Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D149A55949

Title: SMART Detroit MODES

Creator: Michael Hale - **ORCID:** 0009-0004-4055-6078

Affiliation: United States Department of Transportation (DOT) (transportation.gov)

Funder: United States Department of Transportation (DOT) (transportation.gov)

Funding opportunity number: 87 FR 58187

Grant: SMARTFY22N1P1G31

Template: SMART Grants Stage 1 Data Management Plan (DMP)

Project abstract:

The City of Detroit's Office of Mobility Innovation has been selected to receive a $2 million award from the USDOT Strengthening Mobility and Revolutionizing Transportation (SMART) grant program for the Detroit Mobility Optimization through Data for Equity and Safety (Detroit MODES) project.

Detroit MODES will deploy state-of-the-art intelligent infrastructure to create a series of smart intersections, collect key roadway safety data, derive real-time impacts of roadway safety interventions, and ultimately enable the City of Detroit to deliver equitable road safety outcomes where they are needed most, quicker and more effectively. The Detroit MODES project team is comprised highly specialized local and international experts, partners, and AI-driven technology. This program prioritizes working with local talent and partners, and will introduce more equitable access to workforce training opportunities for Detroiters in smart infrastructure.

Detroit MODES has the potential to revolutionize the way Detroit approaches road safety and the adoption of new transportation modalities. We believe that this project can become a model for other cities in the future and will further enhance our position as the global leader in mobility.

Start date: 09-15-2023

End date: 03-15-2025

Last modified: 03-20-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

SMART Detroit MODES

1. Detroit MODES
2. SMARTFY22N1P1G31
3. Michael Hale Jr
4. 0009-0004-4055-6078
5. Michael.hale@detroitmi.gov
6. Office of Mobility Innovation at the City of Detroit
7. mobility@detroitmi.gov - (313) 224-3400
8. https://detroitmi.gov/government/mayors-office/office-mobility-innovation
9. 1/8/2024

1.

Stakeholder Engagement Data

Qualitative data will be collected to identify the needs and ideal experiences of the users of traffic/mobility data (stakeholders). This data will be collected through in-person interviews conducted with stakeholders who have volunteered to take part in the initiative. Interviews will cover topics including current workflows for accessing and utilizing traffic/mobility data, pain points in current processes, identification of ideal future workflows, and how data should be presented, consumed, and utilized to achieve desirable outcomes. Participants will be identified as an "internal stakeholder" (City of Detroit), "external stakeholder," or "community member"; Personal Identifiable Information (PII) will not be collected. Qualitative data collected and the resulting final reports will be staged in secure folders provided by the City of Detroit.

System Data

Primary Data Source Overview

UrbanLogiq will collect and ingest two primary sources of data:

Miovision intersection vehicle, pedestrian, and other traffic count and classification data

Derq vehicle and pedestrian counts, near-miss, traffic accident, and pedestrian intent data

The team member's APIs provide access to raw data produced by its Video based Detection and count device, the Smartsense, deployed and installed in the traffic cabinet. The City of Detroit owns about 350 locations that produce real time data from the intersection.

Documentation of the data can be found here: https://docs.api.miovision.com/

This includes:

· Intersection Alerts

· Intersection Hiresdata (SPaT)

· Turning movement counts, lanes, crosswalks

· Full object classification

· Travel time data

· Intersection snapshots and images

Miovision is investigating producing further data sets that are desired by the city as part of this contract.

More of the data to be gathered in the course of the project includes:

· Data with Digital Object Identifiers (DOIs) for discoverability.

· Video data collected through direct RTSP streams.

· SPaT (Signal, Phase, and Timing) data collected through a UDP stream on the traffic network.

· Count and Event Data shared through the Derq Dashboard export and REST API.

Ingested data will take the form of events in time associated with real-world assets. Using the primary Miovision and Derq sources as concrete examples, vehicle and pedestrian count and classification events will be associated with smart intersections, roadways, or crosswalks where applicable. For instance, these data sources provide details such as vehicle or pedestrian cardinal approach direction, relative turn direction, intersection name, and location.

Third-Party Data Sources

The team is currently assessing a variety of third-party data providers. All of which consist of mobile phone movement data and/or connected vehicle data. The data that is being assessed will consist of speeds, travel times, and volumes. The timeframe of this data will be for a 1-year period.

Existing Data

The team would be reliant on historic data sources already available to the city. Examples of "existing" Department of Transportation data for potential use may include crash incident event reports, turning movement counts (TMCs), or average daily traffic counts (ADT). The team member has worked with dozens of other cities, counties, and States on similar projects, enabling accelerated ingestion and time to insight.

Some existing city asset data (e.g. road segment geometry) will also be ingested as needed.

2.

The nature of the data is diverse, encompassing video snippets, JSON files, and other relevant information from the processing pipeline. The scope involves traffic-related information, including counts, events, and Signal Phase and Timing (SPaT) data. The scale is potentially extensive, considering the system's capability to handle data from multiple intersections.

Locations that are currently deployed in the city already have an Open Data API that collects and stores this data set under previous contracts procured by the city. This will cover any intersection that already has a Miovision device producing the above data sets. Detroit has access to this data as needed.

The preferred method for data acquisition is via automated API service integration. For REST API endpoints, UrbanLogiq plans to use a batch data polling approach. Alternative options include file-based ingestion, via the preferred data formats listed below and in **Appendix A.**

Miovision leverages a combination of REST API endpoints and file export functionality, depending on the permanence of the intersection infrastructure. This is a pilot project, however, Miovision will be making an API data stream available to UrbanLogiq; it will also be making improvements to its API systems.

UrbanLogiq can ingest historical count data sources already available to the city (e.g. consultant reports). Structured data is preferred to accelerate ingestion, e.g. sources such as CSV, Excel, JSON, or XML.

Data is assumed to be in the rough order of magnitude range of ~10-100MB/ day per data source. Data will be compressed upon ingestion into UrbanLogiq systems. It is anticipated that at least 1 year worth of historical data will be ingested for this study, where applicable.

Data is expected to be of high quality and consistency so as to reduce data processing rework. This implies, for example, that the API field entries will remain consistently used throughout the life of the project.

Data for the purpose of providing contextual insights such as demographic information, weather, and other contextual data will be collected in batch updates or API depending on the source.

Details surrounding additional 3rd party data sources or existing DOT data sources will be documented in this document as the project progresses.

See **APPENDIX A - DATA SOURCE LIST** for additional details about the data sources used in this project.

Appendix A --- https://cidetroitmius-my.sharepoint.com/:f:/r/personal/gearat_detroitmi_gov/Documents/2-%20Detroit%20MODES%20-%20SMART%20Grant/DMPEP/Data%20Management%20Plan%20(DMP)/TYME%20DMPEP/DMP%20Attachment?csf=1&web=1&e=VB5c96

3.

Digital Object Identifier (DOI): Primarily used for discoverability, it is not mentioned to contain sensitive information.

Video Data: Collected through direct RTSP streams. Anonymized to prevent the inference of Personal Identifiable Information (PII).

SPaT Data: Collected through a UDP stream on the traffic network, serving traffic signal-related information.

Count and Event Data: Shared through the Derq Dashboard export and REST API. File names include the Intersection Identifier, Data type, and Timestamp for versioning. Count and event data, as well as video data, are designed not to collect any PII.

Most API data descriptions and examples can be found https://docs.api.miovision.com/

The miovision produced data sets related signal status and operation as well full road user counts and object classification. This includes understanding directional movements of each road user as well as the road user type.

Miovision collects 7 different categories today of road users. This includes: Lights, Single Unit Trucks, Articulated Trucks, Buses, Bikes, and Pedestrian Data sets also include travel time information between any node deployed in the city.

Most data are expected to be in the form of anonymized and aggregate counts at specific detector locations like intersections, crosswalks, and road segments. Intersection counts will likely be aggregated into 5 or 15-minute increments. Near-miss, traffic accident, and pedestrian intent data are assumed to be composed of anonymized events or aggregated counts.

3.1

Miovision Does not collect any personal identifiable information. Only the data points of general classification and directional movements.

·   Derq only collects usernames and contact information in its Dashboard, avoiding critical Personal Identifiable Information (PII).

· Video data processed by the system is anonymized to ensure no PII can be inferred.

· The system uses industry best practices, including secure storage with password protection on servers and secure connections for data transfer.

· Authentication requires email/password and mandatory Multi-Factor Authentication.

· Authorization is implemented through Role-Based Access Control (RBAC) to prevent unauthorized access, following a Zero Trust Architecture.

· Data access is provided through the Derq Dashboard with monitoring and logging using Audit Trail methods.

Derq is pursuing SOC 2 certification on data cybersecurity.

UrbanLogiq has its own internal data governance board to manage ethical concerns, should they arise. We have developed a series of enforceable policies and procedures regarding ethical issues, PII, data sharing, security and privacy, networked systems, information exchange, as well as information sensitivity and classification. UrbanLogiq is certified and currently in good standing with ISO 27001 and 27701 standards with attestations for 27017 and 27018.

UrbanLogiq may collect, hold, process, or share personal data, a valuable asset that needs to be suitably protected. Every care is taken to protect personal data from incidents (either accidentally or deliberately) to avoid a data protection breach that could compromise security. Our infrastructure was designed to be secure and GDPR compliant from the ground up.

UrbanLogiq employs the most rigorous data protection handling procedures to all its data, ensuring that storage and transference of data is compliant with GDPR and ISO 27001 and 27701 standards. UrbanLogiq elects to follow General Data Protection Regulation (EU) 2016/679 ("GDPR") guidance to have in place an institutional framework designed to ensure the security of all personal data during its lifecycle, including clear lines of responsibility. This applies to all staff at UrbanLogiq, including temporary, casual or agency staff and contractors, consultants, suppliers, and data processors working for, or on behalf of the company.

4.

Stored UrbanLogiq data can be used in the long term for predictive model training, university partnership research projects, etc., however only with the consent of the data controller.

Each intersection has a cost per year associated with intersection phasing data, count data, and performance data. The city currently has a negotiated rate with miovision per intersection under the existing purchased intersections.

If City of Detroit would like to add more data modules or extend the period of the data collection, we can arrange a scope of duration.

Some more values include:

· Facilitating Benchmarking and Comparisons:

o The data is retained to facilitate benchmarking and comparisons over extended periods of time. This allows for the analysis of trends and patterns in traffic-related information, contributing to a better understanding of traffic dynamics.

· Enhancing Forensic and Root-Cause Analysis:

o The data can be used to enhance forensic or root-cause analysis for specific events. By retaining data over time, Derq aims to provide valuable insights into the factors contributing to incidents or patterns observed in the traffic system.

- Supporting Additional Research Uses:

o Beyond operational use, the data will support additional research purposes. This includes efforts towards performance improvements and further developments to the Derq platform. The data's long-term retention facilitates ongoing research initiatives aimed at refining and advancing the system.

- Compliance with Data Retention Policies:

o The data retention aligns with the contractual agreements, license expiry, and the City's data retention policies. This ensures that data is retained for as long as there is a continued and valid reason to store or process it.

- City-Specific Decision Making:

o The data is made available to registered and authorized users, allowing city-specific decision-making. The city retains control over who has access to the data, enabling informed decision-making based on the insights derived from the data.

- Potential Economic Benefits:

The cost of storing data is manageable, estimated to be under $100 per intersection per year. Additionally, the use of services like AWS intelligent tiering minimizes the impact on resources, making long-term data storage economically viable.

1.

The format of the APIs is JSON, and all data formats can be found here: docs.api.Miovision.com

Within the Miovision platform, the applications already in use by the city, provides data dashboards for data exploration and use. Miovision platform also includes CSV, UTDF, and PDF export options.

More formats include:

- Video Snippets: MPG

- Traffic Count and Event Data: CSV / JSON

- SPaT:

Please see above for details around provisioned data and expected file formats.

Long-term data storage and preservation will utilize Azure commodity storage systems. The team's staff will ensure every effort is made to effectively curate and manage the data beyond the active engagement, subject to requirements. Data is stored in a mixture of data formats within UrbanLogiq's systems.

2.

Derq aims to maximize the utility of its data by employing platform-independent and non-proprietary formats. The choice of such formats ensures that the data can be accessed, shared, and utilized across different platforms and systems, promoting interoperability and long-term accessibility.

The miovision API is available and accessible to the city of Detroit. The openAPI utilizes standard API structures and JSON formats. This allows for data use outside of miovision platforms and tools.

UrbanLogiq core data storage (e.g. traffic counts, collision events) technology chiefly leverages non-proprietary formats that are easily utilizable for other purposes.

Where non-proprietary formats are not used within the UrbanLogiq platform, its systems enable transformation and structured data export into more universally accessible formats.

3.

Derq plans to use metadata standards to describe the data, and at least one metadata file will adhere to the DCAT-US v1.1 (Data Catalog Vocabulary – United States) standard. The metadata file in JSON format, following the DCAT-US v1.1 standard, will enhance the discoverability and interoperability of the data, aligning with federal guidelines for effective data sharing and dissemination.

Miovision utilizes JSON as indicated in our API documentation here: docs.api.miovision.com

UrbanLogiq will adhere to whichever metadata standards are required by the project and its stakeholders, subject to applicable requirements.

1.

Derq's data collection does not involve critical Personal Identifiable Information (PII). The system only collects usernames and contact information in its Dashboard, and this information is not considered sensitive PII.

Miovision does not collect or store any PII data. Miovision collects traffic data only including directional movements and classifications.

UrbanLogiq does not anticipate the collection of sensitive or PII data. This data management plan will be reviewed and updated in the event that changes.

2.

Derq employs various measures to protect privacy and confidentiality:

·   Video data processed by the system is anonymized to ensure no PII can be inferred.

·   Count data and event data are designed not to collect any PII.

·   The system follows industry best practices for secure storage and transfer, including password protection on servers and secure connections.

Miovision has a robust security design and is built and layers. Device level, in transit, and in cloud.

Our security policies and designs can be found here: https://miovision.com/security/

UrbanLogiq will review all incoming data for potential PII and sensitive content. See above sections around its governance board, privacy, and security considerations in data handling.

Data planned for ingestion will not be PII. Intersection traffic counts and vehicle categories are not typically accompanied by identifiable information. While crash or near-miss incident details may be linked to personal information via publicly available sources, we will strip out any PII encountered in these datasets.

Data ingested into our platform is stored in a manner that protects identity and obfuscates identifying factors using techniques like aggregation and record identifier encoding upon ingestion.

2.1

Should the need arise to store such data, relevant parties will be notified in writing and this data management plan updated.

Since Derq does not collect critical PII, deidentification is not a significant concern. Access restrictions are implemented through authentication, authorization, and Role-Based Access Control (RBAC) to prevent unauthorized access.

3.

Data access is provided through the Derq Dashboard, requiring authentication with email/password and mandatory Multi-Factor Authentication. Authorization is controlled through RBAC, ensuring only registered and authorized users have access.

Miovision does user-based authentication and permissions. Miovision has user types:

· Administrator

· Read/write

· View Only w/ Video

· View Only no video

Each user is authenticated through log in.

Data access will be limited to those who require it, adhering to privacy by design principles. UrbanLogiq data scientists and data engineers will work with this data on a regular basis. The UrbanLogiq CTO and trained data engineering staff will control data access provisioning for 3rd parties.

Information will only be made accessible that is necessary for the stated purpose. Unnecessary data, PII, and any data not required will be redacted or removed completely (as appropriate) before transfer or ingestion. Relevant parties will be notified in writing before any usage of data beyond the intended purposes.

Access to data used by UrbanLogiq is tightly controlled via several methods, including but not limited to:

· User access provisioning and deregistration

· User access management

· User policies and responsibility training

· Secure log-on

· Network access control

· Administrator access control

· Source code access control

4.

The responsibility for stewarding and protecting the data falls under the purview of Derq's DevOps/Infra Manager. They oversee data backup, recovery, and ensure compliance with industry best practices for cybersecurity. Additionally, the QA manager is in charge of the performance evaluation and quality assurance step.

The UrbanLogiq data engineering team is responsible for implementing this data management plan on a day-to-day level. Data ingestion (capture), storage, monitoring, and sharing are also the responsibility of the UrbanLogiq data engineering team. Developers at UrbanLogiq are also responsible for the documentation, accessibility, and quality of

the data and metadata.

It is the responsibility of senior management, including the CTO and Data Privacy Officer, to ensure that this plan is reviewed and revised as necessary. Managing backup and restoration activities, as well as areas such as contractual and legal compliance, platform security, privacy, and access provisioning, are primarily the responsibility of the senior developers and technical leaders of UrbanLogiq. UrbanLogiq leadership is ultimately responsible for ensuring that relevant policies will be respected.

1.

Based on this information, the intellectual property rights for the data created or used during the project will initially be held by the recipients, but a comprehensive non-exclusive copyright license is granted to US DOT. The data may need to be shared with the public as per the US DOT Public Access Plan.

The team adheres to all Canadian and United States intellectual property regulations, including both data and the platform supportive systems. We have obtained consent to use all data in our systems from the appropriate data controllers and will remove it when the consent is revoked, or the validity period expires.

All data provided through the UrbanLogiq platform has been obtained legally, either through a contract with the data provider, or directly from clients with consent. Where data is requested for sharing or re-using with additional parties, formal written consent must be obtained, and the appropriate parties will be notified.

As data becomes available for this project, intellectual property rights will be reviewed, and this data management plan will be updated accordingly.

2.

This section is N/A as the contractual rights of the data during operation will be retained through archiving.

Regarding data archiving, the details provided do not explicitly mention whether there is a requirement to transfer rights to a data archive. However, it is advisable to review the grant language or consult with relevant authorities to determine if such a requirement exists.

3.

As for licenses applicable to the data, the comprehensive non-exclusive copyright license granted to US DOT covers various rights, but additional details about specific licenses are not provided.

Data licensing details will be reviewed and updated in this document as additional information becomes available.

Enforcement of terms of use or data citation requirements through a license is not explicitly mentioned, so we may need to refer to the terms of the grant or seek legal advice to clarify this aspect.

4.

It's also recommended to review the Data Management Plan (DMP) for the project, as indicated in the information provided, to check for any additional details or requirements related to intellectual property rights, data archiving, licenses, and legal obligations.

Most of the team's systems are made in-house. UrbanLogiq's code and technology is itself copyrighted (CommunityLogiq Software Inc©). UrbanLogiq complies with relevant open-source licensing requirements where applicable for external systems.

1.

Upon completion of the research project the final data will be placed in a repository for public access, will have the required US-DACT metadata file, and a PID assigned.

2.

A link to the repository will be created at the beginning of the project.

Data will be stored within the systems described and provided as needed in accordance with the relevant considerations and needs. Data will be shared as needed per the guidelines established in this document.

3.

The team will store and archive the dataset and its associated metadata in accordance with and in support of this standard as necessary.

4.

The team will ensure the creation and maintenance of persistent identifiers throughout the data lifecycle.

5.

The team's systems allow it to tag datasets with static identifying values, which may be performed either during data ingestion or using another method to associate the appropriate metadata.