Data Governance and Management of Crowd Sourced Data Applications for Better Transportation Decisions

Adventures in Crowdsourcing: Data Management and Governance for Better Crowdsourced Data Applications (EDC-5 Webinar Series)

January 28, 2020

Subrat Mahapatra
MDOTSHA Office of CHART & ITS Development
PRESENTATION OUTLINE

1. MDOTwide Data Governance Efforts
2. MDOTSHA TSMO Data Governance and Management
3. MDOTSHA GIS Data Governance
4. Planned Waze Data Management
ABOUT MARYLAND DOT (MDOT)

MDOT comprises of six business units and an Authority
- The Secretary’s Office
- State Highway Administration
- Maryland Transit Administration
- Motor Vehicle Administration
- Maryland Port Administration
- Maryland Aviation Administration
- Maryland Transportation Authority

Multi-modal organizational framework for integrated transportation solutions

MDOT Excellerator drives the agency with ten tangible results and performance measures
Strategic, Tactical, Operational Activities at the TBUs generates a highly complex set of data streams...

Many ways to look at the Data Streams...

**By MODE** – Highway, Transit, Air, Water, Other...

**By FUNCTION** - Planning, Engineering, Operations, Maintenance, Performance Management

**By PERFORMANCE AREA** – Safety, Mobility, Asset Management, Administrative, Business Transactions

**By SOURCE** - Internal vs External Data
MDOT DATA GOVERNANCE & DATA HUB

• Data Governance & Data Hub Overview

• Where Do We Want to Be?
  o Vision, Goals, and Objectives
  o Principles to Guide Implementation

• What Did We Learn?
  o MDOT’s Current State
  o Gaps in Data Governance
  o Review of Notable Practices

• What Do We Need to Do?
  o Develop an MDOT Data Hub
  o Take Action to Address Gaps in Data Governance

• How Can We Monitor and Evaluate Data Governance Implementation?
PRINCIPLES & DIMENSIONS TO GUIDE DATA GOVERNANCE IMPLEMENTATION

People
- Staff resources and training
- Organization
- Culture

Processes
- Procedures and documentation
- Accountability and incentives

Technologies
- Data assets
- Information systems
- Tools for Data Governance implementation

1. VALUABLE
2. AVAILABLE
3. VALIDATED
4. SECURED
5. CLEAR
6. EFFICIENT
7. ACCOUNTABLE

AASHTO Reference: [https://data.transportation.org/](https://data.transportation.org/)
TSMO DATA
GOVERNANCE & MANAGEMENT
OPERATIONS & DATA COMMUNICATIONS FRAMEWORK - SYSTEM OF SYSTEMS

Transportation Network

RITIS, IMAP, TBUs Data Hub(s)

WEATHER  TRAFFIC  ROAD CONDITIONS

Transportation Management Agencies/ Companies

Travelers

ATMS/ DSS

MARYLAND DEPARTMENT OF TRANSPORTATION
STATE HIGHWAY ADMINISTRATION
MDOT SHA TSMO DATA GOVERNANCE

Source: 2017 FHWA Data Business Plan - MD Pilot

Mobility Data Programs

Tier 1
- Traffic Volume
- Traffic Speed

Tier 2
- Origin/Destination
- Accessibility
- Connected Vehicle/Automated Vehicle
- Research Data Exchange

Tier 3
- Truck Freight
- Work Zone
- Traffic Signal Timing
- Connected & Automated Vehicle

Multimodal Mobility Framework (MMF)

Data Sources
- MPOs
- CHART
- SHA-GIS
- RITIS
- Other

Planning
Real-time
CROWD SOURCED DATA APPLICATIONS

• **Real time applications**
  • System Monitoring
  • Incident Mgmt.
  • ATM/ICM

• ** Archived data applications**
  • Perf. Evaluation
  • TSMO Decision-making
HOW MUCH TRANSPORTATION DATA IN RITIS?

- **Traffic events:** 690,000 records per day: 0.006 Gb/day
- **Traffic detectors:** 65,000,000 records per day: 5 Gb/day
- **Probe vehicle data:** 8,700,000,000 records per day: 550 Gb/day
- **CCTV, weather, radio, etc:** No Stats Kept: ??? Tb/day

- UMD CATT serves as the MDOTSHA Data Warehouse.
- RITIS provides the data analytics and visualization engines for MDOTSHA Ops, Planning and Performance Management efforts.

*Source: UMD RITIS*
ELEME NTS OF AN EFFEC TIVE ARC HIVE

Data + Tools + Domain Expertise = Insights

Fusion, Statistics, & Integration Analysis & Visualization

Source: UMD RTIS
### Archiving & Distributed Computing Approaches

#### Archiving

**Traditional Relational Databases**
- MySQL
- Oracle
- MsSQL
- DB2
- SyBase
- PostgreSQL

**“NoSQL” and NewSQL**
- Hadoop architectures
- MongoDB
- Couchbase
- Riak
- Memcached
- Redis
- CouchDB
- Hazelcast
- Hbase
- Apache Cassandra

#### Distributed Computing

**Core Modules:**
- Hadoop-common
- HDFS
- MapReduce
- YARN

**Other Modules:**
- Hbase
- ZooKeeper
- Impala
- Storm

*Source: UMD RITIS*
MD & MDOT SHA GEOSPATIAL DATA
GOVERNANCE & MANAGEMENT

GIS Data Layers for Transportation Decisions

- Safety
- Existing Assets
- Traffic
- TSMO Strategies

Maryland's GIS Data Catalog
Hundreds of Services from Partners across Maryland
Policies, standards, and procedures to collect, manage, disseminate, utilize, and archive enterprise data and related applications.

- **Data Development**
  - MDOTSHA Geospatial Data Integrity-Data Submission Policy & Procedure
  - Maryland iMap Data Submission Policy
  - Highway Performance Monitoring System (HPMS)

- **Data Sharing & Security**
  - DoIT Account Management Policy

- **Application Development**
  - MDOTSHA Standard Web Map Configuration in AGOL
  - MDOTSHA Web App Documentation in AGOL

- **Enterprise GIS Data Inventory**
  - Documented in standardized template
Approach to Data Governance

- **Guiding principles**
  - Identify where success can quickly be achieved
  - Identify opportunities for data cleanup, process implementation/documentation

- **Organize by Data Class**
  - Geospatial data
  - Tabular data
  - Tabular with location data
  - Data feeds (e.g. CHART incidents)
  - Data streams (e.g. CAV data)

- **Document policies, standards and procedures for**
  - Data Development
  - Data Management & Storage
  - Data Sharing & Security
  - Application Development

- **Maintain data inventory using standardized template**
NEXT STEPS: Data Governance for Waze type Crowd Sourced Data
Waze provides MDOT various data streams through the CCP Partnership

- Alert data
- Jams data
- Municipal user indicator

Partnership Agreement on Oct 1, 2019 and MDOT SHA joined over 100 other CCP partners.
Traffic Incident Detection

Expanding geographic coverage, resolution, and timeliness of incident detection and clearance

Active Corridor Management
- Traffic Signals Timing
- Road Weather Management
- Work Zones/ Maintenance Activities

Communications & Outreach
- Major Projects
- Special Events/ Emergencies/ Evacuation

Maintenance/ Customer Service Requests
- Potholes, Guardrails, Signposts
- Malfunctioning signals, lights and other ITS assets
Thru’ EDC-5, FHWA and MDOTSHA are collaborating to develop a proof of concept platform to mainstream Waze data for Ops. and performance management. Estimated completion - end 2020.
CONTACT INFORMATION

SUBRAT MAHAPATRA
Deputy Director, TSMO & CATS
Office of CHART and ITS Development
MDOT State Highway Administration

(410) 582-5613
smahapatra@mdot.maryland.gov