

Managing Your Research Data at IUB

Stacy Konkiel
E-Science Librarian
skonkiel@indiana.edu

9/21/2012



INDIANA UNIVERSITY



Overview

- Introduction
- Funding Agency Mandates
- IU Resources
 - Preparing your data
 - Storage (Filesystem & Archival)
 - Access
 - Preservation
- Q&A



Introduction

Stacy Konkiel, E-Science Librarian

skonkiel@indiana.edu

- Background in Scientific Publishing & Digital Research Preservation
- Scholarly Communication Department, Libraries
- Coordinate IU Data Management Service



Introduction

IUB Data Management Service

- Preparing your data
- Storage (HIPAA-compliant)
- Preservation
- Access
- Data Management Plan consultations
 - Lab-based
 - Funder required



Funding Agency Mandates: Overview

A large, dark red, semi-circular graphic element on the left side of the slide, partially overlapping the text boxes. It has a white outline that follows its inner curve.

Data Management
Plans (DMPs)

Data Access
("Open Data")



INDIANA UNIVERSITY

Funding Agency Mandates: Overview

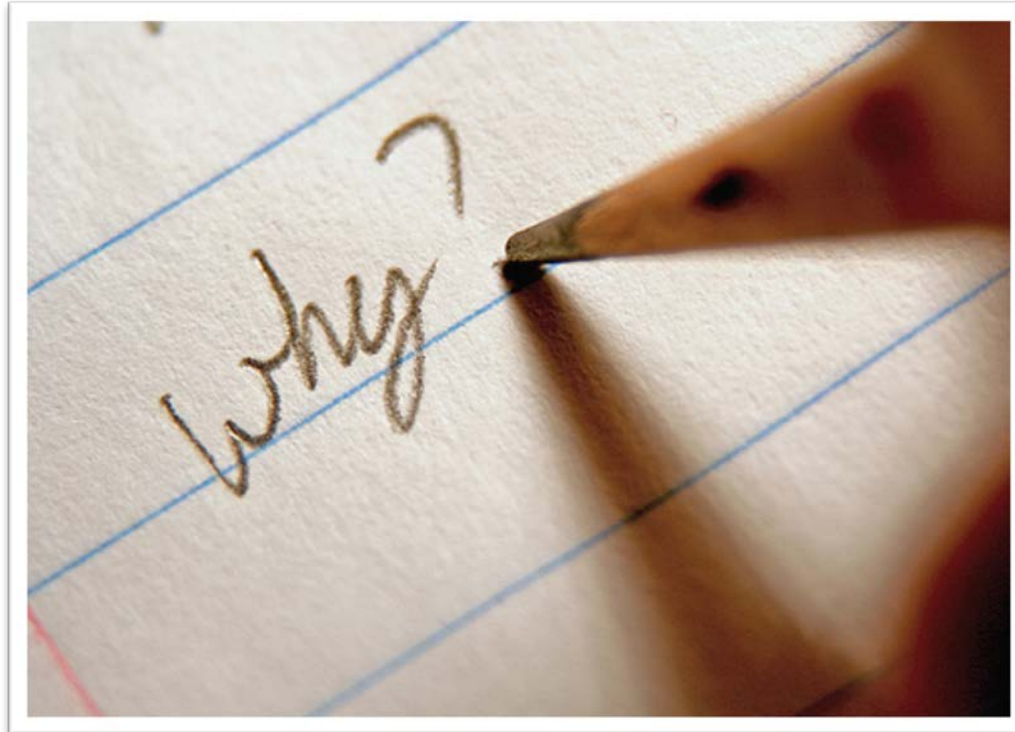


Image: <http://bit.ly/PUztjJ>

9/21/2012



Funding Agency Mandates: Overview

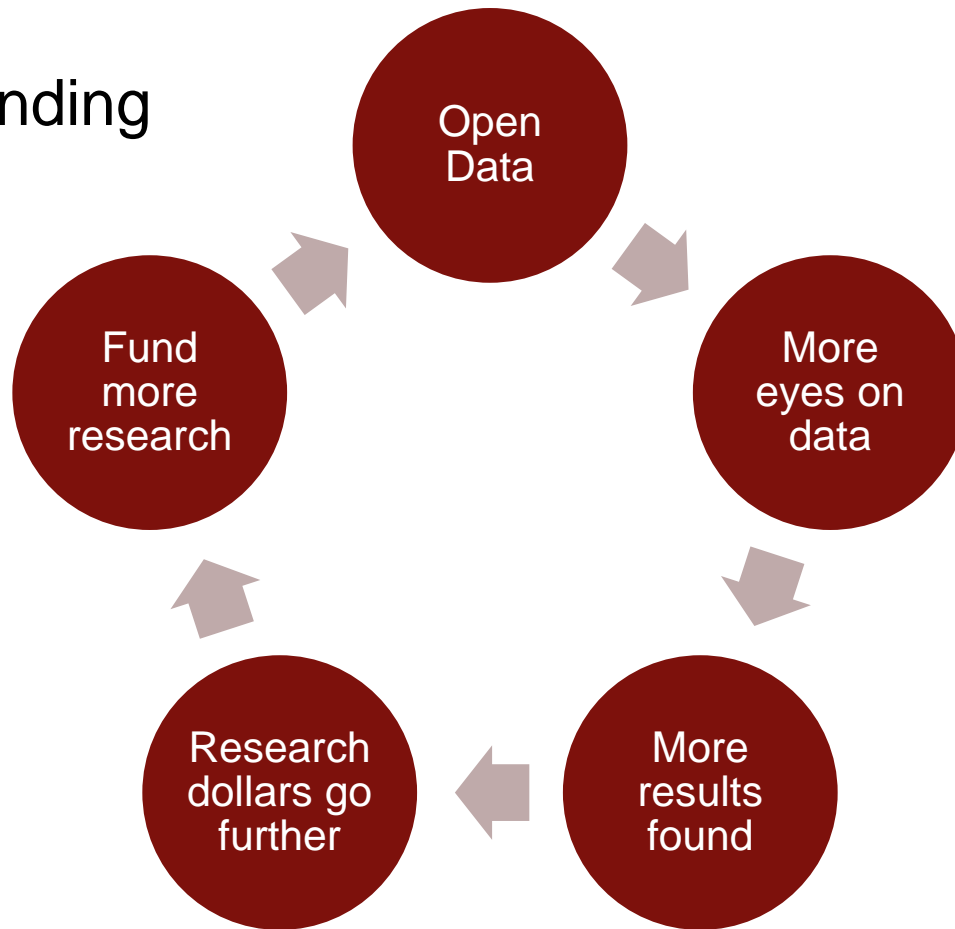
- Organization = Easier Work
- Replicated Data = Safe(r) Data
 - Digital data is more fragile than analog data
- Open Data = More Citations (Piwowar et al, 2010)
- Accelerate discovery

([Houston, 2011](#))



Funding Agency Mandates: Overview

ROI for funding agencies





Funding Agency Mandates: IU Solutions

IUB Data Management Service

- DMP Consultations
- Cyberinfrastructure for data management

Proposal Development Services (OVPR)

- Finding funding
- Grant writing assistance

Office of Research Administration (ORA)

- Grant compliance

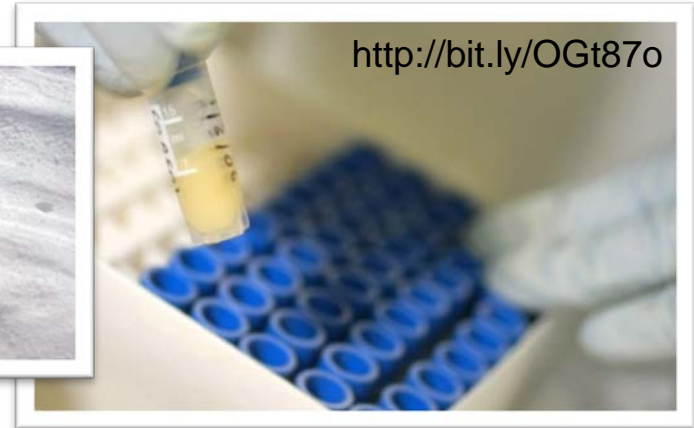
Responsible Conduct of Research (RCR) & DM

- Poynter Center for Research Ethics
- RCR classes via ORA



Funding Agency Mandates: What to address

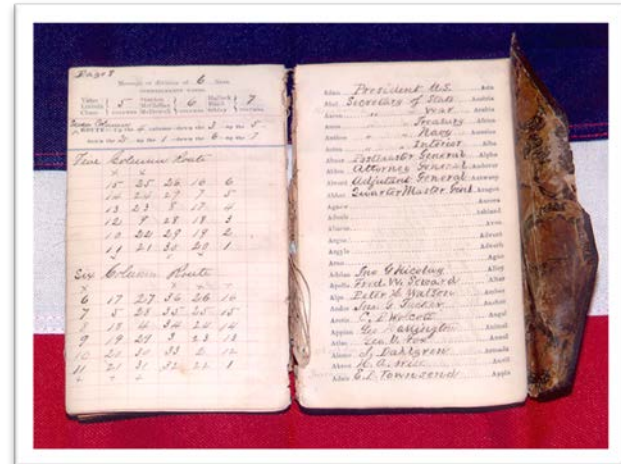
Types of Data



<http://bit.ly/OGt87o>



<http://bit.ly/PBwbA5>



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	9	2		None	None	0	Right	Scale
2	female	Numeric	9	2	The gender of the student.	{00, male}...	None	0	Right	Nomina
3	race	Numeric	12	2		{1.00, hispanic}...	None	0	Right	Nomina
4	ses	Numeric	9	2		{1.00, low}...	None	0	Right	Nomina
5	schtyp	Numeric	9	2	The type of school the student attended.	{1.00, public}...	None	0	Right	Nomina
6	prgtype	String	8	0		None	None	0	Left	Nomina
7	read	Numeric	9	2	reading score	None	None	0	Right	Scale
8	write	Numeric	9	2	writing score	None	None	0	Right	Scale
9	math	Numeric	9	2	math score	None	None	0	Right	Scale
10	science	Numeric	9	2	science score	None	None	0	Right	Scale
11	sect	Numeric	9	2	social studies score	None	None	0	Right	Nomina

<http://bit.ly/OFKpTP>



Funding Agency Mandates:

What to address

Formats



- Community standards
- Commonly used
- Interoperable



Funding Agency Mandates:

What to address

Metadata



Ex. Dublin Core

Title: Christina's World

Creator: Wyeth, Andrew

Date: 1948

Subject: Painting,
American Artists

Format: Painting

Tempera on panel
32 1/4 x 47 3/4"

Provenance: "Stolen in
1999; recovered by the
Museum in 2003."



Funding Agency Mandates:

What to address

Metadata



Image via <http://aymary.wz.cz/members.htm>

Ex. Darwin Core

ScientificNameID:

Bolborhynchus aymara

DecimalLatitude: -23.8169444

DecimalLongitude: -65.4847222

Year: 2005

IdentificationID: 52356

Preparations: tissue, round skin,
other



Funding Agency Mandates:

What to address

Privacy & Confidentiality

- Interrelated with issue of **access**
- Subject to IRB regulations?

Reuse & Distribution Policies

- Also subject to IRB regulations
- How do you want others to...
 - Use your data? (Non-commercial only?)
 - Credit your work?
 - Share your work with others?





Funding Agency Mandates: IU Solutions

Determining Types of Data, Formats, Metadata

- E-Science Librarian
- Metadata Analyst & other library metadata specialists

Privacy, Confidentiality, Reuse & Distribution Policies

- [IU Research Policies](#)
- Legal Counsel



Funding Agency Mandates:

What to address

Data Storage

- 3 copies in separate locations
- **Yes**
 - Stable, short and long term storage for life of project+
 - Attention to sensitive data issues
 - Departmental/University tech support
- **No**
 - Local storage (on lab computers, personal laptops, thumb drives)



Funding Agency Mandates: IU Solutions

Data Storage (“filesystem” storage)

- Research File System (UITS)

100 GB

Individual and team accounts

HIPAA-aligned, not encrypted

Regularly backed up

Small files that are updated frequently



Funding Agency Mandates: IU Solutions

Data Storage (“filesystem” storage)

- **Data Capacitor (UITS)**

Big Red, Quarry, Mason

10 TB

Not backed up, regularly cleared

Large files



Funding Agency Mandates: IU Solutions

Data Storage (“filesystem” storage)

- **XSEDE** (multi-institution)

Visualizations

Software

“Science Gateways”

Big Data

For life of project (some persistent storage in SGs)



Funding Agency Mandates: IU Solutions

Data Storage (“filesystem” storage)

- **IU CTSI “Service Cores”** (IU, PU, ND)
 - Subject-specific (social and physical sciences)
 - Storage space & capabilities vary
 - Computing and project planning
- **Advanced Information Technology Core (UITS)**
 - Electronic personal health information (ePHI)
 - HIPAA-protected information
 - Biomedical research
 - Paid service



Funding Agency Mandates: IU Solutions

Data Storage (“archival” storage)

- **IU Box**

50 GB

Syncs with local storage

Permanent

- **Scholarly Data Archive (SDA)**

Where you back up RFS and

Data Capacitor files

50 PB

Permanent tape storage





Funding Agency Mandates:

What to address

Access

- What will you share?
- How will you share it?
- What provisions will be made to share data securely?



Funding Agency Mandates: IU Solutions

Access

- **IU Box**

 - Role-based access

 - Share with anyone

 - No HIPAA, FERPA, FISMA, GLBA, ePHI data



- **IU Alfresco Share**

 - Role-based access

 - Multi-institution

 - No ePHI data allowed





Funding Agency Mandates: IU Solutions

Access

- **IUScholarWorks Repository**

Open Access repository

Indexed by Google, et al

No sensitive data unless de-identified, etc



- **WinSCP**

SFTP

Access via username and password

Desktop client





Funding Agency Mandates:

What to address

Preservation

- As formats become obsolete, who will migrate data?
Think .BMP to .JPG to .PNG
- As media degrades, who will transfer data?
- In the future, who will know enough about your research (or even your field) to work with your data?
Preservation metadata = Rosetta Stone



Funding Agency Mandates: IU Solutions

Preservation

- **IUScholarWorks Repository**

- Storage in perpetuity

- Format migration

- Interface for OA data stored on SDA

- File integrity validation, embargoes

- Managed by IU Libraries

- **Scholarly Data Archive (SDA)**

- Permanent, perpetual tape storage

- Mirroring & redundancy



Search

Search input field with 'Go' button

- Entire Repository
This Community

Advanced

Browse

Entire Repository

- Communities & Collections
By Issue Date
Authors
Titles
Subjects

This Community

- By Issue Date
Authors
Titles
Subjects

My Repository

- Login
Register

LEADII Vortex2 Archive Dataset

Show full item record

Metadata table with fields: Title, Author, Date, Date(s) Covered, Geographic / Spatial Information, Methodology, File Information, Location, Type, Embargo release date.

Coverage dates

Geographic info

Methodology

File info

Embargo info

External Files

- List of external file URLs for the dataset archive.

SDA link(s)



Funding Agency Mandates: IU Solutions

Preservation

- **Librarian Consultations**

Best home for research (IU and non-IU)

Preservation and metadata consultations



INDIANA UNIVERSITY

Q&A

Download this presentation at:

> <http://hdl.handle.net/2022/14682> <

Contact

Stacy Konkiel

skonkiel@indiana.edu

IU Data Management Service: <http://scholarworks.iu.edu/data>

IU Libraries Data Management Guide: <http://www.libraries.iub.edu/data>