Data Governance and Big Data: Emerging Challenges
Florida’s Transportation System

- **123,099** Centerline Miles of Public Roads
- **12,103** Centerline Miles of the State Highway System (SHS)
- **4,440** Centerline Miles of Strategic Intermodal System (SIS)
- **7,007** Bridges Maintained by FDOT (2019)
- **15 M** Registered Motor Vehicles
- **30** Urban Transit Systems
- **20** Rural Transit Systems
- **7,419** Miles of Bicycle Facilities on Non-Freeway SHS
- **3,512** Miles of Pedestrian Facilities on Urban Non-Freeway SHS
- **20** Commercial Airports
- **15** Deep Water Seaports
- **2** Active Spaceports
- **2,743** Miles of Mainline Railroad Track
Common Challenges

**Person-to-Person Sharing**
A heavy reliance on getting data from individuals instead of accessing data directly from applications and reporting tools.

**Extensive Manual Processing**
A prevalence of manual, home grown processes for copying and transferring data (spreadsheets).

**Limited or No Standardization**
Extensive amount of effort required to relate information from multiple data sources.

**Lack of Sustainability**
A data driven agency poses new challenges.
Organizational Gaps

- **Data organization inconsistencies across SharePoint and local drives**
  - Technical Group Gaps
  - Network bandwidth issues
  - Users don’t know who to contact

- **Lack of consistent data controls**
  - Data Steward Working Groups
  - Data Custodians
  - Meeting and formal communications
  - Identification of decision-making processes

- **Lack of centralized management of extracts, transformations, and load**
  - Data Integration Gaps
  - Issues completing nightly data downloads
  - Challenges in merging information from multiple sources

- **Insufficient data extract tools**
  - Inadequate data extract tools
  - Inconsistent data entry rules
  - Unclear system data definition

- **Duplicate data entry**
  - System integration timing

- **Master Data Management Gaps**
  - Inconsistent data entry rules
  - Unclear system data definition

- **Legacy data repository does not have all required data elements**

- **Transactions with the data**
  - Engineering culture drives development
  - Engineering culture from collecting data to exploiting data

- **Need to refocus culture**
  - Clearly defined business and technical data roles

- **Transition to modern data driven organization is just getting underway**
  - Distinct of standardization process

- **Data Strategy and Architecture Gaps**
  - Need for “one stop shop” for data
  - Lack of business perspective
  - Inconsistent data identifiers
  - Lack of historical data retention

- **Lack of data classification**
  - Lack of data classification

- **Security and Privacy Gaps**
  - Security Approaches are overwhelmed by security requests
  - Security Administrators overwhelmed by manual labor intensive process

- **Lack of unified enterprise security procedures**
  - Users don’t know who to contact
  - Users don’t know which process to follow

- **Data Quality Gaps**
  - Timely access to internal data
  - Timely access to external data
  - Lack of accurate and timely data entry
  - Re-work related to GIS data

- **Lack of effective master data management**
  - Data Quality Gaps
  - Lack of effective data quality issues

- **Organizational Alignment Gaps**
  - Lack of effective innovation execution process
  - Lack of effective organizational change management process

- **Integration of internal and external data needed**

- **Lack of effective internal and external feedback quality loops**

- **Data Quality and Data Usability**
  - Lack of data confidence issues

- **Data Strategy and Architecture Gaps**
  - Lack of data Usability
Impacts of Big Data

- **Managing Large Datasets**
  - Exponential Increase in Data
  - Exceeds Current Storage Capabilities
  - Rise in Unstructured Data
  - Rise in Number of Data Formats

- **Challenges for ‘Real-Time’**
  - Evolving Nature of Tools and Technologies Needed
  - Common Definition of ‘Real-Time’
  - Irrelevant Architecture
  - Internal Process Adjustments
Impacts of Big Data

- **Security**
  - Relating Data from Many Sources
  - Need for Various Data Collection Strategies
  - Data Inconsistencies

- **Staffing**
  - Shortage of Expertise
    - Big Data
    - Machine Learning
    - Artificial Intelligence
  - Technology Cultural Change
  - Employee Resistance to Change
Florida DOT’s Solution
Reliable, Organized and Accurate Data Sharing

2014: Agency-Wide Analysis

**IT Strategic Plan** – FDOT undertook an initiative to develop an enterprise-wide Information Technology Strategic Plan.

**IT Assessment** - The Office of Information Technology (OIT) sponsored a critical assessment of the Department’s information technology capabilities, personnel and infrastructure.

**Technology Alignment** - The intent of the process was to align the Department’s technology assets with its functional business units.

2015: ROADS Introduction

**ROADS** - In March 2015, the Reliable, Organized, and Accurate Data Sharing (ROADS) Project began.

**Goal** – The goal of the ROADS Initiative is to improve data reliability and simplify data sharing across FDOT to have readily available and accurate data to make informed decisions.
The ROADS Initiative is helping to close the data/information gaps identified by:

**People:** Managing a formal data governance structure to make key decisions related to data/information.

**Process:** Training FDOT on the Data Governance Component Model and implementing standard processes and routines to provide a formal approach to data governance.

**Technology:** Providing common standardized BI/DW tools, technologies and frameworks that will be used across FDOT to make data/information more accessible.
## Benefits to Enterprise Governance

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability</td>
<td>Ensuring information is secure, accurate, reliable and at the appropriate level to empower you do your job better.</td>
</tr>
<tr>
<td>Accessibility</td>
<td>Providing the ability to access relevant business data more quickly and efficiently by knowing where to find it.</td>
</tr>
<tr>
<td>Timeliness</td>
<td>Reducing the amount of time to locate the data you need and more time to analyze the data.</td>
</tr>
<tr>
<td>Productivity</td>
<td>Effectively sharing information across our organization to enable better and faster decisions.</td>
</tr>
<tr>
<td>Integration</td>
<td>Enabling a greater capability to link data together from different Districts, Florida’s Turnpike Enterprise, functional areas and systems.</td>
</tr>
<tr>
<td>Sharing</td>
<td>Removing barriers currently in place that prevent the efficient sharing of information.</td>
</tr>
</tbody>
</table>
Established Data Governance Roles

ROADS Executive Team
- Champions Data Quality Improvement
- Represents Data Governance Stakeholders
- Prioritizes Data Issues
- Sets Data Governance Rules and Procedures

Enterprise Data Stewards
- Leads the Data Steward and Data Custodian Working Group
- Acts as a Liaison between ROADS Executive Team and the Data Stewards and Data Custodians
- Ensures Data Governance Compliance

Data Stewards
- Collaborates across Departments on data governance activities
- Supports Business Intelligence/

Data Custodians
- Supports and implements data governance activities
- Supports and prepares documentation
- Maintains knowledge of data structure
- Resolves data quality defects
- Provides data access to approved users
- Implements data confidentiality and security requirements

Data Warehouse Initiatives
- Defines and maintains business metadata
- Coordinates adequate documentation
- Develops data quality process and standards
- Remediates data quality defects
- Approves data access
- Manages data confidentiality and security classification
Effective data governance programs have many processes.

Each component is critical to the overall success of the program.

The inner components are related.

To achieve success, all components must be addressed.
Effective Data Governance

- Open Platforms
- National and International Data Standards
- Further Enable Self-Service (GIS Platform)
- Build Upon Existing Data Governance Foundation
- Support Innovation and Modernization
- Established Frameworks: Ongoing Reassessment of Tools, Processes and Procedures
- Increase Staff Awareness
- Formal Management of Vital Data and Technology Assets
- Continue Efforts to Promote Data Governance as a Priority
- Promote the Value of Managed Data through ROADS Initiative
Thank You