

PresQT

Planning the Data Curation Network

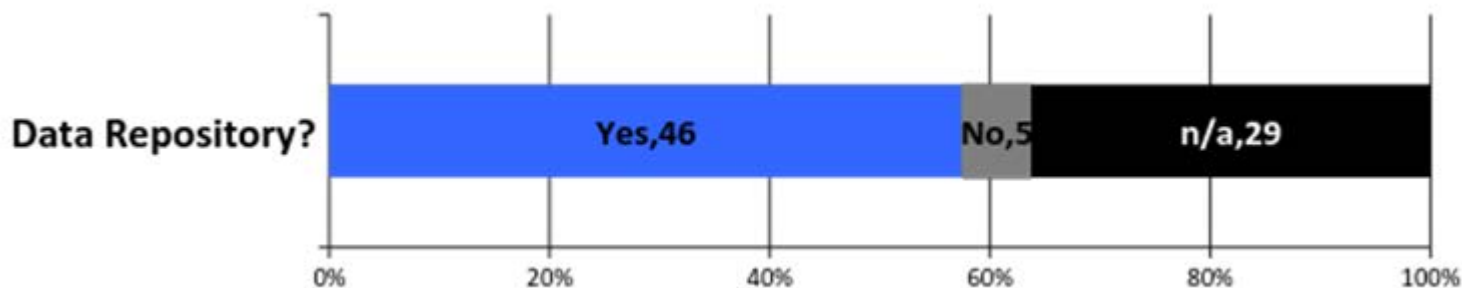
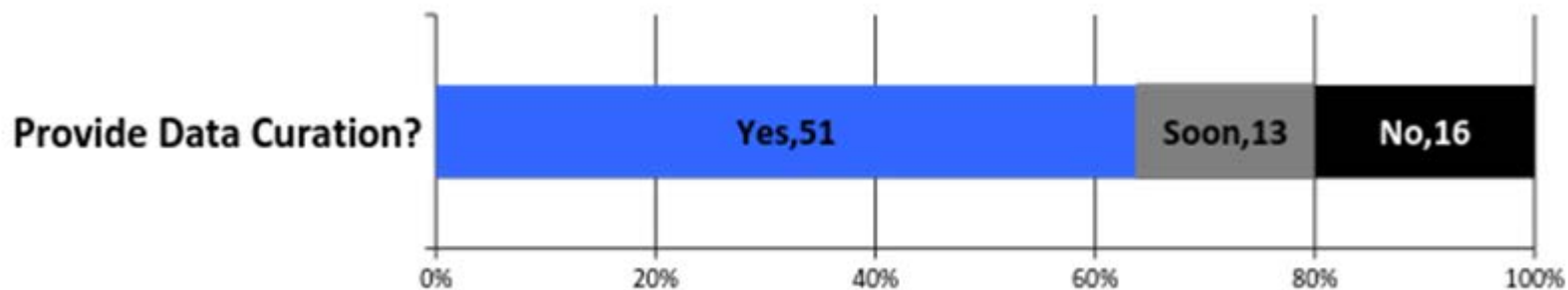
Lisa Johnston, Jake Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, and Claire Stewart.

May 1, 2017

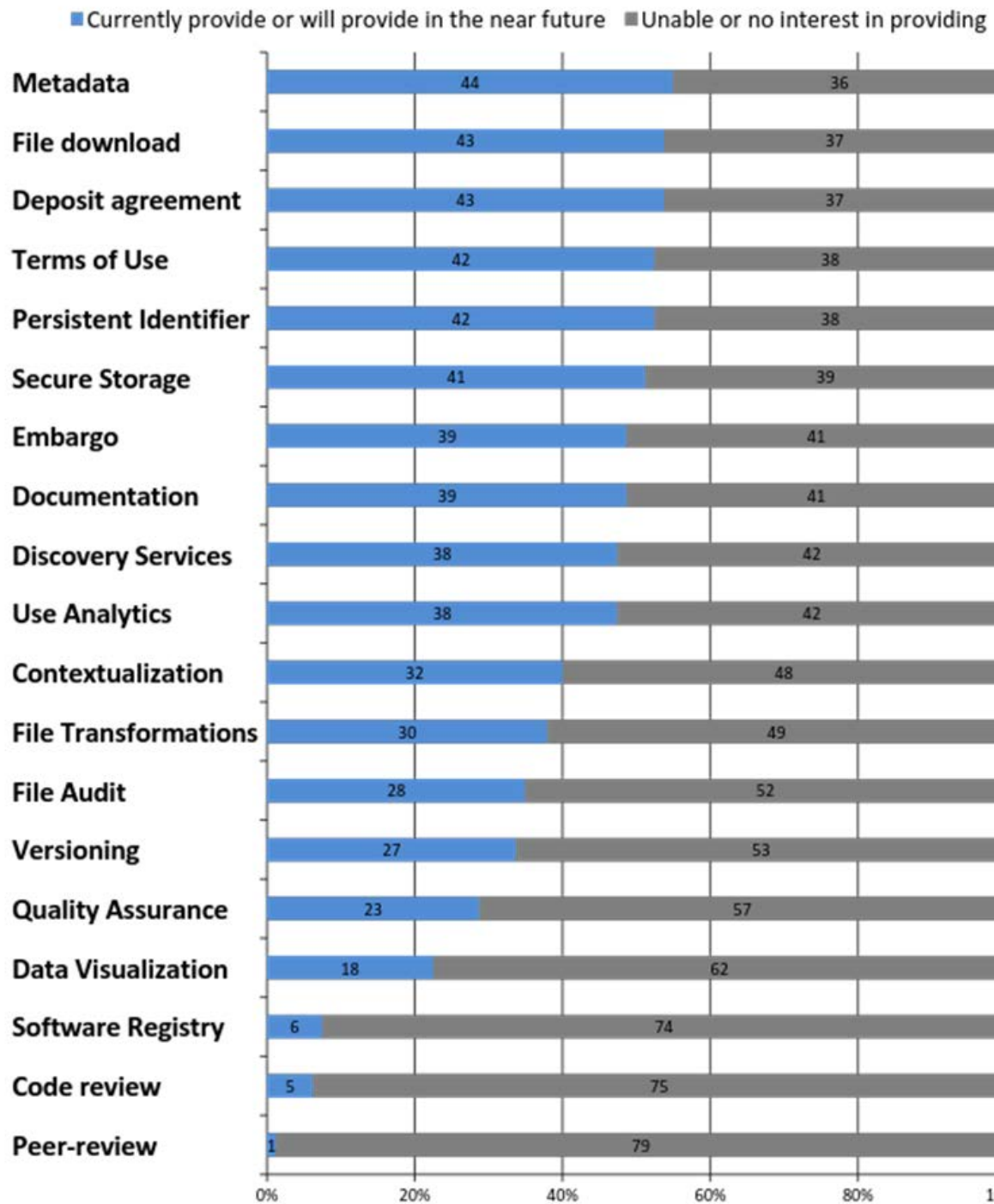
<https://sites.google.com/site/datacurationnetwork>
Data Curation Network

Libraries increasingly support data curation...

Responses to our forthcoming ARL SPEC Kit on Data Curation Services



By offering various types of services...



But library services may not be fully connecting with researcher needs...

Most Important Activities (4 out of 5)

- (Create) Documentation (4.6)
- Secure Storage (4.4)
- Quality Assurance (4.3)
- **Persistent Identifier (4.3)**
- **Software Registry (4.1)**
- Data Visualization (4.0)
- **File Audit (4.0)**
- (Create) Metadata (4.0)
- Versioning (3.9)
- **Contextualization (3.9)**
- **Code Review (3.9)**
- File Format Transformations (3.9)

Not Happening for Majority of Researchers

- **Persistent Identifier** (37% happens)
- **Software Registry** (41% happens)
- **File Audit** (16% happens)
- **Contextualization** (38% happens)
- **Code Review** (38% happens)

Happening, but not satisfactorily

- **Documentation** (26% satisfied),
- **Secure storage** (38% satisfied),
- **Quality Assurance** (14% satisfied),
- **Data Visualization** (12.5% satisfied),
- **Metadata** (29% satisfied)
- **Versioning** (13% Satisfied)
- **File Format Transformations** (29% satisfied)

But library services may not be fully connecting with researcher needs...

Most Important Activities (4 out of 5)

- **(Create) Documentation (4.6)**
- **Secure Storage (4.4)**
- **Quality Assurance (4.3)**
- Persistent Identifier (4.3)
- Software Registry (4.1)
- **Data Visualization (4.0)**
- File Audit (4.0)
- **(Create) Metadata (4.0)**
- **Versioning (3.9)**
- Contextualization (3.9)
- Code Review (3.9)
- **File Format Transformations (3.9)**

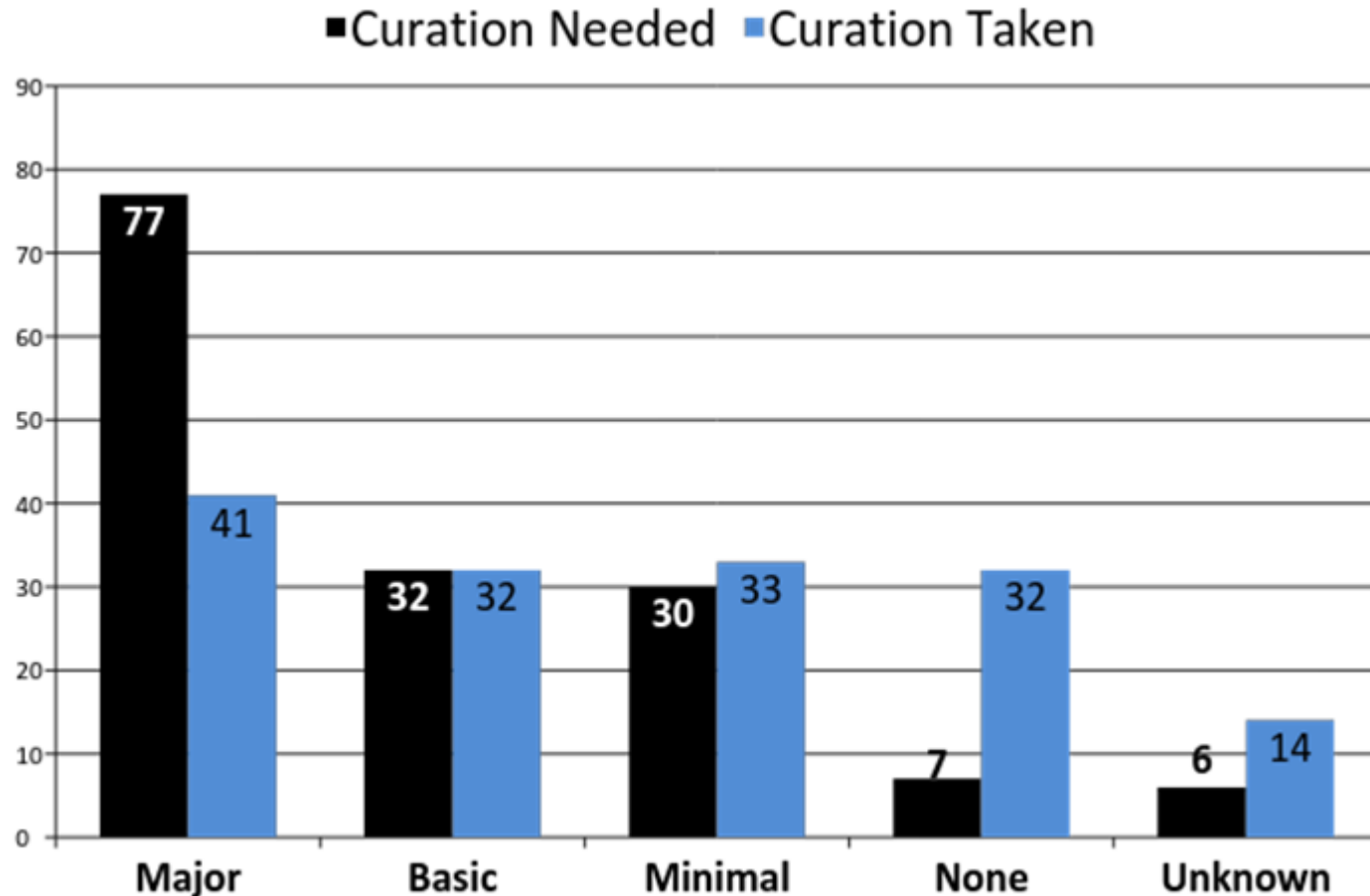
Not Happening for Majority of Researchers

- **Persistent Identifier** (37% happens)
- **Software Registry** (41% happens)
- **File Audit** (16% happens)
- **Contextualization** (38% happens)
- **Code Review** (38% happens)







Happening, but not satisfactorily

- **Documentation** (26% satisfied),
- **Secure storage** (38% satisfied),
- **Quality Assurance** (14% satisfied),
- **Data Visualization** (12.5% satisfied),
- **Metadata** (29% satisfied)
- **Versioning** (13% Satisfied)
- **File Format Transformations** (29% satisfied)

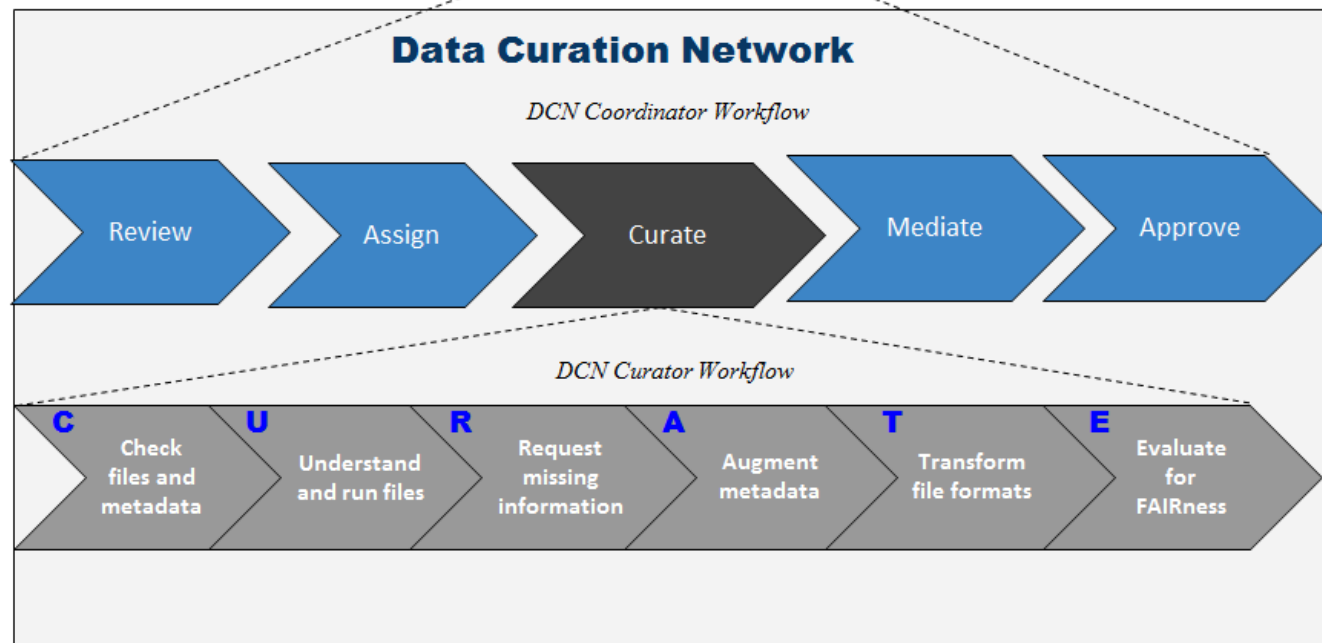
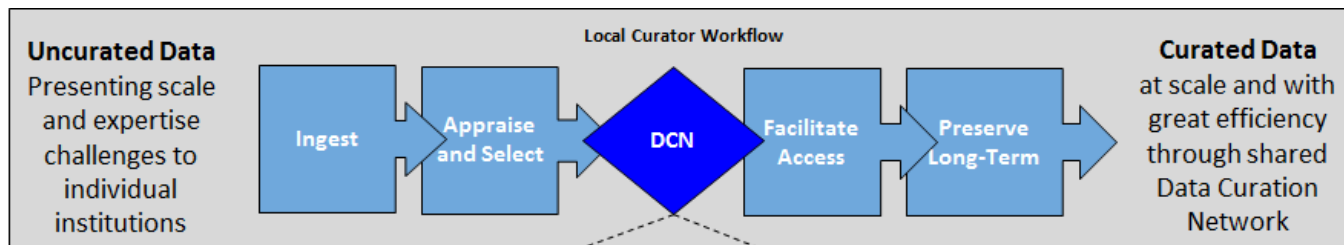
And their data curation needs may be more than one library can adequately support...



Particularly at current staffing levels...

	Repository	Curator Staff Full/Part Time	Subject/Format Expertise
 LIBRARIES UNIVERSITY OF MINNESOTA	Data Repository for the U of M (DRUM)	0 / 7	Excel, GIS, stat/survey (R, <u>STSS</u> , SAS), genomics, public health
 Cornell University Library	eCommons @Cornell	0 / 2	Excel, bio, environmental and earth sciences, EML
	Illinois Data Bank	0 / 3	Excel, stat/survey (R, <u>STSS</u> , SAS), biochem, audio/visual, TEI
	Deep Blue Data	1 / 3	Excel, bio, genomics, databases, health informatics
 PennState University Libraries	ScholarSphere	0 / 3	Excel, GIS, R, Python, PII/sensitive, ecology/ EML, bioinformatics
 Washington University in St. Louis UNIVERSITY LIBRARIES	Digital Research Materials Repository	0 / 3	Excel, GIS, Matlab, Mathematica, RDF/ XML, Qualitative Health

So, we intend to build a Data Curation Network...



To bridge gaps of providing and scaling data curation services.

The Data Curation Network will enable academic institutions to better support researchers that are faced with a growing number of requirements to ethically share their research data.

In the next 3-5 years we will...

1. Develop standards-driven data curation techniques for all types of repository workflows and infrastructure.
2. Expand into a sustainable entity that grows beyond our initial six partner institutions.
3. Datasets curated by the Data Curation Network will be used to advance research and education in ways that are measurably of greater reuse value than non-curated data.
4. Build an innovative community that enriches capacities for data curation writ large.

Project Team



Lisa Johnston (PI)
Research Data Management &
Curation Lead,
University of Minnesota



Jake Carlson
Research Data Services
Manager,
University of Michigan



Robert Olendorf
Science Data
Librarian,
Penn State University



Cynthia Hudson-Vitale
Digital Data Outreach
Librarian,
Washington University in
St. Louis



Heidi Imker
Director of the Research
Data Service,
University of Illinois



Wendy Kozlowski
Data Curation Specialist,
Cornell University



Claire Stewart
Associate University Librarian for
Research and Learning,
University of Minnesota

More Information on the DCN Project

Web:

[https://sites.google.com/site/
DataCurationNetwork](https://sites.google.com/site/DataCurationNetwork)

Twitter

#DataCurationNetwork



Alfred P. Sloan
FOUNDATION