Policies and Regulations

The good stuff is from Carole Palmer and Cheryl Thompson

Data Policies & Regulations

V1: Definitions, types, scope, issues

V2: As an example: *A whirlwind tour of some science-oriented data policies*

V3: Privacy

V1: Definitions, types, scope, issues

(see title of slide)

Definitions

- **Policies** are "…*a set of ideas or a plan for action*
  followed by a business, a government, a political party, or a group of people"
         – Cambridge Dictionary

- "A **policy** is a deliberate system of principles to guide decisions and achieve rational outcomes. A policy is a *statement of intent*, and is implemented as a procedure or protocol."
         – Wikipedia

Types of (data) policies

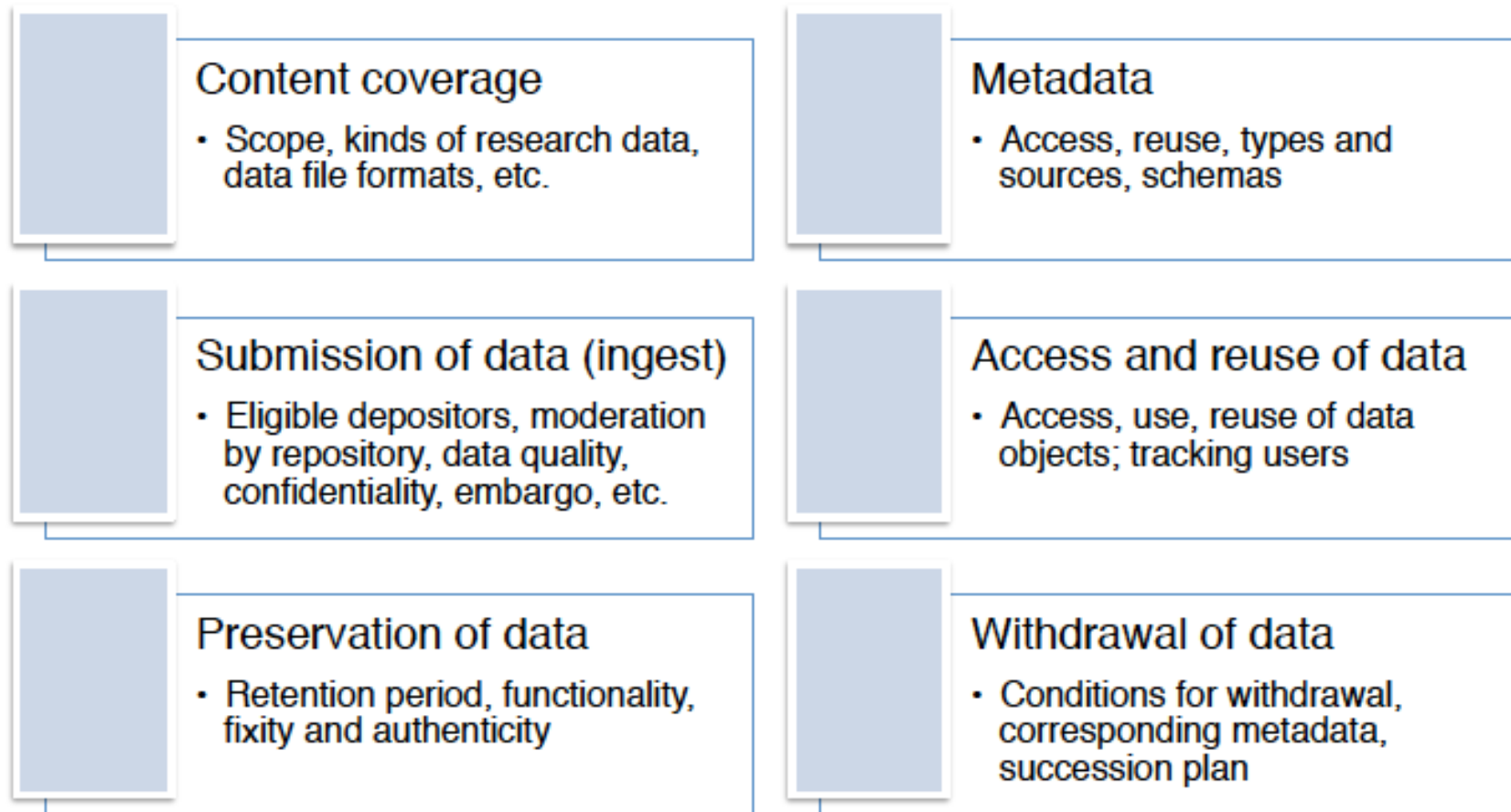**Regulatory**:

…to play you _must_ follow these rules

»   Personal health information – HIPAA

»   Educational records – FERPA

»   Financial information – Gramm-Leach-Bliley Act

Exhortative:

we _advise you_….to adopt a practice or standard, record this information, use this format, etc.

»   NSF Data Sharing policy

»   UIUC Open Access policy

»   American Meteorology Society's Data Archiving and Citation Policy

»   USGS policy on metadata (SM 502.7)

What data policies cover?

Content coverage
- Scope, kinds of research data, data file formats, etc.

Metadata
- Access, reuse, types and sources, schemas

Submission of data (ingest)
- Eligible depositors, moderation by repository, data quality, confidentiality, embargo, etc.

Access and reuse of data
- Access, use, reuse of data objects; tracking users

Preservation of data
- Retention period, functionality, fixity and authenticity

Withdrawal of data
- Conditions for withdrawal, corresponding metadata, succession plan
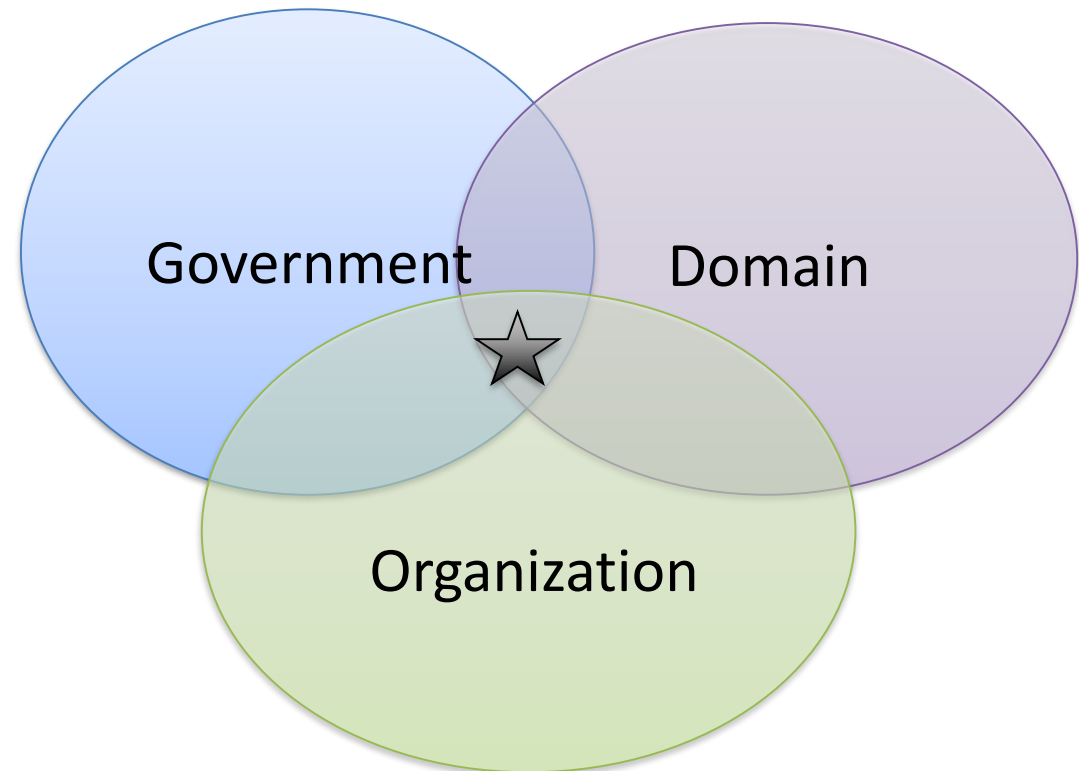
Green, MacDonald, & Rice, 2009

Scope of policies

- Government

- Professional associations

- Consortium

- Company

- Project/Team

Data work is often at the intersection of many policies & authorities

Government   Domain

Organization

The international regulatory maze

Imagine a project with multiple kinds of data and data sources. . .

- One part of that project

  - Is located in one country

  - with funding from two other countries

  - about citizens in yet a third country

  - drawing input from five countries

  - shipping data for analysis to yet another country

  - with the results accessed from any country

- And the other parts of the project are similarly complex

How do you comply with all the security, privacy, intellectual property, format etc. policies?

How do you even discover them?!

An international legal/technical challenge, with important opportunities (and $$$) at stake