The University of Texas at Austin

Data Management Standards

Overview

The Provost's Office and the Office of Financial Affairs recognize the campus-wide need for broader and more robust access to institutional data to facilitate operations and inform strategic decision-making. This need is exemplified in the analytical work now being done in colleges and departments to achieve 4yr graduation rate targets; requests for institutional data (detailed, transactional) are significantly increasing. In addition, the Administrative Systems Master Plan (ASMP), recently adopted by IT Governance, calls for a more robust operational reporting solution *and broader access to institutional data.*

However, wholesale release of detailed unit-level data, and specifically student data, in the absence of thoughtful data governance policies, poses numerous risks to both individuals and the institution. Risks may be legal, political, or reputational and standard practices to avoid risks may not be uniformly coordinated or ensured.

The absence of clear data access policies also introduces potential inefficiencies in our operations, at great potential cost to both colleges and the institution as a whole. These inefficiencies may take the form of duplicated requests for datasets from steward offices, shadow databases proliferating throughout campus, and resulting redundancies in dedicated hardware, software, and personnel resources. Furthermore, accurate definitions, methodologies, and data integrity may be compromised when multiple versions of those data persist.

Institutional data *must* become much more broadly accessible, but with careful consideration of the requester's role (need), type and level of data needed, and risks associated with those data. Therefore, it is imperative that institutional data governance be implemented as soon as possible, including: a comprehensive, detailed data access policy (HOP); a governing body with authority to make decisions regarding institutional data access; and a mechanism for requesting, approving, and releasing institutional data.

Acknowledging broad and comprehensive institutional data is a strategic university asset, this document outlines newly created institutional *Data Management Standards* to govern the following institutional data activities:

- Storage – efficient, protected, accessible
- Access – broad, appropriate, role-based
- Disclosure – purposeful, informed, approved

and includes a framework for establishing:

- IQ as the "Master Data Repository;"
- the Data Management Committee;
- a formal Data Access Governance Policy;
- Role-based Authorization schemes; and

- Mechanisms for requesting (and approving) access to, and use of, institutional data

1. Purpose of Standards

The Data Management Standards ("Standards") establish the structure and roles for Data Access Government at the University of Texas at Austin and provide guiding principles for the work of governance committees in order to ensure:

a. Institutional leaders, (executive, administrative, and operational) have appropriate and timely access to detailed institutional data in order to make informed business and strategic decisions for their colleges, departments, or units in support of the institutional mission.
b. Sound information management practices are incorporated into all aspects of Institutional Data production, reporting, and analysis in order to provide appropriately robust access to accurate information for decision support.
c. Institutional Data is regarded as a highly-valued institutional asset and, as such, is effectively managed to achieve an appropriate level of stewardship.
d. The value of Institutional Data is maximized by making it more readily accessible and by increasing the understanding and consistent use of those data.
e. The roles and responsibilities for effective management of Institutional Data are clearly articulated and understood.
f. Consumers of Institutional Data are adequately educated regarding all university, state, and federal laws, policies and procedures that mandate and regulate the use or dissemination of those data, and risks involving the use or release of Institutional Data – legal, operational, individual, political, or reputational – are adequately mitigated.
g. Institutional Data are stored and used efficiently, eliminating or minimizing any redundancies in hardware, software, processes, reports, or analyses, and the efforts of personnel resources who support them.

2. Scope

These Data Management Standards ("Standards") address the issues of (1) storage, (2) access, and (3) distribution of institutional data. The Standards are specifically applicable to "institutional data" extracted or derived from institutional administrative information systems and residing in the Master Data Repository (MDR) or an associated central administrative data store. The Standards are to be implemented in accordance with, and deference to, existing UT computer, data and acceptable use policies, including but not limited to:

Acceptable Use Policy
http://www.utexas.edu/cio/policies/pdfs/AUP.pdf

Information Resources Use and Security Policy
http://www.utexas.edu/cio/policies/pdfs/Information%20Resources%20Use%20and%20Security%20Policy.pdf

Data Classification Standard
http://www.utexas.edu/cio/policies/pdfs/Data%20Classification%20Standard.pdf

Data Encryption Guidelines
http://www.utexas.edu/cio/policies/pdfs/Data%20Encryption%20Guidelines.pdf

3. Principles

The following principles are set forth as fundamental tenets that form the foundation for future development of policies and procedures for data access governance at UT Austin. It is the responsibility of the Data Management Committee to develop and implement policies and procedures that reflect, refine, and address these principles:

a. **Institutional Data is the property of the University of Texas at Austin and shall be managed as a key asset** - Institutional Data will be managed through defined governance guidelines, standards, policies and procedures that are appropriately communicated to campus stakeholders.
b. **Unnecessary duplication of Institutional Data is discouraged** - Data Stewards (defined below) shall be responsible for making available Institutional Data from official Systems of Record, when reasonable and according to policies and procedures, in a manner that minimizes redundant storage and processing of those data in departmental and collateral repositories. Exceptions are allowed for purposes of business continuity and fail-over, only.
c. **The IQ database will serve as the university's Master Data Repository (MDR)** – current IQ data stores, technical infrastructure, and validation processes will be leveraged and modified as needed to provide appropriate data access, continuity, reporting, analysis, and decision support in a manner that minimizes the need for widespread duplication or redundant storage of Institutional Data.
d. **Institutional Data are to be used only for operational purposes and for advancement of the institutional mission –** Institutional assets are, first and foremost, intended for use in the pursuit of institutional excellence. Therefore, as an asset, and particularly in times of lean resources, the availability and use of institutional data for the advancement of the institutional mission and improvement of operational efficiencies must be prioritized over individual needs, research endeavors, or external requests. To that end, institutional data will not be made available for non-operational research purposes without direct sponsorship by the institution and will not be made available for thesis or dissertation research, publication, or presentation without express written consent by the Data Management Committee on behalf of the Provost.
e. **Institutional Data are not to be released to external entities without formal institutional approval** – Institutional Data should not be used to respond to inquiries from external entities, nor presented to external academic or professional audiences, nor published in any form external to institutional operations without express consent from the institution. The Data

Management Committee will develop and publicize formal guidelines and processes for requesting and approving the release of Institutional Data.

f. **Access to Institutional Data in the MDR (IQ) will be "read-only"** - To the extent appropriate and feasible, updates to Institutional Data will be performed in the System of Record and then automatically proliferated to the MDR via established transformation and load processing. This eliminates redundant processing, increases integrity, and provides better auditing capabilities. Only in exceptional cases, and with approval by the DMC, will analysts or systems be granted direct update capability in the MDR.

g. **Quality standards for Institutional Data shall be defined and monitored** – Data quality standards shall be defined, published, communicated, managed and applied according to reliability and risk levels established by appropriate Data Stewards (defined below). Examples of data quality standards Include data validation rules, timeliness of updates, defined error rates, integrity monitoring processes, etc.

h. **Institutional Data shall be protected** - Institutional Data must be safeguarded and protected according to approved security, privacy and compliance guidelines, laws and regulations established by the university or by state or federal agencies.

i. **Institutional Data shall be accessible according to defined needs and roles** - Institutional Data and Metadata shall be accessible to all constituents in accordance with defined access and use policies and procedures determined by the Data Management Committee and Data Stewards (defined below). Users requesting access to Institutional Data shall be assigned to appropriate roles that have clearly documented guidelines In accordance with all university, state, and federal laws and regulations.

j. **Institutional Metadata shall be recorded, managed, and utilized** - Metadata will be used to model, define and organize Institutional Data and will be published and communicated clearly and consistently in order to maximize the value of Institutional Data to university stakeholders.

k. **Institutional representatives will be held accountable to their roles and responsibilities** - Roles and responsibilities for data management will be clearly defined, and individuals assigned to specific roles will be held accountable to performing data management responsibilities as a part of their regular job duties.

l. **Resolution of Issues related to Institutional Data shall follow consistent processes** - The Data Management Committee shall coordinate the resolution of issues related to risks, costs, access, management and use of Institutional Data, in consultation with the appropriate Data Stewards and Data Trustees.

4. Definitions

The following definitions are applicable to the principles, roles, and processes identified in these Standards and may vary slightly from those used in other IT policy statements or industry vernacular:

a. Institutional Data – any data element captured, extracted or derived from institutional administrative information systems or residing in the Master Data Repository (MDR) or an

associated central administrative data store. Specifically, data captured and stored by source systems, or systems of record, or made available in the MDR or its affiliated databases.

b. Institutional information - a collection of institutional data which may be derived, aggregated, calculated, or contained In any form, Including but not limited to documents, memos, databases, spreadsheets, presentations, tables, charts, graphs, email and web sites.
c. Institutional metadata – detailed information describing both the technical and business characteristics of institutional data, including:
    i. Definitions regarding the purpose, use and context of Institutional Data
    ii. Identification of which system is the official system of record of Institutional Data
    iii. Identification of personnel responsible for management of Institutional Data
    iv. Descriptions of how Institutional Data Is transferred, derived, and stored
    v. Specific security and privacy practices that are used to safeguard Institutional Data
    vi. Risk and compliance classifications for Institutional Data
    vii. Rules concerning retention of records and Institutional Data
d. Source System or System of Record – the system by which any given institutional data originate, are captured, or are initially processed is the "source system." The "system of record" is the authoritative system that defines and maintains a specific subset of institutional data (ex: the registration system is the system of record for student course enrollments). In most cases, the System of Record is also the Source System. However, there are cases in which source system data are further aggregated or refined to produce official institutional information, resulting in a separate authoritative System of Record for specific purposes (i.e. 12th Class Day data).
e. Master Data Repository (IQ) – the central data store to which data from operational systems ("systems of record") are stored for broad-based (read-only) retrieval by campus users for the purpose of informing campus-wide decision-making. Data in the MDR are made available according to the "roles and access" matrix to be developed by the Data Management Committee (DMC). Given that the IQ -Oracle database is well-established with efficient ETL and validation processes for a vast amount of existing operational data, IQ will become the Master Data Repository for UT.
f. Departmental and Collateral Repositories – any of a number of departmental databases or systems that redundantly store institutional data to support remote or independent processes for use in specific units (i.e. not for official institutional or campus-wide use). As referenced in this document, "repositories" refer to large-scale copies of massive amounts of data for further processing, such as on a recurring basis; small analytical databases, flat files, or excel versions of data used for specific analysis are not considered "repositories" by this definition.
g. Subject Area Domains - broad functional/operational areas for which major information systems and databases are defined (i.e. "Research," "Faculty," "Fee Billing," Financial Aid," "Admissions," "Facilities," etc.). Domains generally have one Data Steward (though there may be several) representing the system(s) that form a given subject area.
h. Data Map – the formal reference document that associates each Subject Area Domain with its Data Steward(s) and Trustee (i.e. business experts or "responsible parties").

5. Roles and Responsibilities

a. Data Trustees – VP and Associate VP administrators with portfolio level management responsibilities over Subject Area Domain(s) and their System(s) of Record. Data Trustees are, generally, members of the Business Services Committee; a subset of Data Trustees will provide representation on the Data Management Committee.
b. Data Stewards (custodians) – the operational managers for Subject Area Domain data and Systems of Record. Data Stewards report to a Trustee, though, in some cases, the Steward may also be the Trustee. Data Stewards and their staff have deep knowledge of the System(s) of Record, definitions, processes, and business logic for their domain(s). Data Stewards assist with metadata, modeling, risks assessments, roles/access recommendations, and validation. A subset of Data Stewards will provide representation on the Data Management Committee.
c. IQ Steering / Data Management Committee (DMC) – the primary governing body for Data Access Governance. Previously called the "IQ Steering Committee," membership will be revised and roles redefined to encompass a larger data access management role. The DMC will include representatives from Data Trustees, Stewards, colleges and schools, and executive administrative offices and will report its decisions and policy recommendations through the Business Services Committee, the Operational IT Committee and, ultimately, SITAB. The DMC will be responsible for (a) refinements and addendums to these Standards; (b) development of formal Data Access Governance policies and procedures; (c) development and adoption of Roles/Access rules,; and (d) development of mechanisms for access and use requests to be submitted and considered by the committee. The DMC may form workgroups or task forces as needed to inform policy decisions or facilitate request/approval processes.
d. IQ Sponsors (a.k.a. "standards implementation committee") – the existing operational oversight group for IQ will continue in its current role and also assume responsibility for implementation and coordination of the Data Access Governance structure, maintaining a website with current policies and contact information, facilitating DMC meetings and agendas, and communicating policies to campus constituents.

Proposed Roles and Committees