

Building community Externally: Using the Desirable Characteristics to Evaluate our Repository

Jon Petters

Presented August 6, 2024@

Summit for Academic Institutional Readiness in
Data Sharing (STAIRS)

DATA CURATION NETWORK

DATA CURATION NETWORK

DESIRABLE CHARACTERISTICS OF DATA REPOSITORIES FOR FEDERALLY FUNDED RESEARCH

Jonathan Petters, Virginia Tech

About the Desirable Characteristics (DC-DR)

- From the National Science and Technology Council Subcommittee on Open Science
- Adopted as guidance in several funding agency public access plans
- 14 Characteristics across the three themes of
 - Organizational Infrastructure
 - Digital Object Management
 - Technology
- 6 Additional Characteristics for Repositories Storing Human Data

<https://tinyurl.com/dc-drepo>

Researchers publicly sharing data have a common question:

- Where should I deposit my data?

This can be further broken into two general questions:

- Is there a place/service I should/could use?
- Is this place/service option a 'good' one?

Do Institutional Data Repositories Have these Characteristics?

- Surveyed 17 Data Curation Network repositories in last half of 2022
- Organization Infrastructure - Yes, apart from Risk Management
 - Most do not accept sensitive research data
- Digital Object Management - Yes (mostly)
 - Most/many take curation actions on deposited data but do not take quality assurance actions

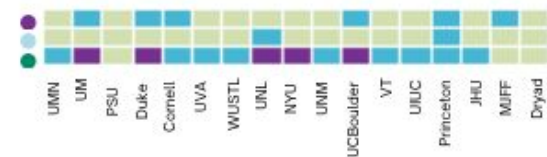
Do Institutional Data Repositories Have these Characteristics?

- Surveyed 17 Data Curation Network repositories in last half of 2022
- Technology - Partially
 - Security appropriate to data, sustainability/preservation are priorities
 - Connecting published data to creators less adopted
- Storing Human Data - N/A
 - Only de-identified human data accepted

Do Institutional Data Repositories Have these Characteristics?

TECHNOLOGY

- Authentication 79%
- Long-term Technical Sustainability 94%
- Security & Integrity 29%



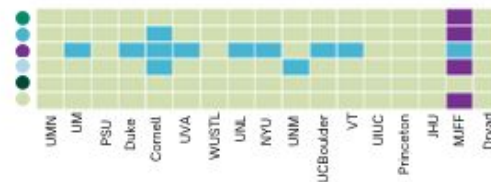
ORGANIZATIONAL INFRASTRUCTURE *

- Free & Easy Access 100%
- Clear Use Guidelines 91%
- Retention Policy 79%
- Long-term Organizational Sustainability 85%



DIGITAL OBJECT MANAGEMENT

- Unique Persistent Identifiers 94%
- Metadata 91%
- Curation & Quality Assurance 74%
- Broad & Measured Reuse 88%
- Common Format 100%
- Provenance 94%



[DCN blog post - S. Connell and S. Wright, March 2024](#)

* Most DCN member repositories do not take human subjects data, so we did not address those additional considerations.

Yes, Institutional Data Repositories Have Many of these Characteristics...

- ...but there can be wide variation in maturity repository to repository (just like domain repositories!)
- Broad similarities in approach of institutional repositories towards their services and role in institution