



Florida Department of Transportation

RON DESANTIS
GOVERNOR

719 S. Woodland Boulevard
DeLand, Florida 32720-6834

KEVIN J. THIBAUT, P.E.
SECRETARY

Questions and Answers #3

ADVERTISEMENT NUMBER: DOT-RFP-20-5008-SIM

DESCRIPTION: ITS Software and Maintenance Services for D5 SunStore and Future Deployments

DATE: April 1, 2020

3.2.1 Support of Existing Deployments

1. Could you provide an example or more detail around how FDOT defines “most granular form”?

Answer: Native format or without aggregation. ATSPM data has 10 Hertz of reporting. The information would be stored at 10 Hertz.

3.2.2.4 Cloud Storage Environment

2. Since the VENDOR will be responsible for creating the cloud storage environment, does this mean the VENDOR will be required to bear the costs of the cloud storage and testing environment?

Answer: Yes for storage and development

3. What would be examples of critical insights users are looking for within the cloud testing environment?

Answer: Research efforts commonly use the SunStore data.

http://www.cflsmartrroads.com/projects/future_projects.html shows some examples.

4. How many key stakeholders does the DEPARTMENT work with? If possible, broken down by Departments?
Agencies?
Universities?
Third parties (public and private)?

Answer: Below are many of the stakeholders. This is not comprehensive

Departments: Planning, Modal, Traffic Ops, Safety, Roadway Design, Construction, Maintenance

Agencies: Brevard County, City of Melbourne, Volusia County, City of Daytona, Flagler County, City of Palm Coast, Seminole County, Orange County, City of Orlando, Osceola County, City of Kissimmee,

Lake County, Marion County, City of Ocala, Sumter County, MetroPlan Orlando, Space Coast TPO, Marion TPO, Lake Sumter TPO, River to Sea TPO, LYNX, Voltran, Sunrail, Space Coast Area Transit
Universities: University of Central Florida, University of Florida, Florida International University
Third Parties: TTS, Live Traffic Data, Connected Signals, FDOT Consultants and Contractors

5. Does the DEPARTMENT anticipate the number of 3rd party vendors and Universities to grow?

Answer: Yes

6. Will the testing environment need to handle large volumes of raw data for insightful data analysis, allowing, for example, for "Schema on Read" to access data in structured and semi-structured formats without having to model data first?

Answer: Testing environment needs for insightful data analysis would likely be a point of discussion with key stakeholders in the potential Cloud Storage Environment task.

7. How will the cloud testing environment be required to scale for high concurrency, large workloads, and queries?

Answer: The intent is to make the data available to others. The contract is TWO based to allow the Department and Vendor to adjust based on demands placed on the system.

8. How does the DEPARTMENT currently share data with key stakeholders?

- a. How is the data provided, for example, via FTP, etc?
- b. How often is data provided or refreshed?
- c. What additional security is in place with respect to sharing of this data?

Answer: Via existing APIs. See metadata for refresh interval. <https://sunstore.cflsmartroads.com/>

9. Does the DEPARTMENT anticipate providing key stakeholders with different levels of access, for example, read-access vs full access?

Answer: Yes

10. In determining their willingness to pay for the ability to access the play space and tools through the cloud service, how important is the ability to provide cost and usage analysis to key stakeholders?

- a. For example, does the ability to charge back each key stakeholder for *only what they consume* from the environment need to be factored into this effort?
- b. Would a stakeholder dashboard showing usage across all tables, queries, and users including metrics trending over time and compared to prior periods, etc. be of interest?

Answer: It is anticipated that costs incurred due to third party activities would need to be the responsibility of that party. The means of accomplishing this would likely be a point of discussion and a potential task.

11. What interfaces does the Department expect the Key Stakeholder to utilize? Will they use their own interfaces (i.e. Tableau, Looker, etc) or forced to utilize another interface? Will they require SQL or other skills to utilize these interfaces? Who will be responsible for attaching third-party interfaces if they are allowed?

Answer: There are already existing interfaces on <https://sunstore.cflsmartroads.com/>. These interfaces allow for some developers to do work for stakeholders. Others are meant for more novice users that just want the data in excel. The cloud environment is anticipated to be for the users looking for larger data sets that conventionally would be required to use harddrives to access information.

12. What other types of semi-structured data would need to be accommodated? (ex. Avro, XML, Parquet)

Answer: No unstructured data sets have been identified. See <https://sunstore.cflsmartroads.com/> for interface information of current data sets.

3.2.2.4. Table 2 Key Data Sets for the Cloud Storage environment

13. What is the total estimated volume of data for these data sets i.e. how many Terabytes of data?

Answer: Current data set storage size information for the existing data sets in the SunStore can be seen in Figures 1 through 3 of the Exhibit A Scope of Services. As noted in Section 4.1.1 of Exhibit A Scope of Services, The VENDOR shall provide a cloud environment that anticipates the data in the will have unknown or changing access patterns for up to 100 TB in data.

14. Can the department please clarify whether it currently stores data in cloud storage such as AWS S3 or Azure Blob?

Answer: This contract will be the first attempt at cloud storage of these datasets.

15. Can the department please clarify the format for each of the 23 Key Data Sets provided in Table 2?

Table 2 - Key Data Sets for Cloud Storage Environment

Data Set	Description	Existing SunStore Outbound Format
1	Advanced Traffic Signal Performance Measures (ATSPM) Approach Route Details	CSV and JSON
2	ATSPM Controller Event Logs	CSV and JSON
3	ATSPM Controller Event Log Counts	CSV and JSON
4	ATSPM Phase Look Up Values and Codes	CSV and JSON

5	ATSPM Signal Locations	CSV and JSON
6	Basemap Route Roads	CSV and JSON
7	C2C Event	CSV and JSON
8	C2C Event Type	CSV and JSON
9	C2C Floodgate	CSV and JSON
10	C2C Locale Data	CSV and JSON
11	C2C Network Data	CSV and JSON
12	C2C Speed	CSV and JSON
13	C2C Traffic Conditions	CSV and JSON
14	C2C TVT Status	CSV and JSON
15	SunRail Stop Times GTFS	CSV and JSON
16	GTFS/GTFS-RT Aggregator Data	Not in Store yet to be determined
17	SIIA	Not in Store yet to be determined
18	ITSIQA Pedestrian Count Data	Not in Store yet to be determined
19	TMDD	Not in Store yet to be determined
20	ITSIQA Volume	CSV and JSON
21	ITSIQA TMC	CSV and JSON
22	ITSIQA Travel Time	CSV and JSON
23	Origin Destination via Sunguide	Not in Store yet to be determined

3.2.2.6.

16. This section states that *The DEPARTMENT wants to have pipelines of real-time data for users to subscribe to so that the requested data can flow real-time to the user.* Does this mean that as data arrives, whenever it arrives, it needs to be available for query and analysis?

Answer: This anticipates an asynchronous feed to provide in near real-time, similar to a subscriber versus a request and respond system

17. Approximately how many university, third-party (private and vendor) partners does the DEPARTMENT envision providing access to the data?
a. Approximately how many total users across all partners?

Answer: See question 4 for stakeholder information. Most agencies partners have 2-10 users. The majority are closer to 2 users.

18. Is the objective of the data pipelines to enable all key stakeholders' access to the same shared data, simultaneously, enabling very high concurrency?
- a. Does the data pipeline require bi-directional data flow, for example, the ability for data sets to be shared back with the DEPARTMENT?

Answer: Yes, to being able to access shared data, simultaneously, with high concurrency. Unidirectional flow from the Department to others.

19. Is the vendor responsible for data pipeline security and encryption of data?

Answer: Role based access to data sets is in place. For new cloud hosted solutions all data is expected to be available for consumption.

20. Would the Department envision creating its own Data Exchange where the Department provides and controls which vendors and stakeholders have access to data and can see whom is utilizing each resource?

Answer: This is in place with the locally hosted data and expected to continue with any needed updates via this contract, though the flow is unidirectional.