate the first automated transit buses in revenue service in North America. 2. Collect and report relevant safety data, as well as data regarding driver and passenger acceptance of the technology. 3. Collect and report on the power and load requirements from the automated driving system (ADS) technology and effects on range. 4. Advance development of ADS-integrated transit vehicles to accelerate technology commercialization.	The project will develop and demonstrate three 40-foot New Flyer Xcelsior Charge battery electric buses, to be operated by CTtransit on the agency's CTfastrak bus rapid transit (BRT) corridor. The buses will be equipped with ADS supplied by Robotic Research (RR), and will be capable of up to SAE Level 4 automation (SAE J3016) with a safety driver behind the wheel. Connecticut Department of transportation (CTDOT) plans to use this ADS technology to fully automate buses on the CTfastrak, a fully separated guideway with BRT stations and 15-inch boarding platforms. The project will focus specifically on cooperative adaptive cruise control (commonly referred to as "platooning"), to improve operational flexibility, precision docking at BRT station platforms, and safety and energy efficiency gains while the ADS is active.	Post-Award	The project team will assess: 1. Precision docking (e.g., docking location and platform gap metrics) 2. Platooning effectiveness (e.g., spacing between buses, responsive lags in braking and accelerating) 3. ADS general safety performance (e.g., sensor reliability, object detection in the right of way, unscheduled disengagement). 4. ADS impacts on energy consumption and battery range. 5. Connected vehicle technology performance (e.g., communications reliability) 6. CTfastrak operations performance (e.g., schedule adherence, vehicle speeds) 7. Driver and passenger acceptance of the technology

1.1 Change Control

The project team will update the DMP on a quarterly basis, as necessary. CTE will manage the review process with all project partners, and submit revisions to CTDOT, which will be responsible for approving any changes and submitting revised documents to FTA.

1.2 Relevant Documents

Luna, Joseph, Elizabeth Machek, and Sean Pierce. (December 2019.) Considerations for Evaluating Automated Transit Bus Programs. John A. Volpe National Transportation Systems Center. https://www.transit.dot.gov/research-innovation/considerations-evaluating-automated-transit-bus-programs-report-0149

2. Data Overview

The following table presents and overview of the proposed datasets.

ID	Dataset Title	Description	Type / Scale	Collection Method	Data File Format(s)
1	CTDOT ADS Performance Data	Assessing the reliability with which the ADS technology identifies vehicles, persons, or other objects in the CTfastrak right of way; the reliability and effectiveness of precision docking at platforms; platooning reliability, including maintaining programmed vehicle gaps and synchronized acceleration and braking; the reliability with which the ADS technology accurately responds to objects in the CTfastrak right of way; ADS technology recognition of signal phase	nSight After Action Review™ will automatically tag events based on programmed user input conditions	nSight™ data recorder integrated into autonomy hardware	1. Raw sensor data collection 2. Robotic Operating System (ROS) bag – (control and status data) information) 3. Localization data (.csv file)
2	CTDOT ADS Energy Data	This dataset consists of real-time energy consumption	Full operating block (i.e., the duration from when a bus leaves the depot until it returns) energy consumption	New Flyer data collection tool, charger data	TBD

ID	Dataset Title	Description	Type / Scale	Collection Method	Data File Format(s)
3	CTDOT Transit Operations Data	This dataset consists of baseline operations data	Average vehicle speeds, on-time performance, fueling time, passenger loads, maintenance tickets	CTtransit personnel will record this data prior to the demonstration launch, and then throughout it	Recorded statistics
4	CTDOT V2X Data	Communications reliability between on-board units (OBUs) and roadside units (RSUs)	Data transmissions between OBUs and RSUs	TBD (Possible both on OBUs and RSUs)	TBD
5	CTDOT ADS Acceptance Data	Bus operator and passenger acceptance of the ADS technology	Subjective data from surveys and interviews with a sample population	Operators, passengers and the general public (including non-passengers) will be surveyed as part of this data collection. Pre and post deployment.	Combination of printed and web-based survey methods.

3. Data Stewardship

3.1 Data Owner and Steward

Dataset Title	Data Owner	Data Steward	Federal Sponsor
CTDOT ADS Performance Data	Robotic Research	Robotic Research	FTA
CTDOT ADS Energy Data	CTDOT/New Flyer	CTE	FTA
CTDOT Transit Operations Data	CTDOT	CTDOT	FTA
CTDOT V2X Data	Robotic Research/CTDOT	Robotic Research/CTDOT	FTA
CTDOT ADS Acceptance Data	UConn	UConn	FTA

3.2 Access Level

3.2.1 Can all data from this project be shared with the public or is controlled-access required for at least some of the data?

☐ All Public Access ☒ Some/All Controlled-Access

3.2.2 Datasets Requiring Controlled-Access

Dataset Title	Reason(s) for Controlled-Access	Safeguarding Methods and Processes
CTDOT ADS Performance Data	Some autonomous vehicle data collected is proprietary	Robotic Research is the owner and steward and will allow access to non-proprietary data through the secure nSight tool. Restrictions will not interfere with the project team's ability to collect and share data for public consumption, including meeting all project objectives described in this DMP. Proprietary and non-proprietary data are defined in Section 3.4.2.

3.2.3 Informed Consent

UConn will be responsible for informed consent of human subjects and subject to an Institutional Review Board (IRB) process for its operator and passenger surveys. It will issue surveys that will be anonymous, thus the IRB process will be expedited, and the team will file for an exemption from full IRB review. UConn plans to submit approval paperwork on July 1, 2021 and expects only take at most 3 months to get approval, giving the team a possible survey period of August 2021 through the deployment. Because the ADS data from vehicle operations will not involve human subjects, UConn will not require IRB review for that part of the study.

Passengers will have both visual and audible signs at platforms informing them they are boarding an automated vehicle.

3.2.4 Access Requests

Robotic Research will assess requests for access to the nSight™ tool and non-proprietary data on a case-by-case basis. Entities of the federal government, including FTA, FHWA, NHTSA, ITS-JPO, and Volpe will have access. Any others will be evaluated on a case-by-case basis according to need and research intent.

3.2.5 Related Tools, Software and/or Code

The project team will use several software tools to collect, aggregate, and analyze data. They are the following:

Robotic Research's nSight™ platform provides an end-to-end data collection and analysis solution for studying ADS safety, through data collection and automated performance analysis of the systems demonstrated. The nSight suite is composed of the nSight Recorder™ (Onboard data recorder), nSight Uploader™ (Data disseminator to servers), and nSight After Action Review™ (AAR) (Server and Analysis Tool). Figure 1 illustrates Robotic Research's complete data collection, storage, and management process.

The nSight Recorder™ is responsible for collecting raw and processed data, and the nSight Uploader™ is responsible for distributing the data from deployments to the various servers. The ARR tool stores the data in a spatio-temporal database for rapid data analytics as well as provides the web application user interface that serves as the user's portal into viewing events of interest.

New Flyer Connect is New Flyer's fleet management software, which also provides energy management tools for monitoring real-time energy consumption.

Autonomie is an open source software tool developed for the heavy-duty trucking industry by Argonne National Labs. CTE has modified the tool specifically for zero-emission bus energy modeling, and incorporates specifications from all Altoona-tested vehicles. CTE will use Autonomie to model energy consumption over the course of bus operations, with and without the ADS active.

UConn will use tools such as Tableau, ESRI ArcGIS, and Qualtrics for maps, charts, dashboards, surveys, and other analysis.

3.2.6 Relevant Privacy and/or Security Agreements

Robotic Research will require a nondisclosure agreement (NDA) for access to the nSight AAR tool. It will not require review of research outputs provided it falls within agreed upon research scope.

3.3 Re-Use, Redistribution, and Derivative Products Policies

Dataset Title	License Used	Reason(s) for Non-Open License
CTDOT ADS Performance Data	nSight After Action Review	Robotic Research Proprietary
CTDOT ADS Energy Data	New Flyer Connect/Autonomie	N/A
CTDOT Transit Operations Data	N/A	N/A
CTDOT V2X Data	TBD	N/A
CTDOT ADS Acceptance Data	N/A	N/A

3.4 Data Storage and Retention

3.4.1 Storage Systems

Data Storage System Name	Data Storage System Type	Dataset Title(s)	Initial Storage Date	Frequency of Update	Archiving and Preservation Period
RR Amazon Web Services (AWS) Cloud Storage	Applicant – Controlled System	CTDOT ADS Performance Data	After onsite testing begins	Weekly	Five years
CTE Enterprise Dropbox	Applicant – Controlled System	CTDOT ADS Energy Data	One week after battery electric bus validation testing	Monthly	Five years
UConn Web Applications	Applicant – Controlled System	CTDOT ADS Acceptance Data	After surveys begin	Each round of surveys	Five years
CTDOT Trapeze/Transitmaster	Applicant – Controlled System	CTDOT Transit Operations Data; CTDOT V2X Data	When onsite testing begins	Nightly & Real-Time	Five plus (5+) years

3.4.2 Data Storage System Description

Robotic Research will be responsible for raw data storage and availability during the demonstration period, and for the minimum five years following the project's conclusion. Because the data collected may reach quantities that are uneconomical for long-term storage, Robotic Research proposes to store identified event data, defined by parameters agreed to by FTA, for a period of at least five years following the project's period of performance. Robotic Research can also compress raw files to reduce long-term storage requirements if necessary.

Collected data is first stored on the nSight Onboard Data Recorder (See Figure 1). The storage drive is removed from the Data Recorder, and physically loaded onto a docking station connected to an uploader computer. The

uploader computer, through an encrypted connection, can upload the data to various end points simultaneously, including Robotic Research's Large Data Server, where the data is automatically analyzed.

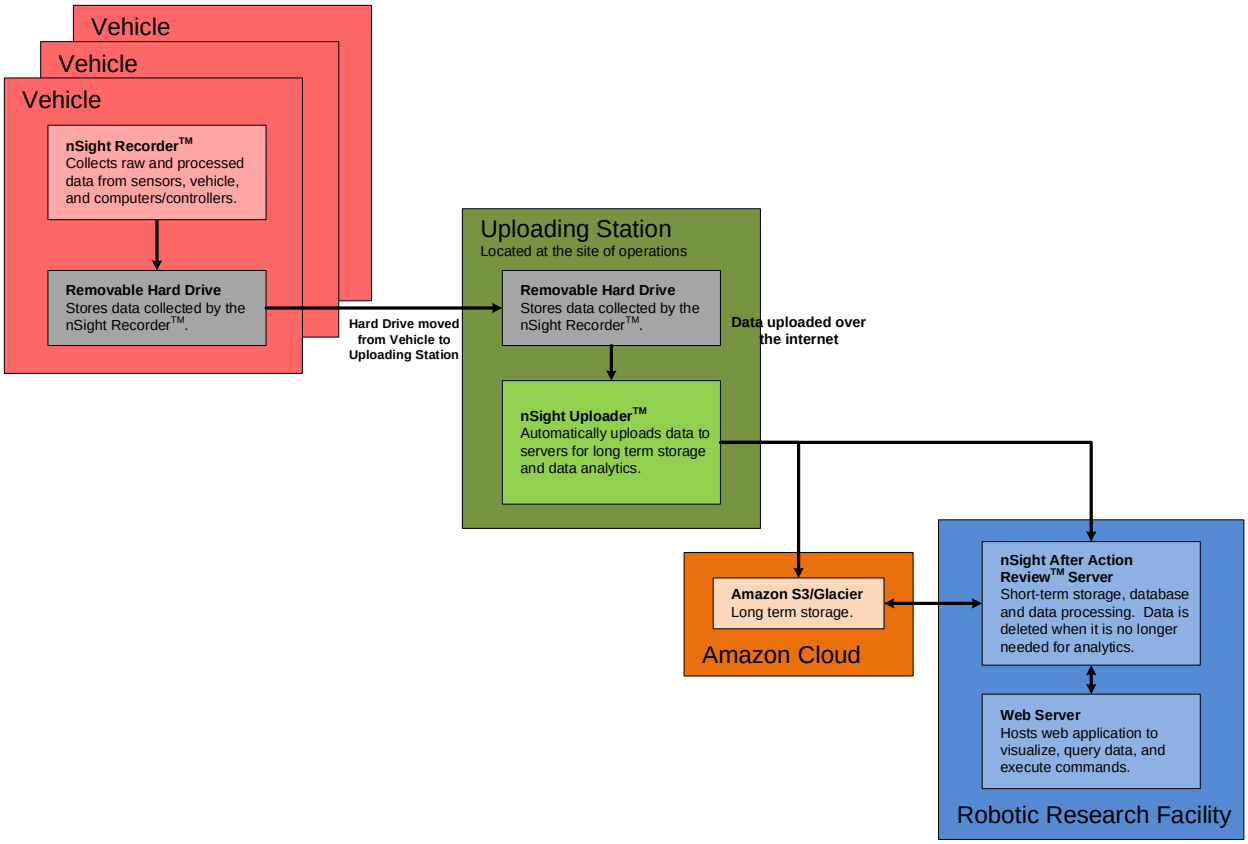

Figure 1: Proposed Data Management Process Flow Using nSight Tools

The data is also temporarily locally cached on Robotic Research's AAR server for post-processing and data analytics. Once the data analysis is completed and stored in a database, the post-processed data is pushed to the long-term cloud storage and the local data cache is removed.

The database can be spatio-temporally queried for rapid data processing. Analysis scripts automatically run on newly-updated data to identify events of interest, statistics, and infographics. For long term storage and real-time retrieval of collected data, a distributed database server setup is utilized with terabytes (TB) storage, using a redundant fail-safe hard drive configuration. The setup is easily expandable, either through the addition of more hard drives, or additional servers.

Robotic Research will establish three general categories for its data storage, analysis, and availability:

1. **Proprietary internal data**, which will remain confidential. These will be almost entirely comprised of contractor format data uploaded to Robotic Research's servers on a periodic basis (daily or weekly). This data in their unprocessed form will not be necessary for any of the proposed analysis.
2. **Non-proprietary processed data**, which will be available to project partners and any other potential authorized users. These data will represent the outputs of all preset event parameters (e.g. platform docking maneuvers, object avoidance, etc.) proposed as ADS research areas in this DMP. If necessary, the

research team can retroactively set parameters for older data and add to the body of processed data. Robotic Research will update this dataset monthly.

3. **Non-proprietary derivative data,** which will be published in project reporting and any other published research. These are the outputs of UConn's analysis, and will be stored on UConn's servers.

CTE will use its enterprise Dropbox as a primary repository for project data, backed up to a solid state drive maintained in its Atlanta office.

CTDOT data is managed by the Department of Administrative Services, Bureau of Enterprise Systems and Technology (BEST). BEST provides a data storage and management platform (Storage area Network (SAN)/Network Attached Storage (NAS)) that is used within the State of Connecticut Government. Its storage service provides fully managed disk storage space and allows for many features to help keep data accessible, safe and easy to manage.

UConn will host survey data on the cloud storage provided by its Qualtrics account, but will export and back it up on its local servers monthly.

3.4.3 Cybersecurity Policies

All systems are protected with a variety of security measures and policies. To prevent unauthorized access each system utilizes a firewall with explicit deny, and rules to only allow ingress and egress traffic as needed for system operation. They keep audit logs for all authentication attempts, and utilize Public Key Infrastructure (PKI) to authenticate any authorized connections for the purposes of test, development, debug, or diagnostics, as well as for inter-vehicle communication between systems. Communication between systems is additionally protected with minimum of HMAC SHA-384 authentication to prevent modification of messages between the various pieces of control software. Remote access is achieved through channels encrypted and secured with PKI, with user accounts operating under the principle of least privilege; granting only the minimum necessary authorization level to perform the work required. This principle is in place for both user, and system level accounts.

3.4.4 Data Security Policies and Procedures

All post-processed data is catalogued and checksummed to ensure integrity. Data critical to analysis of operation of the system is protected with a minimum of AES-256 encryption once it leaves the system, as well as during transmission. Data to be uploaded to the cloud is sent using TLS secured HTTPS, with bidirectional authentication to ensure data integrity and privacy. All access to the cloud file systems are monitored and logged, allowing for internal periodic audits to be performed to ensure no account is accessing resources that it should not be. Additionally accounts are giving the minimum access to data required to perform work.

3.4.5 Back-up and Recovery Policies and Procedures

Backup for Robotic Research data is provided by Amazon Web Services. CTE uses solid state drives to back up its primary Dropbox storage. UConn will back up all survey data on its local servers monthly. CTtransit data is backed up nightly and weekly on local servers. CTDOT has daily incremental backups for all new or changed user data that are on the network drives. Monthly backups take place the first Saturday of every month. Data is written to tapes and these tapes are sent off-site for ten years.

4. Data Standards

4.1 Data Standards

Dataset Title	Data Standard(s)	Data Standard(s) Digital Object Identifier(s) (DOI[s])	Open or Proprietary?	Data Standard(s) Rationale
CTDOT ADS Performance Data	N/A	nSight After Action Review	Proprietary	Data will be collected using the nSight Data collection system and further post processing can be done to comply with desired standards.
CTDOT V2X Data	SAE J2735	CTDOT RSUs, Robotic Research OBUs	Open	This remains TBD, pending Federal Communications Commission (FCC) regulatory action on spectrum allocation for V2X uses

4.2 Versioning

All versions will include the filename with a "_MMDDYYYY_v[#]" suffix to clearly identify document versions.

4.3 Metadata and Data Dictionary

Dataset Title	Metadata Standards Used	Metadata Discoverable (Y/N)	Data Dictionary Discoverable (Y/N)	Metadata and Data Dictionary Access
TBD	TBD	TBD	TBD	TBD

4.3.1 Metadata Description

TBD with relevant stakeholders at a later date.

5. Glossary of Terms

AAR – After-action review

ADS – Automated driving systems

IMI – Integrated Mobility Innovation

OBU – On-board unit

RSU – Roadside unit

DSRC – Dedicated short-range communications

V2I – Vehicle to infrastructure

V2X – Vehicle to everything

U.S. Department of Transportation
Federal Transit Administration

1200 New Jersey Avenue, SE

Washington, DC 20590

www.transit.dot.gov