Pre-Procurement | Procurement | Implementation

Data Ownership Translation Guide

As we've explored in the Data Ownership Fact Sheet, data is a core asset that underpins transit service planning, funding, and performance evaluation. Without clear ownership and access rights, agencies risk losing critical data when their vendors change, bearing significant integration costs, needing to duplicate services, or starting again from "square one."

This guide can help you procure a product that will work for you today and in the future, even in a changing technology landscape.It summarizes best practices for data ownership and provides examples of how you can "translate" these into action.

TRANSLATION GUIDE TABLE OF CONTENTS

1. Establish a team culture that prioritizes data ownership
2. Understand your current data ecosystem
3. Review data ownership best practices
4. Implement in your own documents

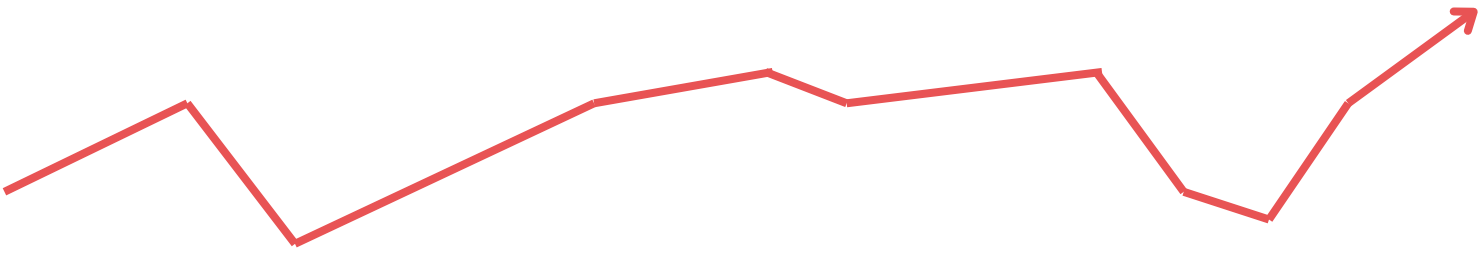
1. Establish a team culture that prioritizes data ownership

Before jumping in, align your team on why data ownership matters:

- Strategic Asset: Data powers service planning, funding justifications, and performance measurement.
- Operational Risk: Without clear language, you risk losing data when a vendor contract ends.
- Control and Agility: Ownership allows your agency to switch vendors, integrate systems, and innovate—without being locked into a proprietary system.

Many transit agencies assume that data generated by their system(s) is automatically owned by the agency. This is a common mistake - access and ownership should be specified in contracts to help your data work for you!

It takes time to build a culture of data ownership if it's not already a characteristic of your team. Leadership can help pave the way by providing staff opportunities to learn and build skills. State Departments of Transportation (DOT), Federal Transit Administration (FTA), or transit conferences can be a great opportunity to participate in workshops on data management – many resources are free or subsidized for transit agencies.



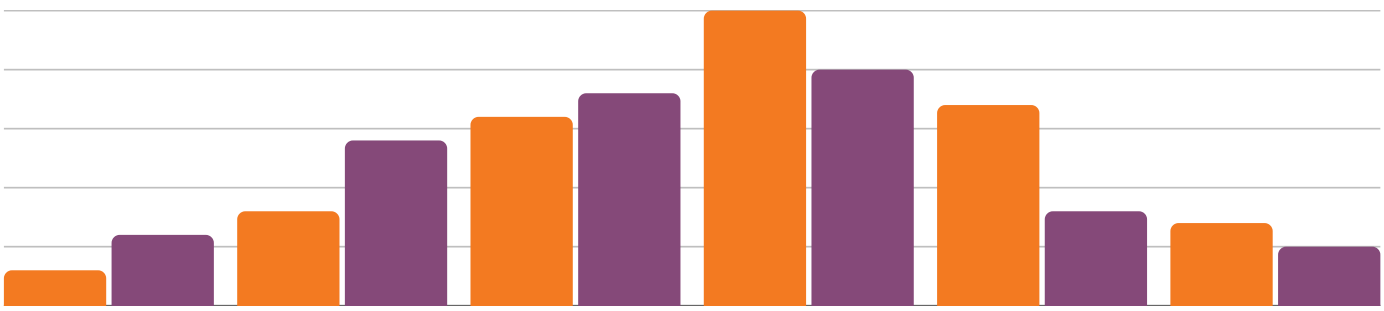
2.Understand your current data ecosystem

Audit Your Current Data Assets

Before you procure new transit technologies, it's important to have a good understanding of your current operations. If you haven't already, conduct a baseline data assessment to identify gaps and opportunities:

Category	Key Questions
Storage	Where is your data stored?
Governance	Who manages it and ensures appropriate levels of access?
Access	Can your team easily access data to export, query and analyze it (e.g., exports are in file formats like CSV, Excel, or JSON)?
Contracts	Do existing contracts and licenses define data ownership, access and usage rights?
Use	Is data available in applicable open standards (e.g. GTFS) or in a well-documented data schema or only available in proprietary formats?
Continuity	What happens to the data if a vendor is replaced, a contract ends, or there is a breach in the security or technology that threatens data integrity?

Tip: Use a spreadsheet or visual map to document vendors, data types, access rights, and contract terms.



Understanding your data helps you manage it. Many agencies find it useful to maintain a simple data inventory / catalog – a list of data sources and what each contains. An example table is:

Category	Data Type	Examples	Sensitivity	Notes
Operational Data	Route and schedule data	Bus/train routes, timetables	Public	Useful for planning and public info
	Maintenance records	Vehicle inspections, repair logs	Internal	Internal use; may include sensitive asset info
Personal Data	Employee records	HR files, payroll, certifications	Sensitive	Protected by labor and privacy regulations
	Rider information	Fare card usage, account details	Sensitive	Includes personally identifiable information
Performance Data	Service metrics	On-time performance, delays	Public	Often used in open data portals
	Ridership statistics	Daily / weekly / monthly boardings	Public	Valuable for public reporting and analysis
Data Management	Data inventory / catalog	Metadata about each dataset	Public	Helps staff understand and manage data assets
	Sensitivity classification	Labels: Public, Internal, Sensitive	Internal	Guides access control and compliance practices

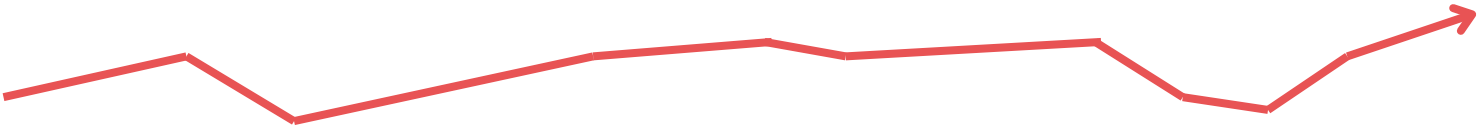
3. Understand your current data ecosystem

Regardless of the size of your agency, everyone can implement certain best practices to help protect your data ownership.

Best Practice #1: Clearly Define Data Ownership in Vendor Contracts

When procuring technology vendors, you should ensure that procurement documents and contracts clearly state that the agency retains ownership of all transit data. It's also helpful to include a 90-day data retention period after the end of the contract to give you time to download and transition your data as needed.

Contracting Tip: Keep things simple; use plain-language contracts that define data-sharing terms, and always claim your agency owns the data.



Best Practice #2: Define Data Accessibility and Usage Rights

Contract terms specify key elements that define rights and responsibilities for accessing, sharing, and using data. You'll want to specify the frequency and type of access that you'd like to have to your data (e.g., programmatically via an Application Protocol Interface (API), monthly downloads as .csv, weekly reports, ability to query ad hoc). Different people in your organization may need access – be sure to identify and specify how many "seats[1]" aka users are needed for the software license. It's also important to clarify if and under what conditions a vendor can leverage, share, or monetize YOUR data, or restrict the vendor entirely.

Contracting Tip: Data ownership doesn't mean much if you can't access your data. Be sure to specify the ways in which you want to access and use your data.

Tip for Best Practices #1 and #2: If you're unsure how to ask vendors about data ownership or accessibility, start with the following questions:

- "Who owns this data?"
- "Can we switch systems later without losing access?"
- "Are we following any privacy or open-data policies?"
- "How is data backed up and stored long term?"
- "What is the cost for data back-up and storage?"
- "What is the data back-up redundancy plan?" (i.e., does the vendor's data center have a redundancy plan and if yes, what is it?)
- "What are the data retention policies and who controls data at the end of a contract?" (i.e. is the data wiped after the contract termination?)
- "What insurance and remedies will be applied for catastrophic system failures?"
- "What happens in the event of business insolvency?"
- "How often can I download the data, query or adjust metrics?"
- "How many "seats" or "users" come with the basic package? How do you define "seat" (i.e., by specific individual or by number of people using the system at one time)? How much does it cost to add more later?"

1. A seat is a license to use the software. For example, if you purchase three (3) seats, it means that up to three (3) people can each have their own account to use simultaneously.

Best Practice #3: Public Data as a Public Asset

Datasets created by transit agencies (e.g. route maps, schedules, bus locations) are public assets. Consider sharing this data in a transparent way such as a downloadable format on your agency website, that meets security, privacy, and regulatory requirements. This transparency promotes trust, accountability, and understanding. It also allows groups such as students, researchers, and open data coalitions to leverage your data for their own learning. Since this data is created with public funds, it should be open and accessible to the public. Remember to institute policies to periodically back up your key data locally.

Contracting Tip: If you are interested in ensuring your transit data can be leveraged as a public asset, when possible, you can ask the vendor:

- "Can this data be accessed by third-parties? What is the process for providing them access?"
- "How many requests can you provide access to in a given period (month, quarter, etc.)?"

Best Practice #4: Collect only what you need

Don't gather sensitive personal information "just in case" – this creates potential risk for you with little reward and makes compliance more difficult. Instead, focus on collecting only the data you need. For any personal data you do collect (names, addresses, travel patterns, etc.), take reasonable security measures:

- Personal Identifiable Information (PII) definitions can vary greatly by state - understand what is considered PII in your area.[2]
- Store data in secure systems (password-protected databases or encrypted files).
- Prepare for the worst and create a response plan to proactively define responsibility for security breaches (i.e., you or vendors). This can help you respond quickly and comply with any legal complications. Ensure all responsible staff are trained and familiar with these plans.

Best Practice #5: Implement Privacy Protocols

Set rules for who can access PII and under what circumstances. For instance, you could limit access to customer contact info to only staff who need it (like customer service or billing staff). Ensure you have data privacy protection protocols – e.g. require staff to keep data secure and confidential.

Tip for Best Practice #4 and #5: Learn more about cybersecurity in transit from the following resources:

- Cybersecurity Issues and Protection Strategies for State Transportation Agency CEOs
- Core Data Principles
- Cybersecurity for Transit- National RTAP

2. https://www.iubenda.com/en/help/109908-sensitive-personal-information-us-states-comparison

4. Implement Best Practices in Your Documents

Now that you've assessed your current data ecosystem and policies, keep these best practices in mind as you review your procurement templates and existing vendor contracts. Your goal is to identify any potential areas that are unclear or could be updated in future iterations. You can also create new templates that incorporate these suggestions into future procurements, contracts, and operating manuals.

Key Concepts to Include in Procurement Documents:

- Data Ownership: The agency owns all non-sensitive data generated through the system or service.
- Access and Portability: Vendors must provide data in a structured, machine-readable format (e.g. CSV, JSON, XML, parquet, etc.) with a well-documented data schema[3] at any time. Formats with unstructured data[4] or semi-structured data[5] can be used for reports on data that is available through a structured data format but shall not serve as the primary means to access that data. Data formats that are strongly tied to specific programs (e.g. XLSX, sas7bdat) are less preferable than formats that are designed for portability (e.g. csv, parquet) as some features may not be universally available or may be incorrectly interpreted across implementations.
- Post-Contract Access: Data must be made available after contract termination for a defined period (e.g., 90 days).
- Audit Rights: Agency can verify how data is stored, secured, and accessed.
- No Proprietary Lock-In: Vendors must provide access to data that does not require using their proprietary tools to download and acquire data.
- Data Security: Vendors must provide a security plan to protect agency data from unauthorized access, loss, or corruption.
- Disaster Recovery: Vendors must provide a recovery plan that outlines procedures and tools to restore critical IT systems, data, and operations after a disruption in the event of cybersecurity or other disaster.

These key concepts don't always make it into contracts, either because the vendor prefers a proprietary system or because agencies don't ask the right questions. The table below is a guide for red flags related to data ownership when you're reviewing contracts:

3. A data schema provides a blueprint of a database or dataset's structure, including fields, tables, relationships, and constraints.
4. E.g. data evident from a chart of contained within paragraphs of text
5. E.g. tables contained within images, PDFs, and HTML

Red Flags	Why it Matters	Better Practice
“Access” instead of “ownership”	Implies limited rights	Explicitly state the agency owns all data
“Proprietary format”	Restricts interoperability	Require exports in open, structured, machine-readable formats (e.g., for Excel - XLSX, CSV, JSON, GTFS) with well-documented data schemas.
No clause for post-contract access	Agency may lose historical data	Add a transition plan and data retention window
Vague data use rights for vendor	May allow reselling or reuse	Define vendor’s rights narrowly and require consent for any reuse

Look Out!

Below are some examples of contracts where these key concepts were NOT included. This can result in a contract that restricts your data ownership rights. Sometimes these nuances are subtle, so we’ve included key indicators to help you recognize when something might not be right. For some topics, we’ve also included examples of stronger terms used in other contracts.

Note: This language has not been reviewed by a legal professional and is intended for illustrative purposes only. Please consult with your legal and procurement teams before finalizing any contract language.

Vague or Missing Ownership Language

Example	Red Flags	Stronger Terms
“Vendor shall provide the Agency with access to operational data as required for the provision of services.”	• "Access" does not mean "Ownership" • No mention of exportability, portability, or long-term rights • Control remains with the vendor	“Agency retains full ownership of all data generated and stored through this contract.” “Agency shall be given the ability to query, sort, fully access data at will.” “Vendor will ensure data is accessible in machine-readable formats (CSV, JSON, etc.) and shall provide complete exports upon request or contract conclusion.”

Pre-Procurement | **Procurement** | Implementation

Vendor-Controlled Access

Example	Red Flags	Stronger Terms
"All data collected by the system shall be stored on Vendor-managed servers. Agency access shall be granted through a web interface during the term of the agreement."	• Data is stored by the vendor, and only viewable—not downloadable. • No guarantee of continued access after contract termination.	"Agency shall at all times maintain title, ownership, rights, and interest in and to the data." "All data collected by the system shall be stored on Vendor-managed servers. Agency access shall be granted access through a web interface and downloadable 90 (ninety) days after contract termination."

Proprietary Format Lock-In

Example	Red Flags	Stronger Terms
"Data shall be available on the Vendor's platform and can be reviewed using the dashboard analytics. Requests for data exports are subject to approval."	• Vendor controls format and usefulness of the data. • Limits integration with other tools or migration to new systems.	"Data shall be available for export in the applicable open data format (e.g., GTFS, TIDES, TODS, etc.)." "Data shall be available for export in a queryable, searchable, machine-readable format (e.g., .csv)." "The agency shall be able to initiate this export / download without approval from the vendor."

Post-Termination Data Access Limits

Example	Red Flags	Stronger Terms
"Upon termination, the Vendor will make a copy of customer data available upon request for a fee within 30 days. After this period, data will be securely deleted."	• No default guarantee of data delivery. • Short recovery window, potentially at additional cost. • Permanent loss if the request isn't made on time.	"Upon termination, the Vendor will provide the agency with all raw data within 30 days. After this period, all closed data will be securely deleted and shall not be used for other purposes."

Monetization of Agency Data

Example	Red Flags	Stronger Terms
"Vendor retains the right to use and distribute data to third-parties for analytical and commercial purposes."	• Suggests the vendor may profit from the agency's operational data • May conflict with public transparency or data privacy goals	"Vendor may use anonymized data for service optimization but not for commercial purposes."

Data Security

Example	Red Flags	Stronger Terms
"The Contractor shall ensure the security of all Personally Identifiable Information (PII)"	• Contractual language is vague- means the security standards may be susceptible to breach. • Best to specify security standards using state or federal privacy standards	"PII shall be stored in a manner consistent with {reference your local or state laws regarding data privacy}." "Access to PII shall only be given according to a policy determined by the Transit Provider or via the Transit Provider's expressed written permission."