Subject: Administrative Data Management and Access Policy

Contents

1. Premise and Purpose

Administrative data captured and maintained at Virginia Tech are a valuable university resource. While these data may reside in different database management systems and on different machines, these data in aggregate may be thought of as forming one logical university resource, which will herein be called the University Enterprise Database (UEDB). The UEDB contains data from multiple operational areas that need to be integrated in order to support institutional research, business analysis, reporting, and decision-making.

This policy establishes uniform data management standards and identifies the shared responsibilities for assuring that the UEDB has integrity and that it efficiently and effectively serves the needs of the university. This policy applies to those data that are critical to the administration of the university, regardless of whether the data are used or maintained by administrative or academic units.

2. Policy

2.1 Data Management Roles and Responsibilities

Chief Information Officer - The university official responsible for overseeing the management of the University Information Resource (***Vice President for Information Technology***). The Chief Information Officer (CIO) has the signature approval authority for the Administrative Data Management and Access Policy. In consultation with the President and the Senior Vice President and Provost, the CIO mediates conflicts and discrepancies between the interest of the data trustees, the University Vice Presidents and the needs and interests of the university.

Data Trustees - Senior university officials (***typically at the level of Associate or Assistant Vice President***) who have planning and policy-making responsibilities for university data. The Data Trustees, as a group, are responsible for overseeing the establishment of data management policies and procedures and for the assignment of data management accountability.

Data Stewards - University directors (***typically at the level of Controller, Registrar, or Director of Admissions***) who oversee the capture, maintenance and dissemination of data for a particular operation. Data Stewards are appointed by the respective Data Trustee. Data Steward responsibilities include the data management activities outlined in this policy and other activities that may be assigned by a Data Trustee.

Data Managers - Information technology staff in a functional area with day-to-day responsibilities for the capture, maintenance, and dissemination of data for a particular operation. ***A Data Manager generally reports to a Data Steward.*** The data management activities assigned to a Data Manager may be specified in this policy or by the Information Technology organization or delegated by a Data Steward or Data Trustee.

Data Experts - Operational managers in a functional area with day-to-day responsibilities for managing business processes and establishing the business rules for the production transaction systems. ***A Data Expert generally reports to a Data Steward.*** The data management activities assigned to a Data Expert may be specified in this policy or delegated by a Data Steward or Data Trustee.

Data Users - Individuals who access university data in order to perform their assigned duties or to fulfill their role in the university community. Data Users are responsible for protecting their access privileges and for proper use of the university data they access.

General System Management Team - A representative group from Information Technology and other Data Stewards which makes recommendations related to data, issues, and standards that affect more than one administrative area. The team will establish and document integration standards for code mappings and crosswalks between administrative applications and systems, and insure that individual responsibilities and procedures are clearly outlined and appropriately communicated.

Data Management Group - A university-wide group (***typically composed of Data Stewards, Data Experts, Data Managers, and interested Data Users***) which reviews data management activities and makes recommendations to Data Trustees. It is the responsibility of Information Technology to see that this group is convened and coordinated.

Information Resource Management (IRM) - A division within the Information Technology organization consisting of data base management and security systems specialists. IRM works cooperatively with the Data Stewards, Data Experts, and Data Managers to specify, implement, and maintain appropriate security controls and authorized access for Data Users.

Information Warehousing and Access (IWA) - A division within the Information Technology organization consisting of data archiving and reporting specialists. IWA works cooperatively with the university community to implement a data warehousing system that collects, structures, and delivers university data to support timely, effective decision-making.

University Information Technology Security Officer - The university official responsible for maintaining a plan for security policies and practices and for keeping abreast of security related issues internally within the university community and externally throughout the information technology marketplace.

2.2 Data Management Principles

Data Ownership - The UEDB is a university resource; individual units or departments may have stewardship responsibilities for portions of the enterprise data.

Data Definition - Data Stewards and Data Experts provide data descriptions so Data Managers and Data Users know what shareable data are available, what the data mean, and how to access and process the data. These data about the data are referred to as data definitions and sometimes called metadata. Data definitions may be stored in an integrated or complementary database known as a ***Metadata Repository***. Data definitions should be based on actual usage, documented and modified only through procedures established by the Data Stewards, and periodically reviewed for currency.

Data Administration - The function of applying formal guidelines and tools to manage the university's information resource is termed data administration. Responsibility for data administration activities is shared among the Data Stewards, Data Experts, Data Managers and the Information Technology organization. Where information is shared among systems, Information Technology will document the process and identify the responsibilities.

Data Integration Model- The Data Stewards, or designated Data Experts and Data Managers, collaborate to establish and maintain a university-wide Data Integration Model that describes all major data entities of the UEDB and the relationships among those data entities. Included in the model are the linkages among data collected or maintained by the various organizational units of the university. The Information Warehousing and Access (IWA) division of the Information Technology organization provides expertise and software tools for data modeling.

Security Administration - The function of specifying, implementing, and maintaining access control to assure that Data Users have the appropriate authorized access needed to perform assigned duties or to fulfill university roles. The Information Resource Management (IRM) division of the Information Technology organization, Data Stewards, Data Experts, and Data Managers shares responsibility for security administration activities.

System Administration - The function of maintaining and operating hardware and software platforms is termed system administration. Responsibility for system administration activities may belong to University Computing Services, Information Technology, or to other organizations within the university.

3. Procedures

3.1 Data Administration

This section describes the data administration activities that are the ultimate responsibility of the Data Steward for a particular functional area. These activities may be delegated to Data Experts and Data Managers as deemed appropriate. (See also the ***List of Data Steward Responsibilities*** included in this document).

3.1.1 Data Capture and Storage

An official data storage location or system-of-record for each data element is identified by the Data Steward.

An official data storage location for valid codes and values for each data element is identified by the appropriate Data Steward.

UEDB data element definitions and codes are managed by the Data Steward to assure they are consistent across all applications or that they conform to pre-established integration standards for code mappings and crosswalks between systems.

Archiving requirements and strategies for storing and preserving archived historical data are pre-determined by the Data Steward for each UEDB data element. Information Technology assists in determining archiving requirements and data storage location for UEDB data.

3.1.2 Data Integrity, Validation and Correction

Data Stewards are responsible for assuring that applications that capture and update UEDB data incorporate edit and validation checks to protect the integrity of the data.

Any Data User may question the accuracy of any data element. The Data User is responsible for helping to correct the problem by supplying as much detailed information as possible about the nature of the problem.

Data Stewards are responsible for assuring data integrity, responding to questions about the accuracy of data, and correcting inconsistencies if necessary.

Upon written identification and notification of erroneous data, corrective measures are taken as soon as possible to:

1. Correct the cause of the erroneous data.
2. Correct the data in the official storage location.
3. Notify users who have received or accessed erroneous data.

3.1.3 Data Collection and Maintenance

Data Stewards are responsible for complete, accurate, valid, and timely data collection. Operational responsibility for data collection and maintenance is typically assigned to the Data Experts.

Delegation and decentralization of data collection and maintenance responsibility is encouraged in order to assure that electronic data are efficiently updated at or near the data source or creation point. Furthermore, data-handling steps that do not add value should be eliminated. Procedures may be added instead to provide new informational status reports to interested parties.

3.1.4 Data Extracts and Reporting

Data Stewards are responsible for specifying business rules regarding the manipulation, modification, or reporting of UEDB data elements. Data Stewards are also responsible for establishing standard UEDB data transformations to create pertinent summary or derived data. Note that summary or derived data are considered part of the UEDB and therefore subject to the same data management standards.

Data Stewards are responsible for specifying proper dissemination of UEDB data; individual Data Users are held accountable for their own use of the data (see the ***Other Related Virginia Tech Policies*** section of this document).

All sets of data extracted or reported from the UEDB should include a notation or display of the time and date they were extracted from the source operational system/s so the currency of disseminated data can be clearly communicated.

Data Stewards work with Data Users to define useful and meaningful schedules for creating standard data extracts. These standard extracts of the data ("data snapshots") are considered part of the UEDB and therefore subject to the same data management standards.

3.1.4.1 Data Views

A data view is a logical collection of data elements, assembled and presented according to a prescribed set of rules. Unlike a data extract that captures data at a fixed point-in-time and often includes moving the data to a secondary physical storage location, a data view is a logical subset of stored data. A data view typically assembles the most current or pertinent data from the primary storage location at the time of access.
Data views are often defined in order to:

- Supply data derived from standardized calculations or analysis,
- Aggregate data from multiple sources,
- Segment data into smaller and more manageable subsets, or
- Segregate data according to confidentiality or restriction characteristics so that access to the resulting subset may be more widely distributed.

Data Stewards are responsible for defining standard data views of enterprise data within the UEDB. These data views are considered part of the UEDB and therefore subject to the same data management standards.

Data Experts, Data Managers, or Data Users may recommend the modification or definition of data views.

3.1.4.2 Data Archiving

Data Stewards are responsible for defining the criteria for archiving the data to satisfy retention requirements. Data Managers work with Information Technology to develop appropriate data archiving strategies and procedures. Note that the capture of historical data into a Data Warehouse does not relieve the Data Steward of the responsibility for maintaining archives of detail transactional data in accordance with legal record retention requirements (see *Policy 2000: Management of University Records*).

3.1.4.3 Data Warehousing

The Information Warehousing and Access (IWA) organization, along with the Data Stewards, are jointly responsible for establishing an informational database known as the Data Warehouse. The Data Warehouse stores sharable historic data from operational systems-of-record, as well as transactional data derived from the operational data and deemed to be useful management information. It supports Data User queries to track and respond to business trends and to facilitate forecasting and planning efforts. Note that the Data Warehouse will often contain summarized data derived from transaction detail and may not contain all the supporting transaction detail stored in the operational system-of-record or in the data archives.

The Data Warehouse design is based on a ***Data Integration Model,*** which is a logical construct that describes entities that comprise the University Enterprise Database (UEDB). The Data Integration Model clarifies the linkages among data collected or maintained by the various organizational units of the university. The Data Stewards work with the Information Technology organization to develop and maintain the Data Integration Model. The Information Warehousing and Access (IWA) organization provides expertise and software tools for data modeling.

A ***Metadata Repository***, which contains both technical and business descriptions and definitions about UEDB data, is a complementary facet of the Data Warehouse. It assists the Data User with understanding the source, meaning and proper use of the warehoused data. The Data Steward works with the Information Technology organization to develop and maintain the Metadata Repository. Data Experts typically supply the business description metadata.

3.1.5 Data Documentation

Documenting UEDB data is the responsibility of the Data Stewards. Some or all of the related tasks may be assigned to Data Experts or Data Managers. Data documentation and definition guidelines are established by the Information Warehousing and Access (IWA) organization and include the following:

- Name and Alias Names
- Business Description
- Data Steward Identification
- Usage and Relationships
- Source and Procedure for Data Capture
- Frequency of Update
- Official System-of-Record Location and Format
- Designation as "Limited-access", "University-internal", or "Public"
- For "Limited-access" Data Elements: Descriptions of the restriction and the access procedures
- Description of Validation Criteria and/or Edit Checks
- Description, Meaning, and Location of Allowable Codes
- Archiving Requirements and Procedures
- Data Storage Location of Extracts
- Quality/Reliability Rating

Documentation for derived UEDB data should also include the algorithms or decision rules for the derivation.

Documentation of data views should include reference to the data elements which comprise the view and description of the rules by which the view is constructed.

Overview documentation for logical segments of the UEDB (databases, files, groups of files) should also be provided to include information about data structure and update-cycles necessary for the accurate interpretation of the data.

Documentation of warehoused data may be stored in a complementary metadata repository. The following guidelines are to be implemented concurrently with implementation of a Data Warehouse:

- IWA specifies a standard data format for receiving and loading data definitions and descriptions into a Metadata Repository.
- Data Stewards provide IWA with data definitions and other descriptive documentation in the prescribed standard machine-readable format.
- IWA is responsible for the data administration function of maintaining a Metadata Repository for the Data Warehouse and for making it readily accessible to all interested parties.
- Changes to any data definition characteristics should be noted to IWA and recorded in the Metadata Repository at the time of the change.

3.2 Access and Security Administration

3.2.1 Data Access Philosophy

The value of data as a university resource is increased through its widespread and appropriate use; its value is diminished through misuse, misinterpretation, or unnecessary restrictions to its access. Furthermore, increased data access and use improves data integrity because discrepancies are identified and errors are subsequently corrected.

As an educational institution with a mission to disseminate knowledge, and as a public institution accountable to Virginians, Virginia Tech values ease of access to information, including administrative data. Permission to view or query data contained in the UEDB should be granted to all Data Users for all legitimate purposes. Update access should be restricted as necessary, but granted to university employees at the location where data are initially received or originates whenever this is feasible. Information specifically protected by law or regulation must be rigorously protected from inappropriate access. Examples include student grades or personnel evaluations that are identifiable with a specific person.

3.2.2 Data Access Categories

As part of the data definition process, Data Stewards assign each data element and each data view in the UEDB to one of three data access categories:

- University-internal
- Public
- Limited-access

*Except as noted below, all enterprise data are designated as **university-internal** data for use within the university.* All university employees have access to these data, without restriction or prior authorization, for use in the conduct of university business. These data, while freely available within the university, are not designated as open to the general public.

Where appropriate, Data Stewards may identify elements or views of the UEDB that have no access restriction whatsoever. *Designated **Public** data may be released to the general public*.

Where necessary, Data Stewards may specify some data elements as limited-access. *Designated **Limited-access** data includes those data for which Data Users must obtain individual authorization prior to access, or to which only **need based** access may be granted.*

When data are designated as Limited-access, the Data Steward should provide the following to the Information Resource Management (IRM) organization:

1. Specific reference to the legal, ethical, or externally imposed constraint which requires the restriction.
2. Description of Data User categories that are typically given access to the data, under what conditions, or with what limitations.
3. Documentation of the process for approving and implementing access.
4. Documentation of the process for maintaining security controls.

Note that a data view can possibly have more open access than that of the underlying data elements that comprise it. For example, removal of person-identifying data elements from a view may result in a view that contains some otherwise-restricted data elements but that the Data Steward may now designate as public or university-internal.

The appropriate Data Steward in collaboration with IRM is responsible for determining and documenting data access procedures that are unique to a specific information resource, view, or set of data elements.

Any Data User may request that a Data Steward review the restrictions placed on a data element or data view, or review a decision to deny access to Limited-access data. The appropriate Data Trustee makes the final determination about restrictions and access rights for enterprise data.

3.2.3 Implementation of Security Controls

The Information Resource Management (IRM) organization and the Data Stewards share security administration responsibilities (i.e., the functions of specifying, implementing, and managing system and data access control). To the extent possible, the Data Stewards work together and with IRM to define a single set of university procedures for requesting and authorizing access to limited-access data elements in the UEDB. Data Stewards and IRM are jointly responsible for documenting these access request and authorization procedures. Data Stewards, with the assistance of Information Technology, are responsible for monitoring and annually reviewing security implementation and authorized access.

All Data Users who are cleared for the limited-access category of UEDB data must acknowledge (by signed statement or other documented means) that they understand the level of access provided and accept responsibility to both protect their access privileges and to maintain the confidentiality of the data they access. Data Stewards are responsible for defining and implementing procedures to assure that data are backed up and recoverable in response to events that could compromise data integrity. Information Technology or other university organizations may assist in this effort.

Data Stewards may delegate specific security administration activities to operational staff.

The University Information Technology Security Officer is responsible for maintaining a plan for security policies and practices and for keeping abreast of security related issues internally within the university community and externally throughout the information technology marketplace.

3.3 System Administration

University enterprise data may be stored on a variety of computing hardware platforms, provided such platforms are fully integrated components of a managed ***University Information System.*** Whenever university enterprise data are stored on any component of a university information system, that system component must have a defined ***System Administration*** function with a designated system administrator whose responsibilities include:

- physical site security
- administration of security and authorization systems
- backup, recovery, and system restart procedures
- data archiving
- capacity planning
- performance monitoring

3.4 User Support and Responsibilities

Data Stewards are responsible for providing user support to assist Data Users with interpretation and use of UEDB data. Data Stewards are responsible for providing documentation of the information resource and also training and consulting services as needed. These responsibilities may be delegated to Data Experts and Data Managers.

The Information Warehousing and Access (IWA) organization assists with training classes on the sub-set of data included in the Data Warehouse. IWA also provides training and consulting on the use of available query and reporting tools. Other university departments may also assist in this effort.

Data users are held accountable for their own use and interpretation of the data and may be required to attend training prior to being allowed to access UEDB data. Data Users are also responsible for reading and adhering to the principles and guidelines of the *Acceptable Use and Administration of Computer and Communications Systems* statement.

3.5 Policy Updates

This policy updates *Policy 2005: Administrative Data Management and Access Policy*, which was adopted September 29, 1999.

As an ongoing document, this *Virginia Tech Administrative Data Management and Access Policy* is maintained and revised as needed by the Information Technology organization with input from the Data Stewards, Data Managers, and Data Experts. Revisions may be reviewed by the Data Trustees and are approved by the Vice President for Information Technology.

4. Definitions

The following is a summary of terms used and defined in this policy.

CHIEF INFORMATION OFFICER is the university official responsible for overseeing the management of the University Information Resource (***Vice President for Information Technology***). The Chief Information Officer has the signature approval authority for the Administrative Data Management and Access Policy.

DATA ADMINISTRATION is the function of applying formal guidelines and tools to manage the university's information resource.

DATA EXPERTS are the ***operational managers*** in a functional area with day-to-day responsibilities for managing business processes and establishing the business rules for the production transaction systems. A Data Expert generally reports to the Data Steward for a given operational area.

DATA INTEGRATION MODEL is a logical construct that describes the data entities that comprise the University Enterprise Database (UEDB) and the relationship among those entities.

DATA MANAGEMENT GROUP refers to an assembly of Data Stewards, Data Managers, and interested Data Users who meet periodically to review and recommend data management processes and procedures relative to the university information resource.

DATA MANAGERS are ***information technology staff*** with day-to-day responsibilities related to the capture, maintenance, and dissemination of data for an operational subject area. A Data Manager generally reports to the Data Steward for a given operational area.

DATA STEWARDS are university directors (***typically at the level of Controller, Registrar, or Director of Admissions***) who oversee data management functions related to the capture, maintenance, and dissemination of data for a particular operational area. Data Stewards are appointed by the respective Data Trustee. Data Stewards generally report to a Data Trustee.

DATA TRUSTEES are senior university officials ***(typically at the level of Associate or Assistant Vice President***) who have planning and policy-level responsibilities for university data and who assign accountability for data management. Data Trustees report to Vice Presidents, who, in turn, have oversight authority for data administration within their reporting structure.

DATA USERS are individuals who access university data in order to perform their assigned duties or to fulfill their role in the university community.

DATA VIEW refers to a logical collection of data elements, possibly from multiple physical databases, which are assembled and presented according to a defined set of rules.

DATA WAREHOUSE refers to a query-only database containing historical point-in-time data and summary information from university operational systems. The data warehouse is used to support business analysis and decision-making.

GENERAL SYSTEM MANAGEMENT TEAM is a representative group from Information Technology and other Data Stewards who make recommendations related to data, issues, and standards that affect more than one administrative area.

LIMITED-ACCESS DATA refers to those elements of the UEDB where, because of legal, ethical, or other externally imposed restrictions, access may not be granted without specific authorization or to which only **need based** access may be granted.

METADATA REPOSITORY refers to a database system that contains descriptive information about the university's enterprise data and administrative systems. The repository is a complementary facet of the Data Warehouse.

PUBLIC DATA refers to the elements of the UEDB that are available to the general public, including people outside of Virginia Tech.

SECURITY ADMINISTRATION refers to the functions of specifying, implementing and managing system and data access control.

SYSTEM ADMINISTRATION refers to the function of applying formal guidelines and practices to the management of a computing resource.

UEDB (University Enterprise Database) is a conceptual term used to identify that body of data critical to university planning, management, and business operations of both administrative and academic units.

UNIVERSITY INFORMATION SYSTEM is a conceptual term used to identify the collection of computer hardware, software, and network connections, which together form the integrated system underlying the logical University Enterprise Database (UEDB).

UNIVERSITY-INTERNAL DATA refers to those elements of the UEDB that may be accessed by all employees of the university, without restriction, for the conduct of university business.

UNIVERSITY INFORMATION TECHNOLOGY SECURITY OFFICER refers to the university official responsible for maintaining a plan for security policies and practices and for keeping abreast of security related issues internally within the university community and externally throughout the information technology marketplace.

5. List of Data Steward Responsibilities

Data Stewards are responsible for the following activities in their respective functional area/s. Data Stewards may delegate these activities to Data Experts, Data Managers, or others as deemed appropriate.

- Establish procedures for defining and changing data elements within the operational systems.
- Work with Information Technology and other Data Stewards to establish and maintain a university-wide Data Integration Model that describes the data entities of the UEDB and the relationships among these entities

- Identify an official data storage location for each UEDB data element and for valid codes.
- Work with Information Technology to determine data retention requirements and archiving strategies for storing and preserving historical operational data.
- Work with Information Technology to insure that data element definitions and codes are consistent across all applications or that they conform to pre-established integration standards for code mappings and crosswalks between systems.
- Assure data integrity, respond to questions about the accuracy of data, and correct inconsistencies.
- Assure data collection is complete, accurate, valid, timely, and that data are maintained as close as is possible to the source or creation point of the data.
- Set business rules regarding the manipulation, modification, or reporting of UEDB data elements and for creating derived elements.
- Work with IWA to establish an informational database known as the Data Warehouse to store historical sharable data from the operational system-of-record.
- Define standard views of enterprise data within the UEDB.
- Work with Data Users to define useful and meaningful schedules for creating standard data extracts.
- Provide data descriptions and other documentation of warehoused data to IWA in a standard machine-readable format for inclusion in a Metadata Repository
- Specify the proper dissemination of UEDB data and security requirements by assigning each data element and each data view to one of the three access categories.
- Work with Information Technology to define and document a single set of procedures for requesting and authorizing access to limited-access data elements.
- Work with Information Technology to monitor and periodically review security implementation and authorized access.
- Work with Information Technology and others to define and implement procedures that assure data are backed up and recoverable in response to events that compromise data integrity.
- Work with Information Technology to provide effective user support through documentation, training and consultation.

6. Other Related Virginia Tech Policies

- Policy 2000: Management of University Records
- Policy 2010: Release of Names and Addresses of Students, Faculty, Staff, and Alumni
- Policy 7000: Acceptable Use and Administration of Computer and Communication Systems

7. References

- Data Administration Guidelines for Institutional Data, Indiana University
- Administrative Data Access Policy, University of Virginia
- Standards, Practices, and Procedures, University of Arizona
- Data Administration Mission, University of Maryland

8. Approval and Revisions

Original Policy 2005: Guidelines for University Administrative Information Resource Management, approved January 5, 1989 by the Administrative Systems Users Group (ASUG).

Approved November 14, 1989 by the Director of the Office of Institutional Research and Planning Analysis, James R. Montgomery.

- Revision 1: Policy rewritten and renamed "Administrative Data Management and Access Policy." Policy reviewed on September 1, 1999 by the Senior Vice President and Provost, P. S. Meszaros, and by the Executive Vice President, M. E. Ridenour.

  Signed into policy by Vice President for Information Systems, E. L. Blythe, on September 29, 1999.

- Revision 2

  Minor changes for clarification and to reflect organizational reporting structure changes.

  Approved April 15, 2002 by Vice President for Information Technology, Earving L. Blythe.

- Revision 3

  Minor changes to reflect organizational structure changes.

  Policy renumbered to Information Technology Policy 7100 from former General University Policy 2005.

  Approved July 18, 2006 by Vice President for Information Technology, Earving L. Blythe.

September 11, 2006: Technical revision to correct referenced policy numbers.

October 23, 2007: Technical revision to correct section 3.2.3 statement to reflect reviewing annually instead of periodically.

- Revision 4

  April 1, 2008: Updates to position titles and/or responsibilities due to university reorganization.