

The Necessity of a Formal Data Governance System

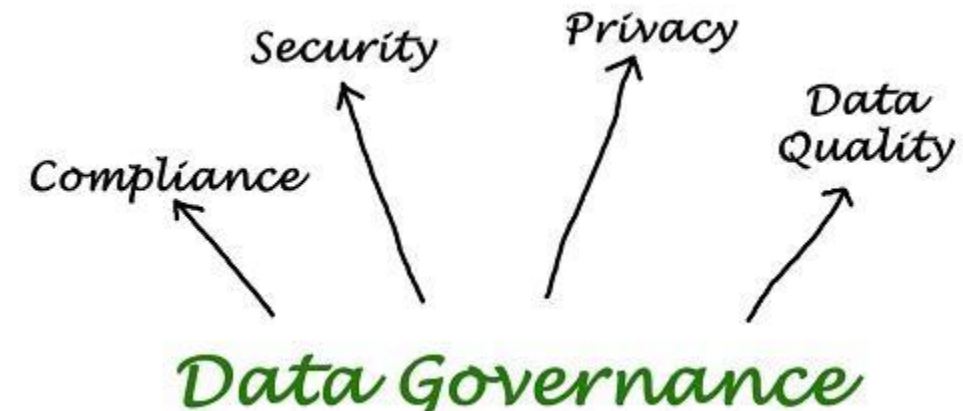


**Dr. Lawrence Redlinger and Ms. Sharon Etheredge
Office of Strategic Planning and Analysis
The University of Texas at Dallas**

Presented at the Rocky Mountain Association for Institutional Research,
October-30-2019

Overview

- Brief History
- What is Data Governance
- Why we should worry about it
- Characteristics of data governance systems: what's "best" for you
- Who needs to be engaged and involved
- Where are we now—assessment
- Changing current practices
- Technological assistance



Evolution of Data Governance: Application Era (1960-1990)

Organizations begin adoption of data processing technology;

Systems built to support transactional business processes;

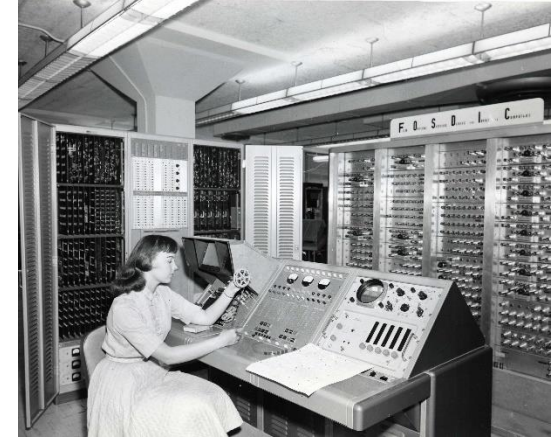
Data was a byproduct of running the business and viewed as having little value beyond the transaction and the application that processed it.

Data governance seen as an IT function focused on data cataloging that was rarely, if ever, seen or used by the business.

Nascent governance attempted by a few organizations through data modeling. Their success was limited for two reasons:

- 1) these efforts were driven by IT and without the broad organizational support and authority to enforce compliance;

- 2) the rigidity of packaged applications further reduced their effectiveness. So the idea of data governance through enterprise modeling was mostly an academic exercise.



Evolution of Data Governance: Enterprise Repository Era (1990-2010)



- Organizations begin to realize that the value of data extends beyond transactions. Organizations realized they could extract value from all of the various data sets they were creating. Decision making increasingly relies on data analysis. Business processes consume increasingly large amounts of data created in distant parts of the organization for a different purpose. This led to a trend of thinking about data for broad use cases beyond the localized context of a transaction.
- Organizations attacked this problem by building large scale repositories, such as data warehouses, that take an enterprise perspective. ERP and ERP consolidation — the notion of having a single, integrated set of plumbing run the business — is driven by the same philosophy. More recently, organizations recognized that not all data is of equal value, and it is more cost-effective pay extra attention to the data that describe core business entities that are widely referenced. This led to the current build out of master data repositories.
- While enterprise repositories yield a lot of benefits, they also have the deserved reputation of being very expensive and risky undertakings. Creating a view of data that supports multiple use cases invariably results in conflicts, and ultimately it's up to the business (the consumers of data) to resolve these conflicts. As a result, data governance gradually but surely came to be recognized as critical to success.
- However, because data related activities are generally carried out on a system-by-system basis, governance is typically siloed around individual enterprise repositories: data governance for a data warehouse or an ERP system, for example, or data governance for master data management. Also, governance is informal, lacking a distinct organizational structure and clearly defined and executed processes.

Evolution of Data Governance: Policy Era (2010-2019)



- In the early days, data governance was strictly considered an IT function. No Longer.
- Systems and data repositories proliferated. Data complexity and volume continue to explode; organizations grow more sophisticated in their use of data, which drives new demand that require different ways to combine, manipulate, store, and present information.
- Organizations form business-led governance organizations to care for data for the enterprise, and create collaborative processes to manage a core set of data deemed critical for the business. Significantly, Organizations took a policy-centric approach to data models, data quality standards, data security and lifecycle management. Rather than envisioning ever-larger and more encompassing repositories, they put processes in place for defining, implementing and enforcing policies for data.
- Emerging from this is a shift in mindset: Organizations take increasing responsibility for data content, and data is widely recognized as one of the most valuable corporate assets throughout the organization. For IT, a policy-centric approach is liberating. It affords more flexibility in designing systems to serve business needs without giving up consistency and control.
- 2018 a “global reckoning on data governance”—massive data breaches
- 2019 data governance is no longer optional. Data security, compliance, privacy, regulatory compliance and risk mitigation are key drivers.

<https://magnitude.com/blog/a-brief-history-of-data-governance/>

<https://www.forbes.com/sites/forbestechcouncil/2019/03/11/the-evolution-of-data-governance/#38d804d05ef7>

<https://blogs.wsj.com/cio/2018/03/19/%E2%80%8Bthe-morning-download-facebook-at-center-of-global-reckoning-on-data-governance/>

“Why are we here?”

DG issues do not arise in a vacuum: Precipitating Factors

- A campus leader went to a session at a conference and has come back with a goal to bring "data governance" to the organization.
- Increase the Quality of the Data
- Create efficiencies in processes (manage data better) (54%)*
- Environmental demands: reporting, compliance and regulations (HIPPA, FERPA, GDPR*). (39%)*
- Data Breaches.
- Organizational isomorphism.
- Adoption of common technical systems.
- Governance siloed around individual enterprise repositories.
- Governance is informal, lacking a distinct organizational structure and clearly defined and executed processes.
- Enhance customer services. (7%)*

<https://ciowatercooler.co.uk/the-data-governance-report-2017-your-copy/>

*percentages in red come from the results in the above source

*The General Data Protection Regulation ([GDPR](#)) is a legal framework that sets guidelines for the collection and processing of personal information from individuals who live in the [European Union \(EU\)](#). Since the Regulation applies regardless of where websites are based, it must be heeded by all sites that attract European visitors, even if they don't specifically market goods or services to EU residents.

What is Data Governance?

Data governance is a collection of processes, roles, policies, standards, and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals.

Objectives:

Processes that ensure the quality and security of the data;

Responsibilities that ensure the quality and security of the data;

Clearly define roles related to the data

Accountability

Who can take what action, upon what data, in what situations, using what methods.

1. Formalizing how data assets are defined, produced, stored, protected, used and destroyed.
2. Establishing standards promoting data integrity and the integration of information systems.

Important Characteristics of DG Definitions

Data Governance Is:

- More about people and behavior than data
- A system that requires and promotes shared agreement
- Formal (i.e. written down)
- Adds value by supporting institutional mission/goals

Data Governance Is Not:

- IT's responsibility
- Solved by technology
- Equally applied across all data assets

Key Features of Data Governance Systems

Documents

Charter / framework

- Principles & values
- Purpose & scope
- Roles & responsibilities

Written & published policies

- Data dictionaries
- Communication strategies

Groups

- Senior leadership [buy in]
- Policy council
- Data steward council(s)
- Information security council/program
- Positions/office to support DG

Individual Roles

- Data stewards
- Data custodians/caretakers
- Data users

Justifications for Data Governance: Cost

A third of Fortune 100 organizations will experience “an information crisis, due to their inability to effectively value, govern and trust their enterprise information.”

Gartner. (2014). “Why data governance matters to your online business,” retrieved August 1, 2016 from <http://www.gartner.com/newsroom/id/1898914s-why-data-governance-matters-to-your-online-business/>

Poor data quality costs the US economy \$3.1 trillion every year

IBM. (n.d.). “Extracting business value from the 4 V's of big data,” retrieved October 1, 2018 from <https://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data>

The average financial impact of poor data quality on businesses is \$9.7 million per year. Opportunity costs, loss of reputation and low confidence in data may push these costs higher.

Forbes (2017). “Poor-quality data imposes costs and risks on businesses,” retrieved October 22, 2018 from <https://www.forbes.com/sites/forbespr/2017/05/31/poor-quality-data-imposes-costs-and-risks-on-businesses-says-new-forbes-insights-report>

Justifications for Data Governance: Risk

Fines Imposed by Federal Student Aid

Fiscal Year	Clery/Part 86 Imposed Fines	IPEDS Imposed Fines	Other Imposed Fines	Total Imposed Fines
2010	\$42,000	\$225,000	\$48,653,500	\$48,920,500
2011	\$195,000	\$144,500	\$4,868,500	\$5,208,000
2012	\$212,500	\$158,500	\$624,000	\$995,000
2013	\$812,000	\$56,000	\$5,204,137	\$6,072,137
2014	\$438,000	\$111,250	\$6,750	\$556,000
2015	\$500,000	\$39,250	\$14,130,000	\$14,669,250
2016	\$307,500	\$57,000	\$79,462,500	\$79,827,000
2017	\$2,542,500	\$1,500	\$382,500	\$2,926,500

Source: Postsecondary Education Participants System (PEPS)

Data as an Asset

By 2020, Gartner predicts that 10% of organizations will have a highly profitable business unit specifically for productizing and commercializing their information assets.

By 2021 companies will be valued on their information portfolios: “Those in the business of valuing corporate investments, including equity analysts, will be compelled to consider a company’s wealth of information in properly valuing the company itself.”

Examples:

- Web sites grant access in exchange for personal data (email address, etc.) –these data have value and can be leveraged or even sold
- Data are purchased (Names of prospective students, Library databases, Various datasets e.g. U.S. News, Academic Analytics)
- Data are sold to vendors for discounts or services
- Lost data carry costs - data breaches

Data Governance Maturity Model	Level 1 Informal	Level 2 Developing	Level 3 Adopted and Implemented	Level 4 Managed and Repeatable	Level 5 Integrated and Optimized
Data Governance	Attention to Data Governance is informal and incomplete. There is no formal governance process.	Data Governance Program is forming with a framework for purpose, principles, structures and roles.	Data Governance structures, roles and processes are implemented and fully operational.	Data Governance structures, roles and processes are managed and empowered to resolve data issues.	Data Governance Program functions with proven effectiveness.
Culture	Limited awareness about the value of dependable data.	General awareness of the data issues and needs for business decisions.	There is active participation and acceptance of the principles, structures and roles required to implement a formal Data Governance Program.	Data is viewed as a critical, shared asset. There is widespread support, participation and endorsement of the Data Governance Program.	Data governance structures and participants are integral to the organization and critical across all functions.
Data Quality	Limited awareness that data quality problems affect decision making. Data clean up is ad hoc.	General awareness of data quality importance. Data quality procedures are being developed.	Data issues are captured proactively through standard data validation methods. Data assets are identified and valued.	Expectations for data quality are actively monitored and remediation is automated.	Data quality efforts are regular, coordinated and audited. Data are validated prior to entry into the source system wherever possible.
Communication	Information regarding data is limited through informal documentation or verbal means.	Written policies, procedures data standards and data dictionaries may exist but communication and knowledge of it is limited.	Data standards and policies are communicated through written policies, procedures and data dictionaries.	Data standards and policies are completely documented widely communicated and enforced.	All employees are trained and knowledgeable about data policies and standards and where to find this information
Roles and Responsibilities	Roles and responsibilities for data management are informal and loosely defined.	Roles and responsibilities for data management are forming. Focus is on areas where data issues are apparent.	Roles and responsibilities are well defined and a chain of command exists for questions regarding data and processes.	Expectations of data ownership and valuation of data are clearly defined.	Roles, responsibilities for data governance are well established and the lines of accountability are clearly understood.

Principals of Data Governance (generic)

Consistency of data in its sourcing and in its vocabulary, definitions, and taxonomies

Quality which is proactively assessed and standards applied

Ownership and accountability defined across the data lifecycle and recorded in the information asset register

Business alignment which ensures that data is regarded and treated as a key business asset

Access to relevant users, kept secure through access control

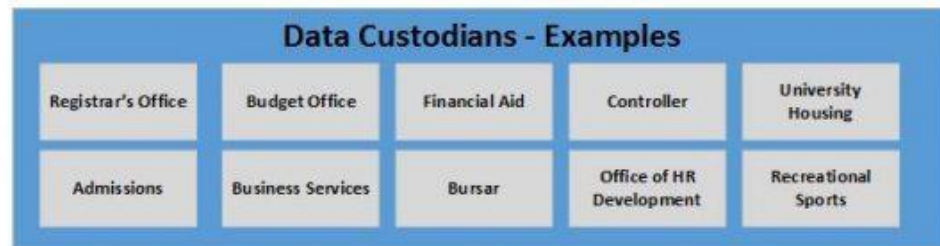
Providing trusted insight

Principles and Values

Example: University of Wisconsin-Madison

Accountability	Determining who is responsible for the management of data at UW Madison as well as holding them to our outlined standards.
Agility	All of our processes should adapt when necessary
Change	
Management	New processes demand new and changing staff at UW. We're committed to ensuring smooth transitions and well informed decisions.
Consistency	All decisions made will be applied consistently across campus.
Metrics Driven	We monitor ourselves against measurable goals on a regular basis and use the results to determine courses of action.
Stewardship	Determine formal roles for those in charge of data. This does not mean that everyone on campus is not responsible despite formal roles.
Transparency	We will make it clear how and when decisions are made and when processes are created. We also strive to ensure that decisions and processes are audited to support compliance based requirements.

Structure – University of Wisconsin-Madison



Data Governance Steering Committee

- provides executive level guidance to the program
- promotes Data Governance across UW-Madison
- allows for / facilitates data-driven decision making
- determines priority and budget of major data-related projects.

Data Stewardship Council

- determines operational structure of the program
- drafts, communicates, and recommends approval of data-related policies
- implements, budgets, and monitors data-related programs across UW-Madison.

Principles and Values

Example: Stony Brook University

Values			
Shared Assets Data and information are shared organizational resources that constitute valuable assets.	Stewardship Employees of Stony Brook University have a responsibility for the curation of data. They serve as caretakers of data to ensure data are collected, stored, and maintained under the premise that others will access and use them over time	Quality To ensure data retain value, quality of data is actively monitored and maintained	Privacy & Confidentiality Maintenance of individual privacy and confidentiality of educational and personal records represent not only legal requirements but also primary outcomes of data management.
Principles for Data Governance			
Organizational Effectiveness	Transparency	Communication	Compliance
Auditability	Integrity	Accountability	Standards

Structure – Stony Brook University



VP Council

VP Council (Project 50 Forward Steer Co)

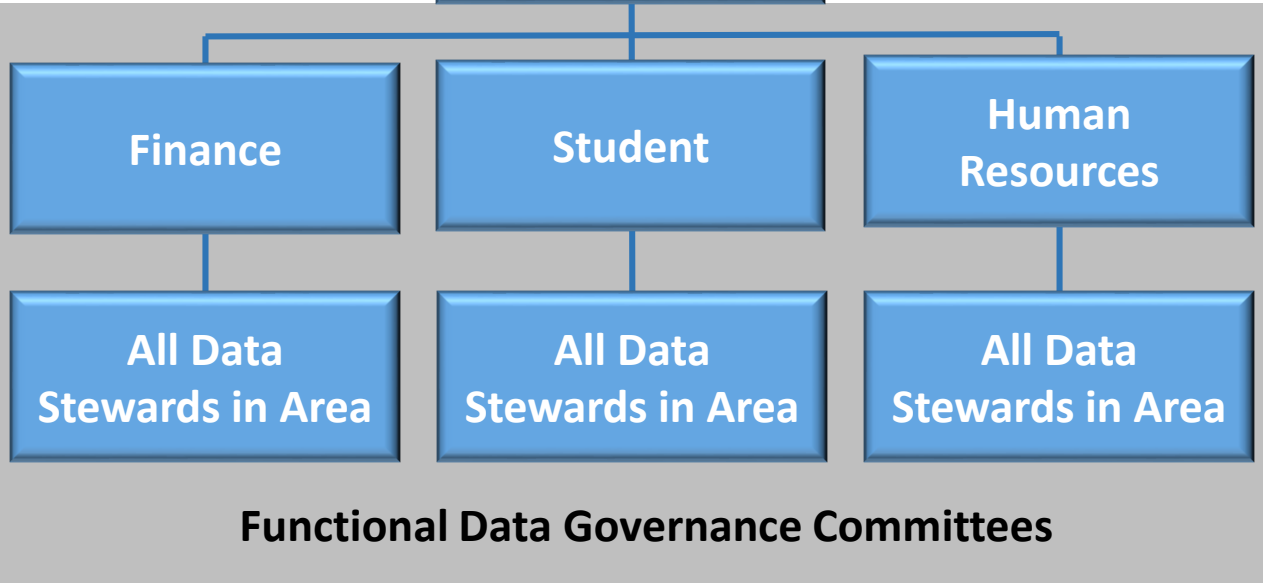
- Executive sponsors of project
- Establishes authority and purview of data governance system



Data Governance Council

Data Governance Council

- Recommends and implements institutional policy for data governance
- Sets priority for



Finance

Student

Human Resources

All Data Stewards in Area

All Data Stewards in Area

All Data Stewards in Area

Data Governance Council Membership Examples

UW-Madison	Stony Brook
Chief Data Officer	Chief Institutional Research Officer
Director of Univ. Communications	Analytics and Enterprise Data Officer
VP for Teaching & Learning	University Controller
VP for Diversity	Chief Enrollment Management Officer
AVC Business Services	University Registrar
AVC Legal Affairs	Chief Financial Aid Officer
Assoc. Dean Biomedical Informatics	Provost's Office designee
VP Libraries	VP Student Affairs designee
CISO	VP Administration designee
Campus Records Officer	VP Human Resources designee
Assoc. Dean Education	VP Information Technology designee
Faculty/Dean Representation	VP Research designee
	SVP Health Sciences Designee
	University Senate designee
	Chairs & Vice Chairs of FDGCs (6 people)

Montana State University

Data Governance Council Members

Position	Represents
Chair	UIT
Member	UIT – Enterprise Systems Group
Member	Planning and Analysis
Member	Human Resources
Member	Institutional Research, Analysis and Programming
Member	Student Success
Member	Registrar
Member	Information Technology
Member	Legal Council
Member	UIT-Security
Member	Academic Affairs
Member	Finance
Non-Voting	Alumni Foundation
Non-Voting	Information Technology
Non-Voting	Information Technology

Policy-Making Body - Data Governance Council

Prioritizes decisions regarding data to address most relevant needs of organization

Reviews, evaluates, and reports on data governance performance and effectiveness

Ensures that annual performance measures align with data governance and business objectives

Reviews and approves data governance policies and goals

Ultimately is accountable for business data use, data quality, and prioritization of issues

Makes strategic and tactical decisions

Defines data strategy based on business strategy and requirements

Who “Owns” the Data?

- Institutions own the data. Institutional data resources, by definition, practice, and intent, are a University asset.
- Consider carefully use of the word “ownership” with data.
 - “Owners” are typically the responsible managers that collect or are the primary users of a data asset. They are really (see below) stewards of university resources.
 - “Ownership” connotes individual control and property vs. “Stewardship” or the caretaking of shared resource.
- Individuals or Departments provide stewardship. Data Stewards are the University Officers having explicit policy-level responsibility for managing a segment of the University’s information resources.

Generic Data Stewardship Definitions

- Data Stewardship is concerned with taking care of data assets that do not belong to the stewards themselves. Data Stewards represent the concerns of others. Some may represent the needs of the entire organization. Others may be tasked with representing a smaller constituency: a business unit, department, or even a set of data themselves. –Data Governance Institute (n.d.)
- Data stewardship is the most common label to describe accountability and responsibility for data and processes that ensure effective control and use of data assets. –Knight (2017)
- Data stewardship is the operational aspect of an overall Data Governance program, where the actual day-to-day work of governing the enterprise's data gets done. –Plotkin(2014)

Types of Data Stewards

Business Data Steward

- Accountable for data owned by business area
- Work with stakeholders to make recommendations on data issues
- Manage metadata for their data
- Champion data stewardship for their areas

Technical Data Steward

- Provide expertise on applications, ETL, data stores, and other links in information chain
- Assigned by IT leadership to support data governance

Domain Data Steward

- Business steward for widely shared data
- Work with business stewards as stakeholders to achieve consensus

Project Data Steward

- Represent data stewardship on projects
- Funded by projects
- Work with business data stewards to obtain info and make recommendations about data stewarded by business stewards
- Notify business data stewards about data issues raised by the project

Operational Data Steward

- Provide support to business data stewards
- Recommend changes to improve data quality
- Help enforce business rules for the data they use

Data Stewardship University of Michigan



Data stewards are responsible for supporting the creation, curation, and protecting the confidentiality, integrity, and availability of university data. U-M policies, federal and state laws and regulations require that some types of data be handled differently than other types.

...

Specifically, it is the role of data stewards at U-M to protect the confidentiality, integrity, and availability of university data. Effective data stewardship is the process for maximizing the value of data as an institutional resource.

5 Levels of Stewardship

[Data Steward](#) - Executive officers with policy-level responsibility for managing a major area of the university's information resources.

[Delegated Data Steward](#) - Senior university officials with policy-level responsibility that have been designated by a data steward to serve as the delegated authority for a specific data area. The responsibilities of delegated data stewards are the same as those for data stewards.

[Data Manager](#) - University officials and their staff that have operational-level responsibility for the capture, maintenance, and dissemination of data for specific data areas.

[Data Management Integration Coordinator](#) - University staff who are responsible for facilitating and resolving shared data management issues among central offices, schools and colleges, and Michigan Medicine.

[Data User](#) - University departments, individual university community members, or university affiliates that have been granted access to institutional data in order to conduct university business.

Data Steward Responsibilities

- Oversee management of selected data assets
- Participate in data governance and carry out decisions
- Assist in creation and maintenance of data dictionaries, metadata
- Document rules, standards, procedures, and changes
- Ensure data quality and manage specific issues
- Communicate appropriate use and changes
- Manage access and security

Data Stewardship Council/Committee(s) (Functional)

Coordinate data stewards in related area

Set and review definitions, data quality rules, creation/usage rules, determines official version of metadata

Consider and approve changes & additions to code sets

Ensure dictionary standards are followed in area

Review data quality in functional area; identify practices promoting data quality identify areas for improvement and monitor improvements

Respond to inquiries about process, content, limitations and uses of data, especially in cross functional settings

Elevate issues that require resolution

Communicate proceedings, including notice of changes and decisions

Data Users

Recognize that institutional data and information derived from it are potentially complex. Make efforts to understand the source, meaning and proper use of the data through training sessions, utilizing data dictionaries and knowledge of supporting system processes.

Include information about the data source and criteria when distributing data, reports and ad hoc analytics to guard against misinterpretations of data.

Respect the privacy of individuals whose records they may access. Unauthorized disclosure or misuse of institutional information stored on any device is prohibited

Ensure that passwords or other security mechanisms are used for sensitive data

Report data quality issues to appropriate data steward

Administrative Office / Positions Supporting Data Governance

In general, offices and positions dedicated to supporting data governance are still emerging in higher education

Chief Data Officer	Data Governance Program Manager
<ul style="list-style-type: none">• Purdue University• University of Florida System• University of South Carolina Columbia• University of Rochester• University of Wisconsin Madison	<ul style="list-style-type: none">• Purdue University• Stanford University

Maturity Models: Assessing Current State of Data Governance

To implement data governance, if not to sell it to senior leadership, assessment of the current state is important.

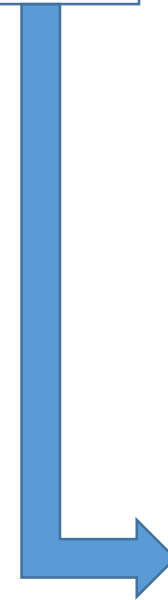
Uses a maturity model to quantify the existing state; allows for measurement of progress in a future state

Baseline

Dimensions



Current
2015



Maturity



Target
2017



How Long Will It Take to Create a Data Governance Policy?

- Organizational dynamics and readiness for data governance
- How process oriented is the organization to be impacted
- How broad reaching is the policy
- How are you defining a policy: as a broad generalized set of principles or with a level of specificity including quality standards and validations or somewhere in between
- How many people are involved in the creation of the policy (across how many aspects of the company)
- Are the right people involved in making the decisions regarding the policy content
- Is the policy being created to make a current practice more formally established or is it to cause sweeping business process changes
- What is the perceived criticality and urgency of the policy
- Is the policy being created in response to a legal mandate with specific parameters already defined or are you starting with a blank slate
- Can the benefit of policy compliance be easily conveyed to those impacted

Assemble Your Group

Data governance requires support of senior leadership and functional leadership

Identify

- **Senior leaders who will sponsor**
- **Functional leaders and their potential for collaboration(includes available bandwidth, interest, capability, willingness)**

Charge for Data Governance (First 9 Months)

Examine existing governance structures

- active and inactive groups and lines of responsibility
- existing processes, practices and procedures that significantly impact data management and stakeholders.

Identify and articulate

- Roles of cross functional groups
- Functional roles in business units (e.g. data owner, data custodian, report owner) will also be identified and articulated.

Draft formal governance structure for university data management

- Principles, mission, and goals
- Post on a website to codify roles and responsibilities.

Formalize a process for prioritization

Charge for Data Quality Improvement

Examine existing practices for ensuring data quality within the Data Warehouse and may extend to quality assurance practices in functional systems.

Articulate and publish practices for developing, maintaining, and communicating

- data definitions (such as robust data dictionaries)
- transparent source information
- update schedules
- error check practices and
- clean-up procedures

Develop and Implement a Communication Plan

Develop a communication plan for

- **How new capabilities for business intelligence go beyond initial reporting functionality.**
- **Availability**
- **Use limitations, and opportunities**
- **including needs identification for documentation, training, workshops, etc.**

Develop, document, and adopt reporting standards

Gather Input for Planning Process

Answer these Questions:

1. What do we want data governance to accomplish?
2. What are the current data governance mechanisms?
3. What is missing (gap analysis)?
4. Who should be involved?

Anticipate responses that may be counterproductive

- e.g. “IT should control data governance”
- “We don’t have a problem”

Technology Applications for Data Governance

Technology can support data governance

- Data dictionary management
- Data quality analysis
- Master data management
- Issue and process management

Technology will not

- Build organizational structures, responsibilities, accountabilities
- Mend dysfunctional organizations
- Implement organizational or cultural change

Issues to Consider When Selecting Technology

Initial cost and annual cost

Alignment with DG needs

- **Metadata management**
- **Integration w/ reporting tools**
- **Data quality**
- **Security/user roles**

User community

Ease of implementation and impact on IT

Takeaways

Data governance is more about people and policies than data

All higher education change management principals apply

Establishing processes and written documents are essential, along with:

- **Leadership support**
- **Broad-based consultation, including faculty**
- **Opportunity for consultation**
- **Representation**

Software can help, but it won't fix broken processes or organizations

Starting data governance is hard work; sustaining it is harder

Resources

https://www.infogix.com/wp-content/uploads/2018/09/Infogix-Data-Governance-101-web_EB_ESC_08_0901.pdf

<https://library.educause.edu/topics/information-systems-and-services/data-governance>

<https://er.educause.edu/articles/2019/8/10-best-practices-to-ensure-data-governance-efforts-will-fail-plus-10-best-practices-for-success>

<https://ciowatercooler.co.uk/resources/DataGovernanceSurvey2017.pdf>

<https://www.ibm.com/analytics/data-governance>

<https://www.youtube.com/watch?v=sHPY8zIhy60>

<https://blog.iandisoft.com/data-governance-roadmap-success/>

<https://blog.iandisoft.com/six-steps-data-governance-success/>

https://spg.umich.edu/sites/default/files/policies/601x12_2.pdf

<https://www.sciencedirect.com/book/9780124103894/data-stewardship>

<https://ualr.edu/chiefdataofficers/>

<https://www.educause.edu/about/mission-and-organization/governance-and-leadership/member-committees/higher-education-chief-data-officers>

<https://www.u-planner.com/en-us/blog/understanding-why-universities-need-a-chief-data-officer>

<http://business.gmu.edu/executive-development/programs-for-individuals/chief-data-officer-program/>

<https://magnitude.com/blog/a-brief-history-of-data-governance/>

https://magnitude.com/long-create-data-governance-policy/?_ga=2.230318618.706745463.1571258592-1539269096.1571258592

<https://www.forbes.com/sites/forbestechcouncil/2019/03/11/the-evolution-of-data-governance/#38d804d05ef7>

<https://www.forbes.com/sites/ciocentral/2017/11/16/he-who-rules-the-data-rules-the-world-a-brief-history-of-data-governance/#2d8096d739b5>

<https://it.toolbox.com/guest-article/the-evolution-of-data-governance>

<https://blogs.wsj.com/cio/2018/03/19/%E2%80%8Bthe-morning-download-facebook-at-center-of-global-reckoning-on-data-governance/>



A copy of this presentation can be found at <https://utdallas.edu/ospa/institutional-research/>

Appendix

More Data Governance Definitions

- Data Governance is a collection of practices and processes which help to ensure the formal management of data assets within an organization. In the university context, effective data governance will:
 - Ensure establishment, maintenance, and delivery of secure, confidential, ethical, trustworthy, stable, reliable, and accessible collections of institutional data for shared access by the university community
 - Maximize the value received from the data asset by increasing the understanding and use of the data
 - Provide an integrated view of the functions of the university
 - Improve direct access to data by end-users in accordance with institutional policies, ethical and privacy norms, and state and federal privacy and security laws and regulations
 - Support the university's strategy to incorporate information technology as an integral part of decision-making, competitive positioning, and delivery of services
 - Establish decision rights with respect to university data that ensure accountability.
 - A successful data governance program helps to support the university's primary missions of teaching and learning, research, clinical care, and public service. –**University of Michigan** (2019), (<https://it.umich.edu/governance/data-governance>).
- “a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods” –Data Governance Institute (2014)
- “the execution and enforcement of authority over the management of data and data-related assets” -R. Seiner (2014)
- “specification of decision rights and an accountability framework to ensure appropriate behavior in the valuation, creation, storage, use, archiving and deletion of information” -Gartner IT Glossary
- “formalizes behavior around how data are defined, produced, used, stored, and destroyed in order to enable and enhance organizational effectiveness” –**Stony Brook University** (2016)
- “adds value to our administrative and academic data systems by the establishment of standards that that promote data integrity and enables strategic integrations of information systems” –**Vanderbilt University**
- “the discipline which provides all data management practices with the necessary structure, strategy, and support needed to ensure that data are managed and used as a critical University asset” –**U of Rochester**

“The Compelling Case for Data Governance”

Educause

What data are available?

How sensitive are they?

Who is responsible for the data?

Where did they come from?

What do they mean?

Are they trustworthy?

Where are they stored?

Who has access to that data?

What are the risks associated with the data?

What regulations apply?

How do we know everything is under control?

What are best practices that

- 1) prevent the misuse of institutional data assets by unauthorized parties but also
- 2) encourage more effective use of these same data assets by the institution itself?

Master Data Management is often confused with Data Governance

Master data management (MDM) is the process of making sure an organization is always working with, and making decisions based on, one version of current, 'true' data—often referred to as a "golden record."

Master Data Management (MDM)

Comprehensive method to link all critical data to a common point of reference

Example: All screens, documents and systems showing a student's address derive from a common location.

Data Governance (DG)

Formalized system for how people make decisions about acquisition, production, storage, distribution and analysis of data.

Example: Group decides on a definition for home address and agrees on a common source field.

Stony Brook Roles and Responsibilities Matrix

	Data Governance Council	Functional Data Governance Committees	Data Stewards
Standards and Policies	Define, Establish, Monitors, Audit, Verify, Develop, Revise	Cross functional implementation, coordination	Functional implementation
Data Quality	Identify, Adopt enterprise wide DQ tool Big picture	Prioritize levels Monitor area Identify needs	Review audit reports, Coordinate clean up Initial prioritization
Metadata	Establish standards	Ensure cross functional alignment	Implement Maintain
Metrics	Review, Identify, Monitor	Monitor area Identify are priorities	Monitor Remediate

Stony Brook Data Governance Maturity Model

Initial Results –Spring 2016

Stony Brook Data Governance Maturity Model 2.0 Results Summary

BY FUNCTIONAL DATA TYPE

OVERALL (Area Averages)	DATA			ROLES &	
	GOVERNANCE	CULTURE	DATA QUALITY	COMMUNICATION	RESPONSIBILITIES
Level 1-Informal	11%	18%	8%	14%	10%
Level 2-Developing	41%	30%	34%	39%	32%
Level 3-Adopted and Implem	11%	14%	9%	14%	3%
Level 4-Managed and Repea	3%	3%	6%	2%	7%
Level 5-Integrated and Optir	0%	0%	2%	0%	0%
Not enough information	34%	35%	41%	32%	48%

Data Governance Maturity Self Test

Capability #1: Organizing Your Data Governance Initiative

Grading Key

Please indicate **how well** each statement describes your institution.

1 2 3 4 5
Not at all Poorly Somewhat Fairly well Well

Institutional Strategy

The institution has a formal data plan which articulates the steps to be taken to better leverage data and which names those responsible for executing those steps.

Score: _____

Campus leaders recognize data governance as an enterprise-level capability, requiring shared governance across the institution.

Score: _____

Subtotal:

Data-Driven Culture

Campus members recognize that data is an institutional asset and as such is not owned by individuals or departments, but is owned by the institution.

Score: _____

Leaders promote a culture of data informed decision-making, pressing campus to base resource decisions and choices on fact, not intuition.

Score: _____

Subtotal:

Data Committee Structure

We distinguish between strategic and operational data issues and match the seniority and expertise of people to the data governance activities we task them with.

Score: _____

The roles and responsibilities of each committee and its members are clearly defined, recorded, and revisited regularly.

Score: _____

Subtotal:

Organizational Continuity

A designated individual(s) oversees data governance and management, working in tandem with unit-level data stewards.

Score: _____

We have an established and clearly documented process for onboarding new members of the data governance group at the leadership and implementation levels.

Score: _____

Subtotal:

Implementation Framework

Data stewards are appointed and known within all operational units, and have clear policies outlining their roles and responsibilities.

Score: _____

Our data domains are clearly mapped, providing full coverage of enterprise information across different functional areas and silos.

Score: _____

Subtotal:

Copy These Scores to Page 3

Data Governance Maturity Self Test

Capability #2: Operationalizing Your Data Governance Initiative

Grading Key

Please indicate **how well** each statement describes your institution.

1

Not at all

2

Poorly

3

Somewhat

4

Fairly well

5

Well

Definition Strategy

Our institution defines data terms in clusters, with a strategic focus, and in pursuit of higher objectives for the institution.

Score: _____

We have a data definition decisioning framework that includes the right people at the right time, enables opt outs, and provides an agile mechanism for creating shared data definitions.

Score: _____

Subtotal:

Common Data Dictionary

Institutional data-of-record is clearly identified, and definitions are readily accessible from a common or centralized location.

Score: _____

Our data dictionary is accessible and understandable by all campus stakeholders, including all metadata associated with represented terms.

Score: _____

Subtotal:

Data Access Controls

Our institution has a standardized method for segmenting confidential data from public data.

Score: _____

Our institution grants data access on a principled, (semi-)automated basis.

Score: _____

Subtotal:

Data Quality Assurance

We actively monitor for anomalous data in enterprise systems and take steps to identify and remediate the underlying causes.

Score: _____

We hold units accountable for data quality by designating unit-level stewards and monitoring compliance with university-wide standards for data cleanliness.

Score: _____

Subtotal:

Data Consumption

Analytical resources are structured so that campus members are guided to the most useful resources associated with their roles and responsibilities.

Score: _____

End users' needs and usage patterns are monitored and incorporated in planning to determine future investments.

Score: _____

Subtotal:

Copy These Scores to Page 3

Data Governance Maturity Self Test

Copy Numbers From Pages 1-2 Below and Add Up to Discover Your Data Governance Maturity Score

Page 1	Page 2	Score Range	Icon	Maturity Level
Institutional Strategy <input type="checkbox"/>	Definition Strategy <input type="checkbox"/>	20-30		Data Denier Your institution is sitting pretty—pretty behind the times. As the world of higher ed marches steadfastly into the data revolution, you're still juggling multiple spreadsheets and arguing about the number of students enrolled for the coming semester. Time for the journey to begin.
Organizational Culture <input type="checkbox"/>	Common Data Dictionary <input type="checkbox"/>	31-50		Barely Begun You're starting to make headway. Maybe there's an evangelist, and people are starting to take notice. But there's no coherent plan, and no-one is taking steps toward an enterprise-wide data-asset revolution. You need more guidance and guardrails, but at least you're on your way.
Committee Structure <input type="checkbox"/>	Data Access Controls <input type="checkbox"/>	51-70		Gathering Steam You've got your committees and stewards; you've got a plan. Someone's leading the charge and people are plodding along—but mostly at higher ed's typically glacial pace. Although you've got the right folks beginning to line up, you just need the right approaches and structure to push them over the line.
Organizational Continuity <input type="checkbox"/>	Data Quality Assurance <input type="checkbox"/>	71-90		Leading the Pack When it comes to data initiatives, you're ahead of the curve. You've got standards and documentation, stewards and evangelists. Your data is reliable, it's inflecting decisions, and the institution is deriving strategic value from your hard work. It's hugs and smiles—not doom and gloom—when the data folks walk into the room.
Implementation Framework <input type="checkbox"/>	Data Consumption <input type="checkbox"/>	91-100		Data Dominance What data problem in higher education? Your university is humming with the excitement of seamless data flows careering through the institutional ether. Data-driven decision-making feeds off clean, secure data, and silos are a thing of the past. "Student success" and "Institutional Efficiency" have replaced traditional business concerns, and the institution is always looking for new ways to add to central data, and leverage it for innovation. Well done you, Data Hero.

Your Data Governance Score

+ =

Looking to Improve?
The IT Forum has a library of tools and frameworks to support you in sustaining effective data governance. Talk to one of our researchers about the services we provide.

Principles and Values

Example: Brown University



Brown has an ever-increasing need for data, specifically data that are accurate and consistent, accessible, and secure. Such data are vital for tracking and evaluating core activities, such as those that support the mission of the University and ensure our continued accreditation. The University's leaders need timely and accurate information about all aspects of university operations to better inform decisions.

Meeting this demand for data across campus and beyond is a growing challenge. This is partly due to the successful implementation of electronic systems that now contain extensive and new types of data. Accessing this information requires the integration of accurate and consistent information across multiple units, reliable and appropriate data security, and attention to the data culture, including oversight and training.

Guiding Principles of Data Governance at Brown

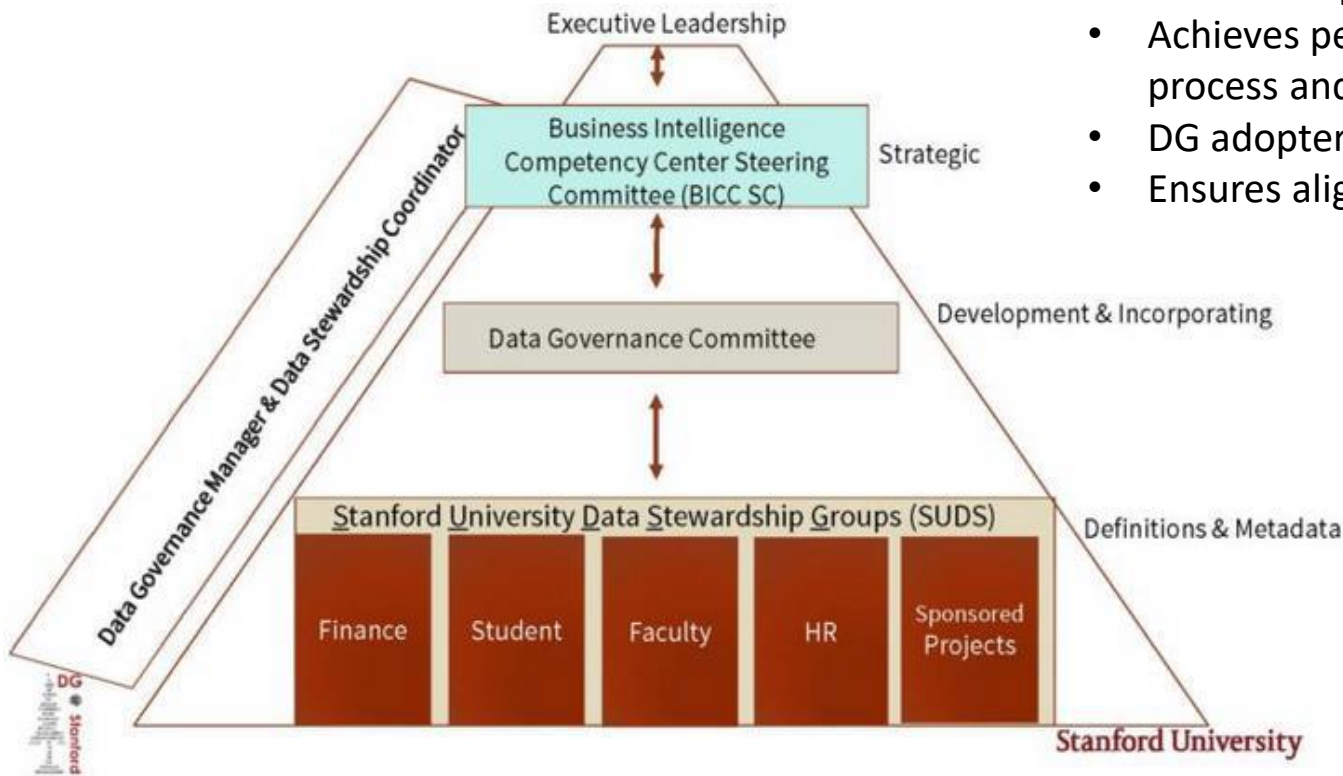
- Institutional data are valuable assets and must be treated as such
- Access to accurate and consistent data is essential to informed decision making across the University
- Data usage and access rules will be articulated and followed
- Data standards can and should be defined and monitored
- The security of institutional data is essential, as is appropriate and timely access
- The privacy of an individual's information will be protected



Structure – Stanford University

BI Competency Ctr. Steering Committee (Senior Leadership)

- Cross-functional oversight & communicates long-term value of BI program
- Achieves peer buy-in, and effects change in business process and data quality
- DG adopters and champions
- Ensures alignment of DG with university goals



Data Governance Committee

- Sets & incorporates DG policies, standards, procedures, roles & responsibilities
- Includes lead steward from each of the data steward groups, plus reps from additional units

Data Stewardship Groups

- Provide metadata infrastructure to support improved decision-making university-wide
- Ensure information integrity
- Build data knowledge
- Meet compliance requirements
- SMEs who define reporting terms and gather metadata associated with their reporting environment

Example Initial Process for Data Collection

With broadly representative planning group (~20 people), conduct focus group with notecards and flipchart



Outcomes

Define data governance as an activity that centers on human behavior more than data

Describe major components of data governance activities

Articulate challenges on their campus and how data governance will address these challenges

Assess their campus culture and organization with a data governance maturity model; select and modify a data governance maturity model for their campus

Identify characteristics of a data governance system; analyze where their own institution has gaps; create an outline for how data governance could fit into existing organizational structures

Discuss how technology may assist but not perform data governance; describe major functions of data governance software applications or solutions

Explain principles of change management in higher education institutions and how they will enable development of data governance on their campuses

Construct an action plan for next steps on their own campus to advance data governance activities