



Data Governance

Understanding the Issues and Rights Associated With
Your Research Data

Scholarly Communications Brown Bag Series
25 April 2012

Geneva Henry
Executive Director, Center for
Digital Scholarship
Fondren Library, Rice University

Overview

- Definition of Data Governance
- Legal/Policy Issues
- Technology Landscape
- Recommendations for Moving Forward with Data Governance



Acknowledgments

- All the participants in the Data Governance Workshop held in Arlington, VA 14-15 December 2011
 - Workshop conveners MacKenzie Smith, Trisha Cruse and William Michener
 - Presenters Carly Strasser, Sarah Pearson and Jonathan Rees
- Rice/HAM-TMC/University of Houston Libraries Scholarly Communications Group
- The Friends of Fondren





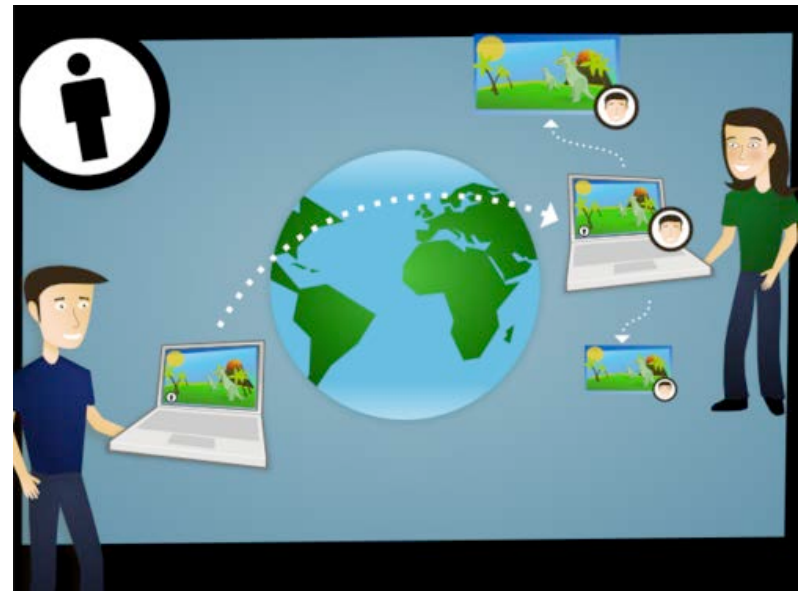
What is Data Governance?



- 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



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



Contracts

- Also known as Data use agreements or data access policies
 - Can implement as “click to agree”
 - Need to be aware of underlying exclusive rights



Contracts

- Benefit: Control
- Risks
 - Perfect control is impossible
 - Control hampers reuse
 - Potential for abuse



Public Licenses

- Requires an underlying exclusive right to be enforceable
- License can be rooted in sui generis or copyright
 - EU enacts sui generis database right, restricting use of content in a database
 - Creative Commons license will cover the content in the database as well as the database itself

cc creative commons
Attribution 3.0 Unported (CC BY 3.0)

This is a human-readable summary of the [Legal Code \(the full license\)](#).
[Disclaimer](#)

You are free:

- to Share** — to copy, distribute and transmit the work
- to Remix** — to adapt the work
to make commercial use of the work

Under the following conditions:

- Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

With the understanding that:

Public Licenses



■ Benefits

- Standardized
- Minimal conditions
- Expansive reach



■ Risks

- Limited reach
- Complexity of copyright law
- Attribution requirement likely does not align well with norms for receiving credit

Waivers

- Surrender all exclusive rights to the data
- Not enforceable in all jurisdictions



Waivers

CC0 1.0 Universal (CC0 1.0) Public Domain Dedication

This is a human-readable summary of the [Legal Code \(read the full text\)](#).

[Disclaimer](#)

No Copyright



The person who associated a work with this deed has **dedicated** the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission. See **Other Information** below.



Other Information

- In no way are the patent or trademark rights of any person affected by CC0, nor are the rights that other persons may have in the work or in how the work is used, such as **publicity or privacy** rights.
- Unless expressly stated otherwise, the person who associated a work with this deed makes no warranties about the work, and disclaims liability for all uses of the work, to the fullest extent permitted by applicable law.
- When using or citing the work, you should not imply **endorsement** by the author or the affirmer.

- Creative Commons created CC0 license

➤ Waiver of all copyright and related rights

➤ Fall-back license that grants all permissions to the licensed work without any conditions

➤ Contains a non-assertion pledge, where the rights holder promises not to assert rights in the licensed work

Waivers

- Benefits

- Legal Certainty
- Interoperability
- No Lawyers 😊

- Risk: lack of control

- BUT citation norms are self-enforcing and will generally work just fine



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



NATIONAL ENDOWMENT FOR THE
Humanities



Policies

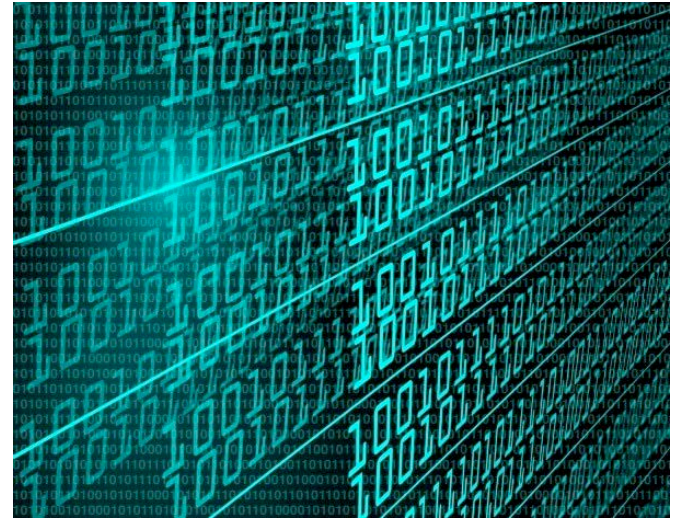
- Research institutions required to have research data policy
 - Policies often tied to funding
 - Failure to comply can result in institution not getting additional grants
 - High risk with non-compliance leads institutions to assert ownership over data



Technology Landscape



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



Recommendations for Moving Forward with Data Governance





Moving Ahead



1. Create a Data Governance Interoperability Panel as an open, participatory and community-driven process.
2. Develop model practices for policies and legal practices related to research data and clarify the legalities of ownership of different types of data.
3. Encourage institutions to provide researchers with the resources needed for good data management planning (e.g., infrastructure for data preservation, data identifiers, or appropriate waivers or licenses).

Moving Ahead



4. Define metadata standards to describe data types and properties (including terms of use), driven by development of a data taxonomy.
5. Provide more education about data governance. Can start with authoring a Wikipedia article on the subject and collecting locally-developed teaching materials that promote data curation to researchers.
6. Refine existing tools such as DMPTool to include policy-awareness such as suggesting specific waivers or licenses available to researchers to share their data.



Summary

- Unlike other types of research outputs, primary research data is not yet well-understood as a research asset or intellectual property.
- Need to address important issues such as:
 - How data fits into the scholarly record
 - Who has responsibility for the data and can set policies for its management
 - What is the relationship between policy and actions by researchers
- The conversation must continue



THANK YOU!



ghenry@rice.edu

