Pre-Procurement > Procurement > Implementation

Data Ownership Factsheet

In today's transit technology landscape, data is a critical asset. From real-time fleet locations to ridership patterns and fare media usage, the information generated by a transit system can provide insights for informed decision-making and strategic investments.
While data can be generated, collected, or used by other parties, like your technology vendors, it is essential that you establish both legal and practical ownership over your data. Data ownership gives you access, control, and use of your business intelligence. After all, you don't want that data to simply disappear one day to the next!

What is Data Ownership?

Data ownership is the legal and practical rights over data generated, collected, and used by transit systems. Legal data ownership is specified in contracts with vendors and their end user agreements and licenses, while practical data ownership is your ability to easily exercise those legal rights.

Why does agency data ownership matter?

Owning your data is essential because it lets you decide what data can be shared, how, when, and with whom. Data ownership is often described in "legalese" which can be difficult to understand. We will explore examples of contract language in Data Ownership Translation Guide.

At its most fundamental, legal data ownership ensures that you control the access and management of your data (data licensing and governance). It applies to all sorts of technology, including on-board systems, software as a service (SaaS)[1] products dashboards and data analyses, etc.

From a user perspective, practical data ownership means that you have complete access to and power over your data when you need it to. Benefits of good data ownership policies include:

Different types of data

Open Data – anyone can use the data free from cost or restrictions (other than attribution).
Example: trip planning apps using an agency's GTFS feed

Internal Agency Data – owned and governed by the transit agency with sharing subject to appropriate data governance and privacy considerations.
Example: Automatic Passenger Counters (APCs) record boardings/alightings per station

Proprietary Data– owned and controlled by a specific entity, typically a vendor, and cannot be used without permission. This could include paying for a license and often means the agency does not have complete ownership rights.
Example: Battery range mgmt. and route optimization provided by an electric vehicle company

1 SaaS is a way to deliver software services, where you pay to use it instead of buying it and installing it yourself. For example, scheduling software may be a SaaS model where you pay an annual license fee.

- Protect yourself against vendor "lock in": If your vendor disappears or you want to switch things up, data ownership means you don't have to start again at square one
- Enable your tech systems to "talk" to each other: If you control the data, you can decide how it feeds into other systems (See our **MDIP Factsheet** for more!)
- Establish a single "source of truth": If each of your technology systems use proprietary data, then there can be competing and inconsistent information
- Improve stakeholder collaboration: If partner departments or regional agencies can access each other's data in standard formats, then you can more easily collaborate on shared goals.
- Strengthen public transparency: If transit data is publicly available, then public trust in your system will also improve. See the TransitCenter's The Data Transit Riders Want.

What data should you own?

With very few exceptions as noted below, you should have legal and practical ownership over all data consumed, entered, or produced by any technology product.

Table 1: Data that should always be agency-owned

Category	Examples	Additional Considerations
Open data that is provided free from restrictions	GTFS Schedules, GTFS Realtime arrival predictions	A transit agency should have the sole discretion as to how its data is licensed and governed, even if provided under an open license free from restrictions.
Rider-generated data	Ride requests, rider profiles, rider location data.	A transit agency should have clear license agreements with riders that also highlight data governance, how their privacy is protected, and a rider's ability to delete their data from the system.
Staff-generated data	MDT input, Driver logins, radio transcripts, software logins and activity	A transit agency should have clear user agreements and training for staff that includes how their privacy is protected.
Personally identifiable information (PII)	Identification of and data about specific drivers or riders that is attributable to them such as age or income (i.e., for reduced fares)	A transit agency should have a strategy for responsible PII data governance and should meet local & state statutes related to privacy.

Table 2: Data that may not be agency owned under certain conditions

Category	Does not need to be agency-owned if...
Intermediate data tables that facilitate calculations	It is not the single source of truth of a particular piece of information (i.e., if this table is deleted, you wouldn't lose any information).
Proprietary metrics are unique to that vendor and only help you understand system performance within the broader context of that vendor's technology.	It is provided under an open data license free of cost or restrictions on its use. Note: You should own this data if it is the single source of a particular piece of information.

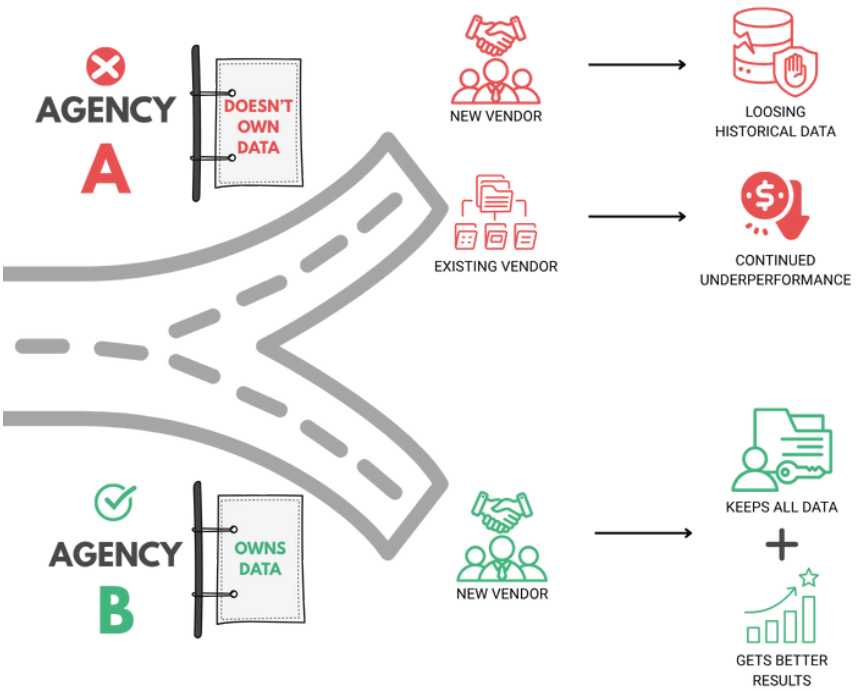
The bottom line is that you should always own data that is the "single source of truth." If there is only one way to get GPS coordinates, you should own that location data.

What does this look like in practice?

Example #1: Providing flexibility when selecting a vendor

Agency A and Agency B have contracts for a CAD/AVL system that are expiring this year, and they have decided to change vendors when the contract ends because the current vendor is underperforming.

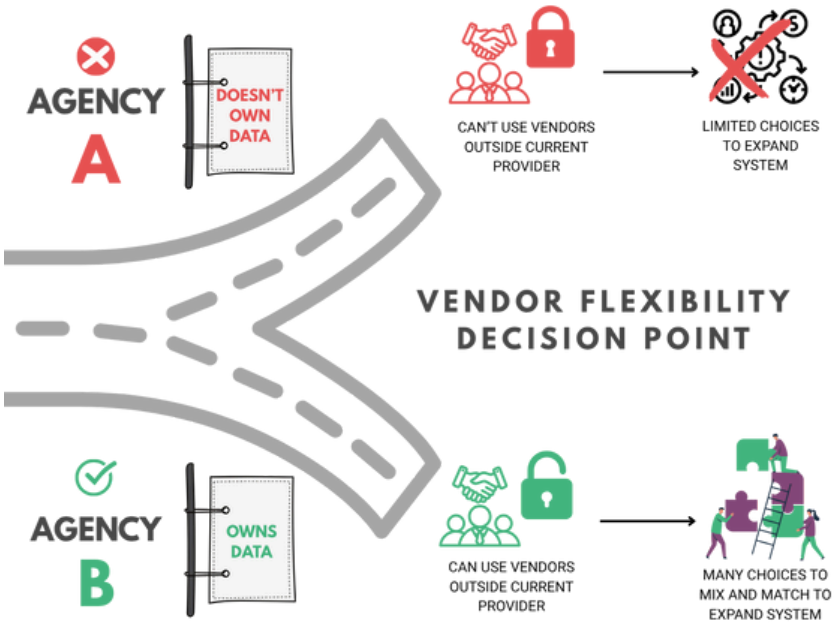
Because they do not own their data, Agency A could potentially have to choose between going with a new vendor and losing their historical data or staying with the existing vendor and keeping their historical data. Because they own their data, Agency B will be able to select a new vendor without losing any historical information, allowing them to analyze their performance at any time in their history. The new vendor will also be able to use the historical data for their new CAD/AVL system.



Example #2: Enabling product interoperability

Agency A and Agency B have an existing GTFS-Realtime vendor. They are now working to purchase and implement a Traffic Signal Priority (TSP) project that depends on real time vehicle location data.

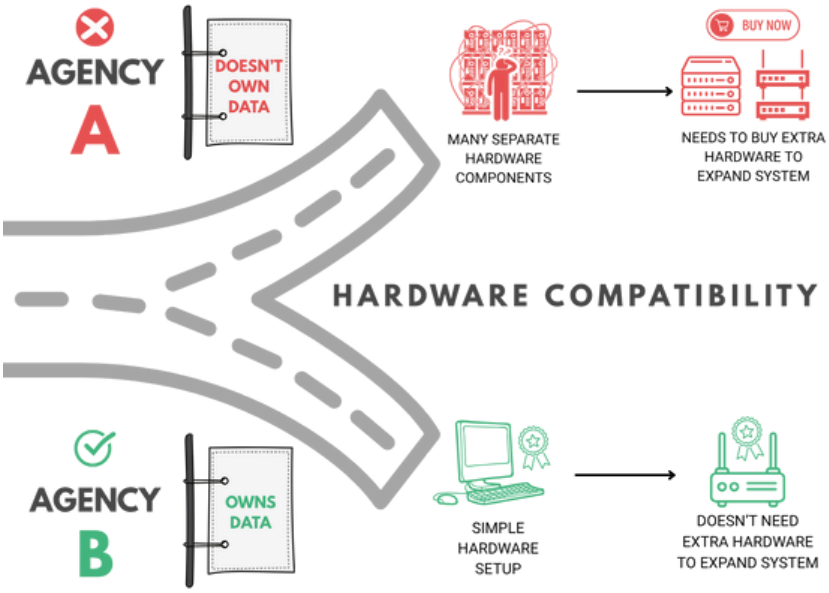
Without data ownership, Agency A might be stuck selecting a TSP solution from their existing vendor, even if that's not what they want. Since Agency B owns their data and is using open data standards (i.e., GTFS-Realtime), they will be able to "mix and match" vendors and won't have to invest in additional hardware or pay integration fees to transmit location data.



Example #3: Reducing duplicative technology and conflicting data inputs

Agency A and Agency B have an existing GTFS-Realtime vendor. They are now working to purchase and implement a Traffic Signal Priority (TSP) project that depends on real time vehicle location data.

Without data ownership, Agency A might be stuck selecting a TSP solution from their existing vendor, even if that's not what they want. Since Agency B owns their data and is using open data standards (i.e., GTFS-Realtime), they will be able to "mix and match" vendors and won't have to invest in additional hardware or pay integration fees to transmit location data.



Because they own the data, Agency B can use its router for multiple purposes and share the location data with its other vendors. This means all the technology gets the same information from one "source of truth" and doesn't rely on potentially conflicting location data sources. It also clears up space in the supply cabinet while minimizing duplication.

How is data ownership related to cybersecurity?

Data ownership is not the same as cybersecurity, but the two are related. You are responsible for protecting the data you own. This is especially important if your agency is using any kind of personally identifiable information (PII). It is important to ask vendors - and your IT department/vendor - what cybersecurity measures are in place.

You can implement basic cybersecurity protocols today. These include keeping software updated, using strong passwords, and having backups for important data. Even small agencies can implement low-cost security measures like enabling firewalls and antivirus or using trusted cloud services for email and data storage (which often include security by default).

Advanced cybersecurity protocols may require additional expertise from your IT department or expert consultants. Advanced cybersecurity strategies may be necessary based on the size and complexity of your operations and risk assessment. The TRB Transit Cooperative Research Program's **Transportation Cyber Risk Guide** Volume 1 and 2, offer strategies for a range of risk factors.

Who Oversees all this data?

When you own your data, you get to be in charge! It is important to review your agency's existing data policies related to public record requests, data retention, and data sharing agreements. If you don't have an existing data policy that addresses these key items, we would encourage you to develop one. Some items to consider:

Know your data	Catalogue what data you have. This includes operational data (routes, schedules, maintenance records) and personal data (employee records, rider info from fare cards, etc.). Understanding your data helps you manage it. Many agencies find it useful to maintain a simple data inventory or data catalog – essentially a list of data sources and what each contains. See the **Data Ownership Translation Guide** for an example.
Public Records Requests (PRR)	As a public entity, be prepared to respond to public information requests (FOIA or state open-records requests). Have a protocol for handling these– know what data can be released and what must be redacted (for privacy or security), open data standards can reduce PRR incidents and minimize time responding to them, since the information is publicly available already.
Data Retention	Decide how long to keep different types of data. For example, you might keep some ridership details for only a certain period (to limit liability and storage costs) or retain performance metrics for historical comparison. Research and follow any state record retention laws.

Data Sharing Agreements	If sharing data with outside partners, use written agreements. Clearly state what data is shared, for what purpose, and that recipients must protect it. This is one way you remain in control of how its data is used.
Partnerships for Tech Support	Small and mid-sized agencies don't have to go it alone. Partner with your Metropolitan Planning Organization (MPO), state department of transportation, larger neighboring transit agencies, or municipal governmental departments for technical assistance. Such collaborations may provide expertise, so you don't need to have it in-house. Similarly, consider joining peer networks, FTA Technical Assistance Centers, and State DOT transit offices which often have programs to help rural and small urban agencies.

How do you put this into action?

First and foremost, remember that it is critical to read the contract before signing with a technology vendor and ensure you understand what is being proposed in the agreement. Legal representatives or technical staff (either in-house or hired) can always review the draft contract to ensure you (1) understand all proposed terms/requirements and (2) can negotiate any terms/requirements of disagreement.

Secondly, check out the **Data Ownership Translation Guide** or tips and tricks on how to incorporate these concepts into your own contract as well as red flags to look for.

At the end of the day, you want to own your data and ensure your transit agency is set up for success both today and years down the road.

Want to learn more?

- **AASHTO**
- **RTAP - cyber security cheat sheet**
- **Mobility Data Interoperability Principles**

Review N-CATTs Data Ownership Translation Guide for Best Practices and Learn How to Put Data Ownership in Action. And if you have any questions don't hesitate to reach out to the N-CATT team at helpdesk@n-catt.org