

# Massachusetts Statewide Trails Database Roadmap

*Version 1.0*

*In coordination with:*



*Funded by:*



*Prepared by:*

Applied Geographics (AppGeo), a division of The Sanborn Map Company, Inc.



**The Sanborn Map Company, Inc.**

Boston Office  
33 Broad Street, 4th Floor  
Boston, MA 02109  
[www.sanborn.com](http://www.sanborn.com)

December 26, 2024

## Table of Contents

|  |           |
|--|-----------|
| <b>Section 1: Introduction to the Roadmap</b> .....  | <b>3</b>  |
| Background and Long-Term Vision for the Massachusetts Statewide Trails Dataset .....   | 4         |
| Roadmap Purpose and Use .....  | 5         |
| Updating of the Roadmap.....   | 5         |
| <b>Section 2: Data Supply</b> .....  | <b>6</b>  |
| Overview of a Trails Data Production Supply Chain .....  | 6         |
| <b>Section 3: Statewide Trails Dataset Creation, Management, and Maintenance.....</b>  | <b>12</b> |
| <b>Creating the Data Production Supply Chain.....</b>  | <b>12</b> |
| Define and Fill Required Roles.....  | 12        |
| Define functional and technical requirements and specifications .....  | 14        |
| Define an implementation plan to build (instantiate) the system as software, hardware,<br>networking and services, and costs ..... | 14        |
| Define a budget and process for maintaining the supply chain.....  | 14        |
| Seek opportunities to build and support the system.....  | 15        |
| <b>Managing the Chain – Establishing Data Governance</b> .....   | <b>15</b> |
| <b>Policies</b> .....  | <b>17</b> |
| <b>Standards</b> .....   | <b>18</b> |
| <b>Improving and Maintaining Data Quality</b> .....  | <b>19</b> |
| Define Quality Requirements for Trails Data .....  | 20        |
| Conduct Data Quality Assessment .....  | 21        |
| Implement Data Quality Actions and Tracking.....   | 22        |
| <b>Publishing Data</b> .....   | <b>23</b> |
| <b>Marketing and Public Awareness</b> .....  | <b>24</b> |
| <b>Appendix A: Trail Data Users and Consumers</b> .....  | <b>24</b> |
| <b>Appendix B: Hand-off Supply Chain Model</b> .....   | <b>24</b> |
| <b>Appendix C: Portal-hub Supply Chain Model</b> .....   | <b>24</b> |

# Section 1: Introduction to the Roadmap

Over the past three years the Massachusetts Metropolitan Areas Planning Council (MAPC), MassGIS in the Executive Office of Technology Services and Security (EOTSS), the Department of Conservation and Recreation (DCR), and the Massachusetts Association of Regional Planning Agencies (MARPA) have embarked on a long-term effort to create and publish a statewide geospatial dataset of trails information for the State of Massachusetts. Data is scheduled to be published in June of 2025. One can think of their effort as a trails data production supply chain initiative – a long-running effort to make trails data available in a sustainable and ever-improving way. The collaboration between these four critical stakeholder groups, the MassTrails Coordinating Body, drives the definition of the data supply chain, creates the statewide trails dataset, and develops the supporting policies, standards, and procedures required to publish and manage that data.

The Massachusetts Statewide Trails Database Roadmap document was prepared to provide near-term and long-term guidance and strategies on how best to continue work that leads to a shared statewide trails GIS dataset. It combines long-term strategies and recommendations with near-term tactics that advance the work accomplished to date. Although the document is titled a “Roadmap”, it might more properly be thought of as a handbook. Like a handbook in a snap-ring binder, the intent is that sections can be revised over time with obsolete content stored for reference at the “back” of the binder (i.e., in a virtual section at the end of the document).

Guidance presented in the roadmap as a near-term item will be completed by the MassTrails Coordinating Body between January and June of 2025. These near-term efforts include continued coordination of the MassTrails Coordinating Body, the addition of previously agreed upon standards, policies, and data management definitions into the roadmap, and discussions to address any remaining near-term items. It should be noted, that though the near-term items are key building blocks for the long-term success and use of the trails data, the MassTrails Coordinating Body does have to not to address or implement all near-term items in order to publish the trails data.

Once the data has been published in June of 2025, the MassTrails Coordinating Body will need to address the presented long-term guidance. The process of continuously revisiting and updating the roadmap is an iterative programmatic and governance capacity building effort. The long-term strategy aims to focus the MassTrails Coordinating Body and any new stakeholders on more detailed definitions required to improve, grow, and sustainably manage the published MassTrails data.

## Background and Long-Term Vision for the Massachusetts Statewide Trails Dataset

Trails, shared use paths, and other bike and pedestrian facilities—both on and off road—are an invaluable part of the Commonwealth’s public infrastructure for recreation and transportation. As with any large set of public infrastructure assets, data is needed to support planning, policy analysis and advocacy, as well as management and use of the MA statewide trails network.

A variety of government and other public interest entities throughout the Commonwealth of Massachusetts produce, use, or otherwise rely on trails data. MA municipalities and regional planning agencies (RPAs) often collect and use trails data for [open space and recreation plans \(OSRPs\)](#) and for transportation planning purposes. A number of state agencies collect/produce, manage, use, and in some cases publish trails data. Specifically, these agencies include: the Department of Conservation and Recreation (DCR) which manages state parkland and the vast network of trails it contains; MassDOT which oversees the state road network including managing facilities data on bicycle and pedestrian facilities within the right of way; and MassGIS who’s advanced 911 data provides State Police with accurate geospatial information for search and rescue purposes, including information on state trails systems.

The challenge is that while each of these state and local entities plays a role in the use and collection of trails data, there is currently no centralized source of truth for statewide trails data from the public sector.

To take on that challenge, the Department of Conservation and Recreation’s [MassTrails program](#) has funded MA RPAs—led by the Metropolitan Area Planning Council (MAPC) and Franklin Regional Council of Governments (FRCOG)—to work with MassGIS and other state agencies to aggregate, manage, and publish a statewide trails database. Trails data for this initiative will be drawn from a variety of sources, standardized, and collaboratively maintained for publication online as a public map and dataset available for download and programmatic consumption.

This work began informally in 2020 with initial efforts among RPAs to develop a statewide trails data standard. Over the course of multiple MassTrails grants, a v1 data standard and research on potential users and use cases for a statewide trails dataset have resulted, as well as an initial internal database set to be published for the first time by June of 2025.

The ultimate vision for this initial effort is for a consortium of trails data stakeholders led by state and local government to contribute on an ongoing basis to the governance and operation/maintenance of the statewide trails database, with a focus on providing “official” data produced or verified by government sources. Such a dataset is intended to

provide a counterbalance to the private sector and “crowdsourced” trails data sources over which the public sector has little control. It is also intended to support public sector and public interest use cases, such as transportation and recreation planning; trails analysis and advocacy; trails asset management, maintenance and operation, and emergency response; and public interest technical use cases, including supporting the development of digital maps and applications that can benefit trails stakeholders and users. It is also envisioned that this dataset can contribute to national efforts, such as the USGS National Digital Trails Initiative.

## Roadmap Purpose and Use

The roadmap is broken into 3 major sections: (1) an introduction and overview (this section); (2) a description of trails geospatial and associated data and supply chain models; and (3) operational considerations. The table of contents should quickly guide a roadmap user to whatever section or subsection is needed.

This roadmap is intended to provide guidance on establishing a programmatic framework within which the MassTrails stakeholders can build capacity to manage and govern their data production supply chain. Rather than a checklist of requirements, this roadmap will serve as a list of recommended focus areas and considerations for assigning roles and responsibilities, allocating resources, and defining future objectives. Topics covered in this roadmap start with a vision for statewide trails data. Data supply chains, like any program of continual data delivery and improvement, can take different configurations. The roadmap then examines two different supply chain models, compares them and recommends an overall supply chain configuration. The roadmap then turns to practical considerations. Coordinating data production and publication requires some form of governance. Therefore the roadmap includes guidance on what is required to implement the data supply chain and then manage it. The roles defined in the supply chain discussion are revisited as part of governance. Policies and standards are important parts of data governance and, in this instance, the supply chain too because they shape the process itself and are the basis of standardization, easing aggregation and promoting data quality. A practical approach to data quality is discussed after this. Operational topics make up the final three sections of the roadmap, including processes for publishing data, making others aware of the trails GIS dataset, and configuring hardware, software, and network access needed to gather, review, improve, and publish trails data.

## Updating of the Roadmap

The roadmap, particularly its operational guidance, is a living document. This is inherent in its purpose and in its organization. Indeed, some parts of the roadmap cannot be written yet because they depend on prior actions and will need to be written or further developed in the future as Massachusetts’ statewide trail data aggregation efforts evolve. These elements are present in the roadmap with statements concerning their completion in the

future. As the statewide trails data program evolves, the roadmap will evolve too. Superseded sections should still be retained as dated entries (initial publication date, supersession date) in an appendix to the roadmap.

Another method that ensures there is a clear history of how the document has evolved over time is a Change Log. The change log provides a detailed and dated record of all modification made to the document including additions, deletions, and updates. The change log also helps in maintaining version control by providing a chronological list of changes. This allows users to revert to previous versions if necessary and ensures that the most up-to-date version of the document is always available. This change log should live at the beginning of the roadmap document to provide an immediate and transparent record to the reader. The MassTrails Coordinating Body will need to determine what consists of a change that needs to be logged. An initial recommendation is to log significant changes that document the structure of the trails data production supply chain. If more granular changes are needed over time, the MassTrails Coordinating Body can adjust their change log approach. A sample change log template with an example is provided as follows:

| <b>Version ID</b> | <b>Change Date</b> | <b>Change Section</b>                 | <b>Change Description</b>                        | <b>Change Contact</b>               |
|-------------------|--------------------|---------------------------------------|--|-------------------------------------|
| 1.1               | 12/23/2024         | <i>Define and Fill Required Roles</i> | <i>Added MassGIS as the owner of the Roadmap</i> | <i>Stephen Larrick<br/>Cy Smith</i> |

The change log will also contain a Version ID. When a change is made, a new sequential Version ID added in the change log and will also be added in the title page.

## Section 2: Data Supply

### Overview of a Trails Data Production Supply Chain

In the business world, a production chain describes how goods get produced from raw materials to finished product. The concept of a *data* production supply chain is essentially similar, describing where data comes from (derived from observations made by a person or a device), how it gets gathered, transformed, standardized, shaped into necessary form(s), and brought together to create a dataset. For this work, the product of the chain is a statewide trails dataset in geospatial format and associated tables or other data formats. At a minimum, the dataset contains lines (vectors) representing trails. As in any GIS data, the vectors carry attributes that further describe them. In a more expanded form, the trails dataset might contain:

- Trail vectors (minimum viable product)
- Trailheads and access points
- Parking and staging areas

- Polygons of trail surfaces
- Trail vectors organized into a linear reference system
- Simple descriptive attributes of trails
- Complex hierarchical attributes of trails
- Historical (retired, replaced, decommissioned) vectors, trailheads, etc.

The current effort (2024-2025) is to gather trail vectors and simple descriptive attributes, compiling them into a statewide form to share through a web-based application. MassGIS has developed an initial version of an ArcGIS Hub for trails data that enables multiple trails data stewards to edit and update their trails data in a single location.

## Defining Trails

There are many definitions of trails. These range from federally designated and recognized trails (cf. the federal trails geospatial data standard, FGDC-STD-017-2011) to almost any route that is regularly used by non-motorized users or motorized off-roadway vehicles.

In this version of the roadmap, trails are defined as: *A prepared surface recognized by a landowner or manager as a purposefully created route for non-motorized travel (including on roadways) or motorized off-highway travel.*

This broad definition of a trail may change over time (see Policies and Standards, below).

## Trail Data Users and Contributors

Trail data has many users and many sources (Figure 1). Understandably, many entities are both users and creators of trail data. This makes defining a data production supply chain somewhat more complicated, because those who create trail data do so for their own purposes, not necessarily to share with others. Trail data consumers also have very diverse purposes – the needs of data users may differ greatly. The goal of the statewide trails dataset is to meet common needs and provide a base upon which others may enrich the data should they choose to do so.

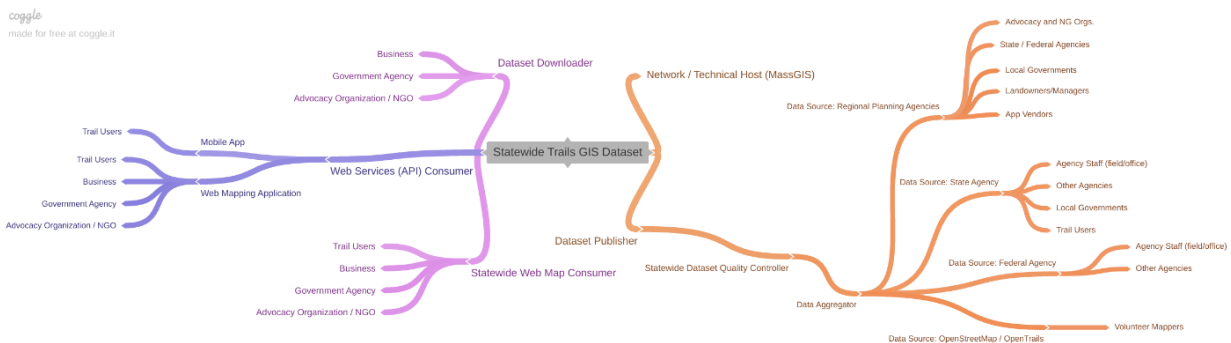


Figure 1. Trail data users (left side of chart) and creators (right side of chart).

## Roles in the Data Production Supply Chain

As Figure 1 shows, one can separate two very general roles involved with trails data: producers / creators and consumers. These simple roles are inadequate for use in designing and managing the creation and sharing or publication of statewide trails data. A fuller set of roles appropriate to the data production process is:

- Owner – Controls access to data
- Authority – Stands behind data as being correct (e.g., official data)
- Data Steward – Takes Responsibility for dataset as a whole: integrity, quality, and availability, often a coordinator of other roles
- Creator – Creates and edits the data in a dataset
- Aggregator – Combines data to populate the dataset
- Publisher – Makes a dataset available to others in one or more forms

Roles separate out duties within the data production program. Consequently, the organization or person filling each role must be clear. More than one role could be taken by an organization or person, but this should be done explicitly rather than just letting duties accrue haphazardly.

Roles can also appear at multiple levels. For instance, there are multiple data creators shown in Figure 1. Each organization shown in that figure might have its own data steward and publisher too. At a higher level, the statewide dataset would still need its own data steward role, aggregator, and publisher. One aspect of governance, discussed in a later section, is sorting out the roles to be filled.

In the near term, the statewide trails data program needs to clearly define a few key roles. The statewide Data Steward, Aggregator, and Publisher roles need to be filled. The Creator, Owner and Authority roles are implicit in this initial work because each regional planning agency is acting in those three roles for its jurisdiction. The Data Steward and Aggregator roles are being filled by the MAPC and the Publisher role by MassGIS.

## Data Supply Chain Model

Figures 2 and 3 show two different data production supply chain models. Figure 2 illustrates a *Hand-Off* model. Figure 3 illustrates a *Portal- Hub* model. Both are practical but they differ greatly in complexity and in the effort that might be required to manage them effectively.

## Hand-Off Data Supply Chain

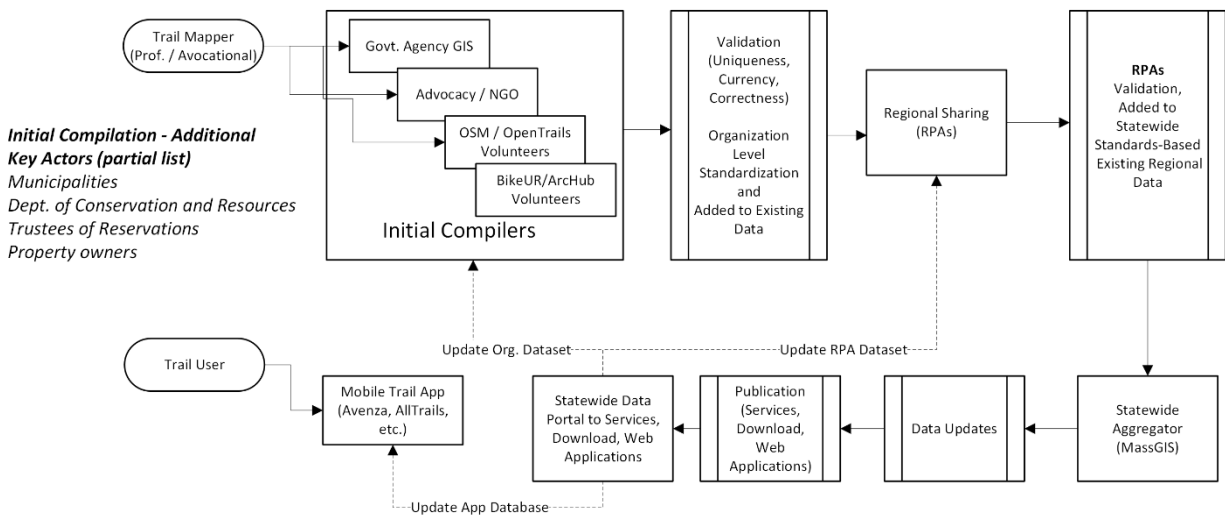


Figure 2. Hand-off Production Chain

The Hand-Off supply chain model (Figure 2) starts from a trail mapper / data contributor who provides trails data to some sort of initial data compiler. The mapper / contributor could be a member of the compiler organization or external to it. The compiler organization adds the contributed data to their own dataset in its organization-specific schema and format. They do so by validating the contributed data, standardizing it to their local schema, and unifying it with their existing data. At this point, the contributed data is standardized and accepted to just a local data schema.

## Portal - Hub Supply Chain

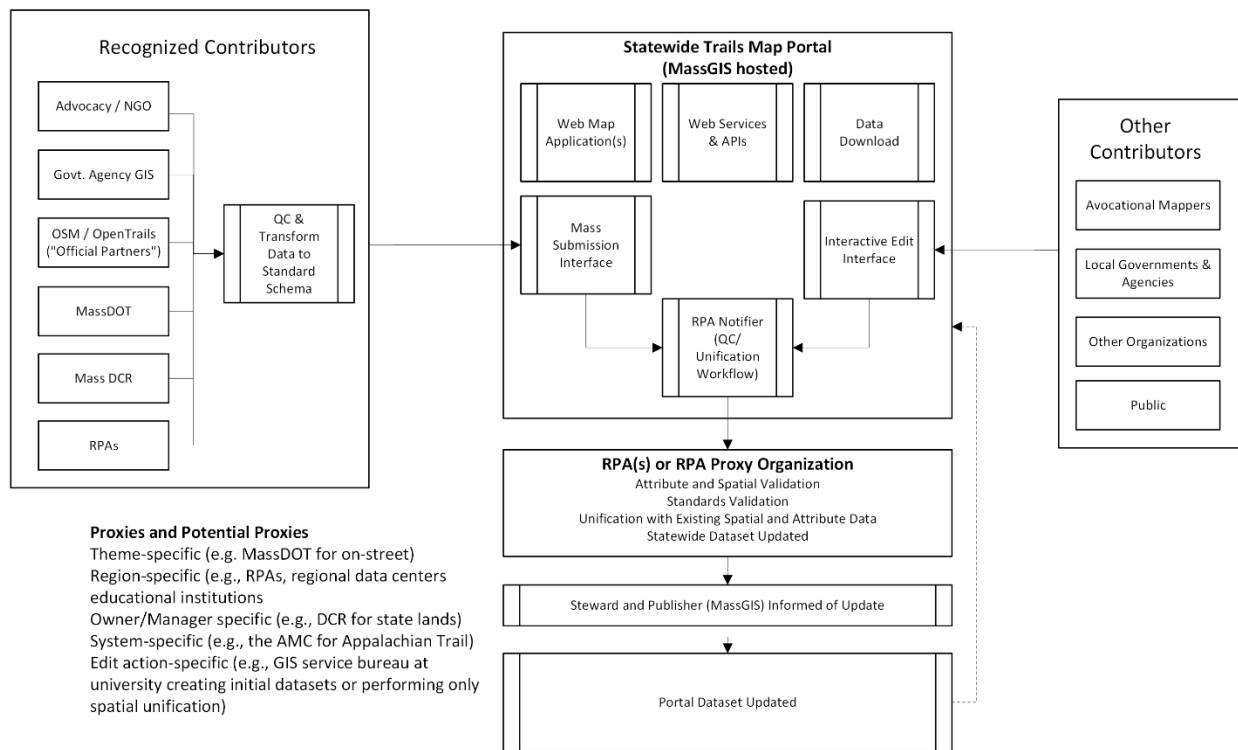


Figure 3. Portal-hub supply chain model.

The next step in the Hand-Off chain is to provide datasets in the local schema to a regional planning agency (RPA; or other organization acting in a larger aggregator role). The RPA is then responsible for aggregating data in its region (Aggregator), taking overall responsibility for its regional data (Steward), and vouching for the correctness of their data (Authority). The RPA is vouching for the correctness of data that it did not necessarily create or own. Furthermore, the RPA data schema may differ from both the contributor schema and the statewide schema. So, transforming the data may be essential when received and again when shared to the statewide compilation process.

The RPA then hands-off their regional data to a statewide aggregator (MAPC or MassGIS in the initial work). The organization in the Aggregator role statewide must again unify and standardize data from the RPA or other regional aggregators.

The next step in the Hand-Off is to the statewide Publisher role. This role shares the data through web applications, download services, web services, APIs, or other means. The supply chain becomes a loop by providing standardized, unified data to users like the

public, other organizations and businesses, and even the subsidiary local and regional participants in the supply chain.

The weakness of the Hand-Off supply chain is that it involves a lot of hierarchical relationships that need management, and that data production can entail redundant effort. On the management side, consider what might be involved in creating a new statewide dataset for annual release. The statewide Steward role must alert RPAs or other aggregators to the desired contribution deadline. The RPAs or other aggregators must alert all their contributors and give them an appropriate deadline for contributing local schema datasets. Local organizations may then need to manage their contributors for timely information. Redundant effort occurs because the same actions happen at multiple levels: stewardship, aggregation, and quality control all happen at the local, regional, and statewide levels.

The Portal-Hub supply chain (Figure 3) avoids some of these weaknesses. In this supply chain, a portal (hub) is the aggregating entity. Trusted contributors can provide data to the hub in a pre-defined format. The hub then incorporates some automated processes to do simple quality control and then alerts a technician at the statewide Aggregator role organization(s) that a dataset needs full quality control. The Aggregator role, including quality control and unification with existing data, could be filled by a single organization such as MassGIS or by each RPA handling trail data within their jurisdiction or by a third-party like a university GIS service center or a vendor.

The Portal-Hub supply chain also allows for individual or ad hoc proposed additions or edits. Rather than always arriving as formal datasets from known, trusted, contributors, these proposed changes are made in a web-based interface where users can provide mark-up, load spatial data in common formats (e.g., KML) or provide data in the standard format used by trusted contributors. The proposed edits again follow a common quality control and aggregation process.

The weakness of the Portal-Hub supply chain is that it depends on a technically efficient and capable hub and portal supported by strong management of it. However, this can be the responsibility of just a single organization. The Hand-Off supply chain relies on many links to form the chain and weak links at different levels can affect the overall outcome. The portal-hub supply chain has fewer links overall. So long as a weak link doesn't break the chain, it may be more easily identified and strengthened.

Advantages of the Portal-Hub approach is inclusive of trusted contributors at any level, lessening workload on RPAs or any other regional aggregators. It also makes timing of updates easier: the Publisher can state a deadline to all contributors since they all work at the same level rather than in hierarchical fashion. Finally, this data production supply chain model is flexible. Quality control and aggregation can be performed by regional partners, by a single statewide organization, or by a third party, or by any combination of

these. The Portal-Hub production and data production supply chain is the preferred model for building and maintaining the statewide GIS trails dataset.

Two different data production supply chain models could be used in the statewide trails dataset program. The Portal-Hub model is stronger, more flexible, and has fewer potential points of weakness. The Hand-Off model has been used for many years in data systems though, underlying many data programs developed before networked workflows were well-supported by technology. It is incorrect to say the model is not a successful one; however, the model requires more stringent oversight.

## Section 3: Statewide Trails Dataset Creation, Management, and Maintenance

Earlier sections of this roadmap defined why the statewide trails dataset supply chain is needed and selected a model for data production and maintenance. This section turns to practical guidance and protocols for the statewide trails GIS data. The statewide trails GIS dataset is already being built in a near-term initial effort. The roadmap subsections that follow distinguish between near-term (ongoing, current) work and long-term needs and actions wherever necessary to do so. For example, the data supply chain model is a long-term action, since there is already a production process for the initial dataset compilation and publication.

### Creating the Data Production Supply Chain

***Section to be prepared by MAPC and MassGIS***

The Portal-Hub supply chain model is needed to create a long-term sustainable method for keeping statewide trails data complete and up to date. This section outlines the steps needed to build the long-term supply chain. The current, near-term, work is described where relevant. Use this section as a checklist or process guide for the long-term system, tracking the needs and subsequent actions taken to build the data supply chain.

### Define and Fill Required Roles

The trails data production supply chain involves actors in several roles. These are shown in the table below for the current near-term work and for the long-term data production supply chain model. As the scope of trails aggregation and publishing scales from local, to regional, to state, and to a national level, roles will need to be adjusted and reassigned to meet the needs of the stakeholders. The roles presented in the following table are currently defined and assigned to support the aggregation, publishing and ongoing maintenance of

trails data at the statewide level. As the roles are assigned to specific entities, revise the table.

| <b>Role and Description</b> | <b>Description</b>                                       | <b>Near-Term</b>       | <b>Long-Term Model</b>   | <b>Comments</b>   |
|-----------------------------|--|------------------------|--|---|
| Owner                       | Controls the system and access to it                     | MassGIS                | TBD – see governance section                                     |   |
| Authority                   | Stands behind data as being correct/official             | Data contributor (RPA) | TBD  |   |
| Data Steward                | Takes Responsibility for dataset as a whole              | MassGIS                | TBD  |   |
| Creators – Contributors     | Creates new or updated data                              | RPAs                   | RPAs + Recognized Contributors; Other Contributors               |   |
| Creators – Revisors/Editors | Maintains datasets by incorporating new or updated data  | RPAs, MassGIS          | RPAs and/or RPA Proxy Organizations                              | These act as revisors to ensure quality control; the hub itself may perform the mechanics of merging and updating data values |
| Aggregators                 | Combine data to populate the dataset                     | RPAs, MassGIS          | Hub itself (technology) plus RPAs and/or RPA Proxy Organizations |   |
| Publishers                  | Makes a dataset available to others in one or more forms | MassGIS                | MassGIS using Portal-Hub technology                              |   |

|       |  |  |                                 |  |
|-------|--|--|---------------------------------|--|
| Users | Consumers of dataset publication (whatever form) | RPAs, MassGIS, current collaborators<br><br>Planners, MassTrails program, general public | General public, all other roles |  |
|-------|--|--|---------------------------------|--|

Beyond this table there are groups who might provide advisory feedback, consume data or downstream data products, and contribute subject matter expertise to the MassTrails Coordinating Body and trails stakeholder groups. These groups may include:

- Conservation Groups
- Trail Users
- MARTAB – Massachusetts Recreational Trail Advisory Board (DCR)

### Define functional and technical requirements and specifications

Building the data supply chain requires defining its functional and technical requirements and using these to write some form of system description and specifications. For the near-term effort this is largely in place: a workable arrangement of sharing data, aggregating it into the current standard, and publishing it in place.

In the long-term this subsection of the roadmap should contain actual needs statements (requirements) and specifications in some form, or at least pointers to these documents.

### Define an implementation plan to build (instantiate) the system as software, hardware, networking and services, and costs

In the near-term, this work has already been done.

In the long-term, creating the Portal-Hub supply chain should follow common practices used in creating any technology solution. These include a description of the system, how it will be configured, what needs to be acquired to instantiate it, and the costs. Revise this subsection with references to the actual architecture and “build” of the supply chain.

### Define a budget and process for maintaining the supply chain

In the near-term, this work has already been done or is not relevant.

In the long-term, creating the Portal-Hub supply chain should follow common practices used in creating any technology solution. These include determining ongoing costs for the system, staffing needs and how those positions will be funded, and other costs. Revise this subsection with reference to the actual cost estimates, sources, etc.

## Seek opportunities to build and support the system

In the near-term, this work has already begun.

The point of this subsection is to track potential partners, grantors, and participants who might support building the supply chain. Revise this subsection with notes, ideas, opportunities to follow and their outcomes.

## Managing the Chain – Establishing Data Governance

### ***Section to be prepared by MAPC and MassGIS***

The statewide trails database is a collaborative work. Data governance is an essential part of the work because it ensures policies and standards are being met and that role-based functions are performed. Coordinating bodies need to be granted the authority to make decisions about the trails data, the standards and policies, and the data publishing guidelines that are required to make this information available to the public. In the near-term, MAPC, MassGIS, DCR and the RPAs are the key members of the trails data coordinating body. These four groups are accountable for the trails data governance design, implementation and compliance. MassGIS will help convene these four groups, as well as additional key stakeholders such as MARPA and MassDOT GIS into a work group. Work group members will have access to the GIS system as admins and will convene for governance purposes on a quarterly basis.

However, in the long term, a coordinating body needs to be created, likely as a standing work group under the Massachusetts Geographic Information Coordinating Council (MaGICC) coordination umbrella. In the long term, collaboration and coordination are intended to deliver the most utility to statewide stakeholders. Because of this the coordinating body will need to expand to include other trails contributors and subject matter experts.

***This section will evolve as the work moves forward. For now, it is an outline of what is needed to put governance in place.***

Near Term

- Convene a small working group of select stakeholders to meet and coordinate efforts prior to data publishing.
  - MARPA / RPAs
  - DCR MassTrails
  - MassDOT
  - MassGIS (working group coordinators)
- Review near-term guidance against current trails standards, policies, and implementation status.
- Identify missing components and definitions, come to agreement on solutions
- Prepare for larger working group sessions with the expanded MassTrails Coordinating Body.
- Review data for publishing, and confirm that the minimum viable product meets the base requirements of the data.

### Long Term

- Create a Coordinating Body (long-term)
  - Coordinating Body Membership
  - MassGIS
  - RPA Representatives (3)
  - MassDOT (1)
  - MARTAB
  - DCR (1)
  - Conservation Organizations (2)
  - App Industry (1)
  - Federal Agencies (1)
  - General Public / Volunteer Geographers (1)
  - OpenStreetMap US (1)
  - Others
- House the Coordinating Body – Administrative Home (choose one)
  - Part of MassGIS Coordinating Committee (Working Group / Subcommittee)
  - Independent of MassGIS Coordinating Committee
- Create a Charter
  - Duties of Coordinating Body
  - Meeting frequency
  - Membership selection and replacement
  - Duties of members / roles within coordinating body
  - Responsibilities of the Coordinating Body (at least these)
- Creating and revising policies
- Creating revising standards
- Establish Data Governance Roles and Duties
  - Roles
    - Owner – Controls access to data

- Authority – Stands behind data as being correct (e.g., official data)
    - Data Steward – Takes Responsibility for dataset as a whole: integrity, quality, and availability, often a coordinator of other roles
    - Creator – Creates and edits the data in a dataset
    - Aggregator – Combines data to populate the dataset
    - Publisher – Makes a dataset available to others in one or more forms
  - Fill governance roles
- Funding Plan and Sources
  - Cost estimations
    - Near-term costs
    - Long-term startup costs
    - Long-term sustaining costs
  - Identify funding sources
  - Define funding plan

## Policies

### ***Section to be prepared by MAPC and MassGIS***

Establishing robust policies ensures the integrity, quality, and security of data throughout its lifecycle. Data supply chains involve multiple stakeholders and require clear policies to ensure seamless and secure data flow. By defining and implementing governing policies, organizations can achieve improved data quality, enhanced security, increased efficiency, better decision-making, and process compliance. ***Some processes that could be formalized as policies have already been established by MAPC and MassGIS. In the following section, review and expand on the required programmatic and operational policies required to govern the trails data supply chain.***

***This section will evolve as the project progresses. For now, policies needed in the near-term and long-term are provided in outline form.***

- Near-Term Policies
  - Programmatic
    - Define data partners and establish a data partnership agreement
    - Define how data is used/ shared across public agencies
    - Define how private landowners participate in the trails dataset
    - Define the intended audiences and approved uses of the data (and define a disclaimer)
    - Define whether/ how data is permitted to be used commercially
    - Define whether/ how commercially aggregated data is used in the statewide dataset
  - Operational/ Technical

- Define when/ how often data should be delivered
  - Define who is entitled to contribute/ edit data in the data portal
  - Define how data portal accounts are requested/ administered
  - Define how public trails are confirmed to be open
  - Define how trails on private property are confirmed as open/ accessible to the public
  - Define how conflicts are resolved among contributed data
  - Define how data is attributed
- Long-Term Policies
  - Programmatic
    - Define how statewide trails data governance aligns with statewide data/ IT governance
    - Define how MassTrails data standard aligns with relevant national data standards and initiatives
    - Define how the Coordinating Body supports implementation of policies and standards (e.g. technical support, grant-making, etc.)
  - Operational
    - Define when/ how often data should be delivered in a mature supply chain
  - Policy Framework
    - Policy Development
    - Convene policy working group
    - Define authority (MAPC/ MassGIS/ Coordinating Body) for policy approval
    - Define scope, outcome(s), and content of policy
  - Policy Approval
    - Coordinating Body conducts peer review of policy
    - Working group revises policy as needed
    - Authoritative body adopts policy
  - Policy Implementation
    - Coordinating Body provides resources to stakeholders to support policy outcomes
    - Policy working group/ coordinating body conduct periodic evaluations of policy

## Standards

### ***Section to be prepared by MAPC and MassGIS***

Standards are essential to govern a data supply chain at both the state and national level as they ensure consistency and reliability in data management processes. Standards define clear roles and responsibilities for all stakeholders involved, facilitating seamless

collaboration and data flow. Standards also help maintain data quality by establishing guidelines for data collection, storage, and usage. Additionally, they enhance data security by setting protocols for access control and protection against unauthorized access. **MAPC has a well-defined trails data standard. Guided by the prompts in the following section, review the existing standard and explore future opportunities to broaden its components to meet state and national trails data sharing and coordination.**

***This section will evolve as the project progresses. For now, standards needed in the near-term and long-term are provided in outline form.***

- Near-Term Standards
  - Utilize Trails Data Standard Version 1
  - Seek continued/additional user input on value of current standard for their uses
  - Determine through testing/implementation if standard has issues in implementation for specific high priority use cases and if so how to address them
  - Propose near-term revisions (or move this to long-term)
  - Develop proposed standards formal review process
    - Trails steering committee / advisory body responsible for this
    - Process should be open and explicit
- Long-Term standards
  - Define “authority” regarding trail data and management (recurring issue found in workshops and discussions)
  - Collaborate on regional and national trail standards
    - Define how Massachusetts trails dataset will (will not) nest, contribute, or fit with regional and federal schemas
    - Relationship of Massachusetts trails to national and regional trails as data
  - Evaluate prescriptive vs. descriptive standards and schemas

## Improving and Maintaining Data Quality

***Section to be prepared by MAPC and MassGIS building upon Sanborn content***

Data Quality is a core component of any geospatial dataset as the quality of data ultimately dictates the data consumer’s ability to use the data, and trust that it is presenting accurate information. Specifically for trails, the quality of the data, and the trust in the data will impact the safety of the end user. As data is aggregated from various sources and contributors across the state, mechanisms that validate that the quality of the submitted data meets the quality requirements as defined in the data standards are required. ***The following section defines the components that will guide MAPC and***

**MassGIS to define data quality requirements, carry out quality assessments, and then implement the tasks required to clean and then maintain high quality data.**

The primary objective of publishing the trails data is to allow the public, RPAs, and other stakeholders to see an aggregated version of state of the trails data. Without visibility into the composition of a statewide view of the trails data, the published data will never be able to be improved. The MassTrails Coordinating Body needs to publish a minimum viable trails product and then evaluate its completeness, accuracy, and quality.

As the MassTrails Coordinating Body Minimum works toward publishing this minimum viable trail product the near-term quality efforts to ensure three key characteristics: privacy, safety, and completeness. Privacy data quality means that the RPAs have removed or hidden trails that local, state or private entities have requested remain hidden from the public; Safety data quality means that RPAs are publishing correct and complete data to the best of their knowledge; Completeness means that all RPAs have submitted their complete inventory of trails data that is currently available to them. Aggregation and publishing of data from every RPAs might not include 100% trail coverage but includes 100% data coverage across the RPAs.

Additionally, the MassTrails Coordinating Body needs to work towards confirming the following items prior to publishing:

- Unique ID and Geometry, two core requirements, must be standardized and complete. This follows the requirements from the v1 data standard.
- Common attributes are standardized if available. Null values are allowed.
- Duplicate trail records are removed, if possible, so that double counting of trail measures and inventories are eliminated.

In the long-term, data quality will always be a consideration, and the subsections are a guide to improving and maintaining data quality. These subsections should be addressed early in long-term development, expanded, revised, and updated (with superseded sections retained in an archive) as the system evolves.

## Define Quality Requirements for Trails Data

The table below presents 6 components of data quality (Sanborn Geospatial. [\*Business Plan to Improve Geospatial Data Value\*](#). May 30<sup>th</sup>, 2024, p.7-9.) For each component, identify the data quality requirements and standards of the trails data. These requirements and data standards will guide data quality validation and improvement efforts. For example, a data quality requirement for *Completeness* might be “all trail records have an assigned trail name value”.

| <b>Consistency</b>   | <b>Completeness</b>  | <b>Precision/<br/>Accuracy</b>   | <b>Integrity</b>   | <b>Timeliness</b>  | <b>Relevance</b>   |
|--|--|--|--|--|--|
| A measure of the similarities and differences between data stored in multiple datasets or databases. | A measure of how comprehensive the information in a dataset is. Are required or priority fields populated? | A measure of the accuracy of the data. Is the accuracy stated? Does it match end user's needs? | An assessment of whether the data's structure, schema and maintenance workflows meet end user's needs and support interoperability between systems | An assessment of the data's update cycles, and temporality to internal and external deadlines and reporting needs. | An assessment of the accessibility and availability of data required to inform business systems and answer business questions. |

### Conduct Data Quality Assessment

Before one can improve data quality, a systematic approach to *assessing* data quality is important. Based upon initial feedback during the roadmap development workshop the assessment should include at least the following:

- Duplicate trails are removed
- Trails are spatially accurate (based upon a standard for accuracy)
- Commonly needed attributes (based upon a standard) are fully populated
  - Core (required)
  - Common (recommended)
- Trail status (*public\_access* v1 standard field) are defined and correct
- Trails are connected in a linear network
  - Establish network topology rules in GIS database
  - Ensure abutting trails connect at nodes
  - Eliminate self-intersecting and overlapping trail records
  - Remove overlaps of trails from different jurisdictions

Other assessments will come from a consideration of the data qualities sought, as discussed in the previous subsection. This section of the roadmap should track planned assessments and assessment results. These may repeat over time, so some method of flagging current data quality assessment of each part of the dataset will need to be devised.

## Implement Data Quality Actions and Tracking

When an assessment is made, the most important step is to determine what actions are needed to remedy data quality issues. In the near-term, this is left to the judgment of the current team. In the long-term, a systematic approach is part of this roadmap.

The following table presents a template that can be used to define, plan, and implement the immediate data cleansing activities and the longer-term data maintenance activities required to ensure high quality trails data. The template is filled out for each characteristic that has a deficiency (e.g., spatial accuracy, see subsection above). The second table is a hypothetical example of the header and first entry in such a table. As outcomes are achieved, they are changed from active to inactive by adding a progress assessment plus date entry.

| <b>Data Quality Improvements for [Quality Characteristic]:</b> <Describe goals in narrative form> |                                     |                             |  |                       |                                  |                            |                        |                 |
|---|-------------------------------------|-----------------------------|--|-----------------------|----------------------------------|----------------------------|------------------------|-----------------|
| <b>Success Criteria:</b> <Describe success criteria in narrative form>                            |                                     |                             |  |                       |                                  |                            |                        |                 |
| <i>Outcome</i>  | <i>Action</i>                       | <i>Ongoing Action (Y/N)</i> | <i>Progress Assessment (Date)</i>                  | <i>LOE Staff Time</i> | <i>LOE/Cost External Support</i> | <i>Infrastructure Cost</i> | <i>Risk Assessment</i> | <i>Timeline</i> |
| <Describe desired outcome, headed by time frame: “Current”; “Near-term”; “Long-term”>             | <Actions needed to achieve outcome> |                             | <Date and progress toward outcome; may be updated> | <hours>               | <hours and/or costs>             | <costs>                    |                        |                 |

| <b>Data Quality Improvements for Trail Vector Spatial Accuracy:</b> Trail vectors will be spatially correct in their geographic location to within 10m of ground coordinates observable in the field. This accords to slightly more accurate than the production standards for USGS 1:24,000 scale published map data,   |   |                             |  |                       |                                  |                            |                        |                 |
|--|---|-----------------------------|--|-----------------------|----------------------------------|----------------------------|------------------------|-----------------|
| <b>Success Criteria:</b> 90% of randomly selected points on vectors will fall within 10m of their true coordinates on the ground. This assessment may be performed through random sampling followed by machine-learning verification to identify intersection between assessment point locations and NAIP-level accuracy orthoimagery showing trail surfaces so that a system or analyst compares a GIS location with a controlled image; or may be assessed using field GPS techniques. |   |                             |  |                       |                                  |                            |                        |                 |
| <i>Outcome</i>   | <i>Action</i>   | <i>Ongoing Action (Y/N)</i> | <i>Progress Assessment (Date)</i>                                  | <i>LOE Staff Time</i> | <i>LOE/Cost External Support</i> | <i>Infrastructure Cost</i> | <i>Risk Assessment</i> | <i>Timeline</i> |
| 15% of trails in dataset lie more than 10m from true ground location based on assessment method  | For vectors not meeting data quality standard, redigitized with orthoimagery or other | Y                           | 12/12/2024 created data quality improvement statement (this entry) | Unknown               | Unknown                          | None                       |                        |                 |

|  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|
|  | sources, perform field GPS measurements to get true coordinates, or both |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|

## Publishing Data

### ***Section to be prepared by MAPC and MassGIS.***

In the near-term, MAPC and MassGIS are currently collaborating on aggregating and publishing trails data to the public through MassGIS Portal and services. ***This section will evolve as the project progresses. For now, data in the near term needs to be published. In the long-term, the additional considerations, provided in outline form, need to be discussed among the coordinating body.***

- Near Term Publishing
  - Minimum viable trails data product will be published
    - Trails data currently being worked on by MAPC and RPAs between January and June 2025.
    - Publish in June of 2025.
  - Trails data portal
    - Published trails data
    - Trail data metadata
    - Overview of MAPC and RPA trails initiative and plan
    - Messaging on data use and disclaimers
- Long Term Publishing
  - Frequency
    - Calendar-based (yearly, quarterly, etc.)
    - Event-based (when a batch of data is approved after update or creation)
    - Hybrid approaches (calendar always, event-based if very important for some reason)
  - Publicizing updates
    - Mechanisms to alert participants, stakeholders, and general users of publication changes and news
  - Coordination with Other Publishers

- Coordination with state, regional, federal, and non-government organizations with similar aims and publication programs
- Minimize duplicative effort and drive toward single source of truth for each trail

## Marketing and Public Awareness

***Section to be prepared by MAPC and MassGIS – Work with RPAs to develop the best approach.***

Making the professional and public aware of the statewide trails dataset itself and of opportunities for contributing to it is an important part of the supply chain success. In the near-term, state-funding agencies already recognize the value of a statewide dataset. In the long-term, the coordinating council or body must create awareness of what Massachusetts is achieving.

This section should contain an overview of how marketing and public awareness is done for the data program. This could be as detailed as communications and marketing strategies and business plans or as simple as media and agency email lists used to contact interested parties.

## Appendix A: Trail Data Users and Consumers

See attached files: Appendix A - Trail Data Users and Consumers

## Appendix B: Hand-off Supply Chain Model

See attached files: Appendix B - Hand-off Supply Chain Model

## Appendix C: Portal-hub Supply Chain Model

See attached files: Appendix C - Portal-hub Supply Chain Model