Data Stewardship and Data Governance In Higher Education

An ECAR Working Groups Webinar
January 10, 2017 • 3-4 p.m. ET
Today’s Speakers

Michael Fary
Enterprise Data Architect
University of Chicago

Michael C. Kelly
Chief Data Officer
University of South Carolina
Agenda

I. Data Stewardship
   - Overview
   - ECAR resources

II. Data Governance / Data Protection
   - Overview
   - ECAR resources

III. Engagement
Establishing Data Stewardship Models

Mike Fary, University of Chicago
Introduction

- Data stewardship is the vehicle to carry out the required tasks in a coordinated manner
- Bedrock of a data governance program
- Data governance cannot exist without data stewards
- Data is an institutional asset
- Manage it carefully and use it appropriately
Developing a Data Stewardship Program

- Different institutions, different focus, different modes of operation
- Based on institutional culture, available resources, and data management maturity
- Focus on:
  - data policy
  - data definition
  - risk management
Developing a Data Stewardship Program

Possible Program Directions

- Align with compliance and risk objectives
- Align with a product rollout
- Align with other campus initiatives
Developing a Data Stewardship Program

First Steps

- Assign responsibility
- Define the steward role
- Publish a roster of data stewards
- Develop the approach
What Makes a Good Data Steward?

- Interested in the data environment and associated processes
- “Go To” person of how data is collected, maintained, and interpreted
- Has “tribal knowledge” of all things data in a particular data domain
Common Characteristics of Good Data Stewards

- Listen to and understand issues related to data quality and connected processes
- Want to find solutions to improve the overall data landscape
- Willing to provide guidance in pursuing better data quality
Common Characteristics of Good Data Stewards

- Know which resources can best assist with data issues
- Understand that not all data issues can be resolved with quick fixes
- Possess the ability to see the “big picture”
What Data Areas Do Stewards Oversee?

- Student
- Academic/Faculty
- Alumni
- Sponsored Research
- IRBs
What Data Areas Do Stewards Oversee?

- Administrative
  - Finance
  - Facilities
- HR
- Legal
- Safety/Security

Academic research not a focus of this discussion, but does require stewardship.
Responsibilities of a Data Steward

- Operational Oversight
- Data Quality
- Privacy, Security, and Risk Management
- Policies and Procedures
- Education and Training
Skill Sets of a Data Steward

- Relationship management
- Facilitation
- Communication
- Process definition
- Change management
- Problem solving
- Policy development
Data Stewardship at U of Chicago

- Data Stewardship Council Founded 2008
- Presidential Approval & Encouragement
- Representation
  - Administrative Units
  - Academic Units
  - Office of the President
  - Office of the Provost
U of C Data Stewardship Council

- Monthly Meetings
  - Forum for Common Data Related Issues
  - Discussions of Data Related Activities
    - New Systems/Services
    - New Policies

- Network Meetings

- Work Groups

[Link to U of C Data Stewardship Council website](dsc.uchicago.edu)
U of C Data Stewardship Council

Past Activities

- Data Usage Request Process
- Classification Matrix & Website
- Enterprise Data Models
  - Student
  - Financial
- Email Archival & Preservation
U of C Data Stewardship Council

On the Horizon

- Enterprise Data Models
  - Alumni Relations & Development
  - HR
- “Data Awareness” Training for HR
- Space Identification Standardization
- Org Naming Standardization
In an era of smart phones, cloud computing, big data, and data-driven decision making, it has never been more important for universities to develop and update policies and procedures to promote the quality, privacy, security, and preservation of institutional data. Ultimately, the goal should be to foster a culture of good stewardship where faculty, students, and staff use, share, and store data in a responsible way.

George Trone  
Chairman, Data Stewardship Council  
Senior Associate Secretary of the University  
The University of Chicago
Summary

Successful data stewardship is the key to effective data governance.
Compelling Case for Data Governance

Mike Kelly, University of South Carolina
Resources

- **Data Protection Primer for Higher Education**
  - ECAR Working Group Paper EWG1603
  - June 20, 2016

- **The Compelling Case for Data Governance**
  - ECAR Working Group Paper EWG1501
  - March 19, 2015
Introduction

- DG is the starting point for managing data
- Data Stewardship is the bedrock of a data governance program
- Shared imperative
- Different strategies for administrative data vs. research data
Overview

A conversation starter among stakeholders and executives.

Three questions:

- What is data governance?
- Why is data governance necessary?
- Who should be involved, and how?
What is Data Governance?

- Clarify what data exist – now / future
- Create systems & processes
- Identify data managers
- Roles & responsibilities
- Program, not a project

Guidelines for:
- Classifying data
- Assigning accountability
- Maintaining data integrity
- Controlling access
- Handling and sharing data
Why is Data Governance Necessary?

- Data is an asset – requiring care & due diligence
  - Maximizing benefits – positive deliverables
  - Reducing risks – minimizing potential negative consequences
Governance Matters

Data differentiation

- Administrative / operational data
- Research data
Maximizing Benefits

... of data, via data governance

- Official vs. ad hoc definitions
- Clear responsibilities
- Structured data
- Unstructured data
- Capacity for analytics
- Competitive advantage
## Reducing Risk via Data Protections

<table>
<thead>
<tr>
<th>Affirmative Protections</th>
<th>Misuse Protections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Governance</strong></td>
<td><strong>Privacy Rights</strong></td>
</tr>
<tr>
<td>Ensuring data conform to quality and standards upon entry and throughout use in operations and analysis; includes a decision framework around what data should be collected and protected</td>
<td>Ensuring personal awareness of rights and exercise of personal choice in use of data</td>
</tr>
<tr>
<td><strong>Integrity Assurance</strong></td>
<td><strong>Information Security</strong></td>
</tr>
<tr>
<td>Monitoring to ensure data are complete, accurate, and timely and conform to data standards</td>
<td>Application of controls to protect data and systems based on data classification; addressing storage media, hardware, databases, operating systems, network and transmission, and physical and cloud data centers</td>
</tr>
<tr>
<td><strong>Data Classification</strong></td>
<td><strong>Incident Response</strong></td>
</tr>
<tr>
<td>Determining and categorizing the sensitivity of data and necessary protections</td>
<td>Procedures established and tested for responding to a breach in a timely manner, consistent with policy</td>
</tr>
<tr>
<td><strong>Access Management</strong></td>
<td><strong>Data-Sharing Agreements</strong></td>
</tr>
<tr>
<td>User identification, authorization, and authentication to access systems and data</td>
<td>Documented agreements for exchange and use of data, internally between units and with external vendors and others</td>
</tr>
<tr>
<td><strong>Culture of Data Value</strong></td>
<td></td>
</tr>
<tr>
<td>Transforming institutional mindset to value data as an asset, under the direction of data stewards</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Business Continuity Protections</th>
<th>Data Life-Cycle Protections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Backup</strong></td>
<td><strong>Collection</strong></td>
</tr>
<tr>
<td>Routine, scheduled backups of critical data to alternate hardware or cloud; may entail off-site storage by a third party</td>
<td>The process of bringing data into the life cycle in a deliberate and thoughtful way</td>
</tr>
<tr>
<td><strong>Systems Redundancy</strong></td>
<td><strong>Change Management</strong></td>
</tr>
<tr>
<td>Technical architecture, including virtualization, to ensure critical systems remain available</td>
<td>Thorough, scheduled stakeholder notification and prior acceptance testing of changes to databases and systems to avoid data corruption and loss</td>
</tr>
<tr>
<td><strong>Disaster Recovery</strong></td>
<td><strong>Archive</strong></td>
</tr>
<tr>
<td>Procedures established and periodically tested for restoring systems from backup in the event of a natural or manmade disaster</td>
<td>Removal of records from production systems and retention for a specified time period for access as needed</td>
</tr>
<tr>
<td></td>
<td><strong>Deletion</strong></td>
</tr>
<tr>
<td></td>
<td>Permanent removal of records from systems, hardware, and media</td>
</tr>
<tr>
<td></td>
<td><strong>Disposal</strong></td>
</tr>
<tr>
<td></td>
<td>Permanent, unrecoverable destruction of records and storage media/hardware</td>
</tr>
</tbody>
</table>

*Data Protection Primer for Higher Education: Environmental Considerations, Culture, and Practices*
Who Should Be Involved, and How?

- Operational area managers (i.e. Data Stewards)
- Institutional executives
- University community
- Dedicated leader/personnel
Resource & Cost Requirements

- **Personnel**
  - Participants from core operational, administrative, and IT areas
  - ? dedicated leadership ?

- **Data Governance Tools**
  - Data dictionary
  - Workflows
  - MDM / RDM
  - Data quality & integrity tools
Getting Started with a Framework

- Charter for data governance program
  - Organize
  - Compartmentalize
  - Formalize vocabulary

- Frameworks &/or Maturity Models
Questions?

- The Compelling Case for Data Governance (March 2015)
- Establishing Data Stewardship Models (December 2015)
- Data Protection Primer for Higher Education: Environmental Considerations, Culture, and Practices (June 2016)
- EDUCAUSE Data Governance and Chief Data Officers Constituent Group
- ECAR Working Groups Home
Save the Date!

Tuesday, February 7 • 3-4 p.m. ET
Focus on Research Data

Tuesday, March 7 • 3-4 p.m. ET
Focus on Administrative Data

Tuesday, April 4 • 3-4 p.m. ET
Focus on Academic Data

Find out more at www.educause.edu/ecar/ecar-working-groups