May 2nd, 2:45 PM - 3:25 PM

Research Data and the Duke Digital Repository: Building a Service-Driven Data Curation Workflow

Mara Sedlins
*Duke University*, mara.sedlins@duke.edu

Follow this and additional works at: [https://digital.sandiego.edu/symposium](https://digital.sandiego.edu/symposium)

*Part of the Library and Information Science Commons*


This Poster is brought to you for free and open access by Digital USD. It has been accepted for inclusion in Digital Initiatives Symposium by an authorized administrator of Digital USD. For more information, please contact digital@sandiego.edu.
Research Data and the Duke Digital Repository: Building a Service-Driven Data Curation Workflow

Session Type
Poster

Abstract
This year, Duke University Libraries launched the Duke Digital Repository (DDR), a service that supports the activities of the University’s faculty, researchers, students, and library staff by preserving, securing, and providing access to digital resources. The DDR is built on three program areas: research data, scholarly publications, and library collections. The research data area of the DDR reflects a new focus on research data management services at Duke, which includes preservation-level storage in the DDR, as well as consulting and instruction. The new program of services is supported by four new staff members (two research data management consultants and two digital repository content analysts), as well as a CLIR postdoctoral fellow. Along with two subject specialist librarians and the head of Data and Visualization Services, we form the Research Data Working Group at Duke Libraries, which provides policy, procedure, and technical recommendations on data curation and management issues related to the library’s repository program. We are currently piloting new services, policies, and workflows to promote data sharing and curation, quality metadata, and reproducible research, offering consultations and workshops in research data management best practices. This poster will report on the organizational structures, data sharing policies, metadata requirements, and workflows for data ingest that are currently in place.

Location
Rotunda / Garden of the Sea

Keywords
institutional repository, data curation, research data management, preservation

This poster is available at Digital USD: https://digital.sandiego.edu/symposium/2017/2017/39
Research Data and the Duke Digital Repository: Building a Service-Driven Data Curation Workflow

Research Data Working Group at Duke Libraries: Jennifer Darragh, Moira Downey, Elena Feinstein, Arianne Hartsell-Gundy, Joel Herndon, Susan Ivey, Sophia Lafferty-Hess, & Mara Sedlins

Duke Digital Repository

The Duke Digital Repository (DDR) is a service of Duke University Libraries that supports the activities of the University’s faculty, researchers, students, and library staff by preserving, securing, and providing access to digital resources.

Collection Areas

- Research Data
- Scholarly Publications
- Library Collections

Duke Libraries Information Technology Services

- Data and Visualization Services
- Digital Production and Curation Services
- Software Development and Integration Services

Senior Research Data Management Consultants

- Jennifer Darragh
- Moira Downey
- Sophia Lafferty-Hess
- Susan Ivey

Repository Content Analysts

Research Data Curation Staffing Model

Fall 2015: Digital Research Faculty Working Group convened, made recommendations to expand research data services¹

Fall 2016: CLIR Fellow in Data Curation co-sponsored by Duke Libraries and the Duke Social Science Research Institute

Winter 2017: Four new positions hired to support research data curation at Duke Libraries

Research Data Curation Workflow

Planning

- Initial Deposit Consultation
- Deposit Coordination

Curation Review

- Content Analysts
- Archival Description and Curation
- Public data & entry in DRI
- Submission

Content Review Checklist

- Review the Data Deposit Form to ensure metadata is clear and understandable.
- Perform a visual inspection of data files for the presence of PII or PHI. The liability for the presence of any sensitive information will remain with the depositor.
- Check that the data package includes adequate documentation to allow for others to interpret and use the data files.
- Check to see if files are stored in preservation friendly formats.
- Optional: have subject liaison librarians review the data package to provide any other data documentation suggestions.


Definition of Research Data

Research data are the original sources or materials (born digital or converted to digital) that were created or gathered in the process of your research. They serve as the foundation from which you draw conclusions and produce results/findings (test hypotheses, study trends, provide evidence, refute claims). They may be numeric or qualitative, structured or unstructured. Among many possible forms, data may take the form of notebooks, statistical or spatial data tables, audio or visual recordings, photographs or models. The end-results/findings of your research such as monographs, articles, white papers, or presentations are not considered research data.

Training and Outreach Initiatives

Fall 2016: CLIR Fellow conducted informal interviews with social science NSF grant recipients about data management planning

Winter 2017: Repository launch - presentations at campus Research Computing Symposium

New Research Data Management workshops piloted:

- Data Management Fundamentals
- Documenting Data in Excel with Colectica
- Data Management and Reproducibility: Enabling Open and Transparent Research through Data Sharing
- Consent, Data Sharing, and Data Reuse
- Data Management Tools: The Dataverse Project
- Research Data Management for the Social Sciences
- Data Management and Grants: Complying with Mandates

Ongoing: RDM Consultant meetings with various campus stakeholders

repository.duke.edu