# COV Data Governance Guidelines

## Purpose

The purpose of this policy is to establish a framework for the management and protection of data within executive branch agencies, ensuring data quality, security, and accountability.

## Scope

This policy applies to all executive branch agencies and their employees who handle, manage, or access data.

## Governance Committees

Data is managed by a multi-tiered hierarchy

| **Committee** | **Purpose** |
| --- | --- |
| **Virginia Data Advisory Commission** | The Virginia Data Advisory Commission (the Commission) is established as an advisory commission in the executive branch of state government to advise the Office of Data Governance and Analytics (the Office) on issues related to data sharing, including open data, data analytics, and data governance. The Commission shall (i) set, plan, and prioritize data sharing performance goals for the Commonwealth, (ii) review agency accomplishments, and (iii) recommend solutions that will establish the Commonwealth as a national leader in data-driven policy, evidence-based decision making, and outcome-based performance management.  [Virginia Data Advisory Commission Charter](https://covgov.sharepoint.com/:w:/r/sites/TM-GOV-OfficeofDataGovernanceandAnalytics-External/Shared%20Documents/General/CDO/Data%20Governance%20Framework/Virginia%20Data%20Advisory%20Commission/Virginia%20Data%20Advisory%20Commission%20Charter%20-%20DRAFT.docx?d=w6c626f353cd94fd5a21bd5d90586ed75&csf=1&web=1&e=oJ8TVW)  *Note: The Data Sharing and Analytics Advisory Committee has been disbanded.* |
| **Executive Data Board** | The Executive Data Board is chaired by the Chief Data Officer who also selects its membership. The Executive Data Board membership consists of executive leadership, or their designees, from executive branch agencies engaged in data sharing and analytics projects. The Board has the following responsibilities:   * Translate the Commonwealth’s data-driven policy goals and objectives to agency performance targets * Allocate appropriate agency resources to support data governance, sharing, and analytics initiatives * Provide to the Virginia Data Commission any reports on the Board’s recommendations and work as required by the Commission |
| **Data Governance Council** | The members of the Data Governance Council are selected by the Executive Data Board. The Council, however, consists of employees of the state agencies represented on the Executive Data Board. The Data Governance Council is chaired by the Commonwealth of Virginia Chief Data Officer (or designee) and has the following responsibilities:   * Liaise between state agency operations and the Chief Data Officer * Advise the Chief Data Officer on technology, policy, and governance strategies * Administer data governance policies, standards, and best practices as set by the Executive Data Board * Oversee data sharing and analytics projects * Review open data assets * Govern the Commonwealth Data Trust * Report progress, compliance, and performance to the Executive Data Board |
| **Data Stewards Group** | The objectives of the Data Stewards Group are to:   * Promote and facilitate the secure and appropriate sharing and use of data assets in support of data-driven policymaking, evidence-based decision-making, research, and analysis * Maximize the value and utility of Commonwealth data-related investments and assets * Promote increased sharing of data between state agencies and localities to provide tangible operational improvements to assist state agencies and localities in fulfilling their respective missions in a coordinated, cost-effective manner * Provide public access to data assets, where lawful and appropriate, to enhance research, innovation, and insight |

## Roles and Responsibilities

Per COV (§2.2-603.F) “the director of every agency and department in the executive branch of state government, including those appointed by their respective boards or the Board of Education, shall be responsible for securing the electronic data held by his agency or department and shall comply with the requirements of the Commonwealth's information technology security and risk-management program as set forth in § 2.2-2009.” In addition, the Director of every department is responsible for the security of the agency’s electronic information, and for establishing and maintaining an agency information security program compliant with this policy and meets all of the requirements established by COV ITRM Security Standards.

## Policy

### Role Definition

* Agencies have clearly defined key roles of Data Owner, Data Steward, and Data Custodian.

### Key Data Management Roles

| **Role** | **Definition** |
| --- | --- |
| **Chief Data Officer (CDO)** | * The Chief Data Officer is responsible for the overall management of the COV’s Data and Information Governance |
| **Agency Senior Management** | * Group in charge of deciding whether to publish the agency’s data assets. A Data Executive supported by a Data Owner has the responsibility for the management of data assigned within their portfolio. |
| **Data Owner** | * The data owner in an organization is anyone who originally created or acquired the raw data, and/or anyone who retains the rights to approve sharing or other actions, can make access control decisions (or restrictions), and can approve memorandums of agreement on data use. |
| **Data Steward** | * Data stewards carry out data owner requirements, translate them into meaningful requirements, and act as subject matter experts on the data and its utility for business use. Where data owners have decision rights, stewards protect those rights and convey them across the enterprise. Data stewards bridge the gap between different data stakeholders, so they need excellent communication skills in order to translate between data owners, platform developers, and data users. |
| **Data Custodian** | * A data custodian is a particular type of data steward. In some organizations, a data custodian is referred to as a technical data steward. The role of the data custodian is closer to the IT side of the house than the mission side, and typically covers tasks such as information flow, operation of APIs, and articulating technical data requirements to platform development teams. |

### Data Catalog

* Each agency must maintain a comprehensive data inventory that lists all data assets under its control, including data sets, databases, files, and data sources.
* The data catalog must list all data assets with descriptions, data classifications, and responsible data stewards.
* The data catalog must be regularly updated and include metadata such as data source, data owner, data classification, and data retention policies.
* The data catalog should be accessible to authorized personnel within Commonwealth agencies for reference and discovery of available data resources.
* Datasets submitted to Archer can be published as the agency’s data catalog or agencies may request ODGA to scan their assets into the Commonwealth of Virginia’s Data Catalog.

### Data Classification

* All datasets must be classified based on its level of sensitivity and criticality of the data
* Data classification labels should include but are not limited to Public, Internal Use Only, Confidential, and Highly Sensitive
* Data shared with ODGA must identify a Tier level (see table below) to ensure the proper security and protection of data.
* Data owners are accountable for ensuring the proper classification of data and data stewards are responsible for communicating the Tier level of data provided to ODGA.

| **Tier** | **Description** |
| --- | --- |
| **Tier 1 Data** | Data that is not protected from public disclosure or subject to withholding under any law, regulation, or contract. Nevertheless, publication of the dataset on the public Internet and exposure to search engines would: have the potential to jeopardize the safety, privacy, or security of a person who may be identified through use of the data; requires subjective redaction to classify the data as Tier 0 data; impose an undue financial or administrative burden on the Data Trust Member; or expose the Trustee or Data Trust Member to litigation or liability. |
| **Tier 2 Data** | Sensitive or proprietary information intended for access or release only on a 'need-to-know' basis, including personal information not otherwise classified as Tier 0 or 1, and data protected or restricted by contract, grant, or other agreement terms and conditions provided on an ongoing basis or as a one-time transfer to Trustee by Data Trust Member for use by Data Trust under this Agreement as detailed in Exhibit B attached hereto or subsequently contributed by Data Trust Member and detailed in the Data Trust electronic metadata registry. |
| **Tier 3 Data** | Sensitive or proprietary information and data elements with a statutory requirement under Data Trust Member's relevant state and federal laws for notification to affected parties in case of a confidentiality breach (e.g. Social Security Number, driver's license number, financial account numbers, personal medical information, etc.) provided on an ongoing basis or as a one-time transfer to Trustee by Data Trust Member for use by Data Trust under this Agreement as detailed in Exhibit B attached hereto or subsequently contributed by Data Trust Member and detailed in the Data Trust electronic metadata registry. Examples of Tier 3 Data may include, but not limited to: Attorney-Client Privileged; Criminal Justice Information; Critical Infrastructure Information; Family Educational Rights and Privacy Act (FERPA); Federal or State Tax Information; or Protected Health Information (PHI) under the Health Insurance Portability and Accountability Act (HIPAA). |
| **Tier 4 Data** | Sensitive or proprietary data where the unauthorized disclosure could potentially cause major damage or injury, including death, to entities or individuals identified in the information, or otherwise significantly impair the ability of the Data Trust Member to perform its statutory functions. Tier 4 Data includes any dataset designated by a federal agency at the level “Confidential” or higher under the federal government’s system for marking classified information. No Tier 4 data shall be knowingly incorporated into the Data Trust. |

### Key Data Governance Policies

* Each agency must establish and adhere to a set of key data governance polices including, but not limited to:

|  |  |
| --- | --- |
| **Policy** | **Minimum Requirements** |
| **Data Stewardship** | * Each data asset must have an assigned data steward responsible for its quality, security and compliance. * Data stewards must ensure data is used for its intended purposes and enforce data access and usage policies. |
| **Data Quality** | * Agencies must establish data quality standards and regularly assess and improve data quality. * Data quality assessments should include data accuracy, completeness, consistency, and timeliness. |
| **Data Security** | * Data security measures must be in place to protect data assets from unauthorized access, disclosures, alteration, and destruction. * Encryption, access controls, and data masking should be implemented based on data classification. |
| **Data Retention** | * Agencies must develop data retention policies to ensure data is retained for an appropriate period. * Data retention schedules must comply with legal and regulatory requirements. |
| **Data Privacy** | * Agencies must adhere to data privacy laws and regulations, protecting the privacy of individuals’ data. * Personal data must be handled with care and in compliance with relevant privacy standards. |

### Reporting and Metrics

* Agencies must maintain a data risk register of data management issues,
* Agencies must establish a system for collecting data governance metrics.
* Metrics should include, but are not limited to, data quality, data risks, and compliance with data governance policies.

**Sample Metrics**

* + **Data Quality**
    - Data accuracy – The percentage of data records that are error-free and consistent with the source
    - Data Completeness – The percentage of expected data elements present in a dataset
    - Data Consistency – The level of uniformity in data values and formats across different systems and sources
    - Data Timeliness – The average time it takes to update and make data available for use.
  + **Data Security**
    - Data access controls – The number of unauthorized access attempts or breaches
    - Data encryption – The percentage of sensitive data that is encrypted at rest and in transit
    - Data masking effectiveness – The number of successful attempts to access masked or redacted data
    - Incident response time – The average time it takes to detect and respond toa data security incident
  + **Data Privacy**
    - Data privacy compliance – The percentage of data assets that comply with relevant data privacy regulations (e.g. HIPAA)
    - Data privacy training – The percentage of agency staff who have completed data privacy training
  + **Data catalog** 
    - Data Inventory completeness – The percentage of data assets included in the data catalog
    - Data catalog usage – The number of searches, views, and downloads of data catalog entries
    - Data catalog updates – The frequency and timeliness of data catalog updates
  + **Data Stewardship**
    - Data steward assignments – The percentage of data assets with assigned data stewards
    - Stewardship activities – The number of data stewardship activities, including data quality assessments, data audits, and policy enforcement
    - Data steward training – The percentage of data stewards who have completed relevant training
* Data Governance Status Reports with key metrics must be submitted to the Commonwealth of Virginia Chief Data Officer, at least annually.
* These reports should include recent successes, areas of improvement and actions taken to address data governance issues.

### Policy Review

This Policy will be reviewed and updated every three (3) years from the approval date, or more frequently if appropriate. Any staff members who wish to make any comments about the Policy may forward their suggestions to the Director of Data Governance, ODGA.

### Related Policies

|  |
| --- |
| **ODGA Policies, Standards and Procedures** |
| Data Protection Policy |
| Metadata Policy |

The Office of Data Governance and Analytics adheres to all Commonwealth Information Technology Resource Management (ITRM) policies and standards for security and architecture [Policies, Standards & Guidelines | Virginia IT Agency](https://www.vita.virginia.gov/policy--governance/policies-standards--guidelines/).

|  |  |
| --- | --- |
|  | **VITA Related Policies** |
| IT Information Security Policy - SEC519 |
| Information Security Standard (SEC501) |
| IT Risk Management Standard (SEC520) |

## Version History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version Number** | **Revision Date** | **Description of Change** | **Author** |
| V1 | 10/31/2023 | Initial Draft | Chris Burroughs |
|  |  |  |  |
|  |  |  |  |