Data Governance in the Scholarly Communication Ecosystem
Understanding the Issues and Rights Associated With Research Data
ARL Research Library Leadership Fellows' (RLLF) Research Projects, Fall Membership Meeting
11 October 2012

Geneva Henry, Rice University
Joy Kirchner, University of British Columbia
Ann Campion Riley, University of Missouri – Columbia
MacKenzie Smith, University of California, Davis
Overview

- Definition of Data Governance
- Legal/Policy Issues
- Technology Landscape
- Case Study Discussion
What is Data Governance?

- Data governance is the system of decision rights and responsibilities that describe who can take what actions with what data, when, under what circumstances, and using what methods.
  - Laws and policies associated with data
  - Strategies for data quality control and management
  - Processes to insure important data are formally managed, including business processes and risk management

- Data governance ensures that data can be trusted and that people are made accountable for actions affecting the data.
Why is Data Governance Important?

- Sharing and integrating research data is increasingly a common practice
  - Funding agency mandates
  - International and interdisciplinary research-intensive collaborations
  - Reproducibility of research results
Issues with Sharing Research Data

- Challenges with being able to share data
  - Legal
  - Cultural
  - Policy

Michael Westley, Creative Commons Licence
Legal Issues

- Data Ownership

- Factual data is in the public domain in the US, but may fall under intellectual property law in another country

- Database vs. its content: one is copyrightable, the other is not

- Some institutions may assert ownership over research data, others may leave with the researchers
3 Legal Mechanisms for Sharing Data

- Contracts
- Public Licenses
- Waivers
Cultural Norms

- Researchers share information with each other, but how it is shared matters

  - Data can be released, but not necessarily reusable

  - Funders would like data to be reused

  - Support for reusable/re-purposable data requires policies to support this
Policies

- Funders starting to develop policies to support data sharing
  - Data management plans
  - Requirements to make data openly available
Technology Landscape

- Need machine-processable information on what can be done with data
  - Media types and formats
  - Metadata
  - Identifiers
  - Persistence strategies

Some rights reserved by sean dreilinger
Data Governance Case Study