



Data Governance Framework

The Path Toward Unified Data Management

Prepared for:

State of Vermont

Kristin McClure

Mahesh Thopasridharan

Bechir Bensaid

Prepared by:

Briljent LLC

Laura Larimer llarimer@briljent.com

Cedric Fillmore

Connor Norwood

Heidi Fox

Susan Clark

Due by: September 30, 2022

This page is intentionally left blank.

Table of Contents

Introduction	1
The Unified Health Data Space Defined.....	1
Goals and Objectives for the Unified Health Data Space.....	1
Expanding the Vermont Health Information Exchange	1
Data Governance and Data Integration Defined	2
<i>Data Governance</i>	2
<i>Data Integration</i>	2
Initial Target Data	2
Need for Governance	3
Data Hub: Aligning the VUHDS to Industry Best Practice	3
Key Principles	4
Conceptual Alignment.....	4
Foundations of Governance	5
Alignment to Strategic Principles.....	5
Objectives for Data Governance	5
Data Governance Framework	6
Quality Data	7
Secure Data	7
Connected Data.....	7
Purposeful Data	8
Communication.....	8
Collaboration.....	8
Culture	8
Governing by Domain	8
Hybrid Federated with Distribution Domain	9
Fully Governed with Distribution Domain	10
Framework Evolution	11
Operationalizing Governance	11
Governance Structure.....	11
<i>Data Governance Council</i>	12
<i>Federated Teams</i>	13
Roles and Responsibilities.....	18
<i>Process Owners</i>	18
<i>Data Domain Roles</i>	18
Data Access Based on Roles	19

<i>Identity Management and Access Management</i>	19
<i>Role-Based Access</i>	19
How the Teams Work Together	20
<i>Intake/Capture</i>	21
<i>Review/Analyze</i>	22
<i>Decision</i>	22
<i>Actions</i>	23
<i>Monitor and Assess</i>	23
Decision-Making Throughout the Governance Process	23
<i>Federated Team Level Decisions</i>	24
<i>Decision Tracking/Logging</i>	24
<i>Policy Development Guidelines</i>	24
The Work of Governance	24
Moving from Use Case to Data Governance	25
Communications Planning	27
<i>Communicate, Communicate Again, and Communicate More</i>	27
<i>Diversify the Communications</i>	27
Considerations	28
Funding	28
Resourcing	28
Process Continuity Through Knowledge Management	28
Domain Maturity	28
Vendor Data Implications	29
Metrics for Governance Success	29
Immediate Next Steps	29

List of Figures

Figure 1: Conceptual data hub	4
Figure 2: Data governance framework pillars	7
Figure 3: Hybrid federated governance framework with distribution domain	10
Figure 4: Fully governed framework with distribution domain	10
Figure 5: Recommended data governance structure	11
Figure 6: Access Perspectives	20
Figure 7: Data governance process	21
Figure 8: Gartner diagram linking business use to key data management steps	25

List of Tables

Table 1: Initial target data..... 3

Table 2: References that include frequently used domains..... 9

Table 3: Suggested Core Federated Teams..... 14

Table 4: Roadmap for data governance 29

Introduction

With a vision for unifying health data from many sources to serve state agencies, external organizations, and its constituents, the State of Vermont has embarked on a project to create a Unified Health Data Space. This project has the opportunity to unite data and organizations and fulfill key public health goals. To make this effort as successful as possible, establishing proper data governance for the Unified Health Data Space is necessary. This document outlines how data governance will support the Health Data Space project, the fundamental pieces of data governance, and next steps the State of Vermont should take to integrate data governance into project planning.

The Unified Health Data Space Defined

The State of Vermont envisions the creation of a Unified Health Data Space that integrates claims and clinical data, as well as other relevant data. The Unified Health Data Space will support health care delivery and coordination, public health operations, health plan operations, Medicaid-driven delivery system reform, and population health management activities. Utilizing an iterative process, the State of Vermont will build the Unified Health Data Space and then grow and mature its technical infrastructure, use cases, governance, and overall value.

The Unified Health Data Space is the foundation on which Vermont will realize its goal for the Health Information Exchange to create one, centrally accessible health record for each Vermont patient. The single-record approach supports optimal care delivery and coordination, bolsters health care operations, and enables a learning health system. Vermont's technical strategy for achieving a central record is to develop and support the use of the Unified Health Data Space.

To enable these outcomes, Vermont is on a path towards including diverse data types within the centrally located health record - clinical, claims, behavioral health, social determinants of health, and more.

Goals and Objectives for the Unified Health Data Space

The development of the Vermont Unified Health Data Space supports the goals identified in the Vermont Health Information Exchange (VHIE) strategic plan. VHIE's system goals include:

1. **Create One Health Record for Every Person** - Support optimal care delivery and coordination by ensuring access to complete and accurate health records.
2. **Improve Health Care Operations** - Enrich health care operations through data collection and analysis to support quality improvement and reporting.
3. **Use Data to Enable Investment and Policy Decisions** - Bolster the health system's ability to learn and improve by using accurate, comprehensive data to guide the investment of time, labor, and capital, and inform policy making and program development.

Expanding the Vermont Health Information Exchange

To accelerate the realization of the Vermont Unified Health Data Space, the existing Vermont Health Information Exchange (VHIE) must be leveraged as a foundation for establishing a unified record and integrating new data sources. This approach is in line with the 2021 HIE Strategic Plan's future focus area "Improving Public Health Capabilities Through Integration with the VHIE." This focus area was described as follows: "To fully participate in the vision of a Unified Health Data Space, public health data and systems must be integrated with the VHIE. In 2022, the Agency of Human Services (AHS) will partner with VITL to design a strategy for bolstering public health management through a partnership with the

VHIE. The integration strategy will consider leveraging VHIE infrastructure to enhance public health operations (an effort to reuse existing technical capabilities), incorporating public health data (e.g., immunizations, birth/death, cancer records) into the health record contained on the VHIE, and reporting for public health surveillance and monitoring. The strategy will build on lessons learned from the COVID-19 response and account for new federal requirements for state-level reporting. The HIE Steering Committee will provide feedback on the public health data strategy and use it to inform ongoing efforts to propel state-wide health information exchange.”

In short, the Vermont United Health Data Space will not be a total reinvention; instead, it will build on existing technical infrastructure, relationships, and successes to improve outcomes and enhance and add to data capabilities.

Data Governance and Data Integration Defined

Throughout this document, the terms data governance and data integration are used frequently. They will also be relevant and often cited phrases throughout the framework’s implementation. Therefore, it is beneficial to have a common, specific definition of the terms with respect to their use in the Unified Health Data Space.

Data Governance

Data governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods. Data governance provides the structure and processes needed for decision-making. It clearly identifies the roles and responsibilities of those who take part in those processes and how they should collaborate.

Simply put, it is a structured means of collaboration to ensure the value and use of data is maximized while also ensuring its quality and security.

Data Integration

Data integration is the process of combining data from different sources into a unified view or format. Making decisions often requires information from more than one source or system. When faced with more complicated questions, the number of needed sources can be difficult for a single person to handle. Data integration can benefit a large number of users when data product teams work together to combine data for use. It minimizes the work for the end user, leaving them to perform more analysis, ask more questions, and get things done. When undertaking integration, data teams standardize how information is combined, ensuring its quality and security.

In the context of the Unified Health Data Space, VHIE data serves as the foundation while additional data will be incrementally introduced to the space, adding context and value for the users of the environment. Additional data may come in the form of a unified health record or newly added information that provides context to a specific population or on a specific question.

Initial Target Data

The Unified Health Data Space will be initiated with the existing data in the VHIE. The following chart shows the initial target types of data as well as the owners of the source data. These are the groups/organizations that have a responsibility to capture the data.

Table 1: Initial target data

Source Owners	Type of Data	System	System Governing Entity
Providers	Clinical	VHIE	VHIE Steering Committee
Payers	All-payer claims	VHCURES	GMCB

Notice the various source owners, systems, and governing entities. As the vision for the Unified Health Data Space is implemented, the number of sources, systems, and potential governing entities will increase. While value can be added with each additional source, each source also introduces new stakeholders, entities, and use cases that can benefit from robust data governance.

Need for Governance

To realize the vision of an integrated health data infrastructure (Unified Health Data Space), Vermont needs data exchange policies and processes that recognize individualized needs while supporting holistic care, system measurement, and improvement. More data sources, each with its own regulatory and statutory restrictions, technical teams, and unique subject matter experts must be included.

The purpose of this document is to provide clear governance models by which health data is stewarded and details behind why, when, and how data may flow through the Unified Health Data Space. Data is reflective of people and should be used to improve health outcomes and advance population health equity. To achieve this, Vermont must operationalize the data strategy for implementing the health data space with the following governance model. The model ensures the appropriate structure for planning and governing the health data space, coordinating services, and appropriate use of the data to advance clinical and public health objectives.

This data governance framework and recommendations document makes frequent reference to the Health Data Strategy, as it provides the strategic focus and guidance behind the activities outlined in this document.

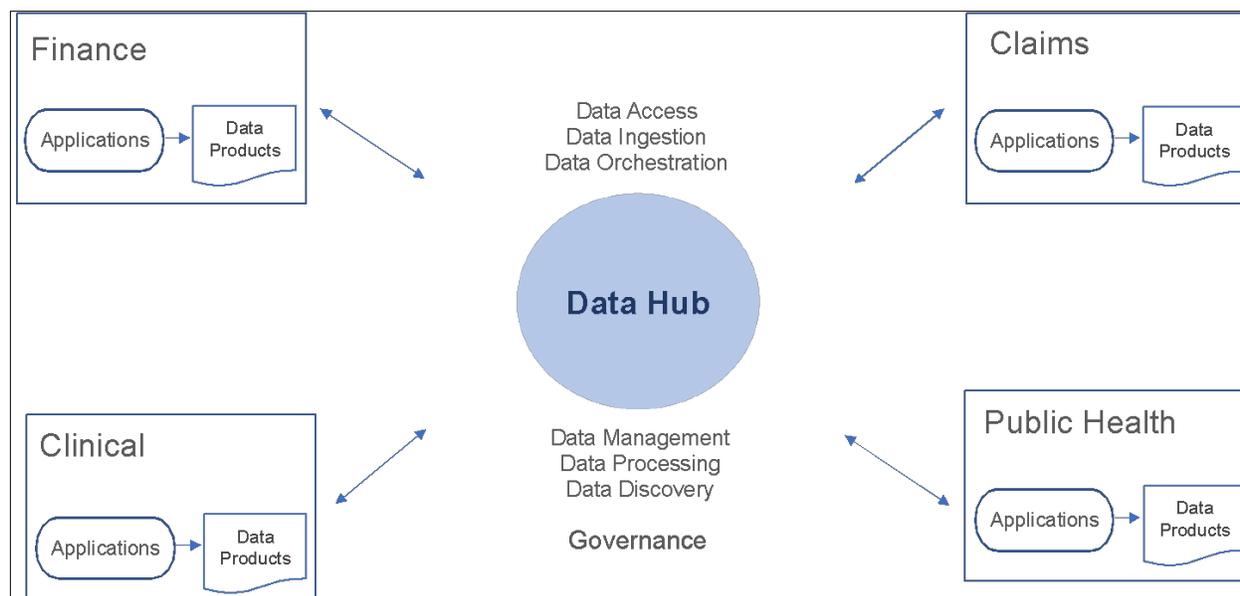
Data Hub: Aligning the VUHDS to Industry Best Practice

Many organizations have tried data management approaches ranging from distributed, at-the-edge, management to centralized, massive systems to end all systems. There are major shortcomings to both approaches, but an emerging hybrid approach is becoming standard practice. This approach allows for a central intermediary point that connects data from across data sources to data consumers. The hybrid approach does not rely on a single technology. It conceptually unites data storage, data integration, and data orchestration (how data moves between systems and applications) and is often referred to as a data hub.

A data hub serves as a single point of access for all data consumers, whether the consumer is an application, a data scientist, or a business user. It also allows for managing data used in various tasks, providing centralized governance, and data flow control capabilities. A data hub connects multiple sources of information, including that sourced from a data warehouse or data lake, to provide users with the required data in a usable structure. Data hubs are great for both data sharing and consolidation. The hub becomes a distribution station, which means it can collect information only to harmonize it and potentially send it to the other end-point systems.

The diagram below illustrates how a data hub can both receive data from and be utilized by multiple stakeholder users or organizations.

Figure 1: Conceptual data hub



Key Principles

The data hub concept subscribes to the following principles which guide how it is both implemented and governed:

Data as a product - Data teams adopt a product lifecycle approach to data. A data product's lifecycle follows agile principles by being short and iterative to deliver quick, incremental value to data consumers.

Data flow orchestration - Data hubs orchestrate data flows to unify, transform, and enrich data, for seamless sharing of data between data producers and consumers.

Operational focus - Drive real-time operations, with continuous bi-directional data movement between source and target systems

Business domain ownership - The people closest to the data control it – the domain experts. Local control increases IT/business collaboration, enables business agility, and decreases time to value.

Data governed where it is - Domains have the autonomy to create and deliver data products while also adhering to global standards which maximize speed and agility.

Conceptual Alignment

The data hub is an industry best practice approach that conceptually aligns with the overall goals of the VHIE and the intended expansion into the Unified Health Data Space. Best practices for each component should continue to be taken into consideration as the Unified Health Data Space matures its data and governance over time. At this time, the best-fit approach for the Vermont Unified Data Health Space is to utilize a data hub.

Foundations of Governance

Traditional approaches to data governance can be quite rigid, rife with heavy organizational structure and controls that limit the expanded use of data rather than accelerating it through evidence-based hypothesis testing and analysis. Adopting a rigid approach does not achieve the appropriate balance of control and agility, leading to either the business bypassing policies to get the data it needs or IT over-centralizing its responsibilities. Though there is no one-size-fits-all approach to data governance, any successful program should have remnants of the following foundational elements.

According to Gartner, there are [seven key foundations for Modern Data Governance](#), all of which are important to driving value and desired outcomes:

1. Alignment to outcomes
2. Trust
3. Transparency and Ethics
4. Accountability and Decision Rights
5. Collaboration and Culture
6. Risk and Security
7. Education and Training

The approach to implementing data governance must be purpose-built to fit the reality of the organization's culture. Success relies on the organization's ability to formalize governance at the speed and level necessary to enable teams to accomplish their goals.

Alignment to Strategic Principles

For the State to use data in investment and policy decisions, data must first be trusted and of high quality. Creating trusted and high-quality data is the responsibility of all parties in the State and in external organizations. Data exists throughout each State agency and beyond. That data is owned and managed by several different groups, each with its own special policies. Creating a unified space that brings data together requires extensive collaboration between those groups. That same collaboration is required to manage the security risks associated with bringing information into a central hub. The governing bodies must be transparent about the purpose of bringing information together; transparency will help as the system is being built and will answer any future questions about the ethics surrounding such integrations. For these reasons, the seven foundational principles are critical to data governance and will be at the core of the approach to the Unified Health Data Space.

Furthermore, the seven foundational principles support data governance as well as the future focus items laid out in the HIE Strategic Plan:

- Governing Use of Expanded Health Records on the VHIE
- Improving Public Health Capabilities through Integration with the VHIE
- Improving Access to Health Information (for people and those who care for them)
- Leveraging the Unified Health Data Space to Support Medicaid Operations

Objectives for Data Governance

In preparation for establishing the Unified Health Data Space, numerous stakeholders were interviewed to establish an understanding of the desired outcomes, purpose, and goals of establishing a data governance program. After analyzing the feedback, the following themes arose from those conversations:

- Ensure **secure access** to all data provided through the Unified Health Data Space. The stakeholders know the challenges and restrictions of the diversity of data on the roadmap for this environment. It will take coordination and clear communication to maintain the security of peoples' health data and to allow the State to continue to operate as good stewards of that information.
- Ensure the quality of data provided in the environment. Combining data in this environment will have little value without confidence in its quality. The more data in the environment, the more challenging maintaining quality becomes. Governance is essential to guide the processes to **ensure quality data**.
- Accurately **unify data** into a single linkable record while maintaining integrity. Having a single record increases the value of data to the State and ultimately its people by introducing the ability to analyze problems beyond a single domain. Leverage unified data for the benefit of public health and individual care.
- **Enhance the usability** of data for social services and healthcare institutions across the State. In the end, data should be highly usable with as few barriers to access and understanding as possible. Data, when available, should have a very specific purpose or utility.

These stakeholder-articulated objectives will help guide the framework and the governance structure.

Data Governance Framework

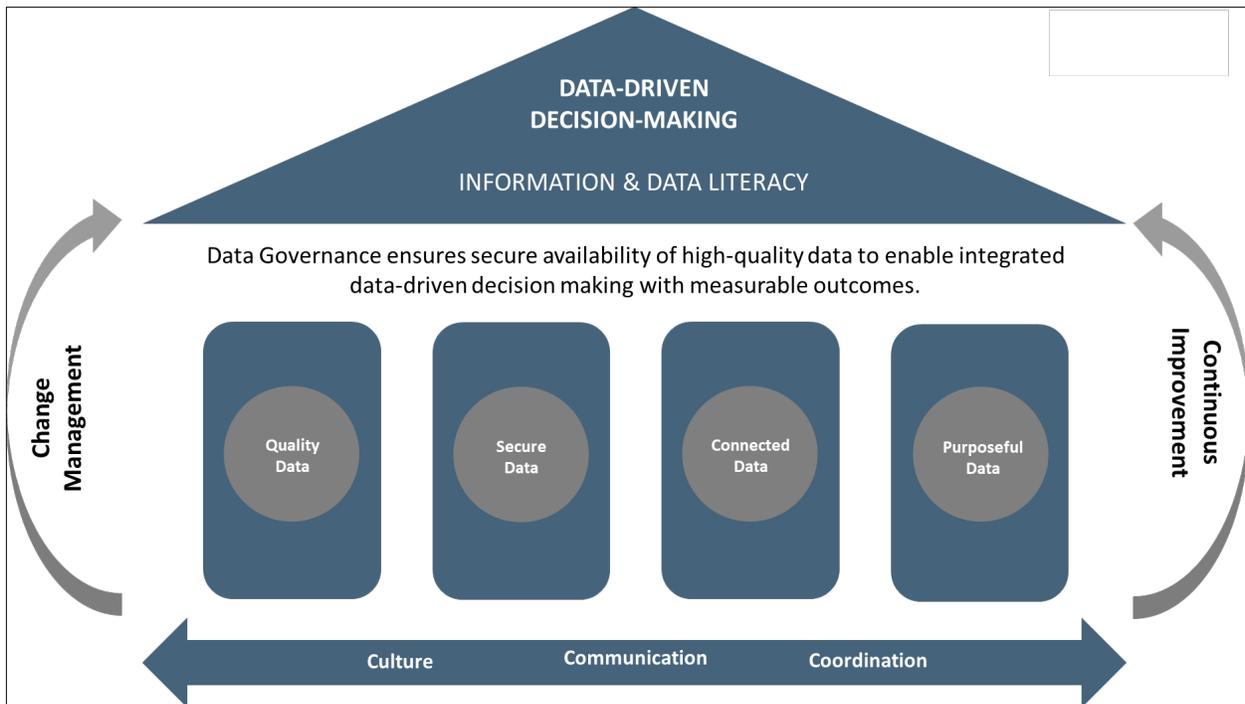
A data governance framework connects the organization's data strategy and goals to the structure necessary to manage the data that will achieve those goals. It contains the rules, processes, and procedures to implement and maintain the program over time. A framework may also have pillars that are representative of the organization's goals for governance and its purpose. These themes should live explicitly throughout the governance program as decisions are made, priorities are set, and projects are planned and implemented.

The data governance pillars for Vermont's Unified Health Data Space are directly tied to the feedback and priorities identified by involved stakeholders. The themes listed above have been condensed to the following pillars: quality data, secure data, connected data, and purposeful data. Therefore, each decision made, and priority determined, should be in pursuit of these key pillars.

There are also key elements of data governance that enable the organization to realize the benefits of each pillar but are not pillars in and of themselves. For Vermont, these elements include communication, culture, and collaboration. For example, collaboration alone is not a sufficient goal in data governance but prioritizing the element of collaboration will enable the Unified Health Data Space to provide quality, secure, connected, and purposeful data.

The diagram below illustrates the four data governance pillars and how they are supported continually by the elements. The following sections describe each of these pillars and elements in detail.

Figure 2: Data governance framework pillars



Quality Data

Data without quality has little utility. Poor-quality data creates significant risks and challenges. Moreover, the ability for data to reflect reality builds trust between the users of the data and the technical stewards of the data. No matter what tools users leverage, users will return to sources of information that have consistent quality they feel they can trust to accomplish their tasks. Policies in this quality data pillar specify the required levels of “fitness” (for example, accuracy, completeness, timeliness, consistency, and validity) for the information to have the best value to the organization.

Secure Data

Providing secure access to data assets is a top priority for the State of Vermont. The State, its agencies, and partners carry statutory responsibility to protect access to peoples’ data. Access rights to data assets are crucial for minimizing risk. Policies in the secure data pillar will help coordinate authorization and access for appropriate parties to data for specific intended uses. This includes data assets that contain personal identifiable information (PII) and protected health information (PHI). Though users need accessible data, that access should be expressly limited by the laws that regulate it.

Connected Data

The primary function of the Unified Health Data Space is to connect data from many disparate systems into a usable view for consumers. Connecting that data will require broad coordination among departments, third-party vendors, technicians, and various subject matter experts. The connected data pillar will focus on developing processes that define the structure of the connected data as well as the processes for identity matching, managing master records, and generating repeatable and auditable procedures for integrating data.

Purposeful Data

Success in data governance is tied to the ability of data to support the organization's overall goals. This holds true for the Unified Health Data Space. Data should be securely connected to support health care delivery and coordination, public health operations, health plan operations, Medicaid-driven delivery system reform, and population health management. However, these areas are already vast, and the scope will grow with the addition of more data sets into the environment. The purposeful data pillar ensures that there are clearly defined goals for each integration/connection of data as well as supporting analysis. This arm of governance reinforces the evidence-based thinking that will drive the culture of the organization. The goal is not to bring data together for the sake of having more data, but for having data with a specific intent and purpose-driven by evidence-based program decision-making.

The four pillars of data governance are enabled by the following elements:

Communication

Communication is important to the implementation of the entire data governance plan and particularly underscores transparency, one of Gartner's foundations of all data governance frameworks. It is the great connector that aligns people to the strategy and is critical to any cultural shift. Communications should flow out regularly regarding the data governance strategy and changes that have been identified by governance bodies as important. Those communications should be tailored to the audience so they understand what they can expect and how the changes affect them. The messaging should be specific and actionable to help people tie the strategy and changes to their daily work. Messages should include the complete picture of the change, including why, what, how, when, where, and who. Processes should be put in place to clearly define who and by what means messages go out to different audience types.

Collaboration

Data governance provided a structured approach to collaborating to achieve broad goals. Specific to the Unified Health Data Space, the use cases are expected to span across departments, vendors, and various subject matter experts. Data governance will provide stakeholders with simple structures that enable collaboration and better coordination of efforts.

Culture

The key to the success of each pillar as well as the entire data strategy is that the State has the talent and capacity it needs to manage, interpret, use, and understand data. This includes culture, development and training, recruitment and retention, and user capabilities and needs. Important initiatives in this area will foster a data-informed culture that is open and shares concepts and ideas appropriately. Efforts should be focused on increasing data literacy and achieving organizational change. Training, communication, reinforcement, and rewards will all need to be aligned with the value of data in mind.

Governing by Domain

How information is categorized and segmented is a large part of determining a governance approach. Categorizing information is one of the first steps in defining boundaries around information ownership and responsibilities for maintenance. Data domains are an industry standard used to segment data. A data domain is typically a logical business grouping of information, one that may likely go beyond the scope of a singular system. Simply put, one domain of data may come from multiple systems. Domains are

typically derived through the operations or functions an organization performs and subsequently the data used to support those functions.

An example of a data domain in the Unified Health Data Space is providers and their data. Though there may be nuanced differences in the detailed content of each provider's data, like between a commercial claim and a Medicaid claim, from the perspective of combining data for a unified view, it is helpful to consider the data more generally as provider data.

In domain-centric governance structures, data is expected to be governed independently by the domain business owners - the group of people who manage the collection and operations of the business functions of the domain. Each domain is therefore responsible for maintaining the quality of its information. This is done with the understanding that the domain's data can then be consumed by other domains within the organization.

As more data domains are integrated into the Unified Health Data Space, it is helpful to refer to any existing standards around domain identification and categorization. Considering and/or conforming to said standards can allow for easier integration across the industry of healthcare systems, making the unified data even more valuable. Table 2 provides references to existing types of data (Medicaid, Public Health, and Social Determinants of Health) and domains frequently used within each of the broad swaths of data.

Table 2: References that include frequently used domains

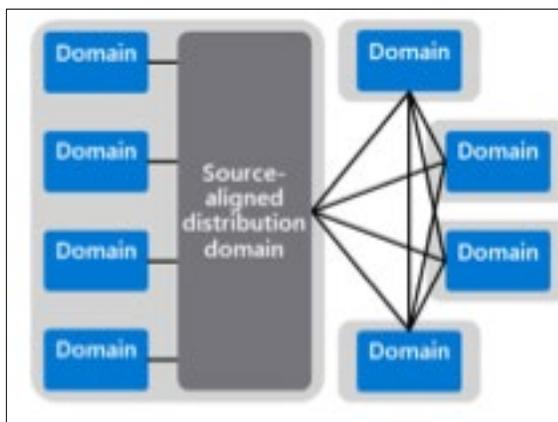
Institution	Type of Data	Reference
CMS	Medicaid Services	CMS Data Taxonomy
PHAB	Public Health	10 Essential Public Health Services
CDC	Social Determinants of Health (SDOH)	SDOH Domains

When considering data governance through a domain-centric data hub perspective, there are two approaches the State of Vermont may utilize. Both approaches leverage a central distribution domain that represents the Unified Health Data Space. The following summarizes the approaches and their characteristics.

Hybrid Federated with Distribution Domain

The first approach is referred to as a hybrid because it has both centralized and decentralized data governance. Some domains are governed by the central domain, while others are governed autonomously. This situation can arise when there are differences in available skills or ability to govern data to the standard; these differences are often based on having the technical staff to manage the data. The ability to govern effectively may be more limited when an organization has more legacy systems that require effort to maintain and extract data. In this structure the distribution domain shares responsibilities to fix data issues for the domain data it governs.

Figure 3: Hybrid federated governance framework with distribution domain

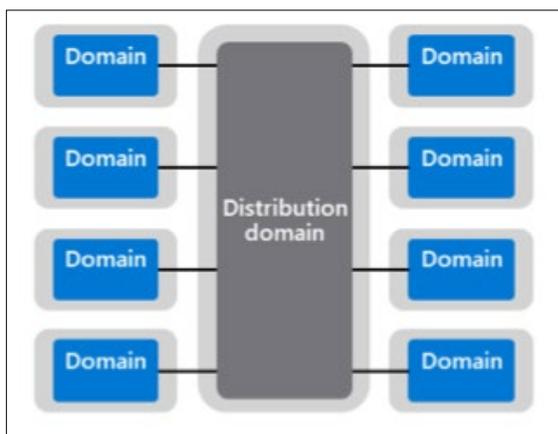


In the hybrid approach, the central distribution governance group takes more ownership of delivering domain data products to consumers, including the central data hub. The federated side of this structure also allows data to flow from one domain to another without going through the data hub (distribution domain).

Fully Governed with Distribution Domain

The fully governed structure with a distribution domain relies more heavily on mature governance that already exists at the domain level of each of the contributing data domains. This structure relegates all sharing of domain data to the distribution hub. However, the onus is on the domain to fix all data issues. This structure implies highly capable domains, meaning they have their own staff to support integrations and maintain quality.

Figure 4: Fully governed framework with distribution domain



In this structure, the central entity has full control of how data is distributed to consumers; data can be blocked or limited if standards are not met. The fully governed structure lends itself to an environment that sees value in the distribution layer as a consumer and contributor, implying that the distribution domain has a deep technical and business value to each domain. A potential drawback can be that relying on the distribution domain to be the primary conduit for sharing data can slow the domains' ability to deliver data down since data would need to integrate into the distribution domain before being made available to other consumers.

Framework Evolution

As the vision for the Unified Health Data Space matures and the technical capabilities of both the contributing domains and the central distribution domain grow, the governance structure may evolve over time. Current constraints may initially relegate Vermont to a hybrid structure. As the domains mature, they will have the ability to fully govern all their data to any globally accepted standards.

This evolution implies a long-term, expanded view of the Unified Health Data Space to serve as the primary source for sharing all agency domain data internally and externally, which is not currently stated in the HIE Strategic plan.

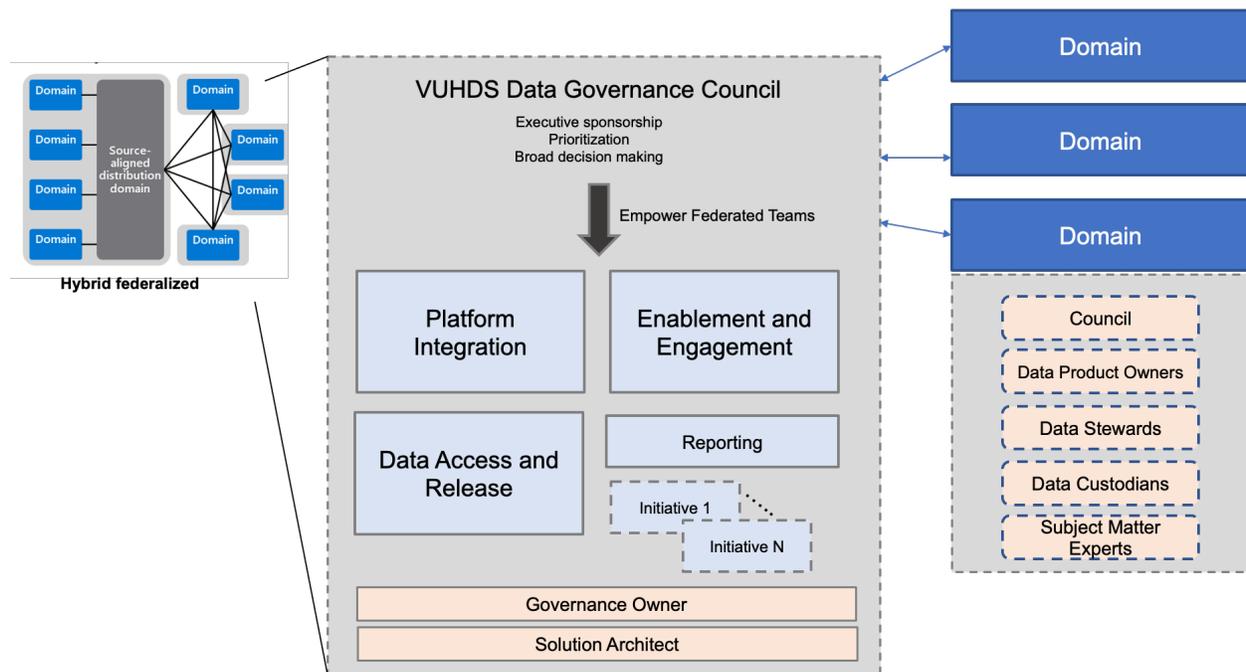
Operationalizing Governance

Collectively, the data governance pillars, elements, and framework all inform the governance bodies that develop policies and processes to operationalize data governance for the State. This section details the governance structure, its components, and how they work.

Governance Structure

Leveraging the hybrid federated governance framework, the following governance structure is recommended for the State of Vermont.

Figure 5: Recommended data governance structure



The recommended structure has the following characteristics:

- Domain-centric ownership
- Operates on the assumption that data is managed by domain owners
- Central function to define standards for integrating and accessing data
- Outcomes-driven operation

The hybrid federalized framework structure has a centralized governance entity that is focused on the distribution of domain data to a range of consumers. In the diagram above, the central distribution domain and the governing entity is the Unified Health Data Space Data Governance Council. The Council empowers Federated Teams.

Data Governance Council

A Data Governance Council provides executive sponsorship, prioritizes the goals, and is responsible for broad decision-making for the Unified Health Data Space. In short, it is the authority that manages the environment. A Unified Health Data Space Council is necessary to include responsible parties who own the data that will eventually live in the environment, which will extend beyond what is currently in the HIE. The Council also acts as an internal authority over how the data is integrated and managed for the environment. This is a different responsibility than what has been established with GMCB and the HIE Steering Committee.

The responsibilities of the Council include:

- Empowering and forming Federated Teams to support the execution and operations of data governance for the Unified Health Data Space.
- Monitoring program effectiveness by tracking the status of governance activities and initiatives. This can be managed by the Governance Owner, whose role is defined below, or by leveraging the Federated Teams.
- Setting priorities for the program by developing and/or evaluating use cases or outcomes for the benefit of Vermonters.
- Setting and approving policies. The Council should provide executive support for policy changes. The Federated Teams and subject matter experts should provide evidence-based support.
- Provide executive sponsorship, including but not limited to funding support, communications, and being an escalation point for Federated Teams.
- Encourage and evangelize data use and data governance across the agency.

The Data Governance Council must formalize its vision and objectives into a charter. The charter will outline the objectives of data governance and the key results desired and to be measured. It is expected that these items will change over time. The changes to the key results should be managed as the Council's priority use case (or governance) backlog. A change to the charter is not required to adjust priorities. If changes to the high-level objectives or scope of the Unified Health Data Space Council occur, the charter should be updated to reflect said changes.

Membership

The Data Governance Council should consist of State domain members whose data is either in the Unified Health Data Space or is slated to be integrated into the environment. Members must fully represent the interest of their domain. In some cases, achieving that level of representation requires multiple member positions per domain as domains may span across departments or organizational groups.

State domain members of the Data Governance Council should be in executive leadership for their respective business area domains. Members should be in a role of operational and policy authority to help support the development and implementation of any new policies, policy changes, processes, and process changes.

The Data Governance Council can also benefit from non-State employee members who can represent the interest of specific populations affected by the policies or data managed by the Unified Health Data Space. In such a case, a domain product owner may add a domain-specific non-State employee to the Council for the purpose of providing first-hand insight or feedback concerning policies or data governance activities that would impact the relevant population. These non-State employee members should be non-voting members.

Governance Operations Managers

The Unified Health Data Space is expected to grow over time. It is recommended to have roles accountable to the Council level that represent the technical needs of the Unified Health Data Space and the coordinated business needs. There should be two roles, listed in the diagram above as the Data Governance Owner and the Solution Architect. These roles are responsible for the ongoing operations of the governance program from two separate perspectives. The details of those roles are described in the section below.

Recommended Meeting Cadence

The Data Governance Council should meet at least every two months to get the program started. This frequency is recommended as the program establishes its initial charter and priorities. It may revert to a quarterly session dictated by the volume, pace, and priority of its backlog of outcomes/use cases.

Federated Teams

There are two types of Federated Teams in the governance structure.

- Core Federated
- Agile Federated

Core Federated Teams

A Core Federated Team represents a persistent function that is required for the operation, maintenance, growth, and expansion of the Unified Health Data Space. Federated Teams may consist of data owners, data managers, data consumers, and information technology managers who can speak to critical data and reporting needs and create plans to meet those needs. The governance program should strive to be agile and flexible in operations to reduce administrative burden and to ensure the ability to pivot quickly as priorities shift. Federated Teams can act flexibly by assembling the appropriate mix of data roles and stakeholders only when there are concrete and specific tasks to accomplish. This reduces the administrative burden while also ensuring the correct mix of stakeholders can be assembled when there is a task at hand.

The formation of a Core Federated Team should initially be driven by the Data Governance Council's priorities, which could include a specific use case or outcome. For instance, if integrating Medicaid data into the Unified Health Data Space is the priority, there is an immediate need to form a Federated Team for platform integration and data access and release. However, there may not be an immediate need to form a reporting and analysis or enablement and engagement Federated Team.

A Core Federated Team's base of responsibilities include:

- Draft and recommend processes and policies to department leadership and/or Data Governance Council.

- Recommend appropriate levels of resources (staff, technical infrastructure, etc.) and ensure that proper planning protocols are in place to support the use case needs.
- Collaborate with cross-functional, cross-department, and external teams to accomplish the objective(s) set forth by the Data Governance Council.
- Encourage and evangelize data use and data governance across the agency.
- Contribute to collaborative knowledge management tools or processes and the development and delivery of agency data literacy initiatives, including education and engagement on the team's specific function and its application of data governance within the agency.
- Drive the operation and integration of data into the Unified Health Data Space.

These are the baseline expectations of any Federated Team. Federated Teams are purpose-built by design and Core Teams should have defined objectives or outcomes to achieve. These objectives should be documented and made clear to all members.

Each Core Federated Team will have one team Lead and can have multiple co-Leads. It is good practice to always have backups for named roles on a Core Federated Team.

Membership

Core Federated Teams may consist of technical specialists, business subject matter experts, or some combination of domain roles defined below.

Recommended Meeting Cadence

Core Federated Teams manage persistent functions and operations of the Unified Health Data Space. The intended nature of these teams is to be agile and provide guidance to other Federated Teams to accomplish their objectives. Core Teams may meet at least monthly but have the flexibility to meet more frequently based on Council priorities and their own needs.

Suggested Core Federated Teams

Table 3 details the core federated teams suggested for Vermont, along with their areas of interest and a description of some of the proposed activities.

Table 3: Suggested Core Federated Teams

Core Federated Team	Area of Interest	Description/ Proposed Activities
Data Access and Release	Data Access	<ul style="list-style-type: none"> ● Define policy details on managing access rights to specific data domains and data source elements. ● Define generic roles that span domains. Perform access audits and report to Council and domain leadership.
	Data Release	<ul style="list-style-type: none"> ● Define process and/or policies around how data quality reviews occur, determining who reviews data to be released to external parties, what data is in scope, etc. ● Operate a data release committee to make those decisions regularly.

Core Federated Team	Area of Interest	Description/ Proposed Activities
Enablement and Engagement	Data Literacy	<ul style="list-style-type: none"> • Develop initial efforts for outreach to people who will assume domain-specific governance roles, educating them on data governance and the role they play. • Prepare skills assessments for targeted groups to understand training needs and gaps. • Build the foundation for a data literacy program on leveraging the data in the Unified Health Data Space that can reach users both internal and external to the state.
	Communications	<ul style="list-style-type: none"> • Develop the Unified Health Data Space communication strategy. • Identify the mediums of communication internal and external, if necessary. • Assign responsibility to Core Federated Team Leads to produce messages around the data governance program. • Produce messages to reach people who would assume data governance roles.
	Change Management	<ul style="list-style-type: none"> • Work with the Communications team to identify power users and change liaisons throughout target domain organizations. • Coordinate forums to gather periodic feedback from state employees “on the ground” and users of the Unified Health Data Space. • Leverage the power users, change liaisons, and other feedback to craft more impactful messaging.
Platform Integration	Data Onboarding	<ul style="list-style-type: none"> • Identify the appropriate members of this group (may need to identify people in domain-specific governance roles first to get the most appropriate candidates). • Define and certify the general data onboarding process. • Audit and report the progress of data domains onboarding to the Data Governance Council.
	Data Matching	<ul style="list-style-type: none"> • Define and manage the data matching processes. • Develop methods for automating the processes. • Audit and report the status of valid data matching to the Data Governance Council.

Core Federated Team	Area of Interest	Description/ Proposed Activities
	Mastering Data	<ul style="list-style-type: none"> ● Reference/Conformed Data Management – ensuring consistency across sources and within domains.
	Data Cataloging	<ul style="list-style-type: none"> ● Determining the broad scope of the cataloging effort. ● Consider existing technology solutions within the agency. ● Determine a solution path. ● Plan a curation approach plan that identifies who will curate data, including education and messaging around the entire effort with the Communications and Change Management teams.
	Business Glossary	<ul style="list-style-type: none"> ● Determining the broad scope of the cataloging effort. ● Determine if an existing repository of terms exists. ● Consider existing technology solutions within the agency. ● Determine a solution path. ● Plan a curation approach plan that identifies who will curate data, including education and messaging around the entire effort with the Communications and Change Management teams.
	Data Lineage	<ul style="list-style-type: none"> ● Determine the initial data domains in scope. ● Identify the representative domain parties to work with the Federated team. ● Determine what lineage information or system exists today. ● determine the level of detail necessary to capture (the needs). ● Consider approaches to gather lineage information systematically or from stakeholders ● Consider working with Change Management to develop a crowdsourcing campaign

Core Federated Team	Area of Interest	Description/ Proposed Activities
Reporting and Analysis	Reporting Standards	<ul style="list-style-type: none"> ● Consider the initial data domains and their user base. ● Identify appropriate members of the Federated team leveraging data governance role assignments. ● Coordinate with HIE and domain power users to understand gaps in reporting and the needs of the target domains. ● Consider defining a unique reporting template style for the Unified Health Data Space. ● Build a backlog of opportunities to add valuable standard reports for the target domains - reports may be limited to domain data but priority is given to use cases that cross other domains.

Agile Federated Team

The Agile Federated Team is a special type of Federated Team that exists to support a specific use case or initiative defined by the Data Governance Council. This type of Federated Team may only exist for as long as necessary to complete the pre-defined objective for which the team was initially formed. It does not persist like a Core Federated Team. The formation of an Agile Federated Team should be driven by the Data Governance Council’s priorities, which could include a specific use case or outcome. These teams are purpose assembled and meet on an as-needed basis with respect to driving progress on its objectives.

Agile Federated Teams are represented in Figure 5 by the dotted line boxes labeled Initiative 1 through Initiative N, showing that there can be a variable number of these types of teams based on the Council’s priorities and the bandwidth of team members.

Each Agile Federated Team will have one Team Lead and can have multiple Co-Leads. It is good practice to always have backups for named roles on an Agile Federated Team.

Membership

Agile Federated Teams may consist of technical specialists, business subject matter experts, or some combination of domain roles defined below.

Recommended Meeting Cadence

The intended nature of these teams is to be agile, so it is expected that after the team is formed the group should meet frequently to orient themselves around the objectives and clearly define the actions necessary. Based on those action steps, the team should meet as frequently as necessary to drive progress on activities and provide status to the Governance Owner, Solution Architect, and Council where applicable.

Roles and Responsibilities

In addition to the various governing bodies that may be established, a data governance program must recognize other critical roles within the team and collaborate with them appropriately. This section outlines key roles and their chief responsibilities.

Process Owners

- **Data Governance Owner** - This role serves as the program manager who is responsible for Data Governance Council operations and is also accountable to the Council. They manage the coordination of all Federated Teams and are responsible for the systems and rules used to ensure data is legal, well-organized, safe, accessible, and valuable. The Governance Owner assists the Federated Teams in reporting status to the Data Governance Council.
- **Solution Architect** - The Solution Architect is the lead engineer who manages the design of the unified data architecture, the physical design for data storage, and the technical solution used to maintain the data. They should have deep expertise in health information interoperability, allowing them to coordinate with domain stewards and custodians to fully manage the data onboarding process.

Data Domain Roles

- **Data Product Owners** - The Data Product Owner has ownership of and accountability to a business operation and/or function that produces and uses data in the organization. These individuals are typically in department leadership level positions and are accountable to maintain data quality and manage other risks associated with the business function. Data Product Owners may not work with their data every day but are responsible for overseeing and protecting a data domain.
- **Data Custodians** – A Data Custodian is an employee of the organization who has administrative and/or operational responsibility over the respective data domain. In many cases, there will be multiple Data Custodians. Data Custodians should have a thorough understanding of security risks impacting the data. For example, storing or transmitting sensitive data in an unencrypted form is a security risk. Protecting access to data using a weak password and/or not patching a vulnerability in a system or application are both examples of security risks. Security risks should be documented and reviewed with the appropriate Data Steward so that he or she can determine whether greater resources need to be devoted to mitigating these risks. These Data Custodians assist the Steering Committee with gaining a better understanding of their security risks.
- **Data Stewards** – Data Stewards are accountable for business controls, data content, and metadata management related to a set of data assets. They work with stakeholders that are impacted by data to develop definitions, standards, and data controls. They may also sponsor data quality, data acquisition, and data entry initiatives. In many cases, business units and operational units that use the same data view it differently. A data steward is a person who makes sure that the data provided and accessible to teams supports all business needs and meets regulatory requirements.
- **Subject Matter Experts (SME)**– SME is a generic term for someone who understands the business and programs that generate or use the data. These individuals are key to understanding business definitions of data elements and interpreting their meaning in relation to business and program operations.

The roles identified above make up some of the foundational roles that will be needed in any data domain of a data governance program, though additional roles may be identified through future discovery processes.

Data Access Based on Roles

Ensuring secure access to data is one of the key objectives of this data governance program. Data access governance is a system that defines how your organization manages and controls who has access to what data assets both internally and externally and then provides that ownership and control. It encompasses the people, processes, policies, and technologies required to manage and protect data access. The goal of access governance is to support organizations as they become larger and more complex by keeping the oversight and control of user accounts simple.

In the context of the Vermont Unified Health Data Space, some future data sources are managed within the State's information technology infrastructure. Other sources rely on third-party vendor systems. Generally, access governance has two components: identity management and access management.

Identity Management and Access Management

Identity management involves registering, identifying, and authenticating a person. The common procedure of supplying a username and password is an authentication process. More advanced processes include multi-factor authentication, where a user may need to supply another piece of information for authentication, like a fingerprint or a specific code. Identity and access management go hand in hand. Once identity services verify who a person is, it's access management that controls what resources a person can access. For the State, the central information technology organization has solutions in place to manage the identity and access management process.

Source systems and applications that capture and use data may or may not integrate with the State identity and access management system. This is not an uncommon challenge. Modern applications have started to leverage open standards for authentication that can interoperate with them. This allows for the authentication function to stand alone as a service.

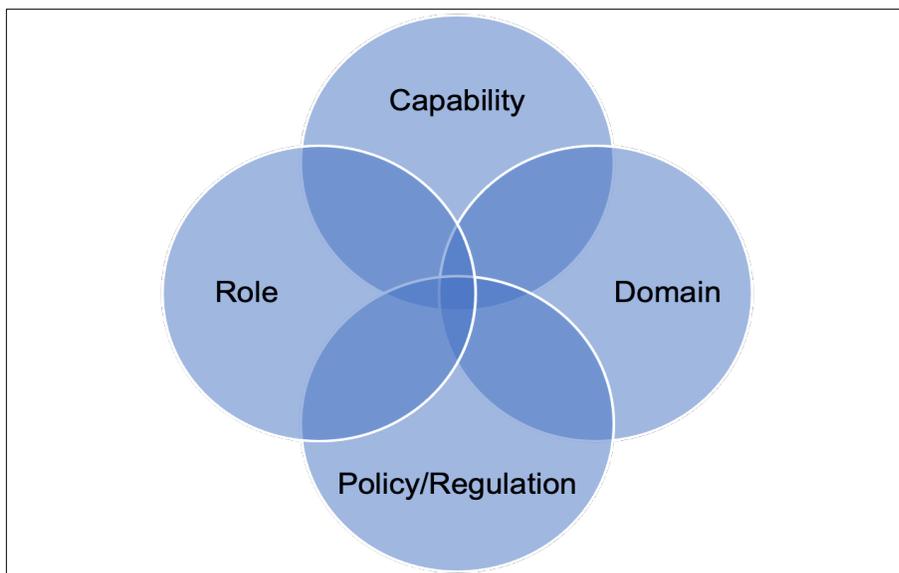
Future implementations at the State and for the Unified Health Data Space should consider leveraging modern solutions to more easily centralize, monitor, and control access to state resources.

Role-Based Access

Stand-alone applications that leverage user accounts allow for role-based access control. In this practice, roles are determined by categorizing user capabilities within the application. For instance, User 1 is a case manager and has access to the case management and eligibility modules of an application. User 1 may not have access to the finance module of the same application.

The Data Governance Council will need to work with central IT teams, HIE teams, and domain-specific data owners to translate any existing access control over to the Unified Health Data Space. The Data Access and Release Core Federated team will take ownership of this function. The users of the Unified Health Data Space can be internal to the state or external. The Data Access and Release team will need to balance what domain the user belongs to (Domain), what data the user should have access to (Role in the organization), what the user needs to do with data (Capability), and whether policy and regulations allow the user access to data (Policy/Regulations).

Figure 6: Access Perspectives



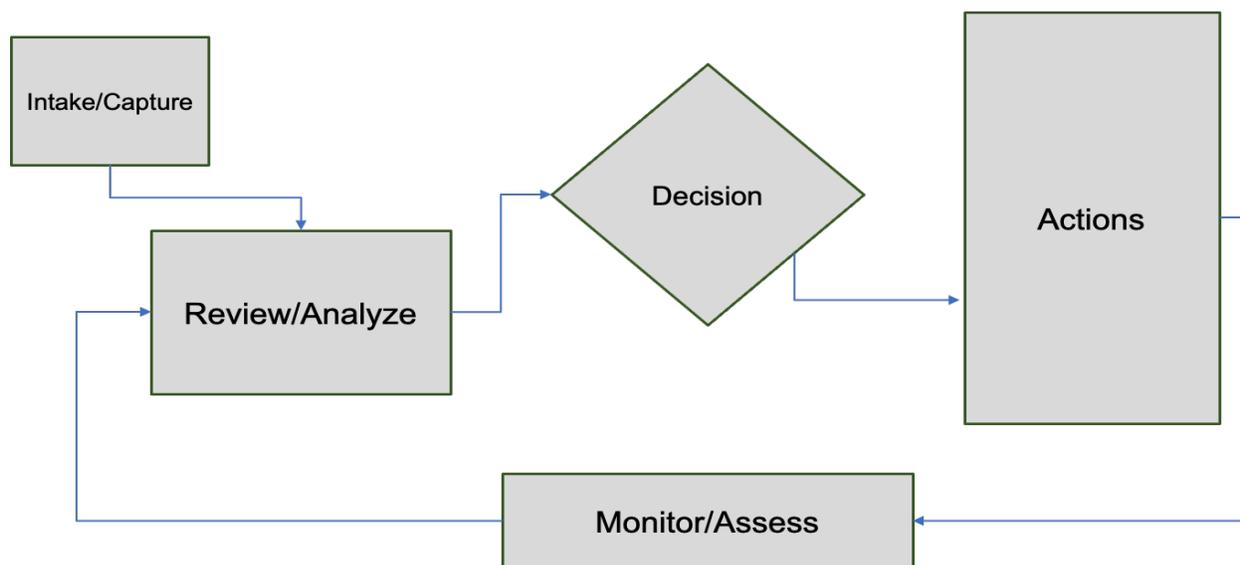
Because source systems will have varying technical capabilities and regulations that dictate who should be able to access information, the Data Access team will need to determine the best recommendations based on the use case for integration.

For instance, a user in the Department of Health Access may have access to individual records in their domain but may only be able to see aggregated information about the population of people receiving Mental Health Services in the Unified Health Data Space. The Data Access and Release team will need to determine the appropriate mix of roles in the Unified Health Data Space environment. Care should be taken to leverage as much system-based automation for auditing and discovery, since access rights can change at any source.

How the Teams Work Together

Each governing body is essentially an agile team working on a backlog of items. The items in the backlog are specific initiatives that involve the expansion or use of the Unified Health Data Space. Core Federated Teams have specific objects that represent common functions within the environment. Core Federated Teams help Agile Federated Teams achieve their goals. *Figure 7* illustrates a summary description of the overall process which is described in detail in the sections below.

Figure 7: Data governance process



This structure allows the Council and its teams to be as agile as possible. When critical items hit the backlog, teams can be adjusted to support these new priorities.

Intake/Capture

The Data Governance backlog will serve as the central, real-time repository for all known challenges associated with data governance practices across the Unified Health Data Space. Maintained by the Council, the governance backlog will ensure the agency is aware of instances where data governance issues are complicating or impeding programmatic delivery. The backlog is maintained and evaluated regularly by the Data Governance Owner as it provides an understanding of common challenges across multiple departments or initiatives.

Items on the governance backlog may come from the following places:

- **Data Governance Council Meetings** - Data governance issues or concerns raised during council meetings will be recorded and added to the backlog.
- **Federated Teams** - The Enablement and Engagement Core Federated team is charged with change management and communications activities. This group will periodically use surveys, focus groups, and other methods to capture feedback from various audiences. Items captured here will be recorded in the backlog.
- **Discovery Sessions** - As Federated Teams work together, special issues may arise that will be captured and added to the governance backlog.

Below are three tips for identifying data governance-related issues:

- **Issues are common.** Data Governance challenges are often not isolated to a single department within AHS. Rather, they manifest in multiple instances across departments.
- **Issues are interrelated.** A problem related to Data Governance in one area most often has an adverse effect in another area.
- **Issues are perpetual.** As the use of data within the agency grows, so too will opportunities to improve Data Governance practices.

Review/Analyze

The Data Governance Owner has the responsibility to ensure every item on the backlog is valid and complete, meaning it has enough context for the Council to make an educated decision. When there is not enough context provided for an item, the Data Governance Owner must work with people in respective data governance roles (Domain Stewards, Domain Custodians, etc.) within the domain of the item to provide the missing context. They will also perform a light root cause analysis. This information allows the council to properly prioritize the items on the backlog.

The appropriate level of clarity includes:

- A detailed description
- What system(s) are involved or affected?
- Does this affect one or more departments?
- Impact description - What happens if this is not addressed?

An item will not move forward in the process without the appropriate level of clarity.

Decision

The Data Governance Council makes decisions on policies, new processes, changes to existing processes, backlog prioritization, funding, and changes to the charter.

When the council makes decisions about what to do with the backlog items, the Data Governance Owner presents items to the body. In some cases, this is either delegated to or supported by others in domain data governance roles. The following criteria should be considered when the council evaluates and prioritizes items on the governance backlog.

- Opportunities that impact multiple departments, programs, or initiatives should be prioritized over opportunities that are unique to an individual entity.
- Opportunities with clear expectations of impact or anticipated results should be prioritized over opportunities with more ambiguous or uncertain outcomes.
- Opportunities with a greater amount of dependent work should be prioritized over opportunities that are not required for additional work to move forward.
- Opportunities that are directly aligned with the stated priorities of the agency or the council should be prioritized over opportunities that are not designated as a priority focus area.

Items can be categorized into the following simple groups:

- **Low Priority** – This category is used in circumstances where the council identifies a data governance need that already has existing processes or policies. The impact of this item is limited to one group, not many. The overall impact of executing on this item is low in comparison to other items. No activities are dependent on this item being worked on or the item does not completely align with the goals of the council.
- **High Priority** – In circumstances where a division identifies a data governance policy need and the opportunity is deemed to be of high priority, the Council should immediately refer the work to the appropriate Working Group(s) for development.

A more robust system of categorizing items can be developed later as the size of the backlog and the diversity of topics grows.

Once an item is determined to be an active governance working item, regardless of its priority level, the council will determine which group will own the item. The owners of the item will be responsible for the execution and delivering specific objectives or outcomes. The council can assign the item to a Core Federated Team, an existing Agile Federated Team, or create a new Agile Federated Team. Recommendations on team members can come from council members or domain governance leadership/stewards.

Actions

Federated Teams manage the execution of all governance items from the backlog. Depending on the nature of the original data governance item, the Federated Team can take a number of actions. Most data governance issues, even business use cases, fall under one of these categories of actions to take:

- New policy or change to existing policy
- Training
- New process or change to existing process
- System enhancement
- New data onboarding
- Data corrections
- Monitoring or Reporting
- Stand-alone project

The Federated Team will coordinate with all respective domain roles to manage the progress toward achieving the team's objectives. It is important to leverage any existing processes within the state to accomplish the goals of the team. The Data Governance structure operates within the structures of the State of Vermont. For example, if there is a need to make a policy change in a specific domain that affects the Unified Health Data Space and there is an existing process to enact policy changes, then leverage that process. There is no need to create anything new; use all the tools available. This is the benefit of collaborating with people embedded at the domain level.

Monitor and Assess

As work continues with the Federated Team, progress is reported back to the council at the same frequency as the council meetings. The Data Governance Owner will ensure proper documentation is complete, perform a final review and acceptance of work products, ensure any necessary changes are communicated to the appropriate audiences, and, finally, close out the governance activity.

The Governance Owner will also determine if the activities in this workstream are related to any of the others in the backlog and update them accordingly.

Decision-Making Throughout the Governance Process

While the Council will make decisions on prioritizing backlog items, many other decisions will be made at various levels throughout the governance process. This section describes how that decision-making should work and be documented.

Data Governance Council Level Decisions

The Data Governance Council makes decisions on policies, new processes, changes to existing processes, backlog prioritization, funding, and changes to the charter. Decisions are all made by majority

vote by all voting members of the council. Non-voting members can provide input to decisions but are unable to vote.

Federated Team Level Decisions

A Federated Team may make recommendations to the council on policies, processes, changes to the council charter, and specific projects or solutions to fund. The Federated Teams can also make recommendations to domain-level governance teams on processes and policies. Federated teams are empowered to make decisions on processes and policies that affect that specific team. The decisions made at this level will also be tracked in a decision log maintained by the Data Governance Owner.

Decision Tracking/Logging

Decision tracking will be integrated into the governance process because it serves as an effective way to recount the rationale behind decision-making. A decision tracking log is a document that captures the situation, the options, the final decision, and the rationale in a single place. All decisions voted on by the Data Governance Council will be logged in a decision tracking log document. The basic scope of decision tracking is to log decisions about policy, processes, data onboarding, funding, and the data governance charter. It is also updated as the council actively maintains and reviews data processes. The Data Governance Owner will maintain a decision tracking log and the document will live along with the council as a historic record.

Policy Development Guidelines

Policies must be developed and deployed in accordance with the following principles:

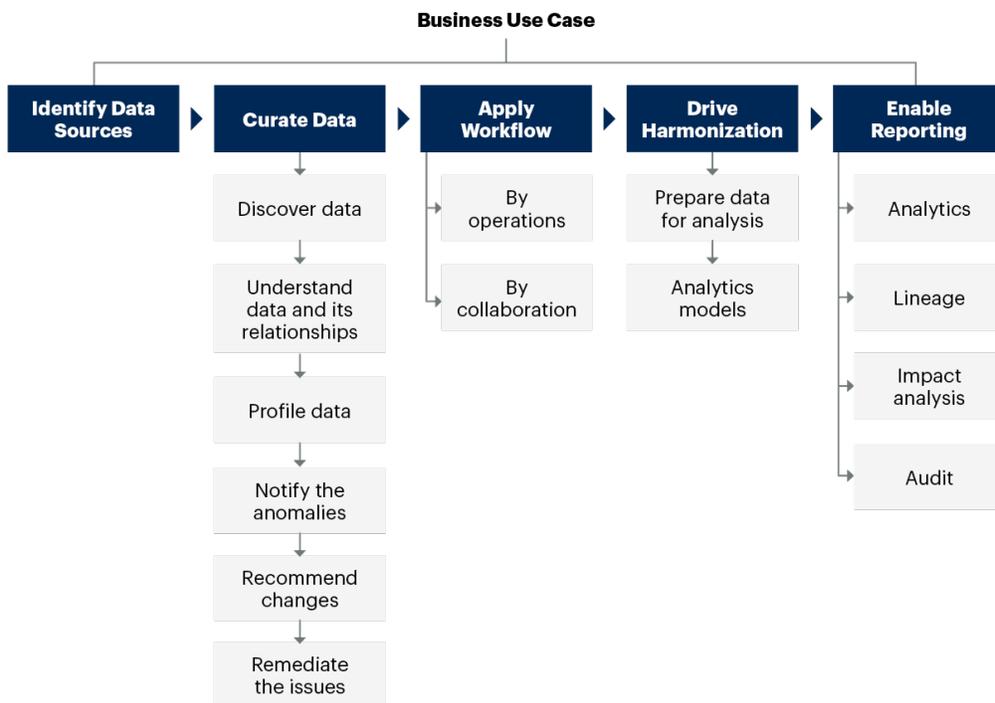
- Policies must be multi-departmental. Policies created must account for the needs of each division to ensure successful deployment across the agency.
- Policies must be comprehensive. Policies created must acknowledge the interdependencies of Data Governance issues.
- Policies must evolve. Lastly, Data Governance policies must be continuously maintained and updated to reflect the ever-changing nature of data use within the agency.

The Work of Governance

While there are many structures, stakeholders, policies, and processes involved in data governance, data governance is not successful unless leveraged to get actual work done and realize new data analysis, insights, and better outcomes. The figure below illustrates how a specific use case or priority would flow through the data governance process:

Figure 8: Gartner diagram linking business use to key data management steps

Linking the Business Use Case to Key Data Management Steps



Source: Gartner
729295_C

Gartner.

Moving from Use Case to Data Governance

Each use case will have its own unique needs, stakeholder groups, goals, and involved data. However, the same high-level process can be applied to all use cases prioritized and pursued by the Council.

1. **Identify Data Sources** - Begin by considering the data sources that will be needed for the prioritized use case. This list may change as more details about the use case are uncovered, but the initial list should be robust. Engaging data owners, subject matter experts, and other relevant stakeholders may be helpful in creating a full list of necessary data sources.
2. **Curate the Data** - Next, the identified data sources should be investigated and curated. The curation process may include:
 - **Discover data** - Discovering data may include seeking out and evaluating new data assets either inside or outside existing United Health Data Space domains.
 - **Understand data and its relationships** - Before using any data to fulfill a use case, the data should be fully investigated such that the relevant Federated Team understands where the data came from and the context around its original source. The curation process should also include identifying any datasets that provide additional and helpful context in combination with the original data and other complementary datasets.
 - **Profile data** - After confirming the necessary data assets, the data should be profiled. Profiling should clarify the structure of the data, provide descriptive statistics, summarize

- types of data and categories seen within the data, and start identifying quality issues and risks associated with using the data.
- **Recommend changes** - Profiling may reveal structural inconsistencies or quality issues that deserve attention. Any changes to the data or metadata should be suggested at this time before moving on to any other data governance processes.
 - **Remediate issues** - Similarly, any issues that have been raised and confirmed as needing to be completed should be remediated during the curation phase.
3. **Apply Workflows** - The workflow step is the crux of the data governance policies and processes that have been established by the Governing Bodies and documents. Workflows may take many forms, namely operational or collaborative.
- **Operational workflows** may include running data pipelines and other necessary data integration processes. These workflows are set to ensure that data is ready and available in the proper place and in the correct format. Operational workflows should also have fail-safes such that if they fail, the right stakeholder is notified, and the problem can be remedied. In the context of a use case, the data identified for said use case would flow through all needed, established operational workflows.
 - **Collaborative workflows** are those that are essential processes but require conversation and collaboration between domains or teams. An example of collaborative workflows includes determining who can have access to the data assets utilized in the use case. Determining access and safeguards is a critical process, but it requires discussions and human intervention to complete the workflow.
4. **Drive harmonization** - After data has been brought into the Unified Health Data Space in a usable structure and the proper limits and rights have been placed around the data, the focus should turn towards actually using the data for the intended use case. Here, driving harmonization means encouraging the proper use and analysis of the data across domains and teams. The relevant roles may need to first **prepare data for analysis**, then **perform analytics** in service of the use case. In most cases, the relevant roles are data engineers and data scientists.
5. **Enable reporting and data use** - Reporting may include many different formats and will be dependent and distinctive to each use case. Summarizing the **analytics** into a digestible format for decision making may be part of the reporting process. If the use case involves agency, stakeholder, or legislative regulations, reporting may be required to fulfill those mandates. Reporting may also involve summary information about the analysis and an **audit** of the data, like the **lineage** of all data involved or any outliers identified during analysis. Finally, reporting may also include more robust and specific analysis, such as **causal impact** or specific analytics models that answer the use case.

Remember that data governance is an iterative process. While use cases may flow through the steps detailed above, there are likely to be feedback loops between the steps. For example, after conducting further data analysis during the harmonization step, data issues can be identified that may need remediation. Changes in the data structure may subsequently require updates to the operational workflows. These changes are normal and expected but should be accommodated through the established data governance process.

Communications Planning

A thoughtful data governance communication strategy and plan will help preserve operational excellence, provide clarity and consistency for data management and governance, and facilitate collaboration across the agency. These artifacts will ensure that leadership and staff know what is being done and why. The purpose is to create both formal and informal feedback loops as the agency's governance needs evolve.

The goals of communication should include:

- Clear and transparent communication of the Vermont health data strategy
- Improving awareness of the data and analytics resources at the agency and beyond
- Arming the State of Vermont with timely and accurate data to communicate evidence during critical policy and decision-making conversations
- Communicating the stories behind the vast amounts of agency data to external stakeholders and policy makers to proactively share State of Vermont's health story

Sustaining and spreading a continuous-improvement culture means fighting inertia and reluctance to change work routines. To do so, focus on solving real problems inside the organization rather than pointing to others' success stories. It also involves several communication activities and achieving small wins.

Clear and frequent communications are key to successful data governance programs and must be tailored for the appropriate audience. Purposeful and consistent communication will provide context and engage employees; stakeholders and leaders are far more likely to engage in change when they understand what it means to the agency, their team, and their job function. Orientation, onboarding, and ongoing communications are three levels to factor into communication and change management plans. Over time, communications should also be tailored to share progress of data governance activities and include progress on the short-, medium-, and long-term results.

Communicate, Communicate Again, and Communicate More

Develop a communication plan that can live as a companion document with the council charter throughout the life of the program. The Engagement and Enablement federated team can take responsibility for developing and recommending changes to the plan. By doing so, communication will be a core piece of data governance and be done frequently. Important items should be communicated with consideration of the following advice:

- Always articulate the Why and the WIIFM (what's in it for me?)
- Consider the sender. Employees want high-level communications from executives/sponsors, but personally impactful messages from their manager
- Don't be afraid to repeat yourself – it takes 5-7 times for your message to resonate
- Make sure to mix up modalities

Diversify the Communications

Furthermore, communication can and should take many forms. By diversifying methods of communication, organizations are likely to reach and inform more employees and stakeholders. Consider the following methods:

- Emails
- Videos

- Cascading Management Communications
- Town Halls, Team Meetings
- Articles in Newsletters
- Intranet/SharePoint/Teams
- FAQs

Considerations

While this document suggests best practices that the State of Vermont may apply to their data governance program, there are additional considerations that may change or inform the format and extent of data governance. The most important of these considerations are detailed here.

Funding

Funding was mentioned in the sustainability section of the Health Data Strategy. The same consideration should be taken for supporting the data governance effort. In any situation where dedicated resources are necessary, funding can become an issue. Data governance is a journey with an expectation of maturity over time. Having resources to frame and initialize the process is critical for data governance to become formally ingrained into how the organization operates. The gap between when it starts and full cultural acceptance and maintenance may require additional resources, including resources with enhanced technical capabilities.

Resourcing

Resourcing goes hand-in-hand with funding. Though governance should not require significant headcount additions, it may require resources to focus on initializing the process. Many resources are already performing data governance-related activities and can be leveraged to supplement teams for implementation projects.

Process Continuity Through Knowledge Management

It is essential to balance understanding the company culture and maintaining a sustainable governance structure. The people define the culture, but the people can change over time.

The goal when establishing data governance should be to build a structure that becomes a part of the culture so it can exist and be maintained as people move in, out, and around the organization. One of the pitfalls to avoid is not building enough data governance principles into policies and processes that will long outlive the tenure of a specific person driving those same efforts in the organization.

As the Health Data Strategy mentions, consider developing a comprehensive knowledge management plan to help the organization persist institutional knowledge that is not always captured.

Domain Maturity

The Data Governance Council can position itself to take on ownership and guide the maturity of domain-specific governance. This approach can help leverage best practices across the agency by borrowing from programs that have matured over time, learning from their successes and missteps. This expands the council's objectives but supports the existing recommended governance structure.

Vendor Data Implications

Consider enforcing or requiring some level of data governance regarding vendor-managed data. This could be as simple as negotiating or requiring access to data, requiring data definitions, requiring data model documentation for systems, and more.

Metrics for Governance Success

Data governance programs must be able to measure their impact on the organization. Measurement is the responsibility of the Data Governance Council and owner(s). The impact and value may be financially quantifiable, but this may not always be true. Measuring efficiency and effectiveness improvements requires benchmarks of the present state and the ongoing activity of measuring and reporting results.

Here are some sample metrics to consider:

- **User adoption** – Number of analytics users (Active*), number of analytics users.
- **Reporting** – Number of reports/dashboards, active* reports, scheduled reports, and queries/jobs executed.
- **Usage/Reach** – Web traffic and site usage for BI tool portal(s) or governance portals (intranet portal, data catalog, other collaboration/governance tools).
- **Data curation metrics** – Number of stewards, the ratio of data elements to owners, the ratio of data elements to stewards, percentage of “completely curated” data elements to total data elements, number of definitions approved, number of new data elements over a set time, percentage of PII/PHI, percentage with defined lineage.
- **Governance actions/decisions** – Policies set/approved, time to approval, number of governance activities, number of activities completed over a defined period, time to resolve/complete activities (actual or average).
- **Cost reduction through efficiency** – This is use-case specific and may be difficult to quantify. It typically involves capturing the current baseline of effort in terms of time and estimated cost, then redoing that same analysis with a new process.
- **Data governance maturity assessments** (annually).
- **Customer satisfaction** (semi-annual to annually).
- **Data quality metrics** – Data accuracy and data completeness, all defined by business owners and stewards. There can be an entire subset of metrics under both topics for the data engineering audience to monitor processes.

Immediate Next Steps

The following are logical next steps to take in the next six to eight months to initiate the journey of data governance maturity.

Table 4: Roadmap for data governance

	Item	Description
1	Draft Data Governance Charter	Develop the Unified Health Data Space
2	Communicate early	Develop a tiered communication strategy that identifies the mediums in the audiences to start communicating the overall message of the data strategy and the initiation of data governance.

	Item	Description
3	Further develop use cases	Initialize objectives and key results (OKR) gathering exercise at the council level to flush out the details of the current use cases.
4	Define appropriate ownership and roles	Leverage the messaging to accelerate identifying product owners, stewards, and custodians. Council members are a great source of initial suggestions. Focus on data in the domains of the prioritized initiatives only.
5	Prioritize	Develop an initial prioritization criterion that allows the council to quickly identify the highest priorities and establish both core and agile federated teams to begin the work of governance.
6	Cultivate a backlog	Synthesize goals from the Health Data Strategy into specific initiatives or business use cases that can be added to the data governance backlog.
7	Establish Transparency	Create a location that is accessible to all state employees (at least AHS) that provides information about the overall data governance program.
8	Gather baseline on data culture	Before initiating an effort to shift a culture that may introduce several new disciplines and tools and get a baseline. There are many different usage stats from the various systems in the current ecosystem to collect baselines. Baseline data will be valuable in showcasing behavioral changes. We are the “data team”; we must show wins/progress with data. Some baseline numbers may not look good. These are still important to capture since the ongoing governance effort is meant to improve those numbers. Share results as part of communications to support the need to move forward with unified health data space use cases and data governance activities.
9	Gauging governance	Determine which metrics make sense to track as the data governance foundation is built. Start with engagement and reach metrics, like usage stats, sentiment around culture, and experience (mentioned above) with data.