

# **Evolution of a Research Data Management and Curation Program**

## **Candid Reflections and Considerations**

**Sophia Lafferty-Hess and Moira Downey**

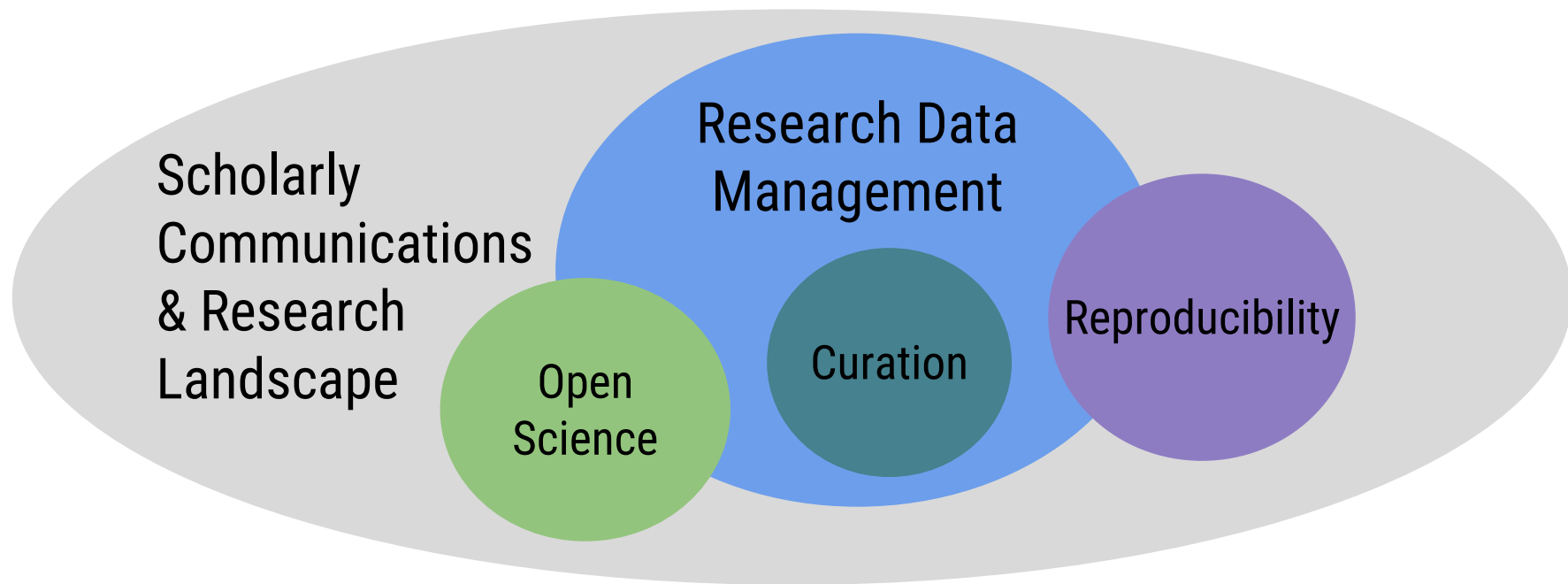
**Duke University Libraries**

**CNI Fall Meeting, 2019**

# Outline

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

# Background



# 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

*Fearon, Jr., D., Gunia, B., Pralle, B., Lake, S., & Sallans, A. (2013). SPEC Kit 334: Research Data Management Services (July 2013). <https://doi.org/10.29242/spec.334>*

# Background

2017 ARL SPEC Kit on Data Curation

64% Providing data curation services\*

\* Response rate - 80 out of 124 member institutions

Hudson-Vitale, C., Imker, H., Johnston, L., Carlson, J., Kozlowski, W., Olendorf, R., & Stewart, C. (2017). SPEC Kit 354: Data Curation (May 2017). <https://doi.org/10.29242/spec.354>

# 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.

***Duke University to Pay \$112.5M***  
**After massive lawsuit, new VP of research set to fix Duke research**



Photo by Courtesy of Duke Photography | The Chronicle

By Ashwin Kulshrestha

September 11, 2019 | 2:29am EDT



if

Insured data surfaced. Maureen Gray for The New York Times

# 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

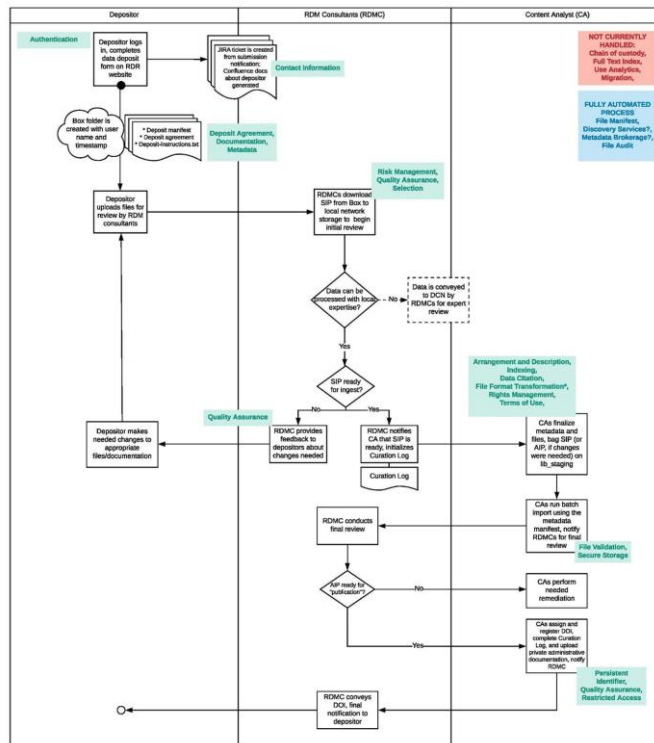
Graduate Student Instructional Program

# Understanding the context

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



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



# Curation Workflow



```
graph LR; Deposit[Deposit] --> Review[Review]; Review --> Ingest[Ingest]; Ingest --> Publish[Publish]; Publish --> End[ ]; Review --> Deposit;
```

## Deposit

- ✓ Submit data, documentation, & metadata

## Review

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

## Ingest

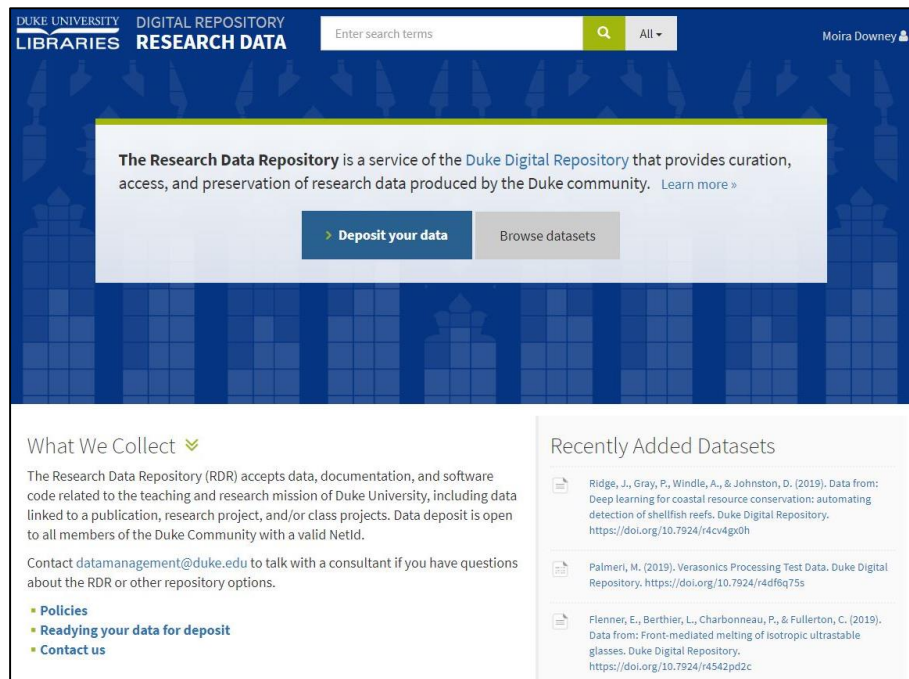
- ✓ Transform files
- ✓ Arrange files
- ✓ Generate metadata
- ✓ Assign license
- ✓ Generate checksums

## Publish

- ✓ Assign DOI
- ✓ Review & test
- ✓ Finalize administrative metadata

# Building a platform and workflow

The workflow guided development of the data repository, including local customizations.



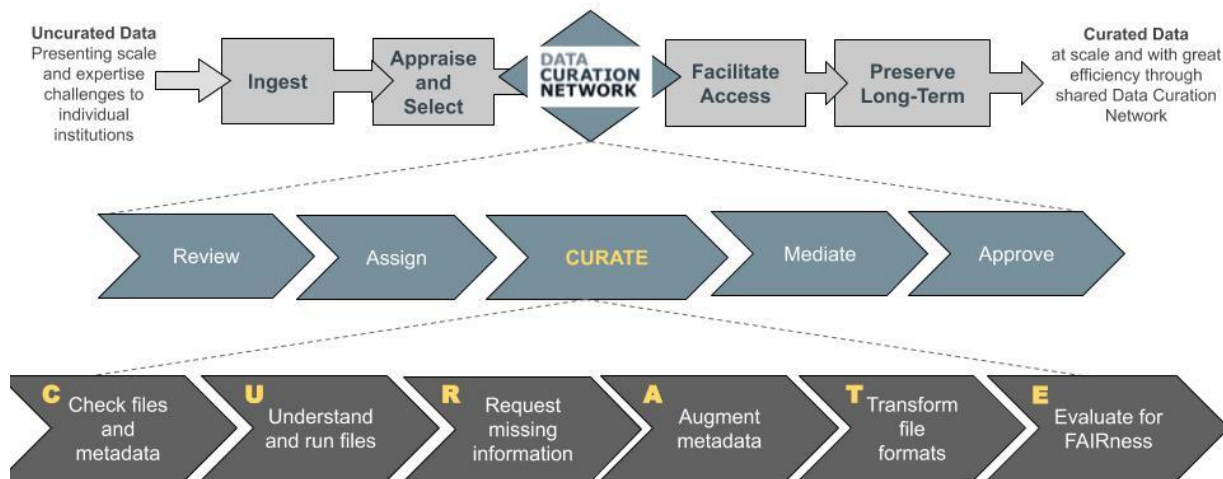
Open source software  
development  
represents our values  
**but** open source  
software development  
is hard.





# Building a platform and workflow

Cooperative models are useful to address capacity and expertise gaps.



Reused with permission from the Data Curation Network

# Communicating with researchers

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



- A place to publish your data!
- Digital Object Identifiers (DOIs) for persistent access
- Review of documentation to optimize reusability
- Check for sensitive or confidential information
- Standardized Dublin Core metadata to improve search and discovery
- File format transformations for longer-term reuse
- Collaboration with depositors to enhance submitted data
- Formal data citations to ensure proper acknowledgement
- Download and pageview analytics for tracking data

# Communicating with researchers

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

Public Workshops	41
Targeted Presentations	7
Public Programming	2
Invited Events & Presentations	17
Online Marketing	lots!

Let researchers speak for themselves. Reduce “library speak” and pitching services when possible.

DUKE UNIVERSITY  
LIBRARIES

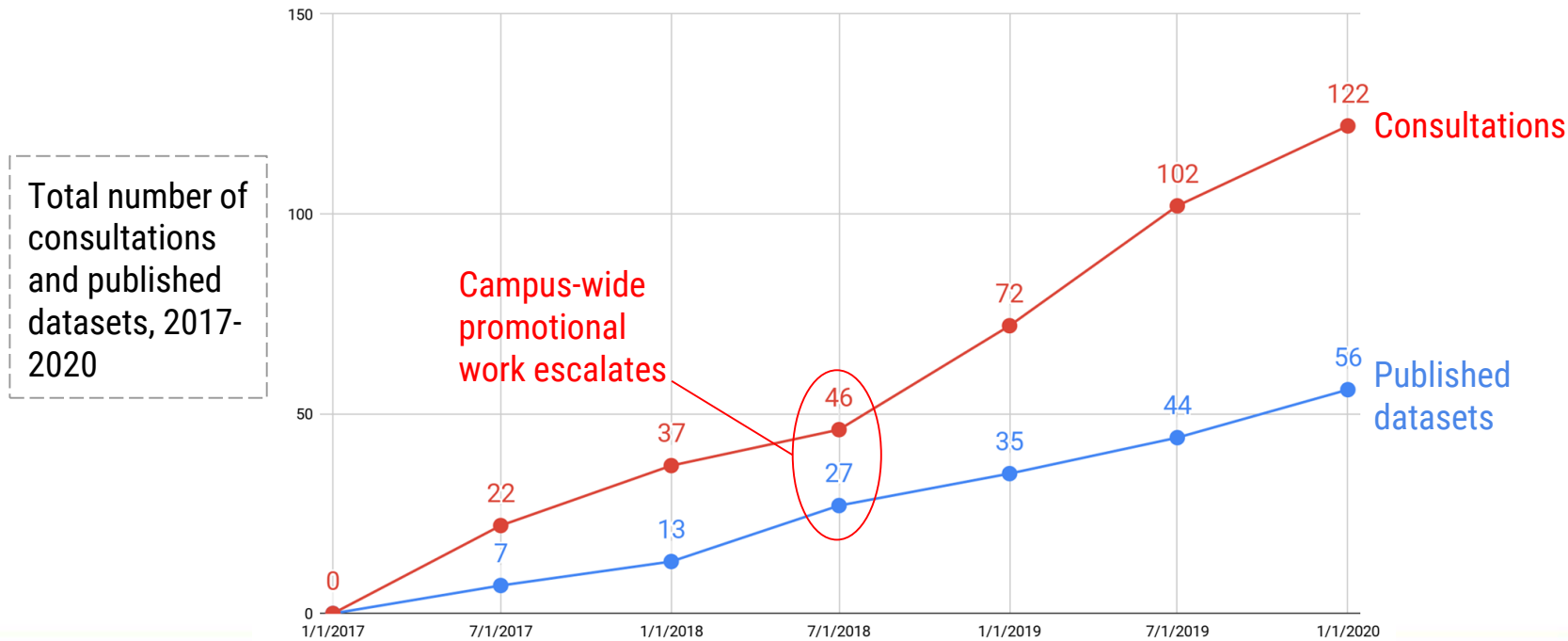
# Taking stock and measuring success



**How do we measure  
our success?**

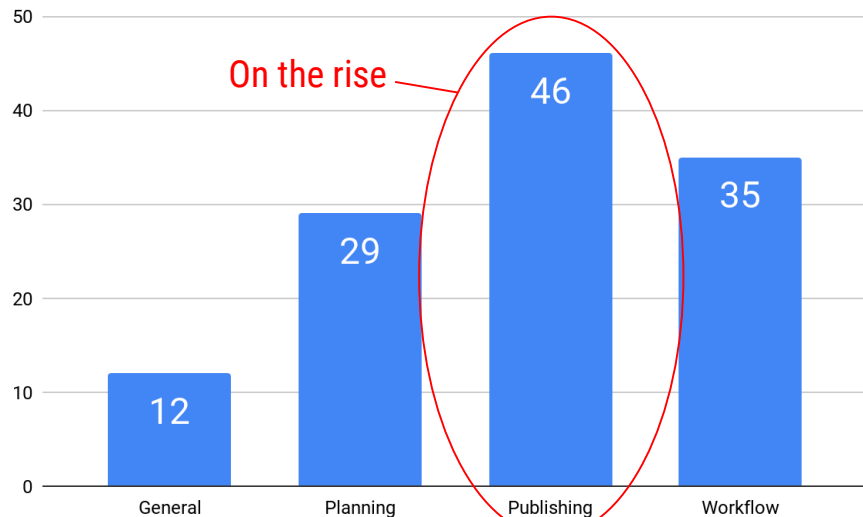
# Taking stock and measuring success

Curation and consults at Duke by the numbers

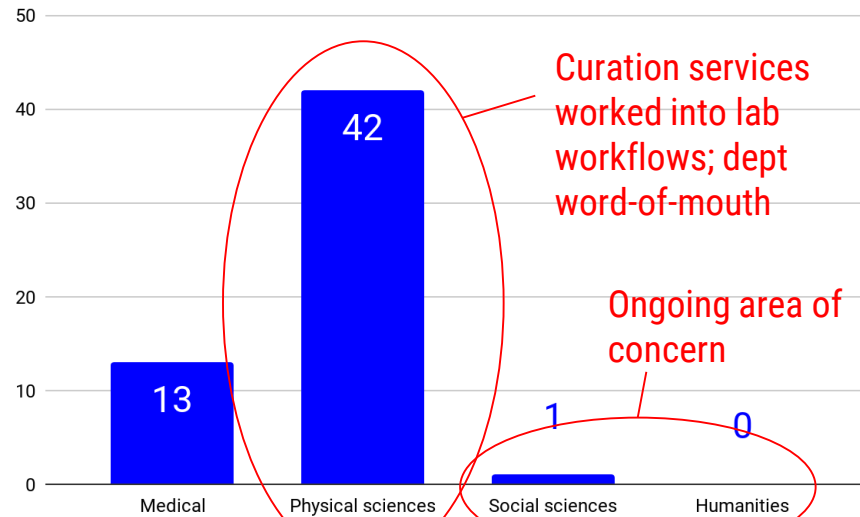


# Taking stock and measuring success

## Curation and consults at Duke by the numbers



Total number of consultations by type, 2017-2019



Total number of deposits by discipline, 2017-2019

# 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)**

The screenshot shows the top of a Duke ScholarWorks article. The header includes the 'Duke ScholarWorks' logo and a navigation bar with links like 'Home', 'About', 'Browse', 'Search', 'Deposit', 'Withdraw', 'FAQ', 'Contact Us', and 'Help'. The article title is 'Making Open Data a Reality' by Prof. Patrick Charbonneau. Below the title is a photo of five people in a lab setting. The article text discusses the importance of open data and reproducibility in research, mentioning the Charbonneau group's work on low-temperature disordered materials and machine learning for crystal structure prediction. It also mentions the group's use of Duke's data management services.

Duke's scholars probe some of the most foundational questions to recognize publications. Professor Patrick Charbonneau, who maintains appointments in both Chemistry and Physics, for instance, has been instrumental in the group's research on low-temperature disordered materials, such as glasses and packing of grains. Prof. Charbonneau and his research group publish articles on the topic of low-temperature disordered materials and teaching a machine to spot a crystal.

Prof. Charbonneau is a proponent of open data and research data sharing and archiving. He has built into his research's workflow a process ensuring that every article published by his group includes a persistent object identifier (DOI) that links to the underlying data, which have been made publicly available in the data repository here at Duke University. This process ensures that the data are available to the community and can be used to validate 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."

As of early 2019, the Charbonneau group has used Duke's



# 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

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

## Research Data Working Group(s)

Karen Barton

Ciara Healy

Ryan Denniston

Elena Feinstein

Joel Herndon

Arianne Hartsell-Gundy

Lee Sorensen

Will Sexton

# Questions?

Contact us:

sophia.lafferty.hess@duke.edu  
@laffertyhess

moira.downey@duke.edu  
@moira\_downey