Evolution of a Research Data Management and Curation Program

Candid Reflections and Considerations

Sophia Lafferty-Hess and Moira Downey
Duke University Libraries
CNI Fall Meeting, 2019
Outline

- Background
- Reflections
  - Understanding the context
  - Building a platform and workflow
  - Communicating with researchers
  - Taking stock and measuring success
Background

Scholarly Communications & Research Landscape

Research Data Management

Open Science

Curation

Reproducibility
Background

2013 ARL SPEC Kit on Research Data Management (RDM) Services

74% Providing research data management services*

* Response rate - 73 out of 125 member institutions

Background

2017 ARL SPEC Kit on Data Curation

64% Providing data curation services*

* Response rate - 80 out of 124 member institutions

Background

2016
Duke Digital Research Faculty Working Group Report

2017
Four new positions hired to support RDM & Curation
Background

2017 - Present

Education Program
Lifecycle Services
Curation Services
Understanding the context

Scanned the environment, met with key campus partners, and built relationships.
Understanding the context

Reflection upon the university culture provides insight into motivations and new initiatives.
Understanding the context

Campus partnerships are key to avoid duplication of efforts and silos and cultivate ongoing relationships.
Understanding the context

Library culture provides opportunities for enhanced collaboration

- Digital Preservation & Publishing Program
- Research Data Working Group
- Digital Preservation Working Group
- Graduate Student Instructional Program
Understanding the context

Library culture provides opportunities for enhanced collaboration **but** it also takes more time and negotiations for resource allocation.
Building a platform and workflow

Built a submission and curation workflow informed by community standards (DCN) and best practices around data curation (FAIR guidelines).
Curation Workflow

Deposit
- ✓ Submit data, documentation, & metadata
- ✓ Check documentation completeness
- ✓ Identify other data/deposit enhancements
- ✓ Assess disclosure risks
- ✓ Assess formats

Review
- ✓ Check documentation completeness
- ✓ Identify other data/deposit enhancements
- ✓ Assess disclosure risks

Ingest
- ✓ Transform files
- ✓ Arrange files
- ✓ Generate metadata
- ✓ Assign license
- ✓ Generate checksums

Publish
- ✓ Assign DOI
- ✓ Review & test
- ✓ Finalize administrative metadata
Building a platform and workflow

The workflow guided development of the data repository, including local customizations.
Building a platform and workflow

Open source software development represents our values but open source software development is hard.
Building a platform and workflow

Cooperative models are useful to address capacity and expertise gaps.
Communicating with researchers

Be explicit about the value your services are adding and highlight things researchers understand.
Communicating with researchers

Employed a multi-modal approach to reach our campus community.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Workshops</td>
<td>41</td>
</tr>
<tr>
<td>Targeted Presentations</td>
<td>7</td>
</tr>
<tr>
<td>Public Programming</td>
<td>2</td>
</tr>
<tr>
<td>Invited Events &amp; Presentations</td>
<td>17</td>
</tr>
<tr>
<td>Online Marketing</td>
<td>lots!</td>
</tr>
</tbody>
</table>
Communicating with researchers

Let researchers speak for themselves. Reduce “library speak” and pitching services when possible.
Taking stock and measuring success

How do we measure our success?
Taking stock and measuring success

Curation and consults at Duke by the numbers

Total number of consultations and published datasets, 2017-2020

Campus-wide promotional work escalates

Consultations

Published datasets
Taking stock and measuring success
Curation and consults at Duke by the numbers

Total number of consultations by type, 2017-2019

- General: 12
- Planning: 29
- Publishing: 46
- Workflow: 35

On the rise

Curation services worked into lab workflows; dept word-of-mouth

Total number of deposits by discipline, 2017-2019

- Medical: 13
- Physical sciences: 42
- Social sciences: 1
- Humanities: 0

Ongoing area of concern
Taking stock and measuring success

Numbers are important, but they aren't the whole story.

"...archiving data this way ensures reproducibility of the underlying science and long-term stability of the work, despite the natural turnover in my research group's composition. It thus brings our research practice in compliance with data management requirements of funding agencies, and helps us respond promptly to external requests for our data. In addition, the very process of depositing gives us the opportunity to validate figures one last time, as we copy-edit page proofs." 

(Dr. Patrick Charbonneau, Chemistry & Physics)
Future considerations

- Partnering with researchers on projects
- Addressing use cases around sensitive and large-scale data
- Integrating tools across the research lifecycle
## Acknowledgements

### RDM & Curation Team
- Jen Darragh
- Moira Downey
- Shadae Gatlin
- Sophia Lafferty-Hess

### Research Data Working Group(s)
- Karen Barton
- Ciara Healy
- Ryan Denniston
- Elena Feinstein
- Joel Herndon
- Arianne Hartsell-Gundy
- Lee Sorensen
- Will Sexton

*Additional thanks to all our collaborators within the library and beyond!*
Questions?

Contact us:

sophia.lafferty.hess@duke.edu
@laffertyhess

moira.downey@duke.edu
@moira_downey