
Data Administration Guidelines for Institutional Data Resources

The University of Michigan
Michigan Administrative Information Services



April 2004

Copyright © 2004 by the Regents of The University of Michigan. This document may be reproduced or reprinted, in whole or in part, without permission as long as the above copyright statement and source are clearly acknowledged. Neither this document nor any reproductions may be sold.

UNIVERSITY OF MICHIGAN
Data Administration Guidelines for Institutional Data Resources
Revised February 26, 1996
Updated April 2004

1.0 PURPOSE

The following guidelines recognize data as a resource of the University and are based on the philosophy that the value of data as an institutional resource is increased through its widespread and appropriate use: its value is diminished through misuse, misinterpretation, alteration, or unnecessary restrictions to its access. The Data Administration Guidelines for Institutional Data Resources will serve to:

- ensure establishment, maintenance and delivery of secure, confidential, trustworthy, stable, reliable, and accessible collections of institutional data in electronic form for shared access by the University community (subject to confidentiality standards).
- maximize the value received from the data asset by increasing the understanding and use of the data,
- improve data management techniques by promoting data consistency, security and standardization throughout the University, and minimizing duplication in capturing, storing, and maintaining data,
- increase data sharing by providing a reliable and secure technical environment for managing data and improving direct access to data by end-users,
- support faster changes to common applications and databases.

While all data captured about University assets are resources of the University, they vary in their relevance to the administrative processes of the University. This set of Guidelines along with the University of Michigan Institutional Data Resource Management Policy are intended to apply to those data which are critical to the administrative, clinical, education and research functions of the University, regardless of whether the data are used or maintained by administrative, academic, or clinical units. Although these data may reside in different database management systems and in different physical locations, these data in aggregate may be thought of as forming a single, logical database, which will herein be called the Institutional Database (IDB). This terminology is not intended to imply that these data now or in the future should reside in a single physical database. Rather, it is a recognition that regardless of where these data reside, there are some general principles of data management that should be applied in order to maintain the value and guarantee effective use of the information resource. Because the Institutional Database is a single, logical database, these principles of data management must be applied uniformly and as part of a coordinated effort.

When implementing vendor provided databases and software, the Data Steward will need to exercise discretion in determining how to conform to the objectives embodied in these Guidelines. A cost/benefit analysis of how best to implement some sections will be necessary, specifically Sections 2.5 Data Storage, and 2.11 Data Documentation.

NOTE: These guidelines use terms whose meanings may not be obvious to all readers. Definitions are supplied in Section 4.0 of this document.

2.0 GUIDELINES

2.1 Data Management Roles

The University is the DATA OWNER of all the University's institutional data; individual units or departments may have stewardship responsibilities for portions of the data.

The MAIS Advisory Committee recommends and the Executive Officers approve overall policy and guidelines for management and access to the institutional data of the University.

Relevant state and federal laws govern access by non-University personnel. University personnel obtaining or attempting to obtain access to confidential data not within the scope of their University responsibilities are subject to discipline.

The Executive Officers of the University are considered the DATA STEWARDS. This role may be assigned to specific senior University officials, or DELEGATED DATA STEWARDS, who have planning and policy-level responsibilities for data in their functional areas. The data stewards or their designees, as a group, are responsible for recommending policies, and establishing procedures and guidelines for University-wide data administration activities. Data stewards, as individuals, have management responsibilities for defined segments of the institutional database.

University officials and their staff who have operational-level responsibility for information management activities related to the capture, maintenance, and dissemination of data are considered DATA MANAGERS. Among the responsibilities of the data managers are any data administration activities outlined in these Guidelines which may be delegated to them by the data stewards.

University officials and their staff who are responsible for facilitating and resolving shared data management issues among central offices and schools and colleges are considered INTEGRATION COORDINATORS.

Individuals who need and use University data as part of their assigned duties or in fulfillment of their role in the University community are considered DATA USERS.

The function of applying formal guidelines and tools to manage the University's information resource is termed DATA ADMINISTRATION. Responsibility for the activities of data administration is shared among the data stewards, data managers, integration coordinators, data users, and the Michigan Administrative Information Services Data Administration (MAIS/DA) area.

The function of maintaining and operating hardware and software platforms (system environments) is termed SYSTEM ADMINISTRATION. Responsibility for the activities of system administration may belong to Michigan Administrative Information Services or to other divisions or departments within the University.

2.2 Data included in the Institutional Database

A data element is considered INSTITUTIONAL DATA, and therefore, part of the IDB, if it satisfies one or more of the following criteria:

- It is relevant to planning, managing, operating, controlling, or auditing administrative functions of an administrative or academic unit of the University.
- It is created, received, maintained or transmitted as a result of educational, clinical, research or patient care activities.
- It is generally referenced or required for use by more than one organizational unit. Data elements used internally by a single department or office typically are not part of the University's INSTITUTIONAL DATABASE.
- It is included in an official University administrative report.
- It is used to derive an element that meets the criteria above.
- It is generated by a University workforce member or agent using any of the above data.

Data elements which meet the criteria for inclusion may be identified as such by a data steward, a data manager, an integration coordinator, MAIS/DA, a user group or a data user.

A data steward(s) must be identified for each data element to be included in the IDB.

MAIS/DA should assist data stewards, data managers, integration coordinators, and data users in the negotiations for inclusion of data elements and for identification of data stewards.

2.3 Institutional Data Model

The data stewards, will be responsible for establishing and maintaining a University-wide INSTITUTIONAL DATA MODEL which describes all major data entities of the IDB and the relationships among those data entities, including the linkages among data collected or maintained by the various organizational units of the University.

MAIS/DA will support the establishment and maintenance of an institutional data model and provide expertise for data modeling.

2.4 Data Collection and Maintenance

The data steward is ultimately responsible for complete, valid, reliable, and timely institutional data collection. Operational responsibility for data collection and maintenance is typically assigned to the data managers.

Further delegation and decentralization of data collection and maintenance responsibility is encouraged in order to assure that:

- electronic data are collected and maintained as close as possible to the source or creation point of the data as identified by the data steward, independent of what office or individual within the University needs the data, and
- each manual or computer process through which data are passed adds some value to the data.

2.5 Data Storage

An official data storage location for each data element should be identified by the data steward.

An official data storage location of valid codes and values for each data element should be identified by the data steward.

Data element names, formats, and codes should be consistent across all applications which use the data and consistent with the University Data Element Naming Standards.

Archiving requirements and strategies for storing and preserving historical data should be determined for each data element by the data steward.

MAIS/DA should assist in determining data storage location and archiving requirements for IDB data.

2.6 Data Validation and Correction

Applications that capture and update institutional data should incorporate both edit and validation checks to assure the accuracy of the data. Where practicable, mechanisms should be put in place to corroborate that data has not been altered or destroyed in an unauthorized manner, and that data that has been transmitted from one location to another has not been improperly modified in the transmission process.

Any authorized data user can question the validity of any element. The data user has the responsibility to help correct the problem by supplying as much detailed information as available, sufficient to permit understanding and diagnosis of the problem.

The data steward or data manager is responsible for data integrity, for responding to questions about the accuracy of the data and for correcting inconsistencies, if necessary.

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

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

2.7 Data Manipulation, Modification, and Reporting

The data steward is responsible for setting policies regarding the manipulation, modification, or reporting of IDB data elements and for creating derived elements, which are also members of the IDB.

The data steward has ultimate responsibility for proper use of the IDB data; individual data users will be held accountable for their specific uses of the data.

All data extracted from the IDB and all reports containing that data should include a record or display of the time and date of data capture. Provision should be made for recording changes from the source to the derived data.

Data stewards will work with data managers and data users to define useful and meaningful schedules for creation of standard data extracts. These standard extracts of the data ("data snapshots") will also be considered part of the IDB.

2.8 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 - which captures data at a fixed point in time and often moves it physically to a secondary storage location - a data view is a logical entity only, which typically assembles the most current or recent data from their primary storage location at the time they are requested.

Data views may be defined in order to:

- 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.

2.9 Data Access

The following philosophy should guide decisions about access to institutional data:

- the value of data as an institutional resource is increased through its widespread and appropriate use to enhance administrative effectiveness; its value is diminished through misuse, misinterpretation, alteration, or unnecessary restrictions to its access.

ACCESS TO INSTITUTIONAL DATA -- the permission to view or to query data contained in the IDB should be granted to data users for all legitimate University purposes, subject to confidentiality rules and relevant state/federal laws.

As part of the data definition process, data stewards will assign each data element and each data view in the institutional database to one of three categories: Public data, Private/Confidential data, or Sensitive data.

Except as noted below, all institutional data will be designated as PRIVATE/CONFIDENTIAL DATA. University employees will have access to these data only if necessary in the performance of their official University duties.

Where appropriate, data stewards may identify elements or views of the institutional database which have no access restriction whatsoever and which may be released to the general public. Examples include data about official University transactions and appointments or promotions that are regularly released to the public. These data shall be designated as PUBLIC DATA.

Where necessary, data stewards may specify some data elements as sensitive. SENSITIVE DATA includes those data for which potential data users must obtain specific authorization prior to access or to which only limited access may be granted. Authorization to access sensitive institutional data will be based on appropriateness to the user's role and the intended use. Designation of data as sensitive will include:

- Specific reference to the legal, ethical, or other constraint which requires this restriction; and
- A description of what categories of data users are typically given access to the data and under what conditions or with what limitations the access is permitted.

Examples of data that are designated as sensitive include, but are not limited to, those which

- pertain to an individual's race, religion, or national origin; or
- are generated in a relationship (such as those that interlock with non-University data systems) that is administratively or legally privileged; or
- are protected health information; or
- describe the state of an individual's physical or emotional well-being; or
- describe the methods or procedures used to safeguard assets or maintain the integrity of administrative data.

A data view does not necessarily inherit the restriction characteristics of the data elements of which it is comprised. (For example, removal of any association with individually-identifying data elements can result in a view containing sensitive data elements being designated as public or private/confidential, and removal of the name itself does not necessarily make data non-individually identifiable.)

To the extent possible data stewards will work together to define a single set of procedures for requesting permission to access restricted data elements in the IDB, and will be jointly responsible for documenting these common data access request procedures.

Each data steward will be individually responsible for documenting authorization policies and data access procedures that are unique to a specific information resource or set of data elements.

Any data user may request that a data steward, or the group of data stewards together, review the restrictions or lack of restrictions placed on a data element or data view, or review a decision to deny access to restricted data.

User access to data must be documented and periodically reviewed and modified for appropriateness. User access to data must be terminated in a timely fashion upon separation from the University or change in job responsibilities that no longer make the access appropriate. The University reserves the right to audit to determine if appropriate user access is being maintained.

As determined by the data stewards, privacy, security, and compliance officials, user activity in a system containing sensitive data may be logged, retained, and reviewed in activity, access and audit logs.

Access to information systems that maintain electronic protected health information (ePHI) or sensitive institutional data will be provided only to those persons or software programs that have been authorized and granted access rights. Access to ePHI should address login monitoring and password management.

When necessary, the MAIS Advisory Committee will, in consultation with Data Stewards and with Deans, Directors, and Department Heads, forward a recommendation to the Executive Officers on data restrictions and requested access rights to institutional data.

2.10 Data Security and Disaster Recovery/Business Continuity

The data steward will be responsible for assigning security classifications and determining access for data validation and maintenance rules for institutional data.

All data users having access to restricted portions of the IDB data should formally acknowledge (by signed statement, either hard copy or electronic) their understanding of the level of access provided and their responsibility to maintain the confidentiality of the data they access.

The data steward is responsible for monitoring and reviewing security implementation and authorized access.

Information systems providing access to data must provide for the ability to uniquely identify a user or entity and authenticate them in some manner prior to permitting access.

Where practicable, information systems containing sensitive information should have user access and activity logging functionality in support of system activity and data access reviews.

All data regardless of where it is stored, will be afforded the same level of protection. Where practicable, sensitive data at rest and in transmission should be encrypted.

The data steward is ultimately responsible for defining and implementing policies and procedures to ensure that data are backed up and recoverable in response to events that compromise data availability or integrity. These policies and procedures shall include, but are not limited to, data backup, disaster recovery, and emergency mode operation procedures. These policies and procedures will also address testing of and revision to disaster recovery/business continuity procedures and a criticality analysis. MAIS or other University agencies may assist in this effort.

2.11 Data Documentation

Documentation of the data elements is the ultimate responsibility of the data steward. Some or all of these responsibilities may be assigned to data managers.

Data elements should follow the University Data Element Naming Standards. The standards consist of rules for defining, documenting, and naming data, including:

- Guidelines for defining data elements
- Major classifications of data
- Standard syntax for naming data
- Suggested formats for data
- Approved abbreviations
- Guidelines for using standards and enforcing their use

Ideally, documentation/definition for each data element should at least include:

- Name and Alias Names
- Description
- Data Steward
- Usage and Relationships
- Frequency of Update
- Source for Data Collection
- Source for Data Maintenance
- Official Data Storage Location and Format (data type and field size)
- Designation as "Sensitive", "Private/Confidential", or "Public"
- For "Sensitive" data elements: Description or specification of the restriction
- Description of Validation Criteria and/or Edit Checks
- Description, Meaning, and Location of Allowable Codes
- Access Rules and Security Requirements
- Archiving Requirements
- Data Storage Location of Extracts

Documentation for derived IDB data elements should 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 IDB (databases, files, groups of files) should also be provided which includes information about data structure and update cycles necessary for the accurate interpretation of the data.

(The following guidelines will be implemented concurrent with implementation of a University Data Resource Dictionary)

Documentation of data elements should be provided to MAIS/DA, preferably in machine readable-format, by the data stewards. This documentation will ultimately reside in a University Data Resource Dictionary.

MAIS/DA will be responsible for the data administration function of maintaining the University Data Resource Dictionary and for making it readily accessible to all interested parties.

Change in any data collection, maintenance and other definition characteristics should be noted to MAIS/DA and recorded in the University Data Resource Dictionary in advance of the change.

2.12 User Support

The data stewards and data managers responsible for each major segment of IDB will define, in consultation with the data users, the extent of support for data access and interpretation which is available to users of these data.

Data stewards will provide user support -- primarily through documentation of the information resource but also, as needed, in the form of training and consulting services -- to assist data users in the interpretation and use of data elements in the IDB. This responsibility may be delegated to the data managers or integration coordinators.

MAIS will provide training classes and consulting for a selected set of supported information access tools and technologies. Other University offices may assist in this effort.

The data users will be responsible for following the guidelines in their use and interpretation of the data which they access.

2.13 Data Availability and Integration

Data stewards are responsible for providing accessible, meaningful, and timely institutional data in machine-readable format for access by users for University use.

Data stewards and MAIS/DA share the responsibility for data compatibility, accessibility, and interfaces among institutional data elements residing in various segments of the IDB.

Data stewards and MAIS/DA will work together toward unification of the various data element coding structures and data storage formats which exist in various segments of the IDB.

Data users will be responsible for following the guidelines in their use and interpretation of the data resulting from their integration with other institutional and departmental data.

2.14 System Administration

Institutional data must be maintained within a single, logically-integrated information system.

Institutional data may be stored on any of many diverse computing hardware platforms (system environments), provided such platforms are fully integrated components of an overall UNIVERSITY INFORMATION SYSTEM.

If institutional data are stored on any component of the UNIVERSITY INFORMATION SYSTEM, that system component must have defined a formal system administration function and have assigned to it a system administrator whose responsibilities include: physical site security; administration of security and authorization

systems; backup, recovery, and system restart procedures; data archiving; capacity planning and performance monitoring.

3.0 PROCEDURES

These Data Administration Guidelines for Institutional Data Resources were reviewed and approved by the Information Technology Policy Committee and the Executive Computing Committee in 1996. They serve as a statement of objectives to manage the administration information resource and apply to all Institutional Database data. These Guidelines should apply to all who capture data and manage administrative information systems using information from the University. Standards and procedures should be developed to conform to the objectives embodied in these Guidelines.

3.1 Updates

As an ongoing document, the Data Administration Guidelines for Institutional Data Resources will be maintained and revised as needed by the data stewards, data managers, MAIS Advisory Committee, and the Executive Officers with support from MAIS/DA. All institutional data users are encouraged to correspond with MAIS/DA describing any suggestions for improving these Guidelines.

4.0 DEFINITIONS

ACCESS TO INSTITUTIONAL DATA refers to the permission to view, query, or capture data contained in the institutional database, but does not necessarily imply delivery or support of specific methods or technologies of information access.

DELEGATED DATA STEWARDS are senior University officials having policy-level responsibility for managing a segment of the University's information resource who have been designated to serve as the data stewards for that segment of the information resource.

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

DATA MANAGERS are University officials and their staff who have operational-level responsibility for data capture, data maintenance, and data dissemination.

DATA OWNER of all the University institutional data is the University itself; individual units or departments may have stewardship responsibilities for portions of the data.

DATA STEWARDS are the University Executive Officers having policy-level responsibility for managing a segment of the University's information resource as designated by the Regental by-laws.

DATA USERS are individuals who access the University data in performance of their assigned duties or in fulfillment of 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.

ELECTRONIC PROTECTED HEALTH INFORMATION (ePHI) is a type of sensitive data and refers to information that is created, received, maintained or transmitted electronically that was created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse; and it relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual.

INSTITUTIONAL DATABASE is a conceptual term used to identify that body of data that is defined as the University's institutional data. INSTITUTIONAL DATA MODEL is a logical construct which describes the data entities which comprise the Institutional Database and the relationship among those entities.

INTEGRATION COORDINATORS are University officials and their staff who are responsible for facilitating and resolving shared data management issues among central offices and schools and colleges.

PRIVATE/CONFIDENTIAL DATA refers to those elements of the Institutional Database which may be accessed by employees of the University only if necessary in the performance of their official University duties.

PUBLIC DATA refers to those elements of the Institutional Database which are available to the general public, including to those outside the University community.

SENSITIVE DATA refers to those elements of the Institutional Database which, because of legal, ethical, or other externally-imposed constraints, may not be accessed without specific authorization or to which only limited access may be granted.

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

UNIVERSITY DATA RESOURCE DICTIONARY is a database system that functions as a repository that contains comprehensive information about the University's institutional data and documentation of University information administrative systems (no such dictionary exists today).

UNIVERSITY INFORMATION SYSTEM is a conceptual term used to identify the collection of computer hardware, software, and network connections which together form the single, integrated system on which the Institutional Database resides.

WORKFORCE MEMBER refers to any faculty, staff, student, volunteer, trainee, or other person whose conduct is under the University's direct control, whether or not the University pays them for their services.

REFERENCES

Standard Practice Guide 201.46 - "Personnel Records - Collection, Retention and Release"

Standard Practice Guide 601.7 - "Proper Use of Information Resources, Information Technology, and Networks at the University of Michigan"

Standard Practice Guide 602.5 - "Development; Alumni Records"

Standard Practice Guide 602.10 - "Development; Development Systems Services"

University of Michigan Hospitals Administrative Policy and Procedures: Confidentiality of Patient Care Information

Bylaws of Board of Regents

University Student Rights and Records Policy (<http://www.umich.edu/~regoff/ferpa/studentrights.html>)

Standard Practice Guide 601.12 - "Institutional Data Resources Management Policy"

Known Institutional Data Subject Areas and Recommended Data Stewards/Data Managers
(<http://www.mais.umich.edu/access/policies.html>)

Part 164 subpart C – Security Standards for the Protection of Electronic Protected Health Information of the Health Insurance Portability and Accountability Act
(<http://www.cms.hhs.gov/hipaa/hipaa2/regulations/security/default.asp>)