INFRASTRUCTURES AND PLATFORMS: DATA CURATION IN THE DIGITAL HUMANITIES

Lydia Zvyagintseva
Forum for Information Professionals, Feb. 7, 2014
Problem

- What is the relationship between research data management and data curation as a set of scholarly practices?

- What does data curation in the digital humanities look like as library service?
The Research Universe

research

research data

research data mgmt

dc

data curation
Data: What is it?
Actually...
Design

- Literature review
- Website review
- Use cases
Definitions (JISC)

- Curation – managing and promoting the use of data; continuous enrichment and updating
- Archiving – ensuring that data is properly selected, stored, maintained over time
- Preservation – long-term storage and accessibility through successive change and obsolescence
Research Data Management

- Larger set of practices that includes institutional policy, planning, funding (local, regional, national, international), strategy development.

- Practical day-to-day of consultation, selection, acquisition, description, access, maintenance, preservation, storage, sharing, marketing, instruction.
Why humanities?

- Provide access to cultural information
- Enable manipulation of data
- Transform scholarly communication
- Enhance teaching and learning
- Make a public impact
In other words...

OPEN ACCESS
Use Cases (Applications)

- DataStaR @ Cornell
- ScholarSphere @ Penn State
- DRYAD
- Figshare
- Readux @ Emory
Repositories

Cornell University Geospatial Information Repository (CUGIR) | GIS data repository | CUGIR website
DataStaR (Data Staging Repository) | Cornell institutional repository | DataStaR
eCommons@Cornell | Cornell institutional repository | eCommons@Cornell
Knowledge Network for Biocomplexity (KNB) | Disciplinary data center | Knowledge Network for Biocomplexity
New York State GIS Clearinghouse | GIS data repository | NYS GIS clearinghouse
Data sharing is vital to ensuring the scholarly record for many reasons:

- **It is required** by NSF and NIH and likely to be adopted by other funders.
- **It fosters the reuse of data**: Allowing researchers access to your data opens up possibilities for new readings, new interpretations, and new discoveries.
- **It benefits research**: Being open about data can lead to sharing of process, workflows, and tools, all of which enhance the potential for replicating the research results.
- **It is good science**: Sharing data frames them as a public good, promoting community, collaboration, and conscientiousness.

---

**Research Data Management Services Team (RDMST)** at the University Libraries

The Research Data Management Services Team (RDMST) at the University Libraries offers consultation services for data management planning to researchers applying for grant funding from agencies requiring a data management plan (DMP).

All data management inquiries should be addressed to RDMST at: UL-RDMST@lib.psu.edu
Influence and impact

- Collection development
- Organization of knowledge (taxonomy, metadata)
- Systems and IT
- Instruction and outreach
- Reference and information seeking
- Leadership and organizational culture
Integrated services

Figure 2: The research data management pyramid for libraries
Thanks!

- lydiaz@ualberta.ca
- @lydia_zv