Data Governance: Essential for Strategy, Policy, and Operations

Shirley C. Payne
AVP for Information Security, Policy, & Records
University of Virginia

EDUCAUSE 2012 Security Conference
May 17th, 2012
Copyright Shirley C. Payne 2012.

This work is the intellectual property of the author. Permission is granted for this material to be shared for non-commercial, educational purposes, provided that this copyright statement appears on the reproduced materials and notice is given that the copying is by permission of the author. To disseminate otherwise or to republish requires written permission from the author.
What is the business case for DG?
What does effective DG look like?
Got models?
How does one develop/enhance a DG program?
Business Case for DG
Data Dark Ages

Centralized Stovepipe Data Stores

Limited Data Distributed To Departments Via Hard Copy Reports
Data Floodgates Began Opening In Early 90’s

Information Warehouse, ERPs, End-user Reporting Tools, Downloadable Spreadsheets

Goal: Make Integrated Data Available To Widest Audience Possible
Basic Data Policies Began Emerging – UVa Example:

- Clarified data ownership:
  - Institution owns administrative data
  - Central admin units are data stewards

- Established data classifications

- Set high level conditions of data use:
  - Use only for institutional business
  - Comply with privacy laws
  - Comply with “reasonable protection and control procedures”
  - Present data accurately
Change Happened...

1990s

Controlled Data Distribution

Mostly Centralized Data

EASY TO MANAGE

HARD TO MANAGE
Change Happened...

1990s
- Controlled Data Distribution
- Mostly Centralized Data

Today
- More Legal, Contractual, & Public Pressures
- More Security/Privacy Threats
- More Storage & Processing Options
- Widely Distributed Data & More Of It

EASY TO MANAGE  HARD TO MANAGE
What Needs Fixing?

- Lack of shared, consistent view of data
What Needs Fixing?

- Overly complex data environment
What Needs Fixing?

- Inadequate data quality
“At bedrock, if regulators, educational leaders, politicians, and others want to use information to

- promote student success,
- lower rates of student attrition,
- improve institutional decisions, and
- promote transparency,

then attention must be paid first and foremost to data quality.”  Richard N. Katz

[1] Yanosky, Ronald
What Needs Fixing?

- High risk of data exposure
What Needs Fixing?

- High IT expenditures
Missed strategic opportunities due to under-leveraged data
“An immense and ever-increasing wealth of knowledge is scattered about the world today; knowledge that would probably suffice to solve all the mighty difficulties of our age, but it is dispersed and unorganized.”

“The overriding principle at the heart of the information strategy is the recognition that the university is a single, integrated organization…[that] can ill afford to forego the benefits of behaving as such.”
What is the business case for DG?

- **Shared Data**:
  - Focused attention on quality
  - Minimized redundancy

- **Reduced complexity**:
  - Consistent definitions
  - Common data protection standards
  - Enabled compliance

- **Higher Quality**:
  - Formal procedures/accountability for data quality
  - Data changes that propagate across institution
What is the business case for DG?

<table>
<thead>
<tr>
<th>Lower exposure risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Minimized storage of sensitive data</td>
</tr>
<tr>
<td>• Known locations of sensitive data</td>
</tr>
<tr>
<td>• Governance across all locations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IT investments better leveraged</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Enhanced utilization of enterprise systems data</td>
</tr>
<tr>
<td>• ROI analyses factor in data value</td>
</tr>
<tr>
<td>• Process reengineering to enhance DG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strategic initiatives enabled by data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High quality, accessible data informs key decisions</td>
</tr>
</tbody>
</table>
Characteristics of Effective Governance
<table>
<thead>
<tr>
<th>Governance</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oversight</td>
<td>Implementation</td>
</tr>
<tr>
<td>Authorizes Decisions</td>
<td>Authorized to Make Decisions</td>
</tr>
<tr>
<td>Enacts Policy</td>
<td>Enforces Policy</td>
</tr>
<tr>
<td>Accountable</td>
<td>Responsible</td>
</tr>
<tr>
<td>Strategic Planning</td>
<td>Project Planning</td>
</tr>
<tr>
<td>Resource Allocation</td>
<td>Resource Utilization</td>
</tr>
<tr>
<td>Doing the Right Thing</td>
<td>Doing Things Right</td>
</tr>
</tbody>
</table>

Exactly what is data governance?

- Decision rights and accountabilities governing data-related matters
- Complements, rather than replaces, data management
- The convergence of:
  - Data Policy & Standards
  - Data Management
  - Business Processes
  - Data Quality
  - Enabling Technologies
  - Risk Management
- Objectives enable use of data as a strategic institutional asset
Elements of Effective DG

OUTCOMES
- Data Risk Management & Compliance
- Value Creation

Enablers
- Organizational Structures & Awareness
- Policy
- Stewardship

Core Disciplines
- Data Quality Management
- Information Life-Cycle Management
- Information Security & Privacy

Supporting Disciplines
- Data Architecture
- Classification & Metadata
- Audit Information Logging & Reporting

[5] Adapted from NASCIO, p. 13; IBM Information Governance Council
OKAY, BUT WHERE DO I START?
DG Capability Maturity Models

Heroics  Continuous Improvement  Ideal State
Use of DG Capability Maturity Models

- Awareness building
- Support for business case
- Scope definition
- Communication
DG/CMM Source Examples

- DataFlux
- EWSolutions
- Gartner
- IBM
- Knowledge Logistics
- MDM Institute
- Oracle
- MIKE2.0
- Microsoft – addresses data privacy, confidentiality, and compliance
Basic

- **People:** poor training & awareness
- **Process:** basic policies & immature procedures
- **Technology:** ad hoc & limited

Standardized

- **People:** formal training & awareness
- **Process:** established & communicated policies & procedures
- **Technology:** foundational & minimal

[7] Microsoft
Risk Management-focused

**Standardized**
- **People:** formal training & awareness
- **Process:** established & communicated policies & procedures
- **Technology:** foundational & minimal

**Rationalized**
- **People:** formal training & awareness w/ compliance metrics
- **Process:** process & procedure improvement
- **Technology:** reduced reliance on manual & compensating controls

[7] Microsoft
Governance-focused

Rationalized
- **People**: formal training & awareness w/ compliance metrics
- **Process**: process & procedure improvement
- **Technology**: reduced reliance on manual & compensating controls

Dynamic
- **People**: formal training & awareness w/ compliance metrics
- **Process**: process transformation & more integrated compliance efforts
- **Technology**: fully automated & integrated controls
Formal Frameworks

- Six Sigma for data certification
- ITIL
- COBIT
- COSO
- ISO/IEC 9000 series
- ISO/IEC 27000 series
- MIKE2.0
Comprehensive framework
Many to choose from, but generally...

Outcomes
* e.g., Risk Management & Compliance

Enablers
* e.g., Policy

Core Disciplines
* e.g., Information Security & Privacy

Supporting Disciplines
* e.g., Classification & Metadata
### Recap

**What does effective DG look like?**

<table>
<thead>
<tr>
<th></th>
<th><strong>THIS</strong></th>
<th><strong>NOT THIS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data decision rights &amp; accountabilities</strong></td>
<td>Defined, repeatable, clearly communicated policies &amp; procedures</td>
<td>Based on heroics, tribal knowledge, individual experience</td>
</tr>
<tr>
<td><strong>DM &amp; DG solutions (people, process, technology)</strong></td>
<td>Managed, institution-level</td>
<td>Localized, if any</td>
</tr>
<tr>
<td><strong>Data consistency, quality</strong></td>
<td>Synchronized across institution</td>
<td>“Siloed”</td>
</tr>
<tr>
<td><strong>Measured, continuous improvement</strong></td>
<td>Embedded in culture &amp; risk-driven</td>
<td>None</td>
</tr>
</tbody>
</table>
Case Study #1

Focus: Policy & Governing Body
Scope:
- All data regardless of location or format
- All individuals encountering IU data
- Excludes research data, health data, IP

Information Governance Policy
http://protect.iu.edu/cybersecurity/policies/ISPP25

Governance leadership:
Information Security and Privacy Risk Council
http://protect.iu.edu/privacy/program/governance
Case Study #2

Focus: DG Strategic Planning
Institutional Data Management and Governance Initiative launched (2007)

Roadmap defined (June 2009)

- Establish data leadership & governance structure
- Develop common data definitions
- Implement shared data analysis tools/approaches
- Improve data presentation to decision-makers
- Implement enabling technologies
Institutional Data Management and Governance Initiative launched (2007)

Roadmap (June 2009)

Institutional Data Council formed (late 2009)

Led by Associate Vice Chancellor

http://idmg.berkeley.edu/index.html
Case Study #3

Focus: Sensitive Data Initiative
Institution-wide Sensitive Data Minimization Initiative (2004 launch)
Ensure sensitive data requested only when essential

Ensure sensitive data access authorized to least # of people

Ensure sensitive data provided only when essential

University Processes & Supporting Systems

- Implement clearer sensitive data use policies
- Communicate responsibilities for data protection
- Implement compliance verification processes

Ensure sensitive data stored only in highly secured devices and file cabinets
Institution-wide Sensitive Data Minimization Initiative (2004 launch)

Lessons learned:

- Executive support critical
- No blame strategy
- Realistic expectations
- Plan/execute for long term
- Strong project core
- Early stakeholder engagement
- Strong communication and governance
- Facilitate compliance with tools and detailed procedures
- Saying “no” is critical
Revitalizing Data Stewardship through Risk Reduction: Managing Sensitive Data at the University of Virginia

ECAR Case Study 8, 2009

www.educause.edu/Resources/RevitalizingDataStewardshipthr/192168
Program Implementation & Enhancement
The Journey
Road Hazards

- No immediate ROI
- Scope too broad

- Powerless leadership
- Politics
Iterative Process

1. Develop Value Statement
2. Prepare Roadmap
3. Plan & Fund
4. Deploy
5. Govern
6. Monitor, Measure, Report
7. Improve

[8] Agosta, Lou
Limit Initial Scope

DATA GOVERNANCE

- Privacy, Security, Compliance
- Data Warehouse, Business Intelligence
- Architecture, Integration
- Data Quality

[9] Thomas, Gwen
Gain Executive Support

- Clearly convey business case
- Seek cultural change
  - Align with institutional goals
  - Be clear about who needs to do what
  - Imbed in existing processes
Establish Boundary-spanning DG Body

Role examples

- Interpret data-related compliance requirements
- Determine which take precedence
- Decide if risks managed through procedural controls or automated ones
- Approve funding and procedural details
- See that policies and procedures are designed and implemented

[1] Yanosky, Ronald
Implement Policies Early

Such as:

- Data managed as enterprise resource independent of specific applications or systems
- Responsibilities and roles
- Data integrity essential; data quality periodically checked
- Data security requirements
- Reporting data exposure incidents
Develop Comprehensive Procedures, e.g.,

- Aligning policies, requirements, and controls
- Establishing decision rights
- Establishing accountability
- Performing stewardship
- Managing change
- Defining data
- Resolving issues
- Specifying data quality requirements
- Building governance into technology
- Stakeholder care
- Communications
- Measuring and reporting value

Formal, Documented, Repeatable

[9] Thomas, Gwen
Research/Select Tools

For Metadata

For Data Delivery

For Data Quality

For Stewardship & Governance

Develop & Report Metrics
Communicate
Approach as iterative process
Limit initial scope
Inadequate involvement = inadequate DG
Implement policies early
Develop comprehensive procedures
Get tools in place
Visibility!
One final thought
### Data Information

<table>
<thead>
<tr>
<th>Data</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured</td>
<td>Unstructured</td>
</tr>
<tr>
<td>Example: ERP data</td>
<td>Examples:</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ Web pages, e-learning systems content</td>
</tr>
<tr>
<td></td>
<td>▪ Email</td>
</tr>
<tr>
<td></td>
<td>▪ Research results</td>
</tr>
<tr>
<td></td>
<td>▪ Audio/video</td>
</tr>
</tbody>
</table>
IG Partners

- Business Data Stewards
- IT
- Research Compliance
- Libraries
- Teaching & Learning
- Intellectual Property Management
- Records Management

And Others
Challenges

- Same road hazards as DG, but even harder to overcome
- Long-term preservation
  - Transitory nature of storage media
  - Tracking data lineage
- More compliance issues
Recap

What’s ahead for DG?

Information Governance
- Expanded organizational boundaries
- Increased compliance issues
- Long-term preservation is key

Data rich and information poor rich
“An immense and ever-increasing wealth of knowledge is scattered about the world today; knowledge that would probably suffice to solve all the mighty difficulties of our age, but it is dispersed and unorganized.”

It has been a pleasure!
References


Case Study Links

Indiana University
- Information Governance Policy
  http://protect.iu.edu/cybersecurity/policies/ISPP25
- Governance leadership: Information Security and Privacy Risk Council
  http://protect.iu.edu/privacy/program/governance

University of California at Berkeley
- Institutional Data Management Governance Initiative
  http://idmg.berkeley.edu/index.html

University of Virginia
- Revitalizing Data Stewardship Through Risk Reduction: Managing Sensitvie Data at the University of Virginia
  www.educause.edu/Resources/RevitalizingDataStewardshipthr/192168
Additional Resources

Cloud Services: Policy and Assessment

EDUCAUSE Data Protection Contractual Language Toolkit
https://wiki.internet2.edu/confluence/display/itsg2/Data+Protection+Contractual+Language

UVa Data Protection Contractual Language Tool –
email Shirley Payne (payne @ virginia.edu)