Open Mobility Foundation

Transitional Architectural Landscape

FINAL - Approved at 9/11/2019 board meeting

Introduction

This Transitional Architectural Landscape is intended to support the initial transfer of code from LADOT to the Open Mobility Foundation (OMF). It defines the initial scope of the technical activities of the OMF and identifies the set of projects to go forward.

As a transitional document, it is not intended to be a full or final expression of the OMF's technology strategy. It reflects many of the ideas and principles embraced by LADOT during initial code development and offers a short-term vision for how that will translate into the OMF. As the OMF's Technology Council and Strategy Committee begin their work, it is expected that this document will be replaced by an official Architectural Landscape that more fully captures the vision and priorities of the Foundation. That document will be reviewed and adopted per the OMF bylaws.

1.0 Core Elements for Launch

1.1 Overview

The Los Angeles Department of Transportation developed an MDS Reference Implementation to host
MDS Agency APIs used to regulate mobility providers. As this work is contributed to OMF's Github repositories for further shared development, it will be subject to a future strategy directed by the Open Mobility Foundation's Board of Directors. This document is intended to guide development of the MDS Reference Implementation in a manner that is consistent with the principles of OMF during the transition period.

1.2 Privacy and Security Principles

Privacy is both an outcome set by a specific agency's policy, and a design principle that affects technology requirements and design at the deepest level. While the OMF does not set policy for public agencies, its deliverables must enable the variety of specific outcomes that jurisdictions and stakeholders may require or desire. The OMF's Privacy, Security, and Transparency Committee will develop a set of OMF Data Protection Principles that will be used to guide the design and evolution of the MDS API interfaces and reference implementations described herein. Appendix B offers two examples of principles and practices from LADOT and the City of Minneapolis, which, along with contributions from other public agencies and private sector experts, can serve as a starting point for this work.

1.3 MDS System Reference Implementation

As part of its work, the OMF will develop a fully open source Reference Implementation of the MDS system. This implementation shall be a "cloud-neutral" architecture using only open source tools, permissive open source licenses, frameworks, and subsystems easily deployed on any public cloud or on private infrastructure.

1.4 Standards and Conformance

As MDS APIs and modules are reviewed and approved by OMF's working groups, in the process described below, they will become part of OMF's MDS "Reference Implementation" used to guide additional and supplemental work. Interoperability with that growing Reference Implementation will be a conformance criterion for additional work intended to be implemented with it. At the direction of the appropriate committees, the OMF may facilitate third party certification of commercialized MDS systems and may create an OMF-Certified logo program to acknowledge third party certification of commercialized product.

1.5 Levels of Interoperability

The goal of the OMF is to promote interoperability between a variety of stakeholders in the mobility ecosystem – mobility providers, applications developers, peer services, municipal functions. Over time, the OMF MDS Reference Implementation and data guidelines will evolve to ensure that all external APIs behave in a consistent manner. This includes role-based access control performance, versioning, database management, and scaling to name a few.

2.0 Deliverables

2.1 Overview

The OMF produces several different kinds of Deliverables (work products). Each Deliverable is developed and released by a Working Group or Committee following the procedure described in section 6 of the Bylaws

2.2 Interface Specifications and Standards

The primary deliverables of the OMF are software API specifications. Each approved API deliverable shall include a formal description using the OpenAPI framework as well as English language documentation describing the usage and semantics of the interface, and a procedure for testing it.

Upon approval as a Deliverable, the OMF website or GitHub repository shall make the Interface Specification materials available via a durable URL. The OMF shall make available a complete directory of the OMF Deliverables on the OMF website, including the URL and MD5 hash for each.

2.3 Reference Implementation Releases

The OMF shall also develop and make available a Reference Implementation of the MDS System, providing an open source implementation of the OMF MDS APIs. The Deliverable shall consist of software source code together with the English language documentation and configuration information necessary for the automated building and deployment of an MDS system.

The MDS System Reference Implementation will change rapidly, since it will include both stable code (based on approved Deliverable interfaces) and new software under development. For this reason, the OMF shall make interim releases of Reference Implementation deliverables available as needed, based on the assessment of the Working Group Steering Committee. Major releases shall follow the normal process described in Section 6 of the Bylaws.

3.0 Approved Projects

3.1 Overview

The first contributions to the OMF shall be the contents of the MDS repositories on the Los Angeles Department of Transportation (LADOT) GitHub[1], together with a copy of the Reference Implementation that LADOT is running at the time of the OMF formation.

The projects listed here are a composite of existing projects deployed by the LADOT and other cities, and projects that are a work-in-progress to support existing near term LADOT initiatives.

The Working Group (WG) assignments listed below are recommendations based on a conservative estimate of available resources. The OMF may choose to revise this Landscape in order to charter additional WGs and reallocate Deliverables.

3.2 Transitional Activities

3.2.1 Transfer the existing MDS API work from the LADOT GitHub to the OMF. This includes managing the sign-up process so that existing MDS developers can choose to be OMF Contributors.
Responsible WG: City Services and Provider Services.
OMF Deliverable: None.

3.2.2 Receive the contribution of the LADOT MDS implementation.
Responsible WG: City Services and Provider Services.
OMF Deliverable: None. (The code shall be freely available but unsupported.)

3.3 First "Baseline" Release

3.3.1 Select tools for and Set up the MDS Reference Implementation framework.
Responsible WG: City Services.
OMF Deliverable: None. (Available to MDS developers for prototyping.)

3.3.2 Port (and contribute) MDS APIs and business logic to the Reference Implementation Framework. The WG shall take advice from the Board of Directors and Technology Council as to the scheduling and functionality for this Baseline. This software release will conform to the requirements of the OMF Data Protection Principles described in section 1.3.
Responsible WG: City Services.
OMF Deliverable: Baseline release.

[1] https://github.com/CityOfLosAngeles/mobility-data-specification

Appendix A. Reference Implementation

Guiding Principles

A. The MDS code and APIs shall be cloud neutral (that is, capable of being run on any cloud provider's system).
B. The OMF MDS Reference Implementation will be suitable for interoperability and conformance testing. In particular, it will verify all API requests for strict syntactic correctness, and provide meaningful error responses whenever possible.
C. The Reference Implementation shall be based on current best practices in architecture and software engineering. The programming frameworks and tools used should be open source, have a useful life of at least three years, and should be supported by a strong community including major industry players.
D. The Reference Implementation will include contributions from many organizations and individuals. It is important that these components can coexist efficiently, securely and safely. For this reason, the Reference Implementation Framework will be prescriptive about the use of certain technologies and design patterns that facilitate the composition of microservices.
E. The Reference Implementation plays an important role in the open source software engineering methodologies of the OMF. Where appropriate, working groups may adopt Continuous Integration/Continuous Deployment (CI/CD) tools for testing new code contributions.

Appendix B. Data Protection Examples

CITY OF LOS ANGELES

CALIFORNIA

Seleta J. Reynolds
GENERAL MANAGER

ERIC GARCETTI
MAYOR

DEPARTMENT OF TRANSPORTATION
100 South Main Street, 10th Floor
Los Angeles, California 90012
(213) 972-8470
FAX (213) 972-8410

April 12, 2019

SUBJECT: LADOT DATA PROTECTION PRINCIPLES

The City of Los Angeles Department of Transportation (LADOT) works to deliver a safe, livable, and well-run transportation system throughout the region. Our vision is for all people in Los Angeles to have access to safe and affordable transportation choices that treat everyone with dignity and support vibrant, inclusive communities. As we work to achieve our responsibilities of safety, congestion relief, equity, and sustainability, we also have a responsibility to protect individual privacy and promote a transportation system free from discrimination and the exploitation of personal mobility data.

The Mobility Data Specification (MDS)[1] is designed to process vehicle data minimally necessary for our stated goals and to apply strong privacy protections and security protocols. For example, we categorize this data as Confidential under the City of Los Angeles Information Handling Guidelines -- which exempts the data from the California Public Records Act[2] -- and we apply strong access controls and de-identification measures to the data.

As part of its Dockless Mobility permitting process, the City of Los Angeles requires Mobility Service Providers (Operators) operating on the streets of Los Angeles to comply with the MDS. Such permitting rules set a consistent standard for the transfer, use, and protection of vehicle data from Operators to LADOT.

LADOT will apply the following data protection standards to all data obtained from Operators to carry out the City's and the Department's data protection responsibilities:

1) *Data categorization*: LADOT designates raw trip data as Confidential Information under the City of Los Angeles Information Technology Policy Committee (ITPC) Information Handling Guidelines. This long-standing policy for the City of Los Angeles governs the obligations of the City to protect all manners of data under its control. LADOT will withhold this Confidential Information as exempt from release under the California Public Records Act.

[1] https://github.com/CityOfLosAngeles/mobility-data-specification

[2] https://static1.squarespace.com/static/57c864609f7456745 7be9b71/t/5bd21654 71c10bf711f24edc/154 04949325 14/Information_Handling_Guidelines.pdf

2) *Data minimization*: LADOT will mandate data sets solely to meet the specific operational and safety needs of LADOT objectives in furtherance of its responsibilities and protection of the public right of way.
    a. *Aggregation, obfuscation, de-identification, and destruction*: Where possible, LADOT will aggregate, de-identify, obfuscate, or destroy raw data where we do not need single vehicle data or where we no longer need it for the management of the public right-of-way.
    b. Methodologies for aggregation, de-identification, and obfuscation of trip data will rely on industry best practices and will evolve over time as new methodologies emerge.

3) *Access limitation*: LADOT will limit access to raw trip data related to vehicles and vehicle trips to what is required for our operational and regulatory needs as established by the City Council.
    a. Law enforcement and other government agencies, whether local, state, or federal will not have access to raw trip data other than as required by law, such as a court order, subpoena, or other legal process. To be clear, the City will make no data available to law enforcement agencies through this process that is not already available to them from Operators now.
    b. Similarly, the City will only allow access to raw trip data by contractors under the LADOT Third Party Master Data License Agreement which explicitly limits the use of raw trip data to purposes directed by LADOT and as needed for LADOT's operational and regulatory needs. LADOT will prohibit use of raw trip data for any non-LADOT purposes, including for data monetization or any third party purpose.
    c. After completion of the Dockless Mobility Pilot, LADOT will create a publicly accessible transparency report discussing the types of third party requests for Dockless Mobility data that LADOT has received and how we have responded to those requests.

4) *Security*: The City will enact appropriate administrative, physical, and technical safeguards to properly secure and assure the integrity of data.
    a. Los Angeles' formal information security program and the comprehensive set of security protections and standards established by the City will govern this data as it does all other city data, including but not limited to security incident and emergency response reporting.[3]
    b. The City will conduct ongoing security testing to audit and improve security protections, consistent with the City of Los Angeles' information technology policies and practices.

5) *Transparency for the public*: The public deserve a clear description of the data used by LADOT and the ways such data is pertinent to the responsibility of protecting the public right-of-way. To that end, LADOT will publish a list of the data types collected via the MDS and the length of time that data is retained.

---

[3] The current version is *City of Los Angeles Information Security Policy Manual* dated March 8, 2017.

a. The City of Los Angeles shares certain information with the public to increase transparency, accountability, and customer service and to empower companies, individuals, and non-profit organizations with the ability to harness a vast array of useful information to improve life in our city.
b. We share data via the City of Los Angeles Open Data Portal. Before we publish any Dockless Mobility data to the Open Data Portal, LADOT will ensure the data is de-identified in accordance with established data protection methodologies.
c. LADOT will not release any Dockless Mobility data on the Open Data Portal until data de-identification and destruction treatments are implemented.

Mobility Data Methodology and Analysis

Overview

In July of 2018, the City launched a motorized foot scooter pilot program that ran through November 30, with 400 e-scooters available for shared use throughout Minneapolis. The City required participating providers to sign a license agreement which established standard data sharing and privacy requirements. The intention in requiring and using this data is outlined in the following goals:

- Maintain individuals' privacy by collecting data responsibly and thoughtfully, and anonymizing and aggregating data
- Provide transparency by publishing aggregated and anonymized data and visualizations to the City's Open Data portal for public interaction
- Determine compliance with applicable regulations as stated in license agreement
- Analyze and report on aggregated trip information; e.g. number of rides, total miles/minutes ridden, average miles/minutes per ride, breakdown by day/week/month/total pilot duration, available motorized foot scooters by day/week/month
- Analyze and report on usage through aggregated origin, destination, and route heat maps
- Inform future policy decisions such as fleet size, distribution requirements, and/or infrastructure planning by looking for trends and patterns from the pilot

Informing our work through data allows us to take an informed and proactive approach to shared mobility, and ensures that we are able to shape those services to fit our desired outcomes in providing safe, equitable, and sustainable mobility options that work for all Minneapolitans.

Looking to the future, Minneapolis hopes to build a suite of dashboards spanning all shared modes operating in the City. This will allow for efficient oversight of existing pilots and programs, better management and pricing of curbside use, as well as better planning for future modes. We also aim to be involved in defining the applicable national data standards and specifications expected from providers to ensure we have enough data to define the vision and successful metrics for shared mobility within the City, but are requiring it in a way that protects individual privacy.

Data Privacy/Sharing in License Agreements

Minneapolis has taken steps to establish clear expectations and regulations for data privacy in license agreements that are required to operate shared mobility systems in City right-of-way. This includes transparency from providers regarding their terms of use, privacy, and data sharing policies, and ensuring users' ability to opt-in to these policies as well as any potential third-party data sharing or access to location-based data. We also include provisions which ensure that personally identifiable information (PII) is not collected by or shared with the City, and that data security practices safeguard any PII collected by providers.

Regarding data sharing, we have ensured that expectations and regulations are clearly established in the license agreement, and that the City is being transparent about its intentions for use of data. The license agreements state what data the City requires from providers, how data is intended to be collected (via MDS or similar API), and a statement of purpose for how data is intended to be used. Also included is language which establishes what data may become publicly available, as well as a requirement of providers to make a publicly accessible API available.

Methodology, Assumptions, and Limitations

At the time the pilot began, a data specification called the General Bikeshare Feed Specification (GBFS) API[1] existed for sharing bikeshare information and providers used this initially for the City's data requirements for compliance. Midway through the pilot, our providers proposed giving the City access to an API endpoint based on the Mobility Data Specification (MDS) API[2] to share additional data with us as required by the license agreement. We leveraged both the GBFS API and the Provider API[3] specification to create a method for pulling data in from multiple vendors using our existing enterprise methods and tenets for data collection, storage, usage, and analysis.

We used the specifications for data provided through the MDS API, which defines both provider and agency endpoints for trips. For our analysis, we used the Provider endpoint and did not make use of the Agency endpoint. MDS also specified the existing GBFS API endpoints should be implemented for real-time availability information, so we consumed data from the GBFS *free_bike_status.json*[4] endpoint. **Appendices A** and **B** list an excerpt of fields provided by both MDS and GBFS, along with if and how the City is using these fields.

Although MDS specifies that no PII is to be sent to any agency, GPS data can be identifiable even when there is no PII provided. As a result, before consuming any trip data, we looked the stated goals of the pilot program and at previous efforts in Minneapolis to anonymize data, researched best practices and methods other agencies had employed both in and out of the state, and consulted with our City Clerk's Office to determine how to consume and store data to meet our goals and provide transparency. The Minnesota Data Practices Act informed our approach to protecting individuals' privacy while enabling us to gain the data needed to support the City's goals and provide transparency. Our intention was to store as little data as possible to be able to meet the goals above, so we analyzed the fields available in both the MDS and GBFS APIs and determined those that would be relevant.

Our immediate need was for compliance and monitoring of motorized foot scooters within the City, so we began by consuming data from the GBFS feed to create a solution for showing availability of motorized foot scooters in the City on a 15 minute polling basis. We later pulled historical MDS trip data to enable aggregate route reporting. We anonymized all data as it was consumed so that no raw data was stored.

Platform

We used a Python frontend and Microsoft SQL Server backend for consuming and storing data. We secured the servers so that only authorized users had access to the data and could not make use of it where there was no business need. We also restricted who had access to the API tokens used for each API. We used several spatial and analytical libraries in Python while consuming data to process and anonymize data in memory so that only processed data was stored. For analysis and visualization, we used R, Python, and Tableau.

We employed methods throughout the lifecycle of this project to ensure it was architected so it can be re-used for both future permitted motorized foot scooters and future expansions of the shared mobility program at the City. The image following shows the general principles we followed, which correlate to our data strategy for enabling consistent, reliable, trustworthy data in the City.

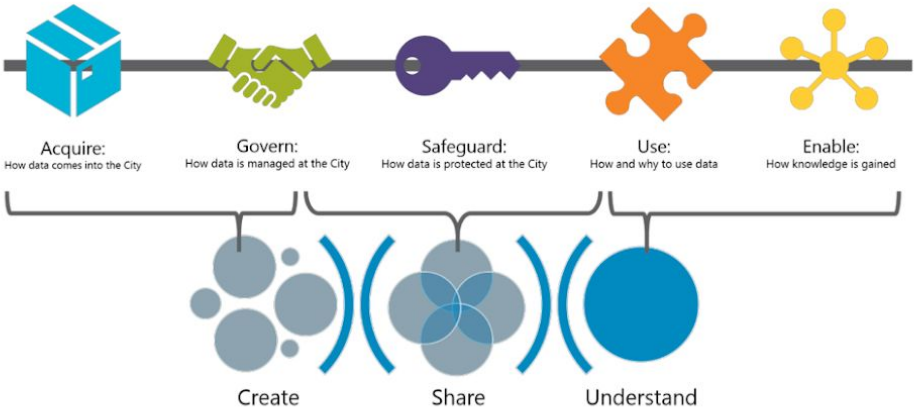
---

[1] See https://github.com/NABSA/gbfs/blob/master/gbfs.md

[2] Developed by LADOT. See https://github.com/CityOfLosAngeles/mobility-data-specification

[3] See https://github.com/CityOfLosAngeles/mobility-data-specification/tree/0.2.x/provider

[4] See https://github.com/NABSA/gbfs/blob/master/gbfs.md#free_bike_statusjson for specifications.



Privacy and Processing Methods

We employed the following methodology to anonymize data:

- All API data was processed in memory using Python, meaning no raw data was stored. Once processed, the anonymized data was stored in a secure database that only authorized users had access to.
- The trip IDs sent from MDS, while already hashed into a unique value intended for anonymization, were discarded. We generated a new unique City trip ID to make the trip harder to link back to the original source data, and stored that value instead.
- If a trip's route had no points or boundaries (e.g. the ride never went anywhere), it was discarded.
- Trip starting, ending, and route polling times were rounded to the nearest half hour at the quarter hours; e.g. if a trip started at 12:04pm, ended at 12:23pm, and a poll time was taken at 12:13pm, those times would be rounded to 12:00pm, 12:30pm, and 12:00pm respectively.
- Using the City's spatial assets for street segments, actual trip start and end points were discarded. Instead, they were binned to the closest of three points on the nearest street centerline: the street segment's start, middle, and end point (*Figure 1*):

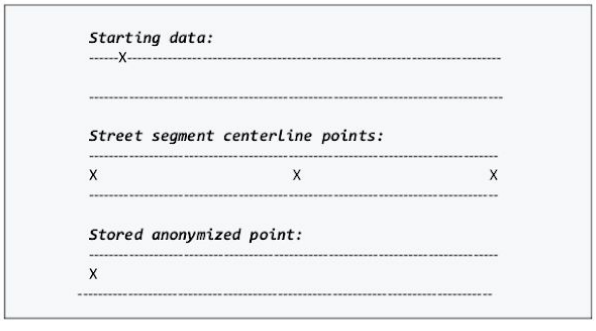


Figure 1: Centerline Anonymization Binning Methodology

This centerline anonymization follows existing methods used around the City to anonymize to the closest street segment's centroid. Because which end of the street the point was on was important for analysis, we binned