Radical collaboration: Joining forces to address user needs around research data management

Open Repositories 2019

Radical Collaboration

…coming together
across disparate, but engaged, domains
in ways that are often unfamiliar
or possibly uncomfortable…
to achieve more together
than we could separately

RLI 296 – focus: research data management

https://publications.arl.org/rli296/

Panelists

Overview/facilitator: Nance McGovern – MIT

Contributors:

Elizabeth Hull - Dryad-CDL partnership

Cynthia Hudson-Vitale - Data Curation Network

Limor Peer - CURE

Amy Nurnberger - RDA

Nance McGovern (she/her/hers)

ORCID: *0000-0002-7733-1516*

Director, Digital Preservation - Massachusetts Institute of Technology
Director, Digital Preservation Management Workshops (dpworkshop.org)
Past President, Society of American Archivists
iPres – co-organizer 2018 and co-chair, iPres Working Group

Twitter: @mcgovern60

Research Library Issues
RLI 296

2018

ASSOCIATION OF RESEARCH LIBRARIES

In This Issue

Why Does Research Data Management Need Radical Collaboration? 3

Radical Collaboration and Research Data Management: An Introduction 6

- Radical Collaboration: Framing the Concept
- Using Working Definitions to Build Understanding
- What Makes Collaboration Radical?
- Coming Together
- Sharing the Table
- Emerging Distributed Digital Practice
- Building (an Inclusive) Community
- Some Guiding Principles for Radical Collaboration

The Radical Collaboration of RDA and What It Means for Developing Institutional Data Management Services 23

- The Evolution of Collaboration
- Encouraging Radical Collaboration
- Bringing Everyone to the Table

Research Library Issues
RLI 296
2018

ASSOCIATION OF RESEARCH LIBRARIES

Archives and Data Management: The Purdue Story 33

Barriers to Collaboration: Lessons Learned from the Data Curation Network 37

What is Open Science, and How Can Radical Collaboration Facilitate It? 44

Example 1. Radical Collaboration in Support of Open Science: Software Citation

Example 2. Radical Collaboration to Preserve Informal Astronomical Communications

Seeking Sustainability and Inclusivity with Transparent Practices for RDM 49

Making Research Data Management a Social Activity

Radical Collaboration: An Archival View 53

Disambiguating Digital Archives and Digital Preservation

Digital Practice and Research Outcomes

Considerations for Radical Collaboration That Engages Archives and Archivists

Forward Together 62

Navigating Our Shared Space(s)

What Could Radical Collaboration Look Like?

Example 1. Constructive Research Methodology Overview

Example 2. Storytelling as a Digital Preservation Strategy

How Can We Achieve Radical Collaboration?

Archives, History, and Technology: Prologue and Possibilities for SAA and the Archival Community

Roots and future of distributed digital practice

American Archivist: Spring/Summer 2018, Vol. 81, No. 1, pp. 9-22

https://americanarchivist.org/doi/abs/10.17723/0360-9081-81.1.9

Digital Practice Roundtable

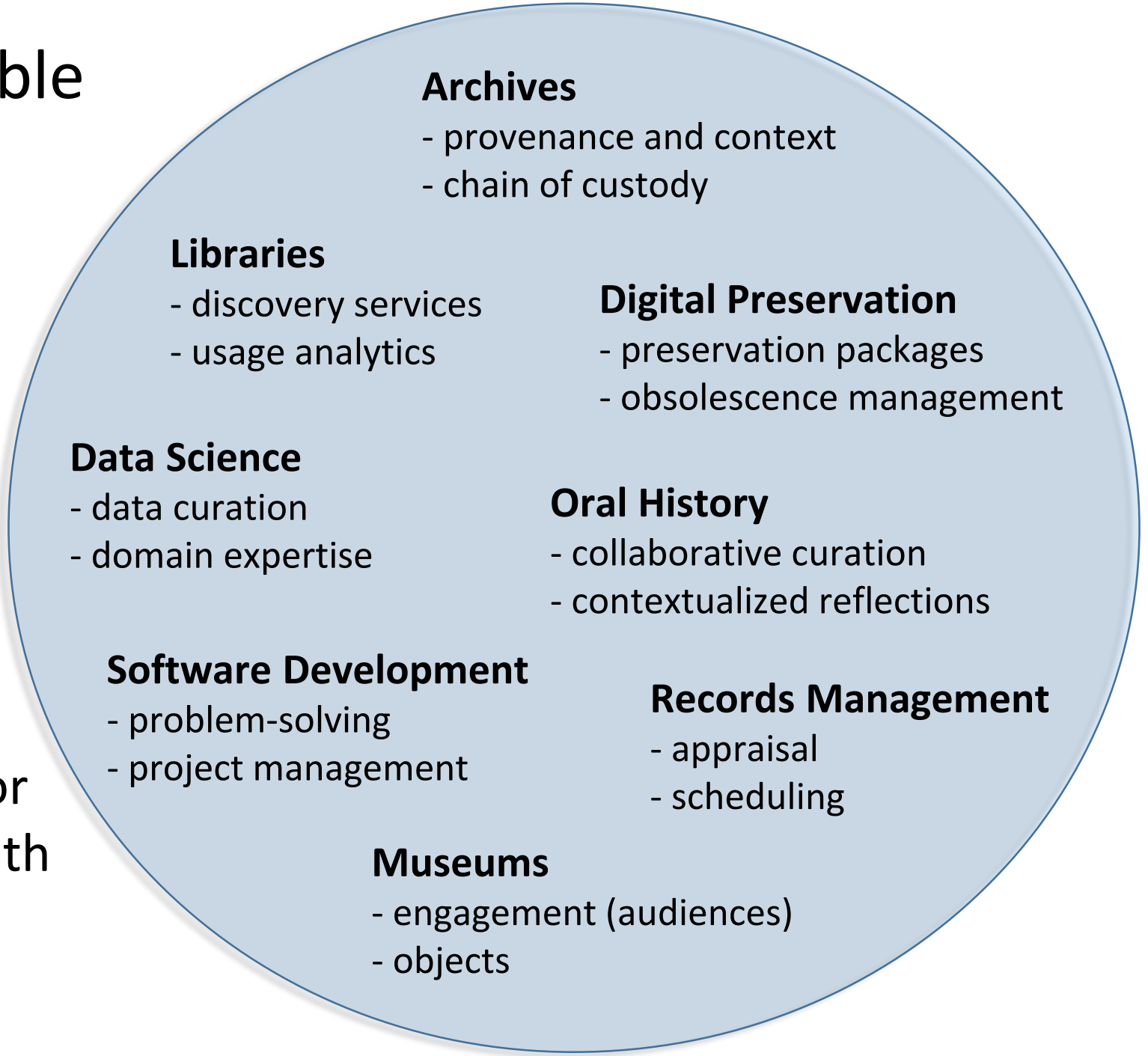
No head – determine together

Reset table for different needs

Add and change seats

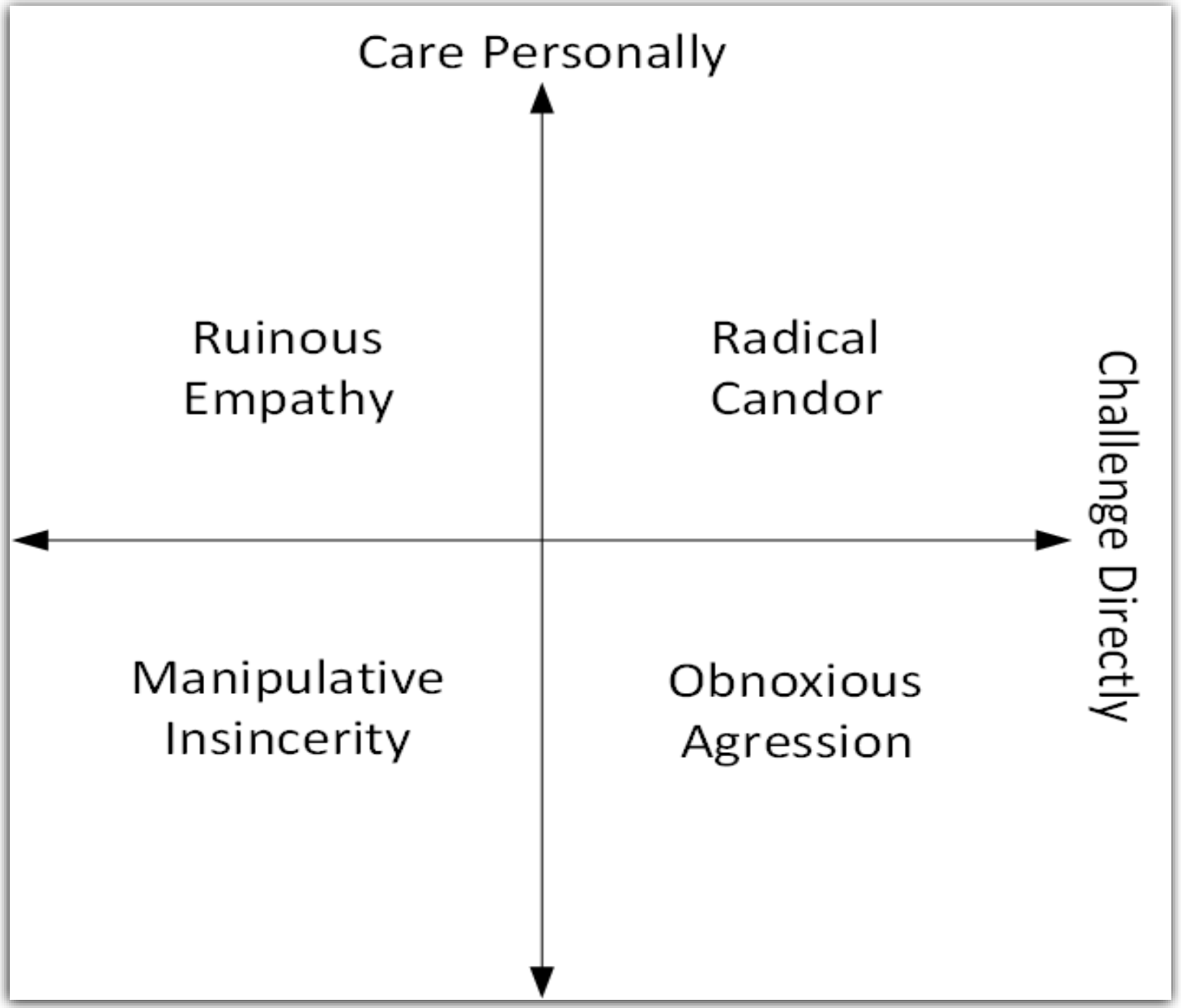
Leverage cumulative strengths

Radical collaboration allows for greatest cumulative impact with awareness of needs and contributions of whole table



Adapt Radical Candor …

An approach for encouraging constructive feedback



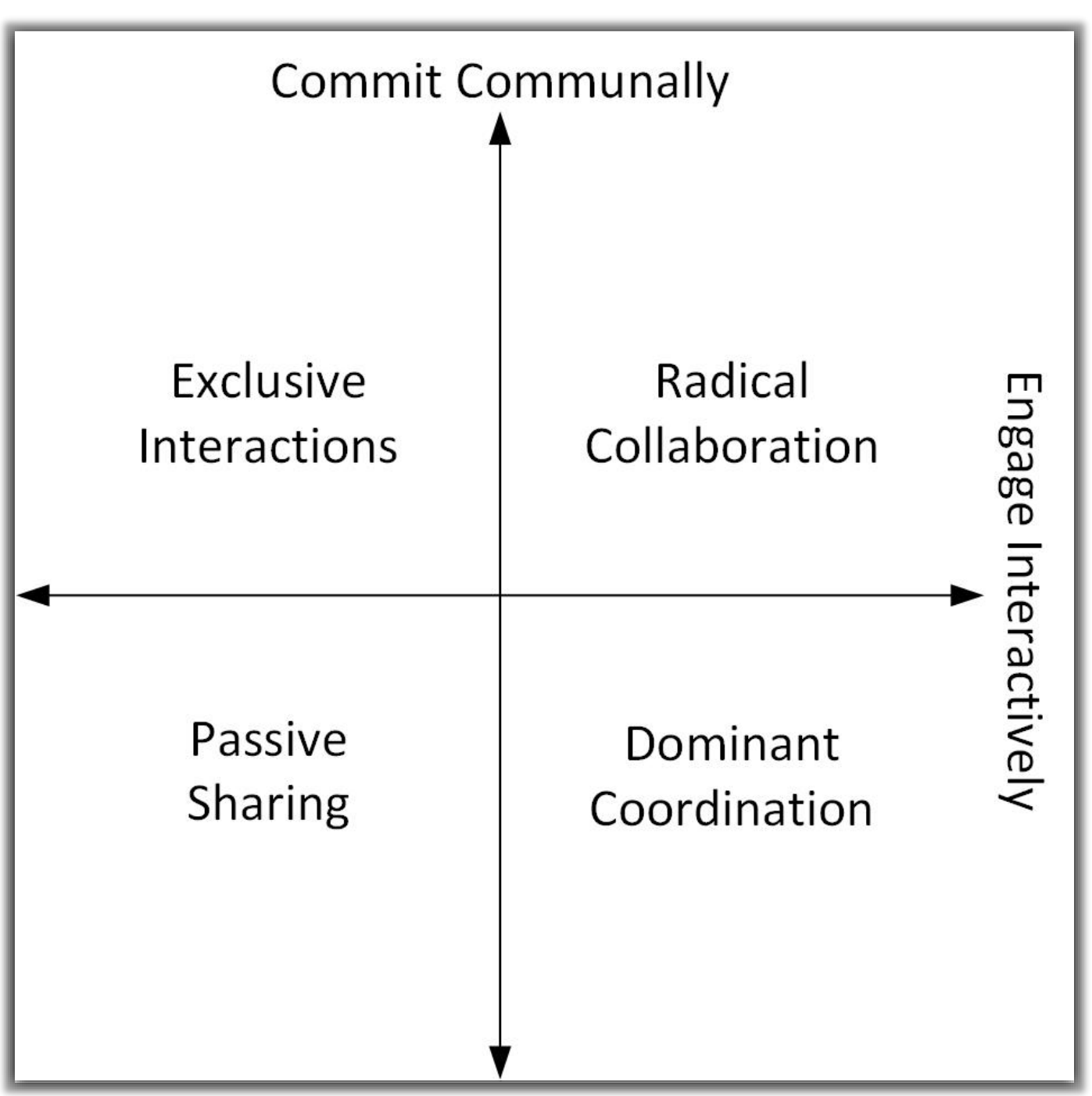
Kim Scott,

Explanatory video: https://www.youtube.com/watch?v=yj9GLeNCgm4

To Radical Collaboration

An approach for encouraging collaborative work across domains

"candid collaboration"



Working Definitions

To build and ensure shared understanding of core concepts

Collaboration: to rely on others to do agreed upon things for or in concert with you and to be relied upon to do agreed upon things for or in concert with others

Collaboration is **not**:

- letting someone know what you did after you did it;
- basic information sharing with no measurable impact; or
- simply allowing someone to be present or observe without means to inform and influence

Inclusion Framework

Social and demographic inclusion

i.e., not excluded based on race, ethnicity, gender, sexual orientation, religion, age, disability, or on any other characteristic or preference

Showstopper: first and foremost ensure that people are safe

Professional inclusion

People from all impacted or related professions and domains are included openly and equitably

Technical inclusion

Technical is not limited to technological

Technology (skills, tools, capabilities) should be available to all

Example: Archives – words are important …

To an archivist:

Archives:

- an **organization** that collects the records of individuals or organizations;
- professional **discipline** of administering such archival collections and organizations – adhere to **archival principles and practice** (e.g., SAA Code of Conduct and Ethics)
- **building** (or portion thereof) housing archival collections

IT use of "archives" and "archiving": often refers to **aggregations** of content, storage of content – is typically not archival from an archivist's perspective and does not equate to preservation (more robust and collaborative concept) e.g., Twitter archives

"digital archivist" may refer to work by anyone on digitized or other digital content of any kind not someone who manages digital records

Historians and other researchers may refer to archiving to mean (see also documentation strategies) ***capture, document, or record*** history and milestones

Question(s)/comment(s) about Radical Collaboration?

Before proceeding with the contributor presentations, one or two questions about the radical collaboration framework we are using?

Panelists

Overview/facilitator: Nance McGovern – MIT

Contributors:

Elizabeth Hull - Dryad-CDL partnership

Cynthia Hudson-Vitale - Data Curation Network

Limor Peer - CURE

Amy Nurnberger - RDA

Elizabeth Hull (she/her/hers)

ORCID: 0000-0002-8661-6265

Associate Director, Dryad – https://datadryad.org

Twitter: @datadryad

"What does it mean to be a community, to operate as a community, to work together towards shared solutions? I would argue that it is much more than just attending a conference and exchanging ideas, that it's really rolling our sleeves up and working together side by side to produce value."

- Kristen Ratan, Coko, 2018 OASPA Conference in Vienna

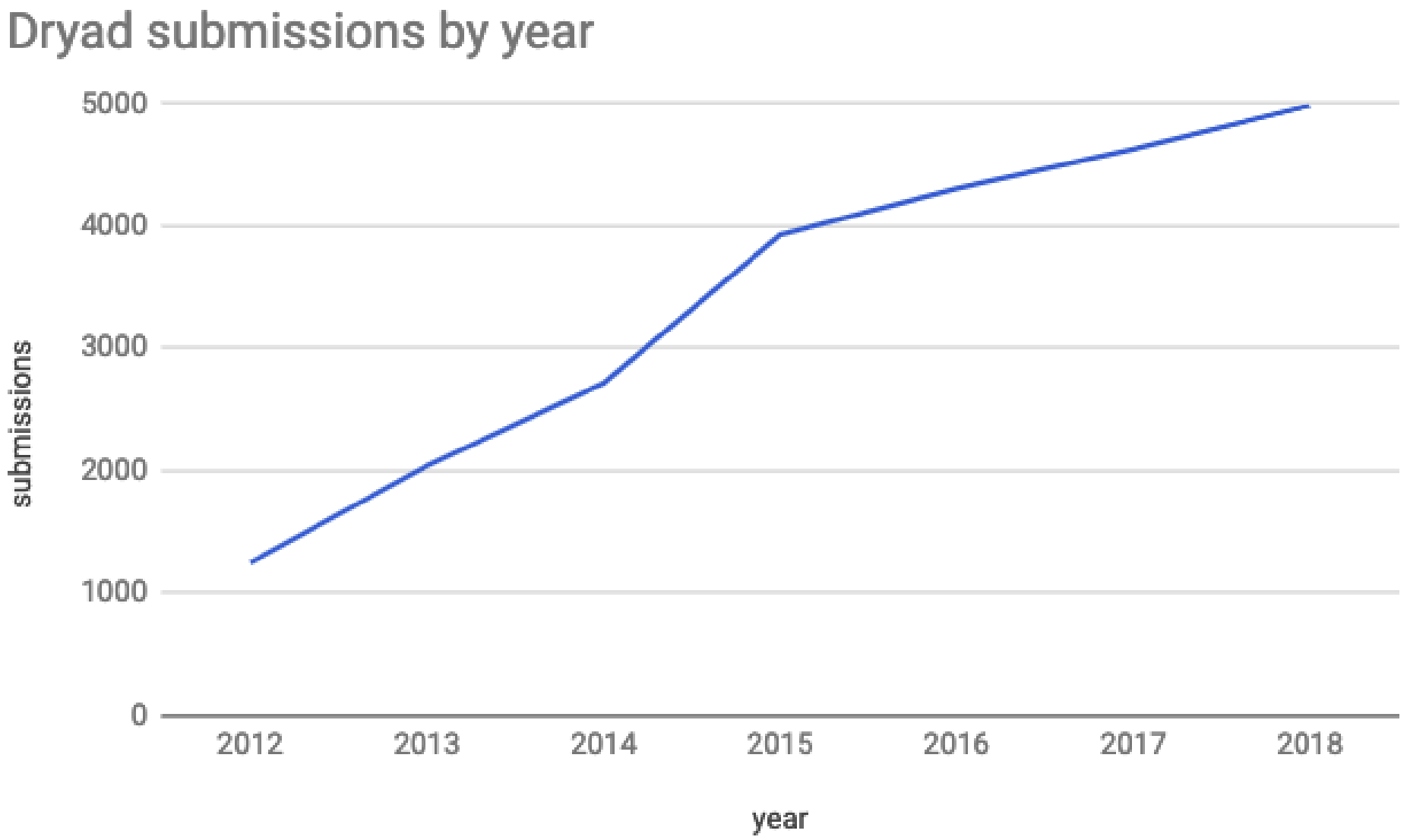
DRYAD

University of California
CDL
California Digital Library

A tale of two data publishing efforts

University of California
CDL
California Digital Library

https://datadryad.org


Dryad submissions by year
5000
4000
3000
2000
1000
0
submissions
2012
2013
2014
2015
2016
2017
2018
year

California Digital Library - https://www.cdlib.org/

CDL provides transformative digital library services, grounded in campus partnerships and extended through external collaborations, that amplify the impact of the libraries, scholarship, and resources of the University of California

Four active programs:
Access & Publishing
Discovery & Delivery
Collections & Licensing
UC Curation Center

DRYAD

University of California
CDL
California Digital Library

California Digital Library - https://www.cdlib.org/

CDL provides transformative digital library services, grounded in campus partnerships and extended through external collaborations, that amplify the impact of the libraries, scholarship, and resources of the University of California

Four active programs:
Access & Publishing
Discovery & Delivery
Collections & Licensing
UC Curation Center

EZID Identifiers made easy

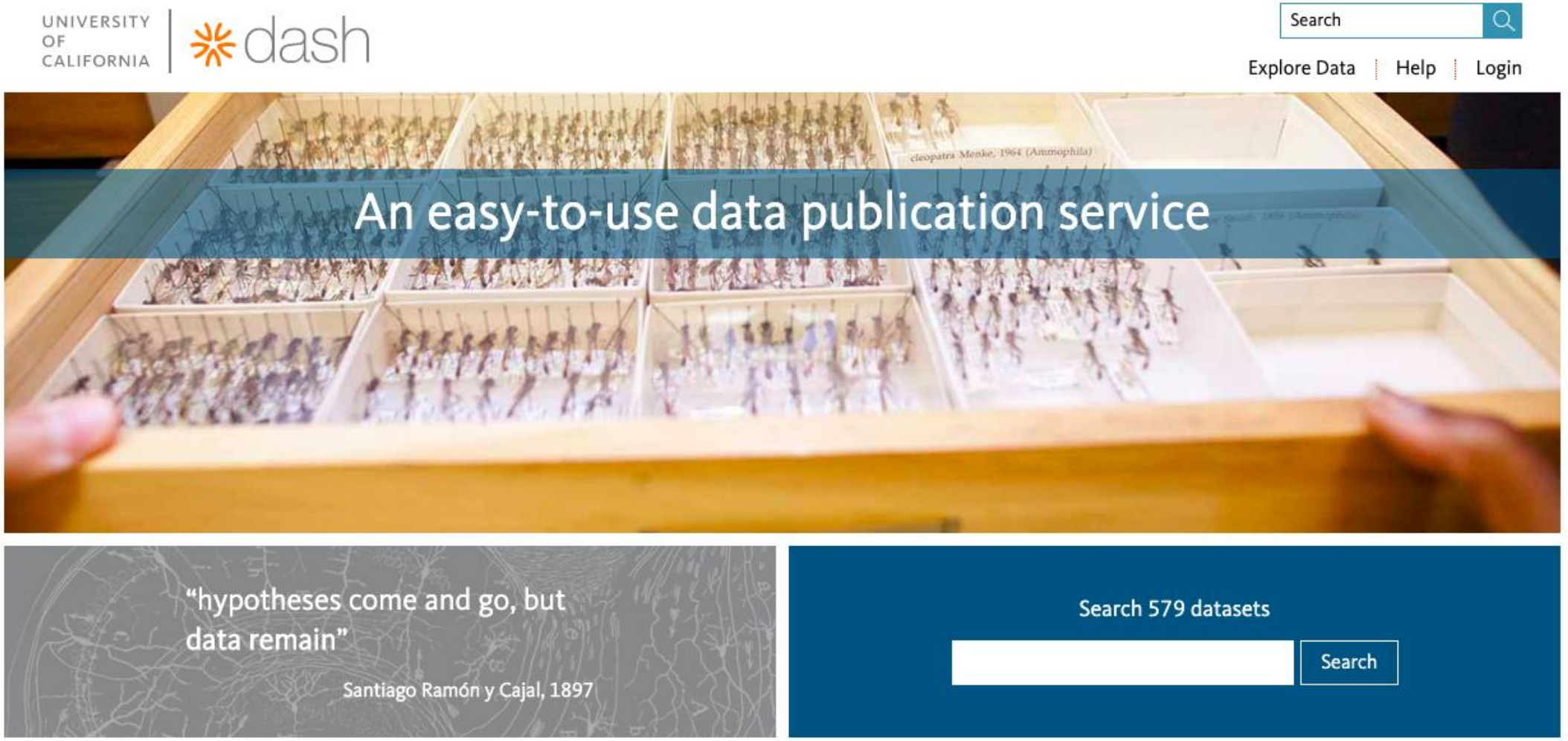
METADATA 2020

ROR

DMPTool

DRYAD

University of California
CDL
California Digital Library



https://dash.ucop.edu

DRYAD + University of California CDL California Digital Library =

What the collaboration is NOT:

- No administrative merger. Dryad remains an independent non-profit
- No significant changes to the operations or mission of either organization

What the collaboration IS:

Two like-minded teams rolling up our sleeves and leveraging our strengths to:

- Drive adoption of curated data publishing
- Develop and sustain a community driven, open source, accessible alternative to commercial offerings

DRYAD

University of California
CDL
California Digital Library

Question(s)/comment(s)?

One or two questions or comments about the Dryad/CDL collaboration?

Panelists

Overview/facilitator: Nance McGovern – MIT

Contributors:

Elizabeth Hull - Dryad-CDL partnership

Cynthia Hudson-Vitale - Data Curation Network

Limor Peer - CURE

Amy Nurnberger - RDA

Cynthia Hudson-Vitale

[she/her/hers]

ORCID: 0000-0001-5581-5678

Head, Research Informatics & Publishing - Pennsylvania State University

Twitter: @cynhudson

cuv185@psu.edu

DATA CURATION NETWORK

leveraging expertise to curate data at scale

Team

PI: Lisa Johnston, University of Minnesota

Co-PIs: Mara Blake, Johns Hopkins University; Jake Carlson, University of Michigan; Cynthia Hudson-Vitale, Penn State University; Joel Herndon, Duke University; Elizabeth Hull, Dryad; Heidi Imker, University of Illinois; Wendy Kozlowski, Cornell University; Hoa Luong, University of Illinois; Tim McGeary, Duke University, Jennifer Moore, Washington University; Claire Stewart, University of Minnesota

Project Coordinator: Elizabeth Coburn, University of Minnesota

With Thanks

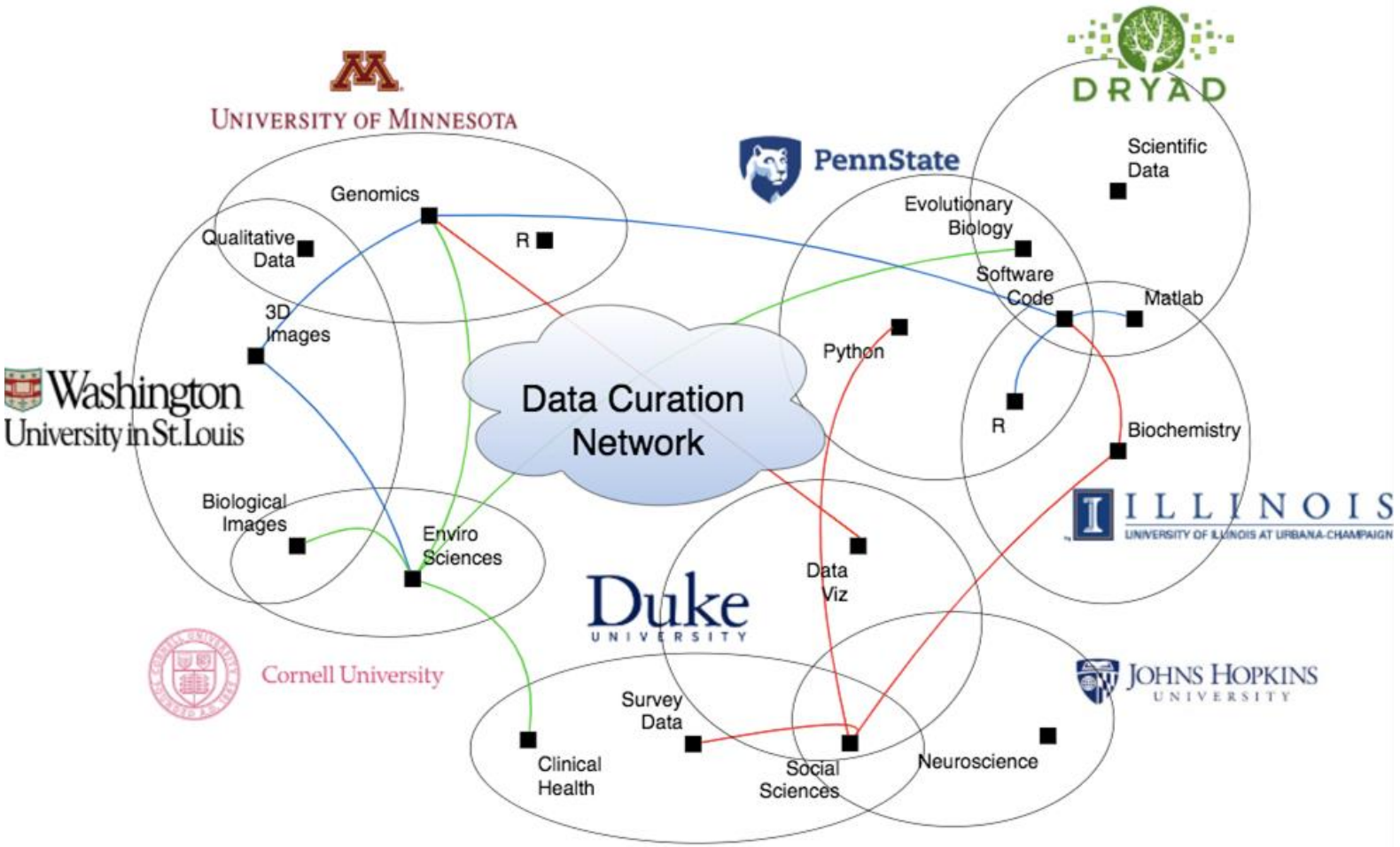
The Data Curation Network is made possible with funding from the Sloan Foundation.

We are also grateful for funding from IMLS for the Data Curation Workshops.

Our Mission

“The Data Curation Network will enable data repositories to better support researchers that are faced with a growing number of requirements to ethically share their research data in ways that make it findable, accessible, interoperable and reusable (FAIR).”

Mission: Expert Curation



Mission: Create and Share Best Practices

DCN Curators will take **CURATE** steps for each data set, that includes:

C **Check** data files and read documentation

U **Understand** the data (try to), if not...

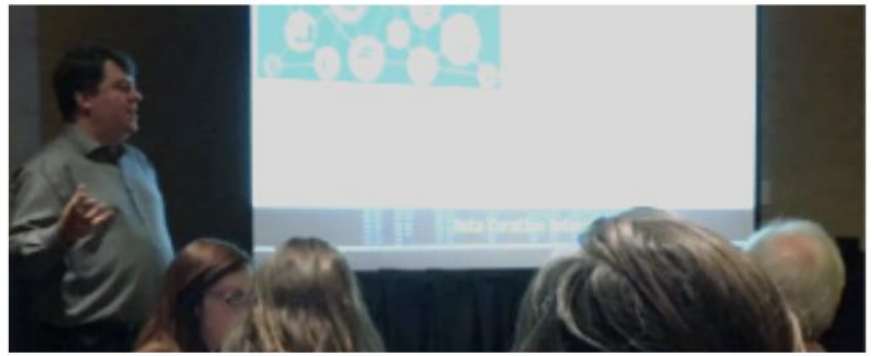
R **Request** missing information or changes

A **Augment** the submission with metadata for findability

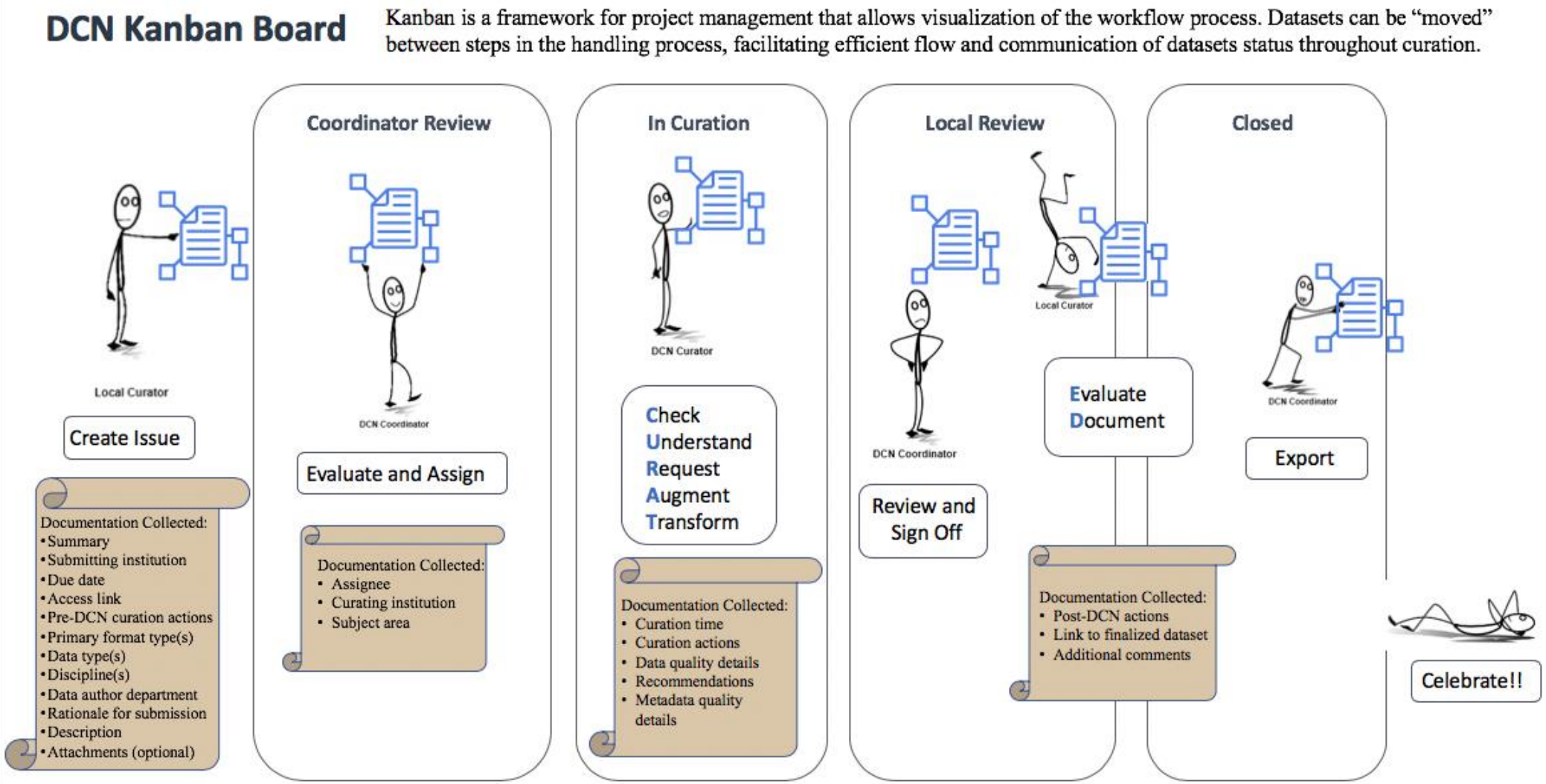
T **Transform** file formats for reuse and long-term preservation

E **Evaluate** and rate the overall submission for FAIRness.

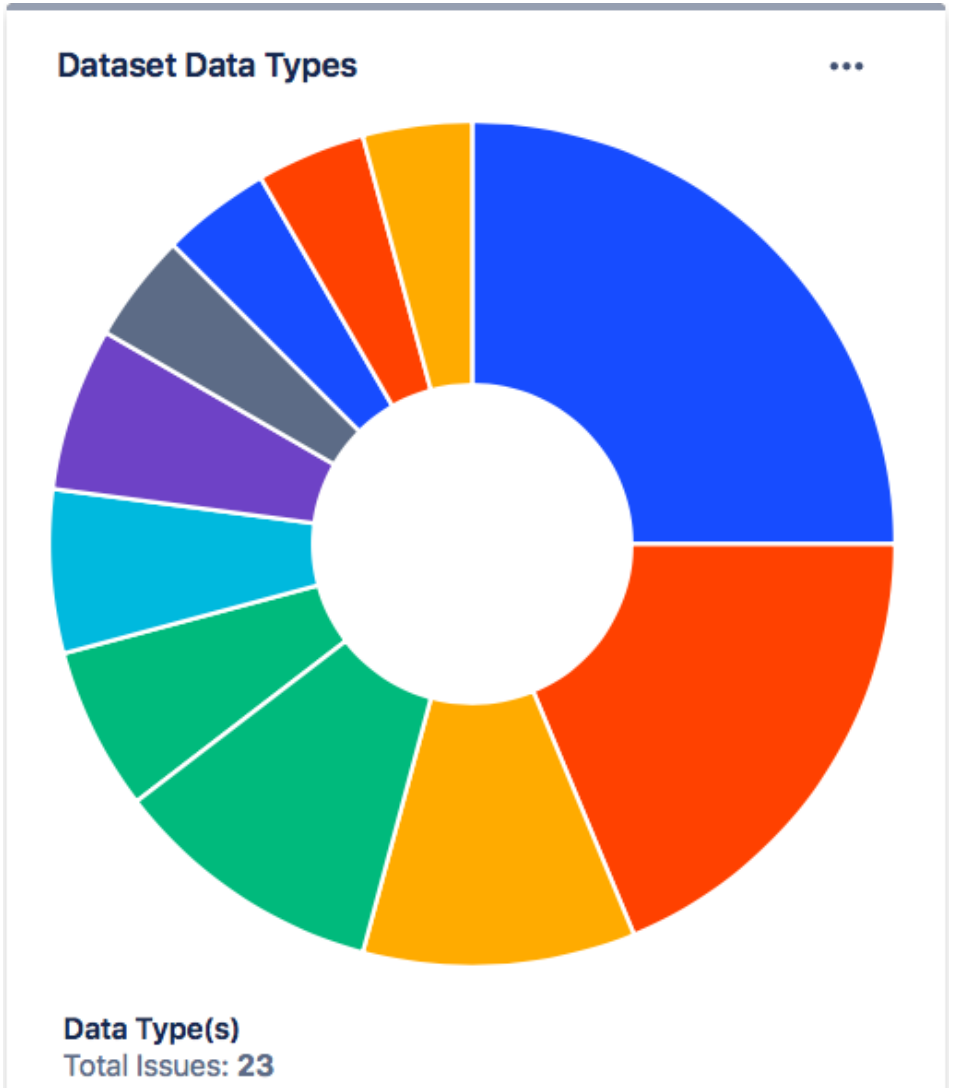
Mission: Build a curation community of practice



Launch of the DCN



Stats to date



Data Type(s)
Total Issues: **23**

tabular	12
code	9
image	5
3D_image	3
human_subjects	3
other	3
database	2
simulation	2
spectral	2
statistical	2
Other...	5

Challenges

- Institutional priorities and culture
- Unvoiced concerns
- Indeterminable or unknown value
- Complex and evolving ecosystem
- Network design
- Antiquated views of libraries
- No sustainability model
- Easier to go alone
- Transparency
- Unbalanced workloads

What's next

- Growing the network
  - Plans to add two institutions in Year 2 and another two in Year 3. Washington University joined this year.
    - Criteria for New Partners:
      - Have a data repository
      - Have data curation staff to fill DCN Representative and DCN Curator roles
      - Adhere to team's code of conduct
      - Preferred: Expand curation expertise and/or bring new diversity to the project
- IMLS Funded Education and training for data curation professionals
  - Data Curation Workshops
    - IASSIST Pilot
    - 3 IMLS workshops
  - Data Curation Primers
    - netCDF primer
    - Jupyter Notebooks
    - SPSS
    - Excel
    - Wordpress
    - Geodatabases
    - Microsoft Access

Resources

https://datacurationnetwork.org/
https://sites.psu.edu/dcnworkshops/

Question(s)/comment(s)?

One or two questions or comments for the Data Curation Network?

Panelists

Overview/facilitator: Nance McGovern – MIT

Contributors:

Elizabeth Hull - Dryad-CDL partnership

Cynthia Hudson-Vitale - Data Curation Network

Limor Peer - CURE

Amy Nurnberger - RDA

Radical collaboration: Joining forces to advance curating for reproducibility

Limor Peer, PhD
limor.peer@yale.edu
ORCID: 0000-0002-3234-1593

Associate Director for Research, Institution for Social and Policy Studies, Yale University
https://isps.yale.edu/research/data

Co-founder, Curating for Reproducibility (CURE) Consortium
http://cure.web.unc.edu/

Open Repositories | Hamburg, Germany | 12 June 2019 | 11:00 – 12:30

CC BY

#OpenRepo2019 @l_peer

A challenge for data repositories: The reproducible scholarly record

Computational reproducibility involves "calculation of quantitative scientific results by independent scientists **using the original datasets and methods**."

Stodden, V., Leisch, F., & Peng, R.D. (Eds.). (2014). *Implementing Reproducible Research.* New York: Chapman and Hall/CRC.

Data curation+

Perform Data Quality Review

Data Quality Review Framework

Peer, L., Green, A., & Stephenson, E. (2014). Committing to data quality review. *International Journal of Digital Curation*, *9*(1). https://doi.org/10.2218/ijdc.v9i1.317

CURE Standards
Journal Data Policy Enforcement

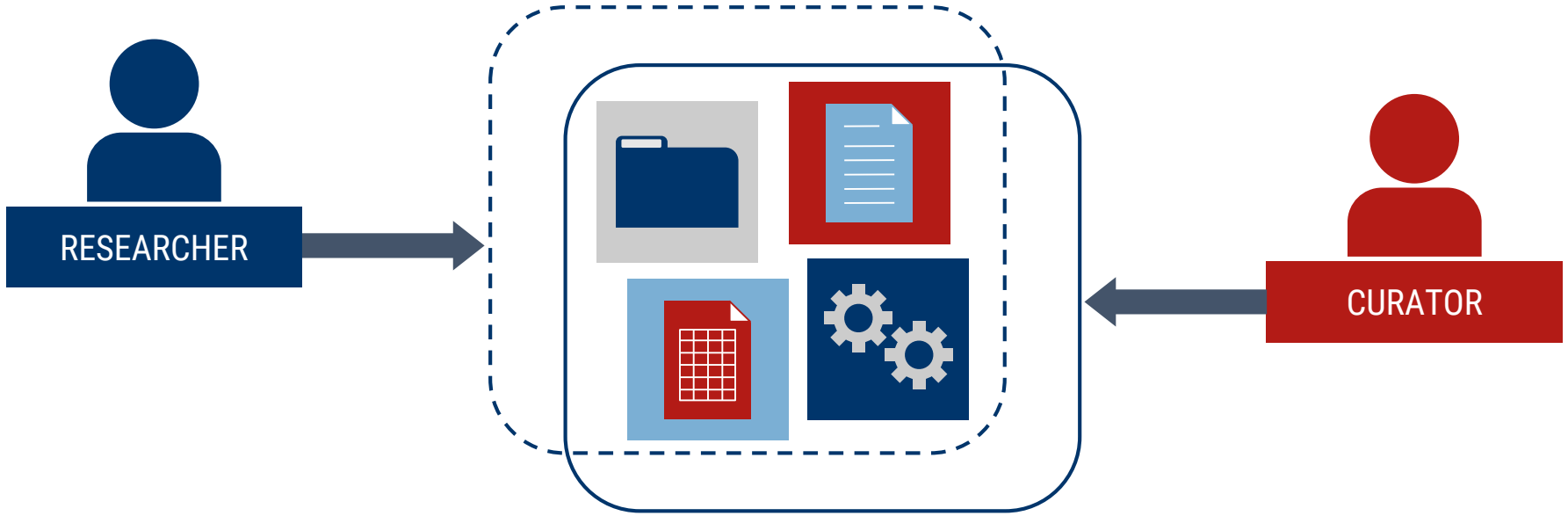
CURE Workflows
Results Reproduction (R^2) Service

CURE Management
Yale Application for Research Data (YARD)

CURE workflows

Results Reproduction (R^2) Service • https://ciser.cornell.edu/research/results-reproduction-r-squared-service/

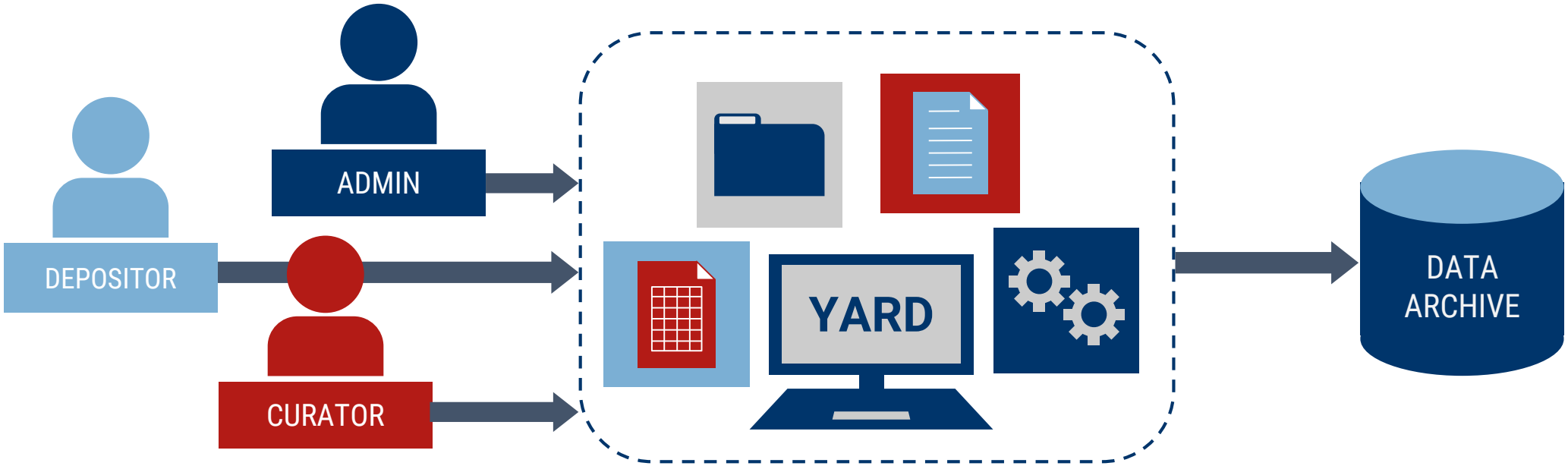
Researchers with papers ready to submit for publication can send their data and code to CISER prior to submission for appraisal, curation, and replication to **ensure that published results are reproducible**.



CURE management

YARD: Yale Application for Research Data • https://yard.yale.edu/

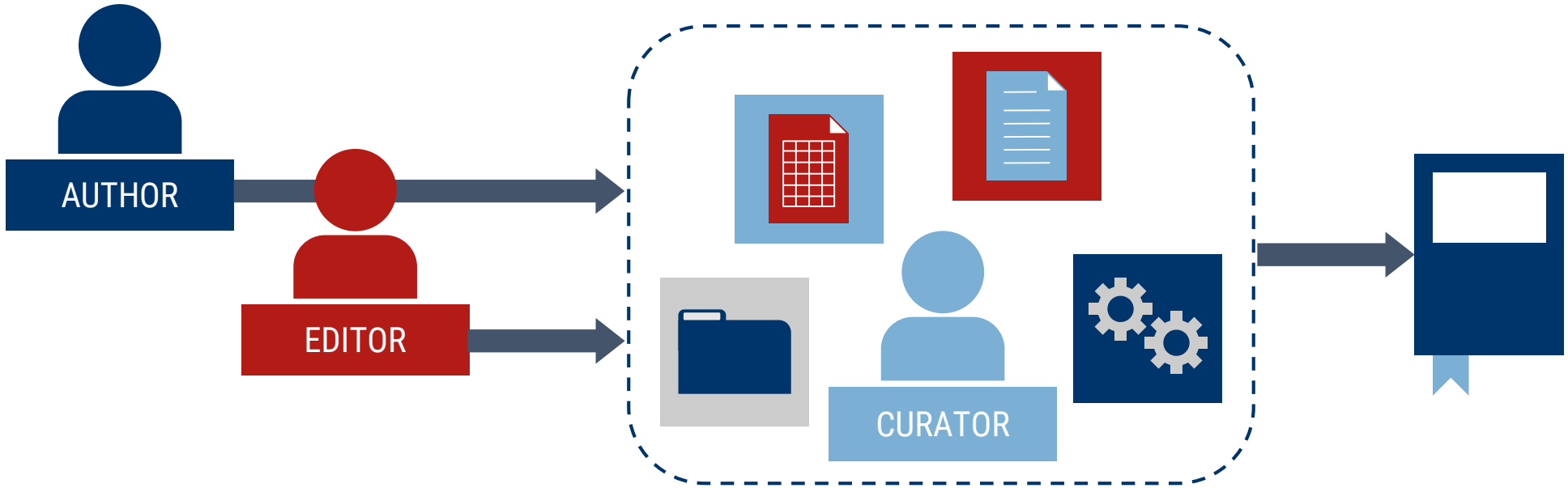
A new workflow tool that allows Depositors, Curators, and Administrators to **submit, review, process, and publish data within one system**, which structures and records data curation and code review workflow actions.



CURE standards

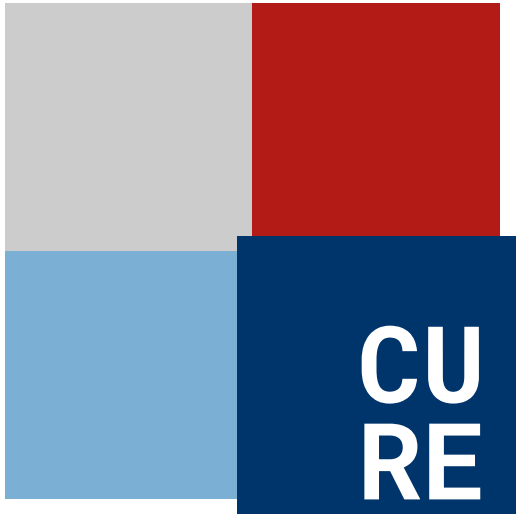
Journal Data Policy Enforcement • https://odum.unc.edu/archive/managementcuration/

The Odum Institute Data Archive works with Odum Institute statisticians to perform data curation and code review tasks on behalf of journal editors to enforce **TOP Level III data policies**.



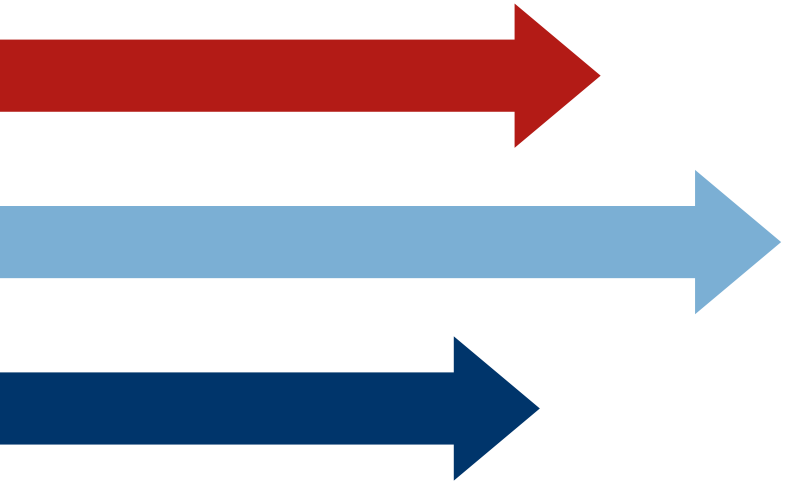
We support **curation of research data** and **review of code** and associated digital scholarly objects for the purpose of facilitating the digital preservation of the evidence base necessary for future understanding, evaluation, and reproducibility of scientific claims.

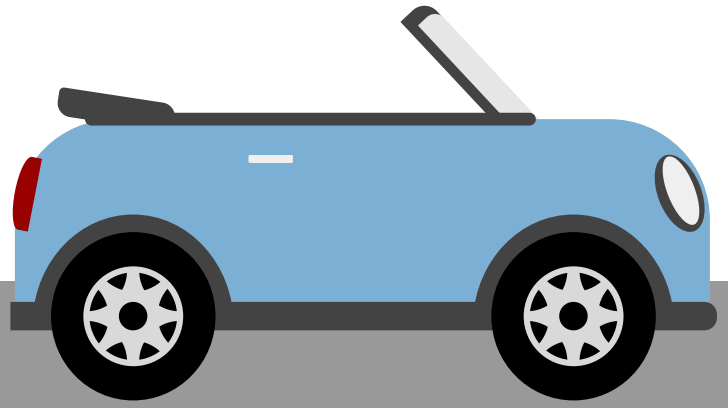
http://cure.web.unc.edu/
https://osf.io/3wkex/



CURATING for REPRODUCIBILITY

CURE highlights





2016

- IASSIST pow-wow
- Articulation of founding principles, sharing practices
- Assemble Advisory group
- Buy-in from home institutions

2017

- Launch website
- IMLS Training Program Planning Grant RE-87-17-0074-14

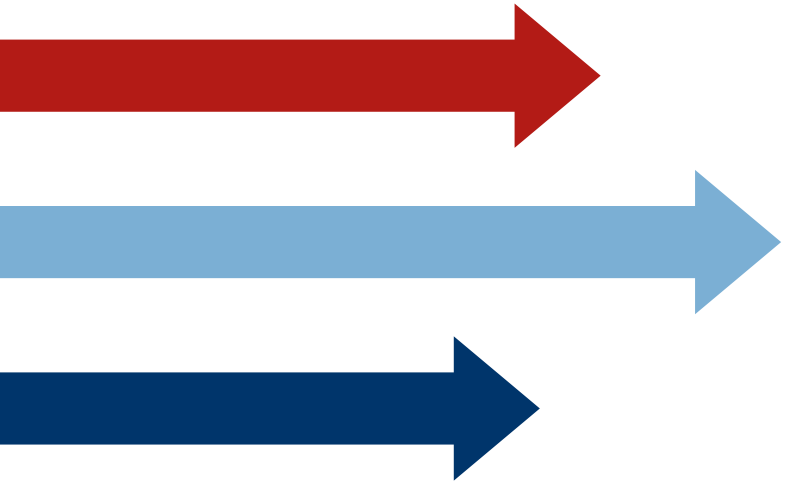
2018

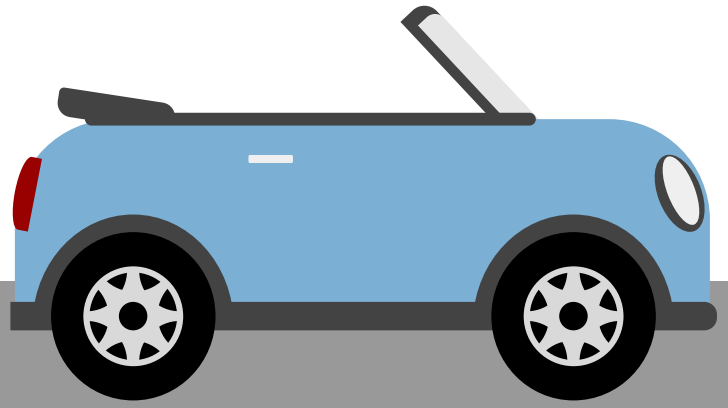
- Workshops

2019

- Workshops
- Community building

CURE radical collaboration





Collaborate to…

- Establish standards
- Share practices
- Promote Curating for Reproducibility

Lessons learned…

Question(s)/comment(s)?

One or two questions or comments about CURE?

Panelists

Overview/facilitator: Nance McGovern – MIT

Contributors:

Elizabeth Hull - Dryad-CDL partnership

Cynthia Hudson-Vitale - Data Curation Network

Limor Peer - CURE

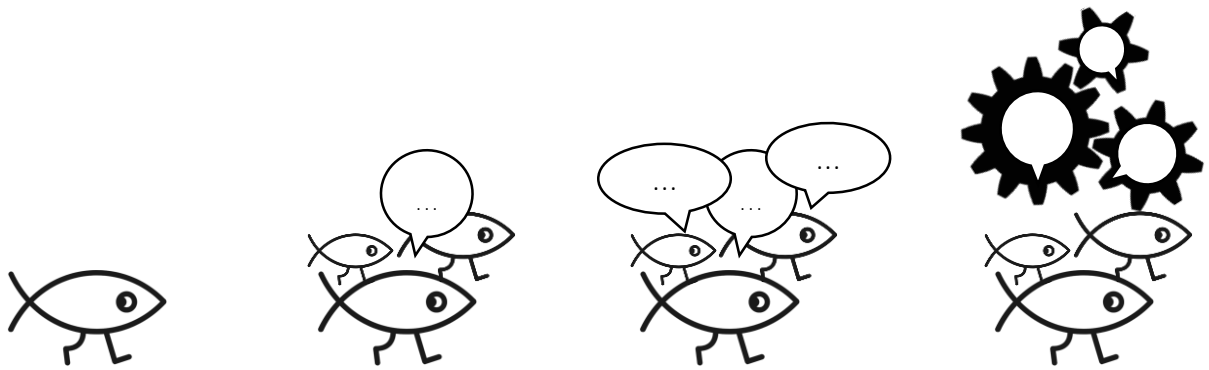
Amy Nurnberger - RDA

The Evolution of Collaboration in RDA and Institutions

Where we start, how we develop to $\sqrt{collaboration}$

OR 2019, Hamburg, DE 2019-06-12
P4B: Joining forces to address user needs around research data management (panel)

Amy Nurnberger MIT Libraries



Introduction

Amy Nurnberger [*she/her(s)*],
ORCID: 0000-0002-5931-072X
Program Head, Data Management Services - Massachusetts Institute of Technology

Adjunct Assistant Professor, Learning Analytics, Teachers College, Columbia University

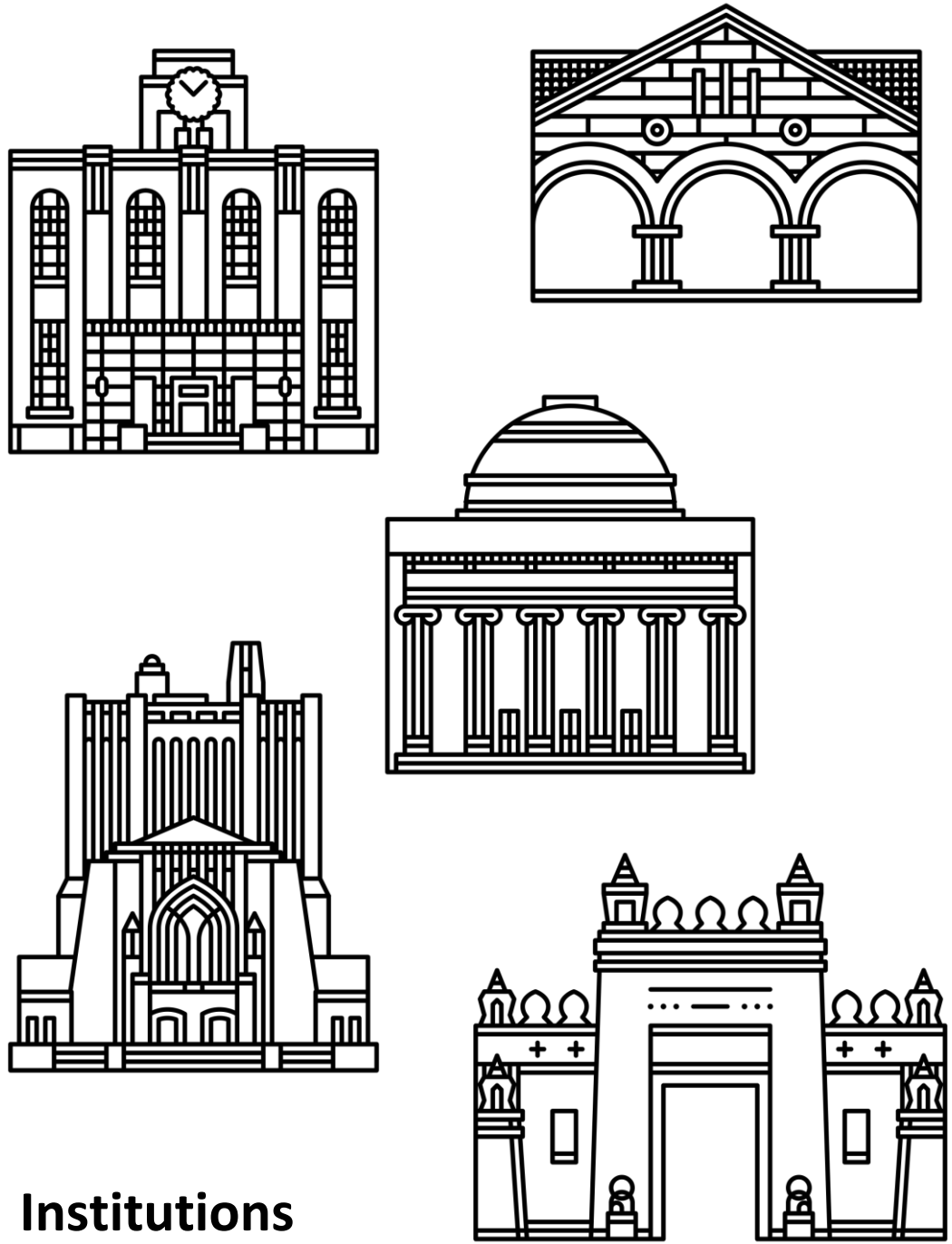
Research Data Alliance
- Organisational Advisory Board, co-chair
- Education and Training on Handling Research Data, co-chair

Twitter: @ANurnberger

These slides based on the article: Amy Nurnberger. "The Radical Collaboration of RDA and What It Means for Developing Institutional Data Management Services." *Research Library Issues*, no. 296 (2018): 23–32. https://doi.org/10.29242/rli.296.3.

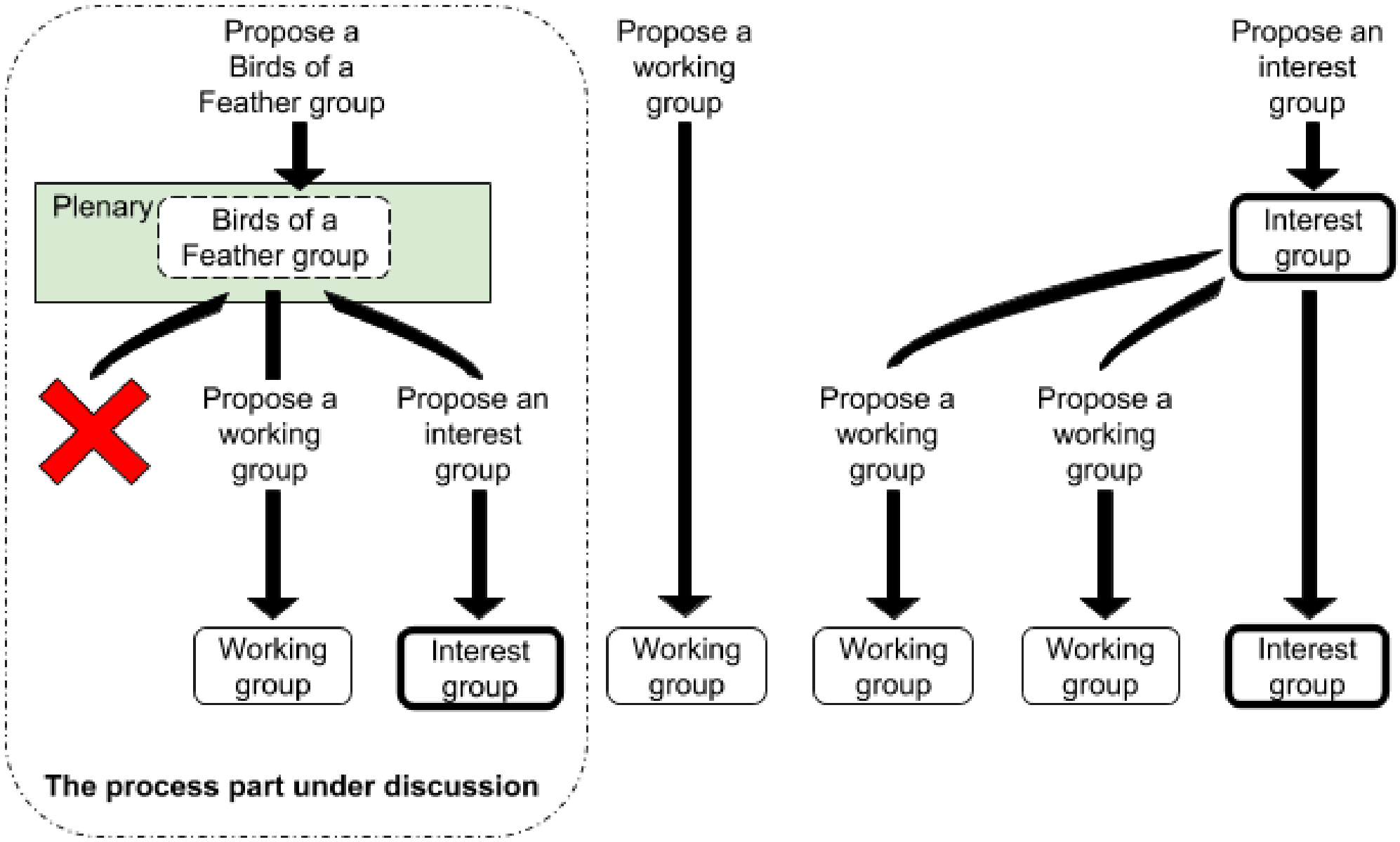
RESEARCH DATA ALLIANCE

Why?



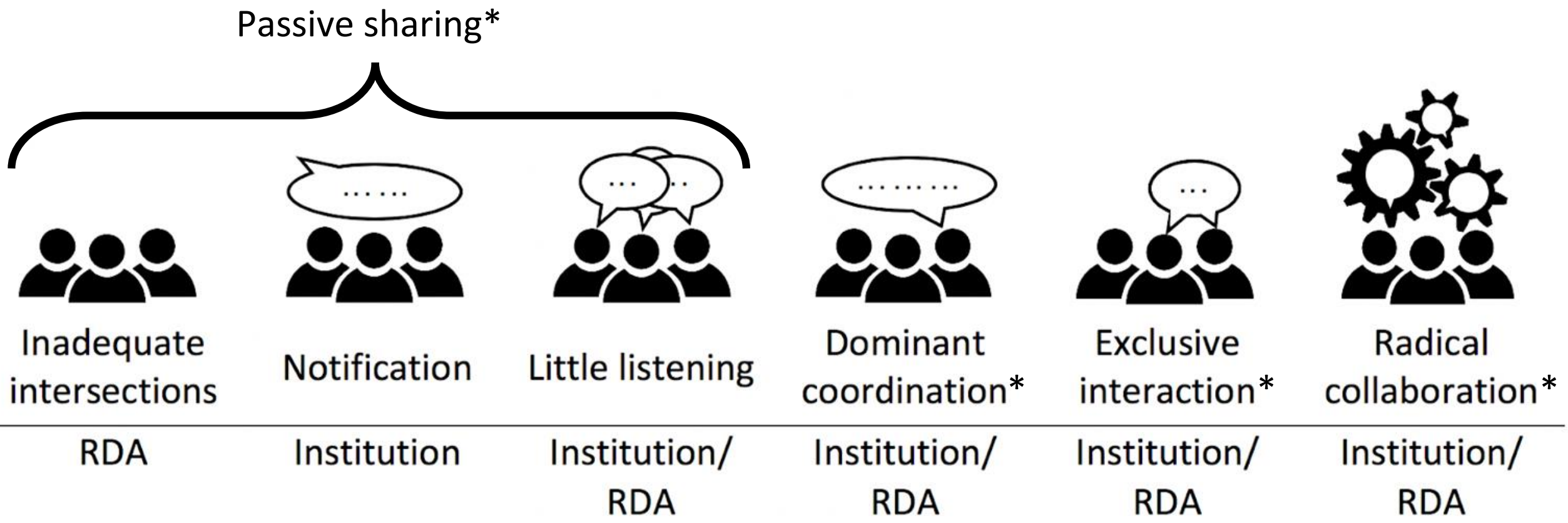
Research Data Alliance (RDA)

Institutions



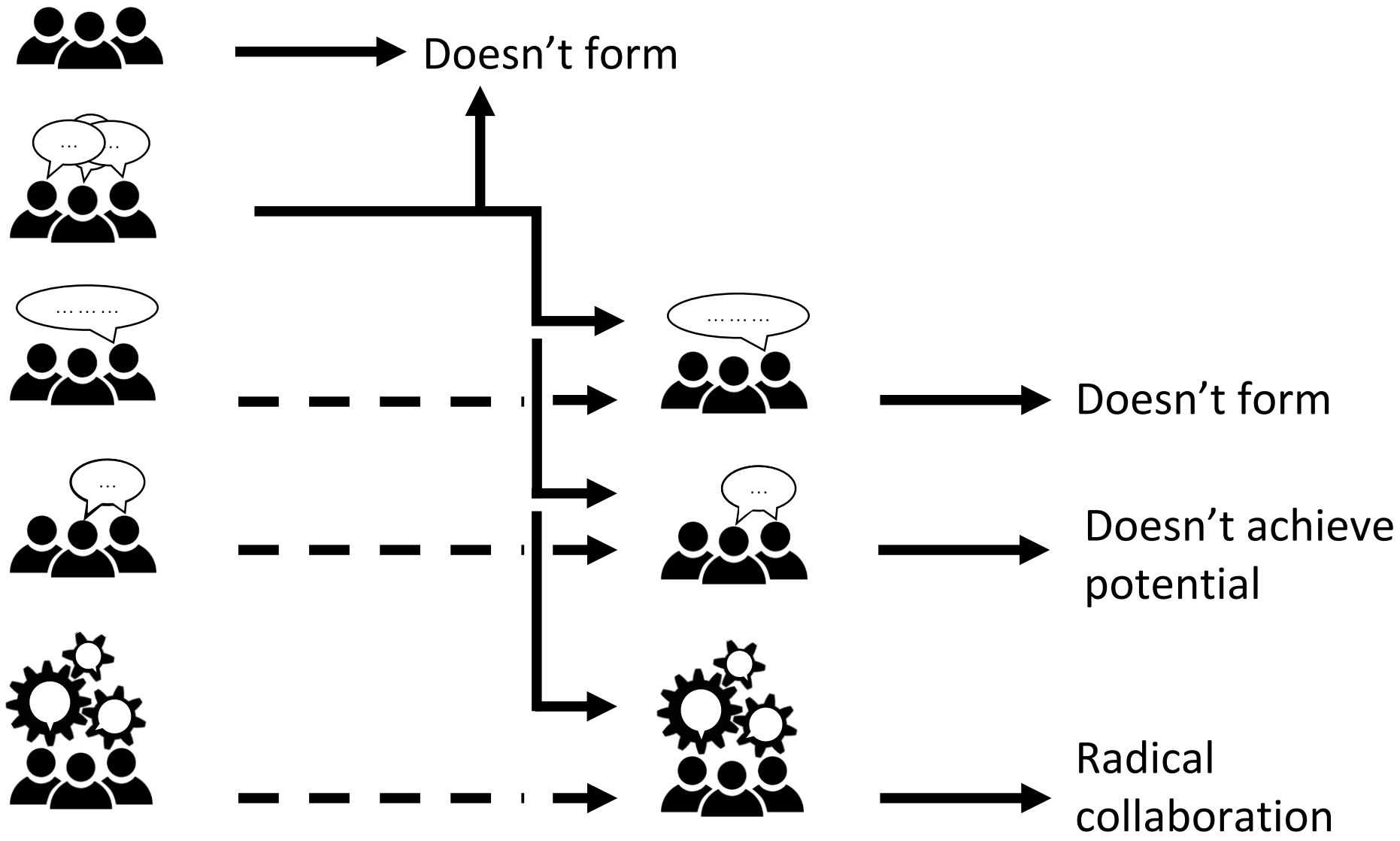
Getting to groups in RDA

Starting points

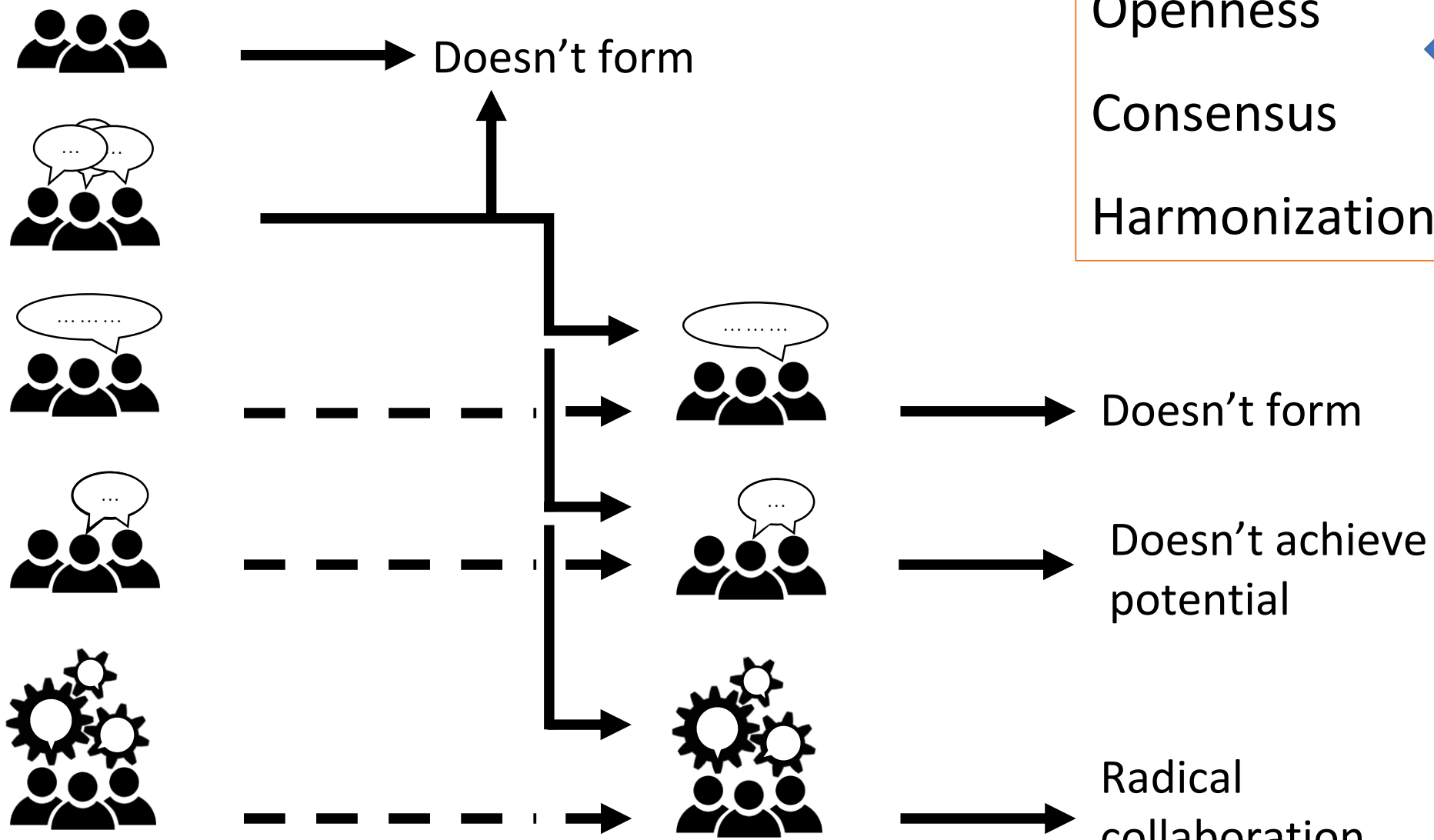


* Nancy Y. McGovern, "Radical Collaboration and Research Data Management: An Introduction," Research Library Issues, no. 296 (2018): X–Y, https://doi.org/10.29242/rli.296.2

From here to there in RDA



From here to there in RDA



Balance

Openness

Consensus

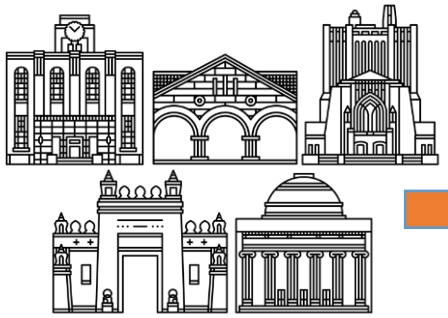
Harmonization

Technical Advisory Board

Responsibility for :

- Balance
- Openness
- Consensus
- Harmonization

RDA RESEARCH DATA ALLIANCE → **T**echnical **A**dvisory **B**oard



→ Often Unclear

Inadequate intersections

RDA

Notification

Institution

Little listening

Institution/ RDA

Dominant coordination *

Institution/ RDA

Exclusive interaction *

Institution/ RDA

Radical collaboration *

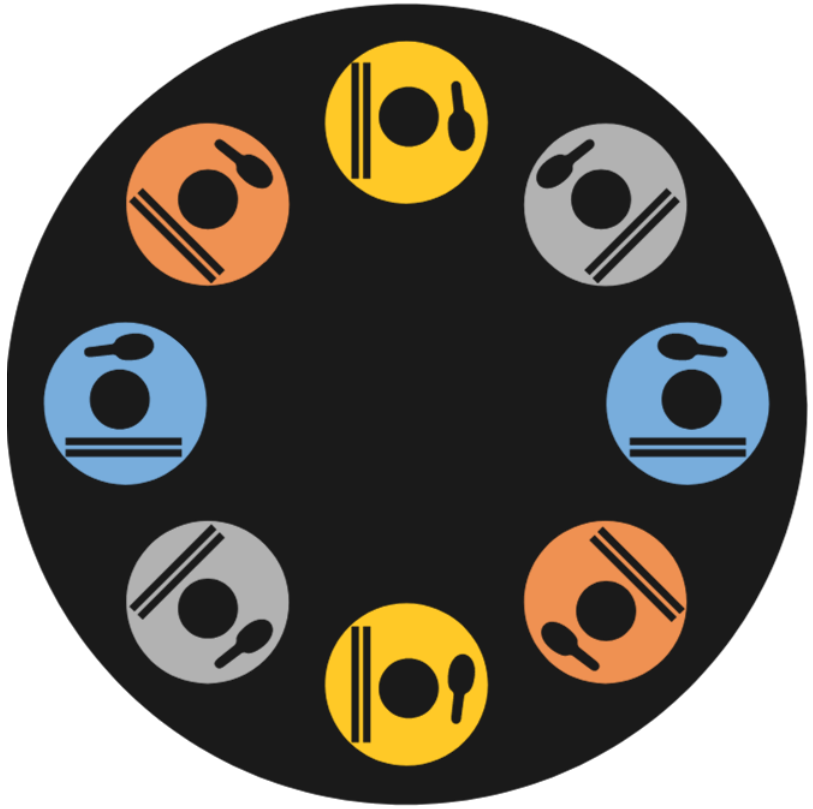
Institution/ RDA

* Nancy Y. McGovern, "Radical Collaboration and Research Data Management: An Introduction," Research Library Issues, no. 296 (2018): X–Y, https://doi.org/10.29242/rli.296.2

From here to there in Institutions

Leaders – Set the table for…

- Balance & inclusion
- Representation
- Inquiry
- Openness
- Consensus
- Harmonization
- Equity



Participants – Claim your place in…

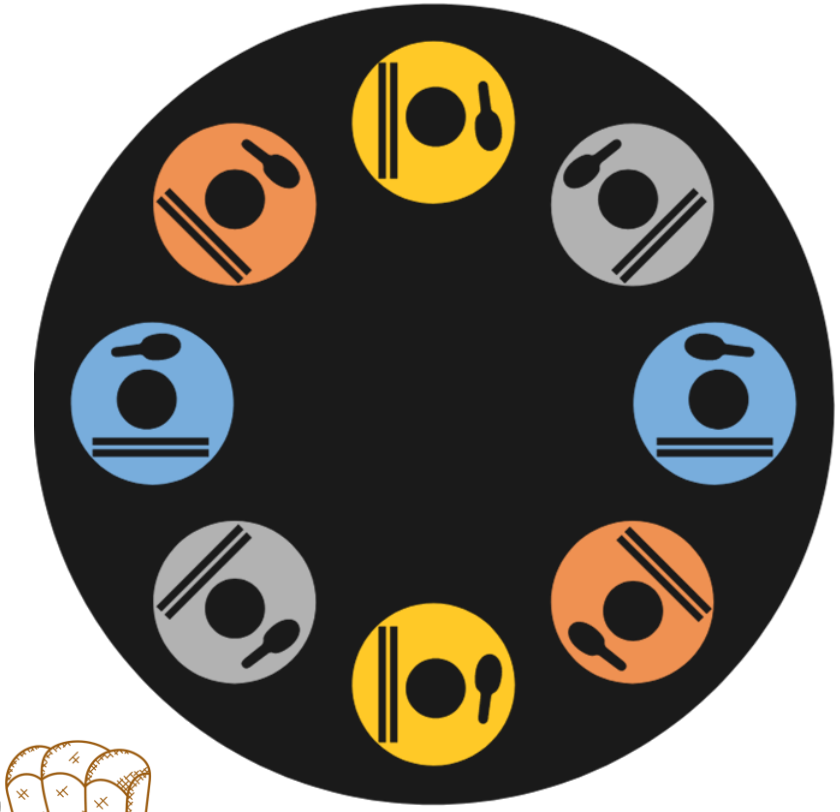
- Involving others
- Setting group expectations
- Co-building…
  - Definitions
  - Mental models
  - Goals

From here to there in Institutions

Leaders – Set the table for…

- Balance & inclusion
- Representation
- Inquiry
- Openness
- Consensus
- Harmonization
- Equity

(= BRIOCHE!)

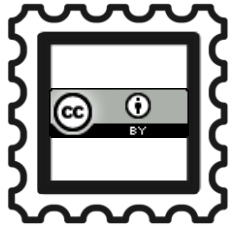


Participants – Claim your place in…

- Involving others
- Setting group expectations
- Co-building…
  - Definitions
  - Mental models
  - Goals

Thank you!

Amy Nurnberger
nurnberg [at] mit.edu
@ANurnberger



Open Repositories, 2019
P4B
Hamburg, DE

^ All icons licensed via NounPro

Question(s)/comment(s)?

One or two questions or comments about RDA?

Topical discussion

Possible Discussion Topics

Communication

Opportunities

Incentives and Inhibitors

Lessons Learned

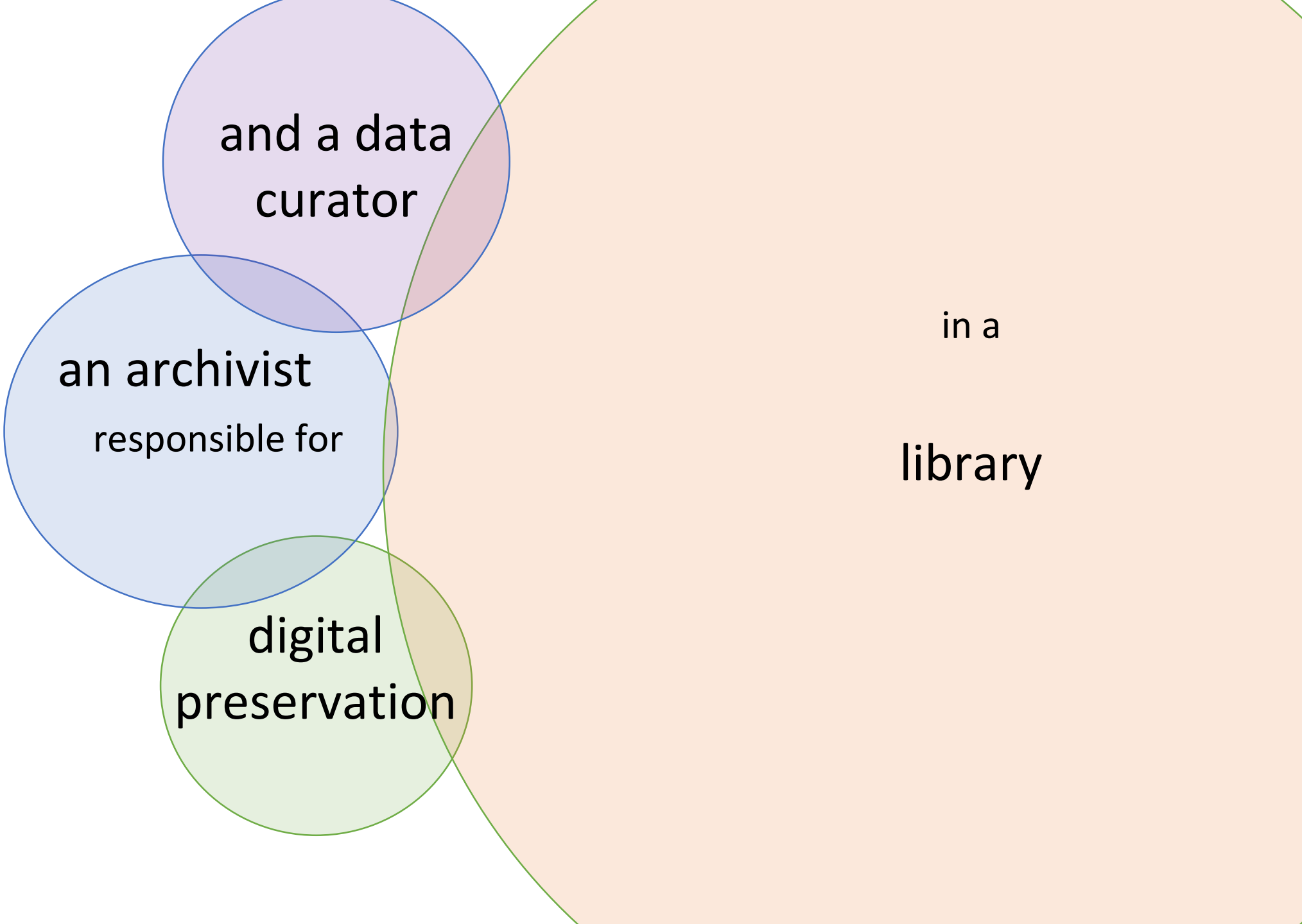
Q&A

Questions for you …

- Questions for the panelists?
- Examples of radical collaboration or possibilities from you to share?
- Professions that you think should be (and might not be) included in:
  - Research data management?
  - Digital practice generally?
  - Project you've worked on?
  - …
- Collaboration challenges you've encountered?

Next Steps Discussion

I am …


and a data curator
an archivist
responsible for
digital preservation
in a
library

Forward Together: collaborating across communities

Be sure everyone who should be or wants to be at the table is there

Welcome unanticipated guests

If you are in the majority, consider what inclusion looks like

Ensure holistic communication for those not at the table

Come to the table with questions not answers – balance inquiry and advocacy

If you haven't learned something at the table, sit longer, ask more questions, and continue to listen…

Forward Together: collaborating across communities

Reset the table as needed to include different perspectives and evolving issues

Never assume you're chairing even if you sent out the invite

Value your knowledge and that of others; bring your whole self and be courageous

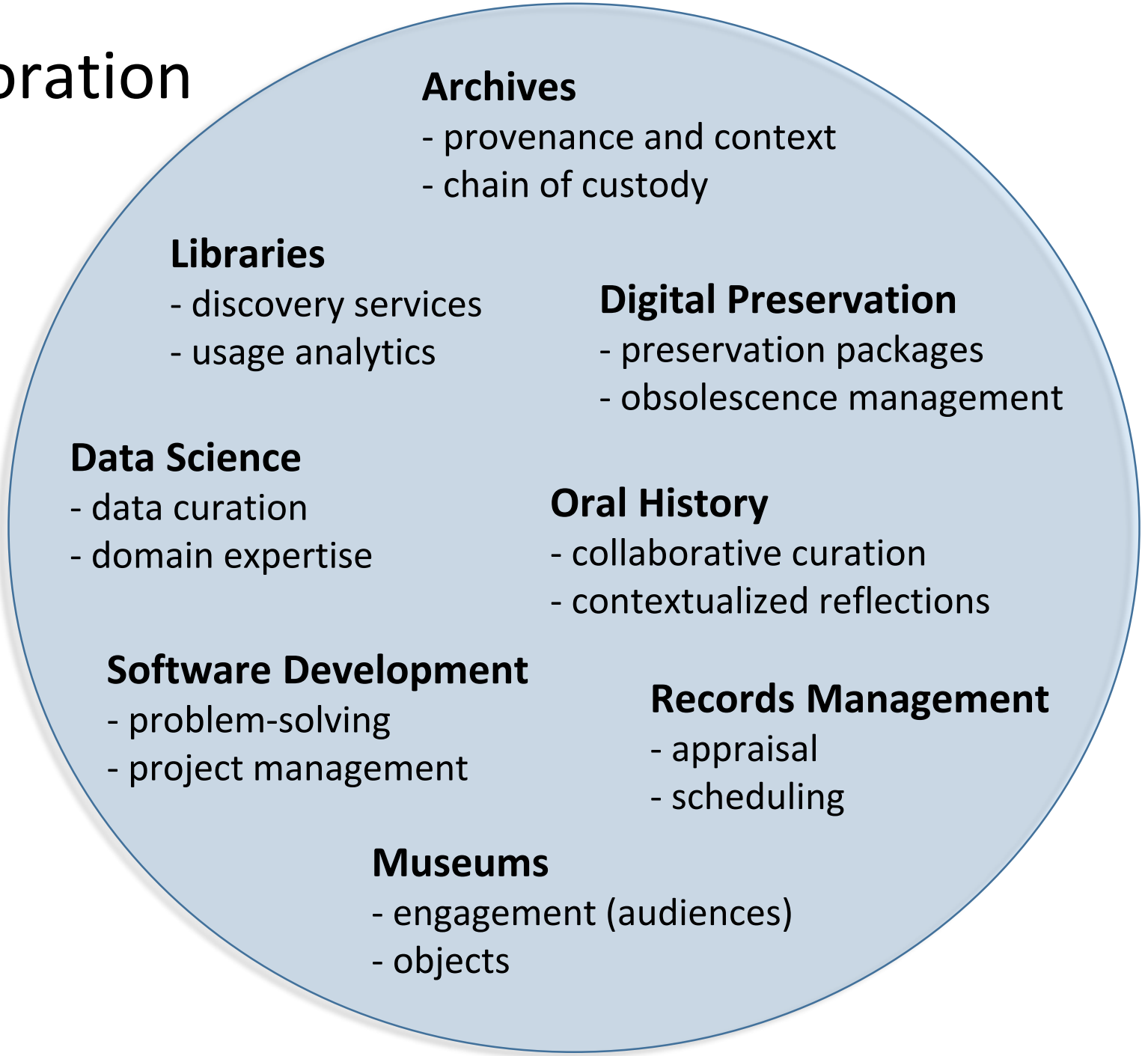
Prioritize the common goal over any single "right" approach

Digital Practice + Collaboration

Who would you bring to the table?

- examples of domain strengths
- round table (with no head)
- stronger together
- professional inclusion
- common interests
- overlapping members
- overlapping objectives

Who's not at the table and why?



Thank you!

Please give **RLI 296 a look and share your thoughts with us:**

https://publications.arl.org/rli296/