How to transform the landscape of analytics with data governance

*Richard R. Burnette III, PhD*
Associate Provost for Metrics, Analytics and Strategic Planning, and Institutional Data Administrator
*Florida State University*

*Braden J. Hosch, PhD*
Associate Vice President for Institutional Research, Planning & Effectiveness
*Stony Brook University*
Data governance is a strategic priority

- Proliferation of data and applications
- Privacy expectations/regulations
- Cloud data mobility
- Fair and ethical use
The 5-second elevator definition

Data governance is ...

• a set of guidelines for how people behave and make decisions about data
What is Data Governance?

John Ladley – Data governance is the organization and implementation of policies, procedures, structure, roles, and responsibilities which outline and enforce rules of engagement, decision rights, and accountabilities for the effective management of information assets.
## 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>
Complementary Elements of Data Governance

Strategic Empowerment

Risk v Reward

Rules & Constraints
Why Do We Need Data Governance?

- Maximize Data Investments
- Gain Deeper Insights
- Promote Efficiency
- Ensure Trust
- Reduce Risk
- Improve the Experience
Principles of Data Governance

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

- **Quality**
  which is proactively assessed and standards applied

- **Responsibility**
  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

- **Secure access**
  to relevant users, kept secure through access control

- **Insight**

What are the Data Dimensions

- Use
- Access
- Management
- Integrity
- Privacy
- Security
What Data are we Governing?

**Administrative**
ERP Data and Operational Data
- SIS, HR, Financials, CRM, Departmental, Organizational Performance Data

**Teaching & Learning**
Instructor and Student Performance
- LMS, Lecture Capture, Clickers, Attendance, Engagement, Grades, Progression, and Course and Faculty Evaluations

**Research**
Data on Researchers, Research & Grants
- Publication and Citation Histories, Proposals Submitted, Research Content & Results, Grant Dollars, PI and Co-PI data, Graduate and Undergraduate Research

**Other Types**
Both Internal and External Data
- Meta-data, Unstructured Data, Geo-location, Event Attendance, Organization Involvement, Social Media, Sentiment Analyses, Survey Data, Business Transactions, Vendor Data
Key features of data governance systems

<table>
<thead>
<tr>
<th>Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Charter / framework</td>
</tr>
<tr>
<td>• Principles &amp; values</td>
</tr>
<tr>
<td>• Purpose &amp; scope</td>
</tr>
<tr>
<td>• Roles &amp; responsibilities</td>
</tr>
<tr>
<td>• Written &amp; published policies</td>
</tr>
<tr>
<td>• Data dictionaries</td>
</tr>
<tr>
<td>• Communication strategies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Senior leadership [buy-in]</td>
</tr>
<tr>
<td>• Policy council</td>
</tr>
<tr>
<td>• Data steward council(s)</td>
</tr>
<tr>
<td>• Information security council/program</td>
</tr>
<tr>
<td>• Positions/office to support DG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Data stewards</td>
</tr>
<tr>
<td>• Data custodians/caretakers</td>
</tr>
<tr>
<td>• Data users</td>
</tr>
</tbody>
</table>
## Common Elements of the Structure

<table>
<thead>
<tr>
<th>Committees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive/Steering – Senior officials focused on holistic goals of institution</td>
</tr>
<tr>
<td>Data Strategy – Occasionally separate from above focused on better data use</td>
</tr>
<tr>
<td>Operational Governance – Responsible for executing policies and procedures</td>
</tr>
<tr>
<td>Data Standards – Maintains data elements and monitors quality and delivery</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Data Officer (CDO) – Often overseas the execution of institutional goals</td>
</tr>
<tr>
<td>Data Trustees – Subject matter owners responsible for data integrity and clarity</td>
</tr>
<tr>
<td>Data Stewards – Responsible for assigning access and assuring standards met</td>
</tr>
<tr>
<td>Data Custodians – Usually IT maintaining security, backups, recovery, availability</td>
</tr>
<tr>
<td>Data Consumers – Functional staff who interact with data and report on data</td>
</tr>
</tbody>
</table>
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
Information Governance Council Purpose

Create a data governance imperative
Promote a data-engaged campus
Create and update policies
Create access control mechanisms
Manage conflict resolution
Promote shared data management
Authorize data movement and storage
Reinforce reporting controls
Define and arbitrate fair and ethical use
Evaluate and assess effectiveness
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

Coordinate
data stewards in related area

Set / review
definitions, data quality rules, creation/usage rules, metadata

Consider and approve changes to code sets

Enforce data dictionary standards in area

Review data quality in functional area; identify practices promoting data quality

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

Expectations should be set for data users. Example formal responsibilities (Stony Brook)

- Recognize that institutional data are potentially complex.
- Make efforts to understand the source, meaning and proper use of the data.
- Include information about the data source and criteria to guard against misinterpretation of data.
- Respect the privacy of individuals whose records they may access.
- Ensure that passwords or other security mechanisms are used for sensitive data.
- Report data quality issues to appropriate data steward.
Keys to Implementation

- Create value statement for DG
- Prepare a roadmap
- Design the program
- Identify rough costs and staffing
- Identify structure and roles on governance committees
- Plan meetings and document via shared media
Additional Keys to Implementation

- Identify early wins
- Prepare policy documents
- Create mechanisms for documenting compliance
- Adopt Maturity Model
- Brand process with visible brand
- Prepare training protocols
- Measure and report on activities
## 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>
<tr>
<td>Issue and process management</td>
<td></td>
</tr>
</tbody>
</table>

- [Collibra](https://www.collibra.com)
- [Informatica](https://www.informatica.com)
- [Melissa](https://www.melissa.com)
- [Oracle](https://www.oracle.com)
- [SAS](https://www.sas.com)
- [Talend](https://www.talend.com)
- [Technikèes](https://www.technikes.com)
## Example Data Governance Maturity Model

<table>
<thead>
<tr>
<th></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><strong>Organizational Structures</strong></td>
<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>
</tr>
<tr>
<td><strong>Culture</strong></td>
<td>Limited awareness about the value of dependable data.</td>
<td>General awareness of the data issues and needs for business decisions.</td>
<td>There is active participation and acceptance of the principles, structures and roles required to implement a formal Data Governance Program.</td>
<td>Data is viewed as a critical, shared asset. There is widespread support, participation and endorsement of the Data Governance Program.</td>
<td>Data governance structures and participants are integral to the organization and critical across all functions.</td>
</tr>
<tr>
<td><strong>Data Quality</strong></td>
<td>Limited awareness that data quality problems affect decision-making. Data clean-up is ad hoc.</td>
<td>General awareness of data quality importance. Data quality procedures are being developed.</td>
<td>Data issues are captured proactively through standard data validation methods. Data assets are identified and valued.</td>
<td>Expectations for data quality are actively monitored and remediation is automated.</td>
<td>Data quality efforts are regular, coordinated and audited. Data are validated prior to entry into the source system wherever possible.</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td>Information regarding data is limited through informal documentation or verbal means.</td>
<td>Written policies, procedures, data standards and data dictionaries may exist but communication and knowledge of it is limited.</td>
<td>Data standards and policies are communicated through written policies, procedures and data dictionaries.</td>
<td>Data standards and policies are completely documented, widely communicated and enforced.</td>
<td>All employees are trained and knowledgeable about data policies and standards and where to find this information.</td>
</tr>
<tr>
<td><strong>Roles &amp; Responsibilities</strong></td>
<td>Roles and responsibilities for data management are informal and loosely defined.</td>
<td>Roles and responsibilities for data management are forming. Focus is on areas where data issues are apparent.</td>
<td>Roles and responsibilities are well-defined and a chain of command exists for questions regarding data and processes.</td>
<td>Expectations of data ownership and valuation of data are clearly defined.</td>
<td>Roles, responsibilities for data governance are well established and the lines of accountability are clearly understood.</td>
</tr>
</tbody>
</table>
Person Roles

- CDO – Chief Data Officer
- CISO – Chief Information Security Officer
- Chief Privacy Officer
- Chief Compliance Officer
- Institutional Data Administrator
- Data Stewards
- Data Custodians
- Data Manager
Key Policies

• Strategic Vision/Policy for Data Use
• Information Privacy
• Data Access and Use
• Data Management (includes 3rd Party)
• Cybersecurity
• Email and Media Use
• Survey Administration
• Data & Device Security
• Fair and Ethical Use
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
Questions?

Rick Burnette
Associate Provost for Metrics, Analytics and Strategic Planning, and Institutional Data Administrator
Florida State University
rburnette@fsu.edu

Braden Hosch
Associate Vice President for Institutional Research, Planning & Effectiveness
Stony Brook University
Braden.hosch@stonybrook.edu