Data Governance: A Primer

Braden J. Hosch, Ph.D.
Asst. Vice President for Institutional Research, Planning & Effectiveness
Stony Brook University

Association for Institutional Research
Data and Decisions for Higher Education
Overview

- Data governance concepts and major aspects
- Selling data governance to senior leadership
- Characteristics of a data governance system
- Maturity models
- Change management in a college or university
- Technological “solutions”
### Outcomes for workshop participants

| 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; and 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 |
What this workshop will not do

• Design your data governance system for you

• Promote specific technological solutions

• Prescribe specific functions, operations, or organization

• Identify how much money to spend
What is data governance?
“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

“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)
<table>
<thead>
<tr>
<th>University</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stony Brook University (2016)</td>
<td>“formalizes behavior around how data are defined, produced, used, stored, and destroyed in order to enable and enhance organizational effectiveness”</td>
</tr>
<tr>
<td>Vanderbilt University</td>
<td>“adds value to our administrative and academic data systems by the establishment of standards that promote data integrity and enables strategic integrations of information systems”</td>
</tr>
<tr>
<td>University of Rochester</td>
<td>“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”</td>
</tr>
</tbody>
</table>
The 5-second elevator definition

Data governance is …

• a set of guidelines for how people behave and make decisions about data
Master data management is often confused with data governance

<table>
<thead>
<tr>
<th>Master Data Management (MDM)</th>
<th>Data Governance</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Comprehensive method to link all critical data to a common point of reference</td>
<td>• Formalized system for how people make decisions about acquisition, production,</td>
</tr>
<tr>
<td>• Example:</td>
<td>storage, distribution, and analysis of data</td>
</tr>
<tr>
<td>• All screens, documents and systems showing a student’s address derive from a common</td>
<td>• Example:</td>
</tr>
<tr>
<td>location.</td>
<td>• Group decides on a definition for home address and agrees on a common source</td>
</tr>
<tr>
<td></td>
<td>field</td>
</tr>
</tbody>
</table>
## Important characteristics of DG definitions

<table>
<thead>
<tr>
<th>Data governance IS</th>
<th>Data Governance IS NOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>• More about people and behavior than data</td>
<td>• IT’s responsibility</td>
</tr>
<tr>
<td>• A system that requires and promotes shared agreement</td>
<td>• Solved by technology</td>
</tr>
<tr>
<td>• Formal (i.e. written down)</td>
<td>• Equally applied across all data assets</td>
</tr>
<tr>
<td>• Adds value by supporting institutional mission/goals</td>
<td></td>
</tr>
</tbody>
</table>
Activity 1 – What data governance features do you have?

List formal and informal structures you have for promoting data governance:

<table>
<thead>
<tr>
<th></th>
<th>Formal</th>
<th>Informal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policies/Practices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsibilities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Why do we need data governance?
Justifications for Data Governance

<table>
<thead>
<tr>
<th>Value</th>
<th>Cost</th>
<th>Risk</th>
</tr>
</thead>
</table>

Justify data governance on your campus based on:

- Value
- Cost
- Risk
Justifications for Data Governance - Value

Educause identifies significant institutional value to higher education institutions from data governance:

• Official vs. ad hoc data definitions
• Clear responsibilities
• Capacity for analytics
• Competitive advantage

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

Justifications for Data Governance – Cost (2)

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

Justifications for Data Governance – Cost (3)

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.

## Justifications for Data Governance - Risks

### Fines Imposed by Federal Student Aid

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

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.”
## Data as an Asset for Universities

<table>
<thead>
<tr>
<th>Generic Example</th>
<th>At Colleges &amp; Universities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web sites grant access in exchange for personal data (email address, etc.).</td>
<td>Data are purchased</td>
</tr>
<tr>
<td>These data have value and can be leveraged or even sold.</td>
<td>Names of prospective students</td>
</tr>
<tr>
<td></td>
<td>Library databases</td>
</tr>
<tr>
<td></td>
<td>Various datasets (U.S. News, Academic Analytics, etc.)</td>
</tr>
<tr>
<td></td>
<td>Data are sold</td>
</tr>
<tr>
<td></td>
<td>To vendors for discounts or services</td>
</tr>
<tr>
<td></td>
<td>Lost data carry costs</td>
</tr>
<tr>
<td></td>
<td>Data breaches</td>
</tr>
</tbody>
</table>
Who owns the data?

Consider carefully use of the word “ownership” with data

- Often represents assignment of responsibility
- Connotes individual control and property vs. caretaking of shared resource

Institutions own the data

Individuals provide stewardship
Activity 2 - Why do we need data governance?

Identify institution-specific examples that help make the case for data governance

- Value – what could you do that you can’t do now?
- Costs – what costs are you incurring because data are not well governed?
- Risks – what risks are you taking because data are not well governed?
Features of Data Governance
# Key features of data governance systems

<table>
<thead>
<tr>
<th>Documents</th>
<th>Groups</th>
<th>Individual roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Charter / framework</td>
<td>• Senior leadership [buy-in]</td>
<td>• Data stewards</td>
</tr>
<tr>
<td>• Principles &amp; values</td>
<td>• Policy council</td>
<td>• Data custodians/caretakers</td>
</tr>
<tr>
<td>• Purpose &amp; scope</td>
<td>• Data steward council(s)</td>
<td>• Data users</td>
</tr>
<tr>
<td>• Roles &amp; responsibilities</td>
<td>• Information security council/program</td>
<td></td>
</tr>
<tr>
<td>• Written &amp; published policies</td>
<td>• Positions/office to support DG</td>
<td></td>
</tr>
<tr>
<td>• Data dictionaries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Communication strategies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Principles and values

Establishing principles and values for data governance assists with:

<table>
<thead>
<tr>
<th>Initial design and implementation</th>
<th>Answering critics</th>
<th>Maintaining focus</th>
<th>Navigating difficult situations</th>
</tr>
</thead>
</table>
Principles 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**

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accountability</td>
<td>Determining who is responsible for the management of data at UW Madison as well as holding them to our outlined standards.</td>
</tr>
<tr>
<td>Agility</td>
<td>All of our processes should adapt when necessary.</td>
</tr>
<tr>
<td>Change Management</td>
<td>New processes demand new and changing staff at UW. We’re committed to ensuring smooth transitions and well informed decisions.</td>
</tr>
<tr>
<td>Consistency</td>
<td>All decisions made will be applied consistently across campus.</td>
</tr>
<tr>
<td>Metrics Driven</td>
<td>We monitor ourselves against measurable goals on a regular basis and use the results to determine courses of action.</td>
</tr>
<tr>
<td>Stewardship</td>
<td>Determine formal roles for those in charge of data. This does not mean that everyone on campus is not responsible despite formal roles.</td>
</tr>
<tr>
<td>Transparency</td>
<td>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.</td>
</tr>
</tbody>
</table>

Source: [https://data.wisc.edu/data-governance/#principles](https://data.wisc.edu/data-governance/#principles)
## Principles and Values – Example

### Stony Brook University

<table>
<thead>
<tr>
<th>Values</th>
</tr>
</thead>
</table>
| **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

<table>
<thead>
<tr>
<th>Principles for Data Governance</th>
</tr>
</thead>
</table>
| **Organizational Effectiveness**  
Transparency  
Communication  
Compliance |
| **Auditability**  
Integrity  
Accountability  
Standards |
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.
Connect Data Governance to Mission

Data governance is a system to improve the effectiveness of the organization, not an activity for its own sake.

Anchor data governance to mission when justifying need or presenting structure.
Activity 3 - Distill university mission

Data governance should be established to support the institution’s mission and/or strategic goals. Colleges and universities have notoriously lengthy mission and goal statements, so it can be a challenge to distill them.

Summarize the main points of your institution’s mission, preferably so that it fits on a slide.
Stony Brook’s framework for data governance outlines a set of principles, structures, roles, and responsibilities to improve the data infrastructure and to advance institutional goals.

Stony Brook has a five-part mission to provide & carry-out:

- Highest quality comprehensive education
- Highest quality research and intellectual endeavors
- Leadership for economic growth, technology, and culture
- State-of-the-art innovative health care, with service to region and traditionally underserved
- Diversity and positioning Stony Brook in global community

Source: https://www.stonybrook.edu/commcms/pres/vision/mission.php
Structure – Generic Example

Executive Steering Committee
- Authorized to change the organization
- Drives cultural change
- Supports the program enterprise-wide
- Provides funding for the Data Governance Program

Data Governance Board
- Made up of high-ranking representatives of data-owning business functions who can make decisions about data for the company
- Assign members of the Data Stewardship Council
- Approve decisions of the Data Stewardship Council
- Approve data-related policies

Business Data Stewards
- Experts on use of their data domain data
- Able to reach out to SMEs to gather information and make decisions
- Typically someone who others come to as the most knowledgeable about the meaning of the data (and how it is calculated)
- Makes recommendations on data decisions and write data-related procedures

Plotkin (2014). Data stewardship: An actionable guide to effective data management
Structure – Stanford University

**BI Competency Ctr. Steering Committee**
- 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
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.
Structure – Stony Brook University

VP Council (Project 50 Forward SteerCo)
- Executive sponsors of project
- Establishes authority and purview of data governance system

Data Governance Council
- Recommends and implements institutional policy for data governance
- Sets priority for

Functional Data Governance Committees
- Implements institutional policy for data governance
- Recommends solutions to specific data issues
- Considers and approve changes to code sets, additions to tables
- Develops solutions to data governance issues
- Communicates with data caretakers in their areas
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

Plotkin (2014). Data stewardship: An actionable guide to effective data management
# Data Governance Council Membership Examples

## UW-Madison
- Chief Data Officer
- Director of Univ. Communications
- VP for Teaching & Learning
- VP for Diversity
- AVC Business Services
- AVC Legal Affairs
- Assoc. Dean Biomedical Informatics
- VP Libraries
- CISO
- Campus Records Officer
- Assoc. Dean Education
- Faculty/Dean Representation

## Stony Brook
- Chief Institutional Research Officer
- Analytics and Enterprise Data Officer
- University Controller
- Chief Enrollment Management Officer
- University Registrar
- Chief Financial Aid Officer
- Provost’s Office designee
- VP Student Affairs designee
- VP Administration designee
- VP Human Resources designee
- VP Information Technology designee
- VP Research designee
- SVP Health Sciences Designee
- University Senate designee
- Chairs & Vice Chairs of FDGCs (6 people)
Data Stewardship Definitions

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)

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.)
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

Plotkin (2014). Data stewardship: An actionable guide to effective data management
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
### Functional Data Stewardship Council/Committees

<table>
<thead>
<tr>
<th>Coordinate data stewards in related area</th>
<th>Set and review definitions, data quality rules, creation/usage rules, determines official version of metadata</th>
<th>Consider and approve changes &amp; additions to code sets</th>
<th>Ensure dictionary standards are followed in area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review data quality in functional area; identify practices promoting data quality; identify areas for improvement and monitor improvements</td>
<td>Respond to inquiries about process, content, limitations and uses of data, especially in cross-functional settings</td>
<td>Elevate issues that require resolution</td>
<td>Communicate proceedings, including notice of changes and decisions</td>
</tr>
</tbody>
</table>
# Stony Brook Roles and Responsibilities Matrix

<table>
<thead>
<tr>
<th></th>
<th>Data Governance Council</th>
<th>Functional Data Governance Cmtes</th>
<th>Data Stewards</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standards and Policies</strong></td>
<td>Define, Establish, Monitors, Audit, Verify, Develop, Revise</td>
<td>Cross functional implementation, coordination</td>
<td>Functional implementation</td>
</tr>
<tr>
<td><strong>Data Quality</strong></td>
<td>Identify, Adopt enterprise-wide DQ tool</td>
<td>Prioritize levels</td>
<td>Review audit reports, Coordinate clean-up, Initial prioritization</td>
</tr>
<tr>
<td></td>
<td>Big picture</td>
<td>Monitor area</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify needs</td>
<td></td>
</tr>
<tr>
<td><strong>Metadata</strong></td>
<td>Establish standards</td>
<td>Ensure cross-functional alignment</td>
<td>Implement Maintain</td>
</tr>
<tr>
<td><strong>Metrics</strong></td>
<td>Review, Identify, Monitor</td>
<td>Monitor area</td>
<td>Monitor Remediate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify area priorities</td>
<td></td>
</tr>
</tbody>
</table>
Data users

Often not considered in data governance systems (but should be).
Example formal responsibilities (Stony Brook)

<table>
<thead>
<tr>
<th>Responsibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
</tr>
<tr>
<td>Include information about the data source and criteria when distributing data, reports and ad hoc analytics to guard against misinterpretations of data.</td>
</tr>
<tr>
<td>Respect the privacy of individuals whose records they may access. Unauthorized disclosure or misuse of institutional information stored on any device is prohibited</td>
</tr>
<tr>
<td>Ensure that passwords or other security mechanisms are used for sensitive data</td>
</tr>
<tr>
<td>Report data quality issues to appropriate data steward</td>
</tr>
</tbody>
</table>
Administrative Office / Positions Supporting Data Governance

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

<table>
<thead>
<tr>
<th>Chief Data Officer</th>
<th>Data Governance Program Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Purdue University</td>
<td>• Purdue University</td>
</tr>
<tr>
<td>• University of Florida System</td>
<td>• Stanford University</td>
</tr>
<tr>
<td>• University of South Carolina – Columbia</td>
<td>• Yale University</td>
</tr>
<tr>
<td>• University of Rochester</td>
<td></td>
</tr>
<tr>
<td>• University of Wisconsin - Madison</td>
<td></td>
</tr>
</tbody>
</table>

Maturity Models
Assess your current state of data governance

Formal assessment of current data governance practices

• Assists with senior leadership buy-in

• Identifies gaps and important implementation considerations

• Extends beyond the informal list we made in Activity 1

• Uses a maturity model to quantify the existing state; allows for measurement of progress in a future state
## Activity 4: Data Governance Maturity Model

<table>
<thead>
<tr>
<th>Data Governance</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal to Data Governance is informal and incomplete. There is no formal governance process.</td>
<td>Developing</td>
<td>Adopted and Implemented</td>
<td>Managed and Repeatable</td>
<td>Integrated and Optimized</td>
<td></td>
</tr>
<tr>
<td>Attention to Data Governance is informal and incomplete. There is no formal governance process.</td>
<td>Data Governance Program is forming with a framework for purpose, principles, structures and roles.</td>
<td>Data Governance structures, roles and processes are implemented and fully operational.</td>
<td>Data Governance structures, roles and processes are managed and empowered to resolve data issues.</td>
<td>Data Governance Program functions with proven effectiveness.</td>
<td></td>
</tr>
</tbody>
</table>

### 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 & 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.
## Stony Brook Data Governance Maturity Model 2.0 Results Summary

### By Function

<table>
<thead>
<tr>
<th>Overall (Area Averages)</th>
<th>Data Governance</th>
<th>Culture</th>
<th>Data Quality</th>
<th>Communication</th>
<th>Roles &amp; Responsibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1-Informal</td>
<td>11%</td>
<td>18%</td>
<td>8%</td>
<td>14%</td>
<td>10%</td>
</tr>
<tr>
<td>Level 2-Developing</td>
<td>41%</td>
<td>30%</td>
<td>34%</td>
<td>39%</td>
<td>32%</td>
</tr>
<tr>
<td>Level 3-Adopted and Implement</td>
<td>11%</td>
<td>14%</td>
<td>9%</td>
<td>14%</td>
<td>3%</td>
</tr>
<tr>
<td>Level 4-Managed and Repeat</td>
<td>3%</td>
<td>3%</td>
<td>6%</td>
<td>2%</td>
<td>7%</td>
</tr>
<tr>
<td>Level 5-Integrated and Optimize</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Not enough information</td>
<td>34%</td>
<td>35%</td>
<td>41%</td>
<td>32%</td>
<td>48%</td>
</tr>
</tbody>
</table>
Baseline Dimensions

- Data Governance
- Culture
- Data Quality
- Communication
- Roles & Responsibilities

Current 2015

Maturity

- Integrated & Optimized
- Managed & Repeatable
- Adopted & Implemented
- Developing
- Informal

Target 2017
Change Management in Higher Education
## Elements to change management

<table>
<thead>
<tr>
<th>Process Representation Deliberation</th>
<th>Executive sponsors Mission alignment Project mgmt./timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem statement</td>
<td>Interested allies</td>
</tr>
<tr>
<td>Research/environ. scan</td>
<td>Interest mapping</td>
</tr>
<tr>
<td>Ideas for solutions</td>
<td>Advocacy from others</td>
</tr>
</tbody>
</table>

### Initiative
Activity 5 – 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)
Case Study – Stony Brook University

Initiative to strengthen university data infrastructure (Jan 2015-Sept. 2016). Charge to examine:

- Data governance
- Data quality
- Communication
## Charge for data governance (first 9 months)

<table>
<thead>
<tr>
<th>Examine existing governance structures</th>
<th>Identify and articulate</th>
</tr>
</thead>
<tbody>
<tr>
<td>• active and inactive groups and lines of responsibility</td>
<td>• Roles of cross-functional groups</td>
</tr>
<tr>
<td>• existing processes, practices and procedures that significantly impact data management and stakeholders.</td>
<td>• Functional roles in business units (e.g. data owner, data custodian, report owner) will also be identified and articulated.</td>
</tr>
</tbody>
</table>

## Cross functional review group

<table>
<thead>
<tr>
<th>Draft formal governance structure for data management</th>
<th>Formalize a process for prioritization</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Principles, mission, and goals</td>
<td></td>
</tr>
<tr>
<td>• Post on a website to codify roles and responsibilities.</td>
<td></td>
</tr>
</tbody>
</table>
Charge for data quality improvement

Examine existing practices for ensuring data quality within

- PeopleSoft
- data warehouse
- other 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
Charge for training and communication

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
Example initial process for data collection

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

<table>
<thead>
<tr>
<th>FOCUS GROUP ACTIVITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>list three current data governance mechanisms at our [INSTITUTION], the systems or applications they cover</td>
</tr>
<tr>
<td>list three aspects of data governance that are absent at [INSTITUTION] or need to be strengthened</td>
</tr>
<tr>
<td>list three things that data governance at [INSTITUTION] should accomplish</td>
</tr>
<tr>
<td>list three roles or structures that should be included in [INSTITUTION’S] data governance system.</td>
</tr>
</tbody>
</table>
Activity 6 – Draft input for planning process

Using the framework below, draft useful responses to be incorporated into local planning

- list three current data governance mechanisms
- list three aspects of data governance that are absent
- list three things that data governance should accomplish
- list three roles or structures that should be included

[Anticipate responses that may be counterproductive]
- E.g. “IT should control data governance”
Technological “Solutions”
# Technology applications for data governance

<table>
<thead>
<tr>
<th>Technology can support data governance</th>
<th>Technology will not</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data dictionary management</td>
<td>Build organizational structures, responsibilities, accountabilities</td>
</tr>
<tr>
<td>Data quality analysis</td>
<td>Mend dysfunctional organizations</td>
</tr>
<tr>
<td>Master data management</td>
<td>Implement organizational or cultural change</td>
</tr>
</tbody>
</table>
Selected Data Governance Applications

- Axon (Informatica)
- Collibra
- Data Cookbook (iData)
- Melissa Data
- Oracle Data Quality Middleware
- SAS Data Governance
Issues to consider when selecting technology

Alignment with DG needs
- Metadata management
- Integration w/ reporting tools
- Data quality
- Security/user roles

Initial cost and annual cost

User Community

Ease of implementation and impact on IT
Final Thoughts
Data governance is not a project, it’s a process

**Project Model**
- Linear
- Implies conclusion

**Process Model**
- Cyclical
- Ongoing
Data governance is only one part of a data strategy

A data strategy is a larger vision for how your organization will work with data.

Data acquisition

Data governance

Data quality

Data access

Data usage & literacy

Data extraction & reporting

Data analytics

Takeaways

• Data governance is more about people than data
• All higher ed change management principals apply
• Process and written documents are essential
  • 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