A Strategy for Managing Freight Forecasting Data Resources

Center for Information Management & Educational Services

Florida State University

Rebecca Augustyniak, MS, Principal Investigator
Amy Finley, MS, Project Director, Co-Principal Investigator
Brian Arsenault, MS, Research Lead

Disclaimer

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation.

Contents

Introduction

The Florida Department of Transportation (FDOT) would like to get a more comprehensive picture of the movement of freight within the state. As part of this initiative, The Transportation Statistics Office (TranStat) at the FDOT anticipates being assigned the responsibility for acquiring, storing, and maintaining freight-related data associated with various transportation modes. TranStat is currently responsible for the collection, storage, and maintenance of highway and traffic data.

To begin planning for their future responsibilities, TranStat contracted with the Center for Information Management and Educational Services (CIMES) at Florida State University (FSU) to provide research and analysis services.

This report proposes a strategy for the creation of an FDOT Freight Data Clearinghouse and discusses its components, staffing, and responsibilities. The strategy proposes to meet TranStat's future needs for acquiring, organizing, storing, maintaining, and assuring the quality of freight-related data as well as providing specifications for access, delivery and reporting.

Project Overview

The first step in the research process was a review of the literature. CIMES conducted a comprehensive search of data storage and maintenance practices in relevant bibliographic databases, electronic resources, and the Web. CIMES staff reviewed the resulting set of best practice literature and resources and documented it in a resource guide[1].

CIMES conducted interviews with stakeholders at the FDOT to determine freight data monitoring needs from the various modal offices (Aviation, Rail and Motor Carrier, Transit, Seaport), the Office of Policy Planning, and the Systems Planning Office. The interviewees discussed how the data sources they use meet their needs and what needs are not being met. After analyzing and reviewing the interviews and data sources, CIMES staff documented and discussed the findings in a data needs analysis and audit report[2].

CIMES also conducted a telephone interview with Cesar Quiroga, Senior Research Engineer and Manager in the San Antonio Office of the Texas Transportation Institute. He is one of the authors of *Guidance for Developing a Freight Transportation Data Architecture*, published in 2011 by the National Cooperative Freight Research Program as *Report 9*.

Finally, CIMES developed a Web-based survey and sent it to freight and intermodal freight contacts at the department of transportation in each of the 50 states and the District of Columbia. The results of the survey were integrated with information from the literature review and the previously conducted interviews. The findings were included in a summary of best practices report[3].

---

[1] *Freight Forecasting Resource Guide*
[2] *Freight Forecasting Data Needs Analysis and Audit Report*
[3] *Freight Forecasting Summary of Best Practices*

Introduction

Research Step		Report
1. Literature Review	→	Resource Guide
2. FDOT Stakeholder Interviews	→	Data Needs Analysis and Audit
3. Expert Interview		
	→	Summary of Best Practices
4. States Survey		

Research Discoveries

A wide variety of freight-related data is collected publicly and privately, at different geographic resolutions. The data support planning, creating and maintaining the transportation system used by businesses to move cargo and freight across multiple modes. Whether due to geography, planning, or policy, each state has different questions regarding the movement of freight within and across their boundaries, and the states vary in the extent of their focus on intermodal freight movement.

The data describing freight movement are more complete in some modes than others. In Florida, coverage of freight movement by sea and air is better than coverage in the rail and motor carrier modes. Gaps in the data arise, in some cases, because the data are not collected. In other cases, the data are proprietary and subject to rules that require aggregation to protect confidentiality. This, in turn, reduces the usefulness of the data in describing freight movement at a state or local level. Because of these gaps, it is not possible to follow freight as it crosses modes.

Florida has no central access point to data purchased and used by the FDOT offices and districts. Duplication of effort and expense arises because of a lack of effective communication of available data or opportunities for sharing data.

FDOT has challenges with providing access to systems that they manage. FDOT staff can easily access systems behind the FDOT firewall, but consultants and other interested parties require VPN accounts to gain access. Access through VPN must be requested and approved, and has been reported to be difficult to set up and maintain. A third-party vendor manages VPN access, charging an annual fee for the account holder. Some systems are behind an additional firewall; this level of access must also be approved, but is managed by the FDOT IT department. The second level access requires a change in password every three months. Failure to change a password on this schedule requires contacting the IT department to gain access again.

Systems on the Internet, not behind the firewall, and housed by FDOT have certain restrictions that require them to be housed and maintained by the IT department. This creates obstacles to content management since all change requests must be submitted to the IT staff rather than being directly managed by the office responsible for the content. This cumbersome process often leads to weeks or months of delay. The Office of Policy Planning has moved the Enterprise Strategic Intermodal System (eSIS), from the Internet to behind the firewall, to manage the content more effectively.

Contracting with a consultant to house and provide system access for both data consumers and content managers seems to provide the most flexible and broadest access configuration. This would require

recurring funding to the vendor for maintaining the system. Lack of funding in a given year could result in lack of access to a system housed with a contractor.

Goals

A strategy to meet TranStat's needs for managing freight forecasting data resources involves two sets of goals. The first set of goals covers the period when the new responsibilities for freight-related data are assigned to TranStat. The goals address this change and the research findings that include duplication of effort, data silos, lack of data awareness, and data gaps. This set of goals also forms the foundation that will support future goals. The first goals are discussed here. Future goals are discussed later in this report.

The first goal is that the strategy must satisfy the planning and information needs of the FDOT offices that are consumers of freight-related data. The modal and planning offices already use many sources of freight data as identified in the Data Needs Analysis and Audit Report prepared by CIMES.

Strategic Goals

• Satisfy FDOT planning needs	• Enhance data usability
• Reduce redundancy	• Enable easy data sharing
• Support sharing costs	• Incorporate spatial components
• Improve awareness of data	• Fill data gaps
• Provide access to data	• Protect proprietary data
• Eliminate institutional silos	

The strategy should reduce redundancy and support sharing costs of acquiring data. Currently some cost sharing is accomplished by negotiating extra licenses for Transearch data from IHS Global Insight for use by FDOT and other planning organizations. However, some redundant data are still being purchased multiple times at the district and state levels.

Another goal is to improve awareness of freight-related data and access to them. A belief shared by the modal and planning offices at FDOT is that there are institutional data silos. It is difficult for employees to know what data are available or how those data could be useful to them. A freight data virtual library with expertise organizing data in meaningful ways and providing access to them could remove barriers that currently keep the data in silos. Improving awareness and access may also help to reduce redundancy and support cost sharing.

The strategy should enhance the usability of the freight-related data, making it both easier to use and more valuable to use for creating new information. One of TranStat's visions for its management of freight-related data is to use technology to represent data in a form so that users can manipulate them quantitatively and spatially to create new information. Ways to satisfy this goal are to adopt standards to enhance data sharing and to use shared and transferable data whenever possible. Another way to satisfy this goal is to determine which data have a GIS component and may be spatially represented.

Finally, one determination found in the research is that there is no way currently to accurately create a picture of the intermodal movement of freight across all modes of transportation in Florida. Parts of the overall image are missing because the data are not collected or are blurry because the data are aggregated to protect proprietary and confidential data to a degree that makes them difficult to use at a state-level resolution. A goal should be to improve the picture by filling these gaps in the data while still protecting proprietary data.

Proposed Strategy

FDOT Freight Data Clearinghouse

The proposed strategy for meeting TranStat's needs and the goals identified above is the creation of a Florida Department of Transportation Freight Data Clearinghouse.

There are several benefits of this strategy. A Clearinghouse provides stakeholders with ready access to reliable, consistent, and relevant information, allowing them to pursue their mission. Clearinghouses facilitate innovation, increase efficiency, and reduce redundancy by actively collecting information, organizing it, and creating reliable access to it. They can also provide stakeholders with tools for networking and collaborating, in addition to services that connect users to relevant and useful data.

Other states are creating virtual libraries for their data. Of the states that responded to CIMES' survey, California and Washington are two examples. California has an online library of GIS data that provides metadata of the files, information on when they were last reviewed and updated, and forms to request the data. Washington responded to the survey that it is in the process of developing an online freight data library to provide a "one-stop repository" that will "support planning, programming, communicating, and other needs."

An important observation regarding these cases of virtual data libraries is that all of the systems are housed outside of each state's IT department. California specifically mentioned that the arrangement was to avoid issues with the IT department. The research also identified the Oregon Freight Data Mart, hosted at a university rather than the state DOT.

Clearinghouse Services

There would be three aspects of or services provided by the FDOT Freight Data Clearinghouse: Information Systems, the Virtual Library, and the Data Warehouse.

The Information Systems aspect would design and manage any networked information communication systems for the Clearinghouse. These may include interfaces for accessing and using the freight data, and also include tools that support collaboration between the various stakeholders.

The Virtual Library service would organize the available data, help users define their data needs, and make intelligent connections between the users and data. It would also monitor relevant information streams for new information products and research results that would fill data gaps. In conjunction with the Information Systems service, the Virtual Library service would provide online access to both useful information about the data, such as metadata and research, and the freight-related data itself.

The Data Warehouse would store and maintain the freight data.

Clearinghouse Staff

FDOT Freight Data Clearinghouse	
Services	**Staff**
Information Systems	System Designer
Virtual Library	Librarian/Educator
Data Warehouse	Freight Data Manager

The Information Systems component of the Clearinghouse would have a Systems Designer responsible for creating the interfaces for using freight data and for creating the networked collaboration and communication tools.

A Librarian/Educator would staff the Virtual Library. The Librarian's responsibilities would include organizing and evaluating data sources, identifying and assessing data needs, and using the organized data to satisfy the data needs. As the primary contact at the Clearinghouse for freight data users, the Librarian would be in an ideal position to make connections between those users to help foster innovations. This individual would also have an explicit role as an educator to raise awareness about ongoing research and developments regarding freight data.

The Data Warehouse would have a Freight Database Manager responsible for designing and managing the data warehouse. This is a technical position that would administer the storage and maintenance of freight data in its various formats.

Information Systems User Roles

The following groups are the potential users of FDOT Freight Data Clearinghouse information systems:

• FDOT content managers	This group would consist of staff responsible for managing assigned content areas or data.
• FDOT data consumers	This group consists of the planning and modal office staff, administrators, and other staff including district staff across the state.
• Consultant data consumers	This group consists of vendors either currently under contract with the FDOT, or interested in pursuing a contract with the FDOT.
• Public data consumers	This group consists of individuals or organizations interested in finding out about or obtaining freight data.

Clearinghouse Responsibilities

The Clearinghouse will have several concurrent responsibilities to fulfill the goals of the strategy for meeting TranStat's needs. This report discusses data identification, collection, and acquisition; data organization, storage, data maintenance, and quality assurance, data access, and data delivery and reporting. Each responsibility will be met by performing several ongoing tasks.

Data Identification, Collection, and Acquisition

Tasks under this responsibility are:

- Centralize the acquisition of freight-related data within the Freight Data Clearinghouse. This will prevent the duplication of both effort and expense that comes from redundant data acquisition.
- Assess data needs to identify and prioritize the freight questions that need to be answered.
- Identify the current data sources used by FDOT.
- Determine if current sources are being efficiently shared, especially if the source is purchased.
- Identify additional data sources required to address identified needs and freight questions.
- Acquire additional data.
- Ensure that additional data sources are efficiently shared.

All of these would be ongoing tasks to be performed by the Clearinghouse.

Identification and acquisition of data will usually come from publicly or commercially available sources. If a data need cannot be met with the available sources, the Clearinghouse may be involved with outsourcing the collection of custom data.

As data are centralized and acquired, those sources that include GIS components would be identified and documented to support future goals.

Data Identification, Collection, and Acquisition	
Tasks	**Responsible Staff**
Centralize data acquisition	Librarian/Educator
Assess data needs	Librarian/Educator
Identify current and new data sources	Librarian/Educator, Freight Data Manager
Acquire data	Librarian/Educator
Ensure efficient sharing	Librarian/Educator, System Designer

Data Organization

Organizing the freight data and useful information about the freight data would make it possible to connect users and data to meet planning and reporting needs. The tasks under this responsibility would also be ongoing:

- Adopt or create standards for data organization.
- Catalog and index data sources.
- Create an online virtual library with tools that search the catalog or index of information about the data sources.

Proposed Strategy

This responsibility involves describing the data sources to create access points for connecting users to data and for grouping similar data together. The standards used should describe what a data source is and what it is about in a meaningful way to prospective users.

In addition to the data, the Clearinghouse would also organize peripheral information such as research and published reports created from the freight data.

Data Organization	
Tasks	**Responsible Staff**
Adopt or create standards	Librarian/Educator, System Designer
Catalog and index data sources	Librarian/Educator
Create virtual library	Librarian/Educator, System Designer

Data Storage

Acquisition of data would be centralized through the Clearinghouse to prevent duplication of effort and cost, but the data storage would employ a hybrid solution of centralized and distributed storage. Purchased data and data requiring special handling would be stored centrally in the data warehouse along with any data that have been identified as benefitting from local storage. Other data that may be useful but that do not need to be stored in the data warehouse would be cataloged but remain at the source to be downloaded when necessary. Data produced by outside vendors for specific FDOT projects may also be stored with the vendor rather than in the data warehouse, but the existence of the data and their description would be cataloged by the Clearinghouse.

Peripheral information such as research reports and other documents related to the freight data would be stored locally using content management tools. Clearinghouse staff could design and create these tools:

- Create data warehouse to store data in a variety of formats.
- Create content management tools to handle peripheral information.
- Create policies for storing sensitive and secure data.

Since IT department issues are a common theme with several states, housing the data warehouse where the Clearinghouse will have easy access to and control over it is an important consideration.

Data Storage	
Tasks	**Responsible Staff**
Create data warehouse	Freight Data Manager
Create content management tools	System Designer, Librarian/Educator
Create policies for sensitive data	Freight Data Manager, Librarian/Educator

Data Maintenance

Maintaining the data in the warehouse would involve keeping it up to date:

- • Update data and keep it current.
- • Archive data that is not current.

The practice of maintaining the data in the warehouse would also necessarily involve the Virtual Library and information systems as the associated catalog records are also updated and the updates are communicated to FDOT stakeholders.

Data Maintenance	
Tasks	**Responsible Staff**
Update data	Freight Data Manager, Librarian/Educator, System Designer
Archive data	Freight Data Manager, Librarian/Educator, System Designer

Data Quality Assurance

According to interviews with FDOT stakeholders and the states' survey, quality assurance usually rests with the source providing the data. It is possible, however, that some sources of data may be more reliable than others. Clearinghouse staff would be involved in evaluating which sources of data have the best quality data.

Various freight data sources cover different time periods or are "frozen in time" because the source is no longer collected. One frozen source is the Vehicle Inventory and Use Survey (VIUS). It was last collected in 2002 but is combined with more recent data to answer questions regarding commodity flows. The Librarian could educate stakeholders regarding data concurrency issues when they arise.

- • Evaluate data sources.
- • Support data concurrency.

Data Quality Assurance	
Tasks	**Responsible Staff**
Evaluate data sources	Librarian/Educator
Support data concurrency	Librarian/Educator

Data Access

There should be a central access point for finding information about freight data. Ideally this would be easily accessible by all the data consumers, such as a public Web site. All users should have the ability to search the Freight data catalog, view record results, download unrestricted data sets, and request assistance or information. FDOT content managers should have access to content management tools for the content area assigned to them. FDOT data consumers and consultant data consumers should have access to restricted data sets that they have been specifically granted permission to view, as well as a

mechanism to request access to restricted data. These requests need to be routed to the appropriate FDOT data content manager to grant or deny access based on established policies.

Determining the best options for the central access point is complicated by firewall, policy, and funding issues. Systems behind the FDOT firewall make it cumbersome for non-staff users to get to the freight information. A public Web site housed with FDOT handicaps anyone responsible for managing the content or the data by placing primary control with the IT department. Hosting the system outside of FDOT would provide the most flexible access, but risks a system shutdown if funding is cut. TranStat will need to evaluate each option to determine the best choice for its needs.

The following is a list of tasks regarding data access:

- Provide a single access point for all users to information about freight data.
- Provide access for all users to search the data catalog, view record results, and download unrestricted data sets.
- Provide a tool for all users to request assistance or information.
- Provide content management tools for FDOT content managers.
- Provide a mechanism for requesting and managing access to restricted datasets.

Data Access	
Tasks	**Responsible Staff**
Provide a single access point	Freight Data Manager
Provide access to search the data catalog, view record results, download unrestricted data sets	Freight Data Manager, Librarian/Educator, System Designer
Provide a tool to request assistance or information	Librarian/Educator, System Designer
Provide content management tools	Freight Data Manager, Librarian/Educator, System Designer
Provide a mechanism for requesting and managing access to restricted data sets	Freight Data Manager, Librarian/Educator, System Designer

Delivery and Reporting

An easy query tool or tools for searching the freight data catalog should exist. Views of search results should allow the user to review the data set description, fields included, update frequency, last revised date, data source, contact information, a link to the data if available and unrestricted, and information about how to request the data set. Reports should fulfill the known and most common recurring information needs of TranStat and the modal and planning offices. There should also be an electronic form for requesting information. Information requests would be routed to the Librarian to manage.

The following is a list of tasks for delivering and reporting data to users:

- Design Query tools for searching the catalog.
- Design Views for query results.
- Create Reports on available data.
- Provide Request form for information or assistance.
- Respond to information requests.

Delivery and Reporting	
Tasks	**Responsible Staff**
Design Query tools for searching the catalog	Librarian/Educator, System Designer
Design Views for query results	Freight Data Manager, Librarian/Educator, System Designer
Create Reports on available data	Freight Data Manager, Librarian/Educator, System Designer
Provide Request form for information or assistance	Librarian/Educator, System Designer
Respond to information requests	Librarian/Educator

Future Goals

Future goals would grow upon this foundation built to support the new responsibilities assigned to TranStat. Once established, the Freight Data Clearinghouse would have a unique perspective as an intersection of expertise, data, information, research, users, and stakeholders that eliminates institutional silos. It may even proactively connect ideas and information resources to make freight related data more usable and useful.

Using the easily accessible, organized data and combining it with the previously identified GIS-compatible data, the Clearinghouse would facilitate collaboration to create new tools for visualizing and manipulating the data to create new information.

As the Clearinghouse identifies data needs, identifies data sources, and monitors research regarding freight data, it can also support collaboration to fill data gaps through acquisition of new data sources or creating new methods for collecting the data directly.