Institutional Data Governance Policy

Vanderbilt University and Medical Center

Effective Date: 07/09/2014
Revision Date: N/A

DOCUMENT CONTROL

Document Title	Institutional Data Governance Policy		
Summary:	This document defines the policies of Vanderbilt University regarding data governance.		
Date of Issue:	07/09/2014		
Version:	Version 1.0		
Contact Officer:	Institution Data Governance Executive Director		
Relevant References:	Vanderbilt Internal References: • Human Resources, Policy #HR-025 (http://hr.vanderbilt.edu/policies/HR-025.php)		
Change History	Version	Who	What
	1.0	Bell, Roberta	07/09/2014: Policy approved by PPIDG (Process Priority and Institutional Data Governance)

CONTENTS

INTRODUCTION

Background

Institutional data are assets maintained to support Vanderbilt's central mission of teaching, research, service, and healthcare. To support effective and innovative management, institutional data must be accessible, must correctly represent the information intended, and must be easily integrated across Vanderbilt University's information systems to support the organization's strategic plans.

The Vanderbilt executive leadership team recognizes the value-added benefits of being able to aggregate information across multiple complex systems and business processes that enable Vanderbilt University to be a world-class University and Academic Medical Center leader. The Institutional Data Governance team is responsible for establishing data governance policies, procedures, standards, and guidelines for ensuring maximum value of our data can be achieved.

Objectives

The Data Governance Policy addresses data governance structure and includes sections on data access, data usage and data integrity and integration. Adherence to the data governance policy and procedures shall;

- Establish appropriate responsibility for the management of institutional data as an institutional asset.
- Improve ease of access and ensure that once data are located, users have enough information about the data to interpret them correctly and consistently.
- Improve the security of the data, including confidentiality and protection from loss.
- Improve the integrity of the data, resulting in greater accuracy, timeliness and quality of information for decision-making.
- Establish standard definitions for key institutional data to promote data integrity and consistency.

The purpose of data governance is to develop institution-wide policies and procedures that ensure that our data meet these objectives within and across our administrative or academic data systems.

Scope

For the purpose of this policy, the term "Vanderbilt University" includes the following areas:

- Academic Affairs
- Administration
- Athletics
- Development & Alumni Relations
- Finance
- General Counsel
- Health Affairs
- Information Technology
- Investments
- Public Affairs
- University Affairs

"Institutional Data" refers to data elements that are aggregated into metrics relevant to operations, planning, or management of any unit at Vanderbilt University, including its medical center, that is reported to Vanderbilt's Board of Trust, federal and state organizations, generally referenced or

required for use by more than one organizational unit, or included in official administrative reporting.

Policy applies to anyone engaged with Vanderbilt University by employment or contract that creates, manages or reports these data referenced in scope above on behalf of Vanderbilt University, or relies on these data for decision making and planning.

The Vanderbilt University Medical Center clinical enterprise is excluded from this policy as are sole possession notes and records that are the personal property of individuals in the university community; research notes, data and materials; data that results from sponsored research projects; and instructional notes and materials.

Who Should Read This Policy

All Vanderbilt employees who use data, regardless of the form of storage or presentation. All senior administrators have the responsibility to understand and implement this policy, including, as necessary, the adoption of specific procedures for their respective areas in furtherance of and in accordance to this policy.

POLICY

Data Access

One purpose of the data governance policy is to ensure that employees have appropriate access to institutional data and information. While recognizing the institution's responsibility for the security of data, the procedures established to protect that data must not interfere unduly with the efficient conduct of our business. The policy applies to all uses of Vanderbilt University data covered by the scope of this policy regardless of the offices or format in which the data reside.

The value of data as an institutional resource is increased through its widespread and appropriate use; its value is diminished through misuses, misinterpretation, inaccuracies, and unnecessary restrictions to its access.

The institution will protect its data assets through security measures that assure the proper use of the data when accessed. Every data item will be classified by the relevant Data Steward to have an appropriate access level. VUIT will provide the technology framework for data access to be provisioned. The Data Stewards are responsible for ensuring the access levels are appropriate. Read-only access to administrative information may be provided to employees for the support of institutional business without unnecessary difficulties/restrictions.

Any employee or non-employee denied access may appeal the denial to the Data Governance Committee. Escalation to the executive management should only be pursued if the Data Governance Committee decision needs to be appealed.

Data Usage

Another purpose of data governance policy is to ensure that institutional data are not misused, and are used ethically, according to any applicable law, and with due consideration for individual privacy. Use of data depends on the security levels assigned by the Data Steward.

Vanderbilt University personnel must access and use data only as required for the performance of their job functions, not for personal gain or for other inappropriate purposes; they must also access and use data according to the security levels assigned to the data. Data usage falls into the categories of *update* and *dissemination.*

Update

Authority to update data that is reported as key institutional data shall be granted by the appropriate data steward only to personnel whose job duties specify and require responsibility for data update. This restriction is not to be interpreted as a mandate to limit update authority to members of any specific group or office but should be tempered with Vanderbilt's desire to provide excellent service to faculty, staff, students, patients and other constituents. Data Stewards shall ensure that adequate internal controls and/or change management procedures are in place to manage 'updates' to key institutional data, their definitions and processes.

Dissemination
Dissemination of data must be controlled in accordance with the security practices set forth by the Data Stewards. Appropriate use must be considered before sensitive data are distributed. Unauthorized dissemination of data to either internal/external personnel is a violation of this policy.

Data Integration

Data integration refers to the ability of data to be assimilated across information systems. It is contingent upon the integrity of the data and the development of a data model, corresponding data structures, and domains. Data model designs should focus on utilizing Master Data Management (MDM) methodologies in order to streamline how data is integrated when applicable.

System-to-system interfaces are a standard practice to move data from one system to another in order to streamline processes that extend across systems and contribute to using data efficiently and effectively.

Operational processes often require systems to exchange information. System interfaces are often developed between systems to facilitate the exchange of such information. The systems that exchange information fall into two broad categories:

- Internal – Systems that are implemented within the Vanderbilt computer systems network. They can either be procured, procured but modified, or custom developed products.
- External – Systems that do not reside on a Vanderbilt computer network. These systems are hosted by vendors and/or through sub-contracts managed by vendors.

Downloading of individually identifiable data from central systems to electronic files for the purpose of uploading or connecting the data to non-central systems (e.g., shadow systems, external vendors) without the knowledge of the data steward. This practice is not supported and introduces risks associated with data integrity, security, and long-term sustainability of information systems that may not be mitigated due to the nature of the practice. Departments and/or personnel responsible for these practices that are found to be in violation of this policy, may result in disciplinary actions up to and including dismissal from employment consistent with Staff Guidelines.

Approval by the Data Steward is specific to each request. Data granted for one purpose is not universally granted for all purposes. Each new use case must be approved by the Data Steward in a new request or an amendment to the original request, even if you already have the data.

Documented agreements regarding data use, retention, and responsibility should exist with the data stewards (and vendors in the case of data integration with external entities) of the systems providing and utilizing data. Data extraction practices that are already in use should be registered and documented agreement developed with the appropriate Data Steward member.

Data Integrity

Data systems and/or processes that are involved in the creation of institutional reports should incorporate data integrity and validation rules that ensure the highest levels of data integrity are achieved. Validation rules within data systems may need to include reconciliation routines (checksums, hash totals, record counts) to ensure that software performance meets expected outcomes. Data verification programs such as consistency and reasonableness checks shall be implemented to identify data tampering, errors, and omissions.

RESPONSIBILITIES

Data Governance Structure

The function of applying policies, standards, guidelines, and tools to manage the institution's information resources is termed data governance. Responsibility for the activities of data governance is shared among the roles listed below. Descriptions of roles and responsibilities below provide the framework of how data governance will be implemented and maintained.

Current membership of each committee and role can be found at:
www.vanderbilt.edu/datagovernance

Executive Sponsors / Process Priority and Institutional Data Governance (PPIDG)

The Executive Sponsors / PPIDG are senior Vanderbilt University officials who are responsible for setting the overall prioritization for institutional business process redesign projects; communicating process transformation priorities across the institution; ensuring project resources are available and adequate to meet established time lines; bringing clarity whenever necessary to project, process and data work; approving data governance policy; appointing members of institutional data governance committees. Executive Sponsors / PPIDG will review and make approval decisions on policies presented by the Data Governance Committee.

Institutional Data Governance Committee (IDG)

The Institutional Data Governance (IDG) committee is the body responsible for developing and submitting to Executive Sponsors / PPIDG for approval the data governance policy on data access, data usage, data integrity and integration, and data security, proposing prioritization of business intelligence work; ensuring that work plans are established and met; and, reporting up to the Executive Sponsors / PPIDG on project status and seeking input on projects that have broad institutional implications related to business intelligence and data. Assignment of personnel to the key roles listed below requires consensus within the IDG committee.

Membership

The Executive Sponsors / PPIDG are responsible for the Institutional Data Governance (IDG) committee membership. Changes to the IDG membership must be nominated to the IDG committee. Upon approval, the Executive Sponsors / PPIDG will review and provide an approval decision based on the recommendation.

Data Standards & Reporting Committee

The Data Standards & Reporting committee is a sub-committee to the IDG committee. This committee carries out policies set by the IDG committee, addresses data quality and integrity, and

sets forth data standardization and standard reporting practices into the institutional reporting environments.

Membership
The Institutional Data Governance (IDG) committee is responsible for the Data Standards & Reporting committee membership. Changes to the membership must be nominated to the IDG committee. The IDG committee will review and provide an approval decision based on the recommendation.

Data Architecture Committee

The Data Architecture sub-committee designs the technology architecture to support the reporting needs specified by the IDG and Data Standards & Reporting committees. The work of this committee is a collaborative effort between the three committees. The Data Architecture committee is responsible for maintaining technology product roadmaps (software & hardware) necessary to support the current and future state reporting requirements.

Membership
The Institutional Data Governance (IDG) committee is responsible for the Data Architecture committee membership. Changes to the membership must be nominated to the IDG committee. The IDG committee will review and provide an approval decision based on the recommendation.

Data Stewards

Data Stewards are appointed by functional area senior leadership to develop data centric policies and carry out the overall administrative data security policies. Data Stewards are responsible for making known the rules and procedures to safeguard the data from unauthorized access and abuse. They authorize the use of data within their functional area, and monitor to verify appropriate data access. They assist institutional data users by providing appropriate documentation and training to support institutional data needs.

Data Managers

Data Managers coordinate and manage the data in the business process that results in the data adhering to Vanderbilt standards. Once data have entered the system, there is a process by which they are validated, transmitted, stored, and archived. The capture and checking are typically based on a functional process or business process. This data manager role oversees adherence to the business process and in some cases develops the process. While there may be several data managers, the Data Stewards will appoint one as primary for each application.

Data Reporters

Data Reporters are individuals within the institution who have an intricate understanding of the data in their area. They establish reporting procedures for institutional data, which may include recommending changes to data entry practices. They are responsible for implementing the decisions of the data stewards in functional areas, assuring that census, backup, and retention plans are implemented according to defined needs. Because data reporters have a hands-on role with data, they monitor or oversee monitoring of data quality.

Functional Security Leads

The functional security leads are responsible for allowing access within the rules and standards set by the Data Steward for their area. The security leads should work with the Data Stewards for each area to document the agreed upon procedures that will be followed to administer security access. It is the responsibility of the functional security leads to routinely monitor access and ensure that access levels are up to date.

PROCESS

Data Governance Standards

The purpose of establishing standards is to ensure that institutional data have a high degree of integrity and that key data elements can be integrated across functional units and electronic systems so that faculty, staff, and management may rely on data for information and decision support.

Institutional data will be consistently interpreted and clearly documented, according to the best practices agreed upon by the IDG Committee, and it will have documented values in all Vanderbilt systems. It is the responsibility of each data steward to ensure the correctness of the data values for the elements within their charge.

Institutional data are defined as data that are maintained in support of a functional unit's operation and meet one or more of the following criteria:

1. the data elements are key fields, that is, integration of information requires the data element;
2. the institution must ensure the integrity of the data to comply with internal and external administrative reporting requirements, including institutional planning efforts;
3. the data are reported on or used in official administrative reports;
4. a broad cross section of users require the data.

It is the responsibility of each data steward, in conjunction with the IDG Committee, to determine which core data elements are part of our institutional data.

Documentation (metadata) on institutional data will be maintained within an institutional repository according to specifications provided by the Data Standards & Reporting Committee. These specifications will include both the technical representation/definition of each element, as well as a complete interpretation that explains the meaning of the element and how it is derived and used. The interpretation will include acceptable values for each element, and any special considerations, such as timing within an academic or fiscal calendar.

All employees are expected to bring data problems and suggestions for improvements to the attention of the appropriate data stewards or any member of the IDG Committee.

Communicating Data Governance Standards

The Data Standards & Reporting Committee is responsible for establishing data standardization and standard reporting practices. The committee will obtain approval from the Institutional Data Governance (IDG) committee when standards are developed and/or modified. A central repository will be maintained and should be referenced for specific guidelines and decision outcomes related to data governance as set forth within this policy. The repository of reporting standards, documented institutional data, and key decisions can be found at:

- www.vanderbilt.edu/datagovernance

OVERSIGHT

Penalties for deliberate violations of this policy will be adjudicated in accordance with applicable disciplinary policies and procedures of the Human Resources Staff Guidelines as applicable.

RESOURCES

- Vanderbilt Policies
    - Electronic Communications and Information Technology Resources
        - Policy #HR-025 (http://hr.vanderbilt.edu/policies/HR-025.php)
    - Vanderbilt Computing Privileges and Responsibilities – Acceptable Use Policy
        - http://www.vanderbilt.edu/info/computing-aup/
    - Federal and State laws and regulations as well as VUMC policies define requirements for protection of patient health information and research health information. Users that access patient or research health information are responsible for knowing and following VUMC policies:
        - IM 10-30.03 Access to Confidential Information
        - IM 10-30.19 Authorization and Access to Electronic Systems and Applications
        - IM 10-30.01 Confidentiality of Protected Patient Information
        - IM 10-30.13 Protection and Security of Protected Health Information
        - IM 10-30.14 Protection and Security of Research Health Information
        - IM 10-30.12 Sanctions for Privacy and Information Security Violations
        - IM 10-30.15 Electronic Messaging of Individually Identifiable Patient and Other Sensitive Information

- Related Federal Regulations
    - Family Educational Rights and Privacy Act (FERPA)
        - http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html
    - Health Insurance Portability and Accountability Act (HIPAA)
        - http://www.hhs.gov/ocr/privacy/

CONTACTS

Executive Director, Vanderbilt Institutional Research Group: roberta.bell@vanderbilt.edu

Director of Data Governance, Vanderbilt Institutional Research Group: daniel.kirby@vanderbilt.edu

DEFINITIONS

Data Element

A single data item. For example, last name is a data element.

Data Dictionary

A reference tool which provides a description of all the core institutional data elements.

Data Dissemination

The distribution of data to either internal or external stakeholders. Included in dissemination is the sending of data to external entities including vendors that provide services for Vanderbilt University.

Data Integrity

The qualities of reliability and accuracy of data values that permit the institution to have dependable data on which to make plans, projections and decisions. Data integrity contributes to the efficient operation of the institution by supporting quality customer service to students, faculty and employees, and helping the institution remain competitive.

Data Integration

The ability of data to be assimilated across information systems, is contingent upon the integrity of data and the development of a data model, corresponding data structures, and domains.

Data Model

A diagrammatic representation of the objects and their properties that are needed within an organization to accomplish its mission. Sometimes represented as an ER (entity-relationship) diagram or a data flow diagram.

Data Value

The set of values that each data element can have. For example, A&S, BLR, ENG, and GPC are a selection of values of the data element named school.

Institutional Data

The data elements that are aggregated into metrics relevant to operations, planning, or management of any unit at Vanderbilt University, including its medical center, that are reported to Vanderbilt's Board of Trust, federal and state organizations, generally referenced or required for use by more than one organizational unit, or included in official administrative reporting.

Metadata Repository

Information about the data in an organization's electronic systems. It is used to catalog the data elements and to enable software development tools and operational systems to assess the data. Data stewards add interpretive information to the repository so that the meaning of each element is clear, and can be use consistently across all systems. Data dictionaries are built from the repository.