Results from Data Governance Exercise

Workshop #2: March 18, 2019

Question #1: Challenges to Sharing

- **List the 3 most important data sets that you need from other organizations today.**

Topic	Names	Count
Too many choices	• Interface: with what interface we should go? Per security, performance, legal… • Formatting data properly for external consumption	2
Why collect	• No one knows what to actually do with data or why they even need it/want it • Undefined outcomes or objectives	2
Too much data! Resources /costs	• Hosting large amounts of data • Resources (people) to manage and maintain data connectivity, storage, etc. • Understanding the cost / resource requirements to get clean, useable, consistent, complete operational data • Workload • Time availability of staff to create data sets to share • Cost of integration and cost of maintenance	6
Inconsistent Access / Challenges to Access (platform)	• network architecture – ports to get inside and outside secure network • Monitoring data. Maintenance and keeping service working • Data platforms that are universal • Platform or various program • The info is on the other organization's network and we don't currently have network integration • Network / firewall issues • Differing file sharing services • Signal software is not center to center • CCTV sharing between agencies • Working on different platforms	10
Access / discovery	• Availability • Access to data • Lack of information about what data exists and which organization or department has it • Access to data for analytical purposes in an easy to use format • Not knowing which organization have what • Knowing who to share with… • Institutional bottlenecks (IT – "need to know security, proprietary, restrictions)	8

Topic	Names	Count
	• Internal – provisioning quick information aggregating key performance metrics	
Inconsistent Structures, Formats, Semantics	• Having an abundance of data but not synthesized in a format to share • Incompatible systems or data formats • Structure of data collected (amount of detail needed or not needed) • Not having all data consolidated • Compatible formats • Understanding data in detail (type, scale, temporal, coverage, estimated vs. observed, etc) • Data sharing or integration takes a long time as a result of data type inconsistency and mapping difficulties • Data formats: what data with what format and what schedule • Useable data format to integrate in their system • Software integration of system capturing same or similar data (i.e., CAD/AVL in use by everyone in region but software is different) • Inconsistent / incompatible data formats coming form multiple vendors or contractors	11
Inconsistent Geographic data	• Our geographic boundaries are not jurisdictional or census tracts or other standard. It makes giving and getting data sets very difficult	1
Data responsibility	• Unique point of contact of lack of data sharing protocol • Culture of cooperation and openness • Process of requesting • Getting permissions to share data through our legal department • Identifying the point of contact. Protocol, person knowledgeable about data source, network connection, firewall, etc. for access to a dataset • Information technology coordination • Organizational IT departments • Appropriate point of contact • Finding the right contact party • Slow response	10
Privacy	• Privacy / privacy data public data • Privacy • I don't know what I don't know! Understanding the data landscape. I need on inventory, schema, and dictionary that stays up-to-data and accessible • Privacy issues	4
Data Restrictions (license)	• Who owns the data • Restrictions on sharing due to contracting agreements / licensing • Private company data sharing restrictions • Third party data sharing • Unknown legal limitations • Policy restrictions	9

Topic	Names	Count
	• Data is licensed so legal agreements must be signed in order to share data • Data sharing agreements (too restrictive or nebulous) for data obtained / procured from private sector data vendors / sources • Legal issues	
Metadata / quality	• Data accuracy • Need to record/retrieve camera footage from ratesign cams, but GDOTs cameras aren't recorded due to policy • Various and conflictions uses of the data • Accurate data • Differing needs / efforts across organization • Up to date data (regarding project implementation)	6

Question #2: Current Data Set Needs

- What are **3 major challenges** to your organization sharing data with other organizations today?

Topic	Names	Count
Planned / Project data	• Updated project implementation • Planned / upcoming transportation project data • RTP project level performance measures • TIP project level performance measures • Unit costs for materials (related to construction) • Current capital investment plans • Coordination for review of projects • Notification from GDOT for projects they are conducting in air jurisdictions (deployment horizons) • Land use plans	9
Asset Management	• Device asset / maintenance data • Regional shared infrastructure resources • Asset inventory data • Maintenance data (cost for maintaining asset) • Natting of IP addresses for ATSPM	5
Network performance **Speed / Travel time**	• Real time arterial performance • Operational data (ridership, hours, miles) • Travel speeds and volumes • INRIX speed • Travel times • Travel speeds • Private data – realtime data from 3rd parties like Motorola solutions • Speed data • Live roadside data from GDOT • Real time data • Local traffic • (location based data) • GDOT Navigator reports • Speed data	12
Signal data	• Real time traffic signal data • High resolution signal data • Signal updates	3
Traffic counts	• Traffic counts at nearby locations • Traffic data • Vehicle volume • Traffic analysis and data (level of service, counts) • Traffic counts and turn movement	5
Analysis / Performance measures	• Equitable justice analysis (including environmental justice) • DASH includes all the above performance measures • Regional performance • Vehicle occupancy / transit ridership	7

Topic	Names	Count
	• Congestion data • ATSPM data • Origin-destination data (region-wide) including trucks and commercial vehicles • Corridor-level travel time reliability	
Transit	• Transit asset • GTFS clean data for transit • Realtime GTFS • Transit info could be interesting • MARTA data	5
??	• Platoon releases from adjacent jurisdiction • SE data?	
Crash / safety data	• Live crash data • Crash / incident data • Consistent crash reporting • Accident data • Crash data • Safety numbers in real or near real time • Real-time crash data with attributes	7
Work Zone / closures	• TIR / construction data • Real time road closure data with accurate location information	2
Geography and demographic	• Elected official district boundaries • Land use and related economic / demographic data • Tax digest information • Road centerline data updated with accurate location information • Address point data updated with location information	5
Camera	• More camera data and resulting analysis • Camera data, streaming from GDOT • Camera feeds • Historical (recorded) camera feeds of rate sign cameras	4
Incident	• Near real-time incident data (accidents, break downs) • Incident data • Roadway clearance time	3
Traffic signal	• Traffic signal timing plans	1
TSMO	• TSMO info	1
Multimodal	• Mobility delays (bike, ped, bus) • Rideshare data	2
Lane closures	• Lane closures (planned and unplanned)	1

Question #3: Future Data Set Needs

- **Given projects under deployment, list the 3 most important data sets you expect to share in the future?**

Topic	Names	Count
Traffic Data	• Traffic signal data • Traffic data • Traffic • Traffic counts by time of day/location/mode • Freight related demand by time of day and by geography • Bottlenecks • Traffic counts and turn movements	7
Travel Time	• BlueTOAD travel time data • Travel times • Travel time data	3
Smart city	• Smart city pilot project data SPaT, smart lighting, signal priority preemption • SPaT • SPaT •	3
CAV	• Autonomous vehicle/shuttle ridership by TOD and amount of travelers • DSRC • V2I data • Connected vehicle message format uniformity • Real time performance measure from CV • AV/CV data • Connected vehicle data • CV data as market proliferation grows • CV data • BSM data stream (basic safety message from connected vehicles0	10
Occupancy, volume	• Vehicle occupancy patterns in express lane corridors • Realtime data in express lanes (trips, volume)	2
Dynamic pricing	• Dynamic pricing data	1
Crash data	• Crash statistics • Revamp of crash reporting forms to include clean/consistent crash data	2
Pavement data	• Pavement data	1
Fleet Management	• Vehicle trajectory information • Computer aided dispatch (connected to traffic control system)	2
Transit	• GTFS, GTFS-realtime • GTFS-realtime	4 +4

Topic	Names	Count
	• Real time transit data • Transit Transit Performance • Regional transit reliability • Farebox data • Regional transit operational KPIs (calculated across data types) • On time performance / connections	
Crowd sourced data & big data sources	• Aggregated user location information for example, a device user that is using a tree map or location app in a device has agreed to share their location data. That data is compiled into an aggregated (anonymized) data set to display locations where users are present. Location data such as roads, walkways, etc. can be derived and use to detect location where map data updates. • User generated content • weather	3
Performance	• Transportation performance • Manage lane performance • Travel time reliability • Future planning performance data such as visual evaluation (large scale not detail) • Comprehensive performance visualization tool (DASH TDM option planit type) • Usage / utilization (modal) • Emissions reductions / inputs • Performance metrics from smart corridor technologies (pre-emption, bus priority)	8
Project, planning and economic data	• Planned project data • Current capital investment plan • Economic impacts at major projects • economic development info (zoning, variances, reviews) • factors to help determine where to locate charging stations (e.g., population, volume, car-ownership, etc.)	5
General	• Innovative data collection	1
Payment / Security	• Security credentialing tokens / information • Decentralized ledger information	2
Video	• GDOT rate sign camera footage	1
Closure	• Right of way closure information	1
Safety	• HSM predictive analytic data for data driven safety • Predictive analytics based on crash data	2
Asset Mgmt	• Connected vehicle OBU/RSU asset management data • Asset inventory using software • Roadway asset information • Asset management / vulnerable infrastructure	4

Topic	Names	Count
Geography / Map	• Community improvement district boundaries	1
Multimodal	• Shared bike • TNC, uber/lyft • Vanpool • Shared scooter • carpool	5