

INFRASTRUCTURES AND PLATFORMS: DATA CURATION IN THE DIGITAL HUMANITIES

Lydia Zvyagintseva

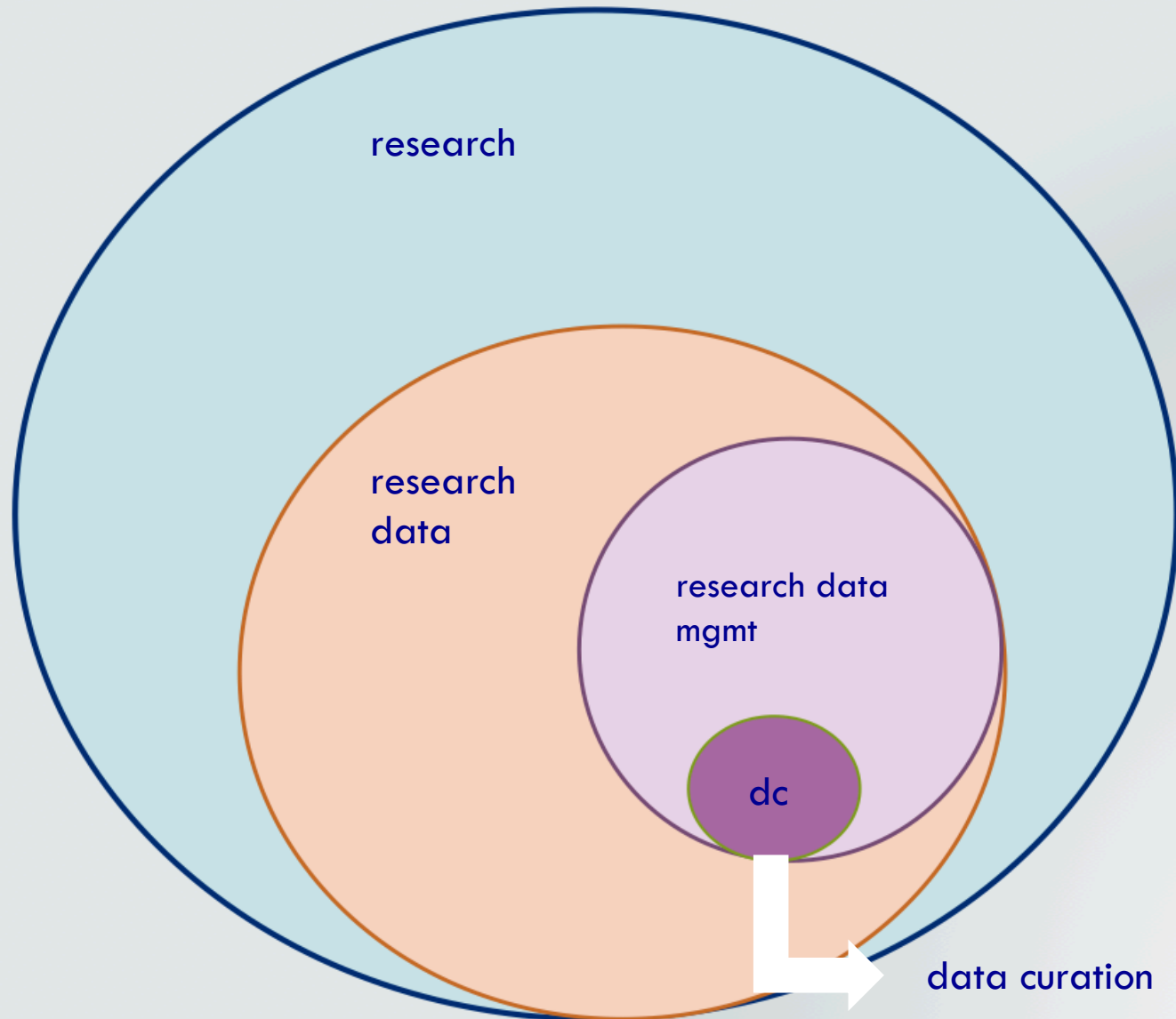
Forum for Information Professionals, Feb. 7, 2014

Problem



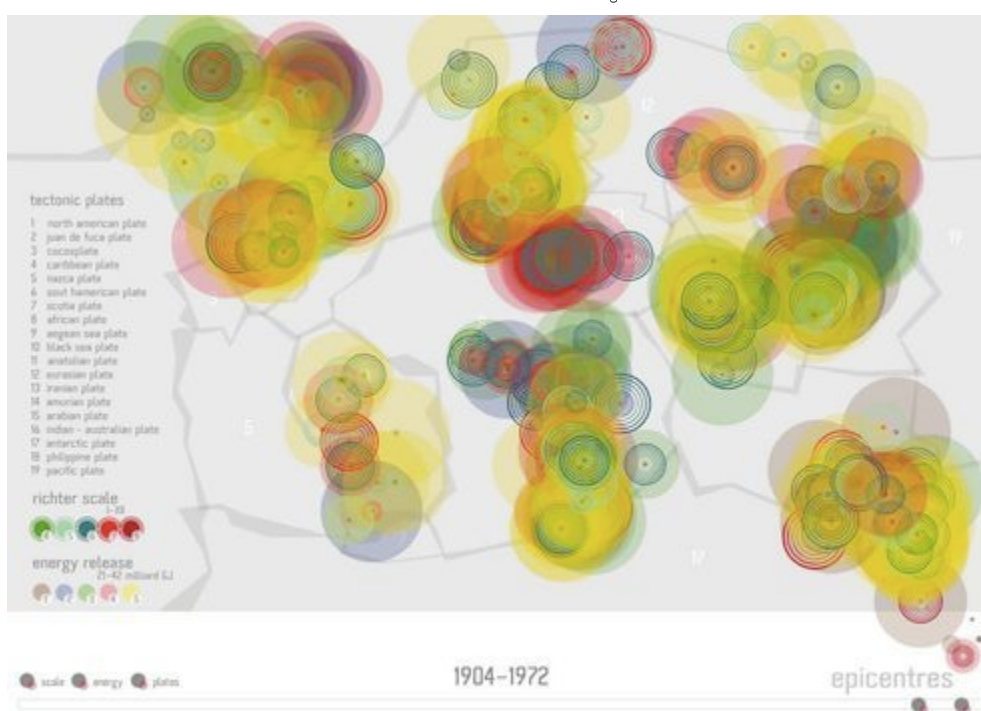
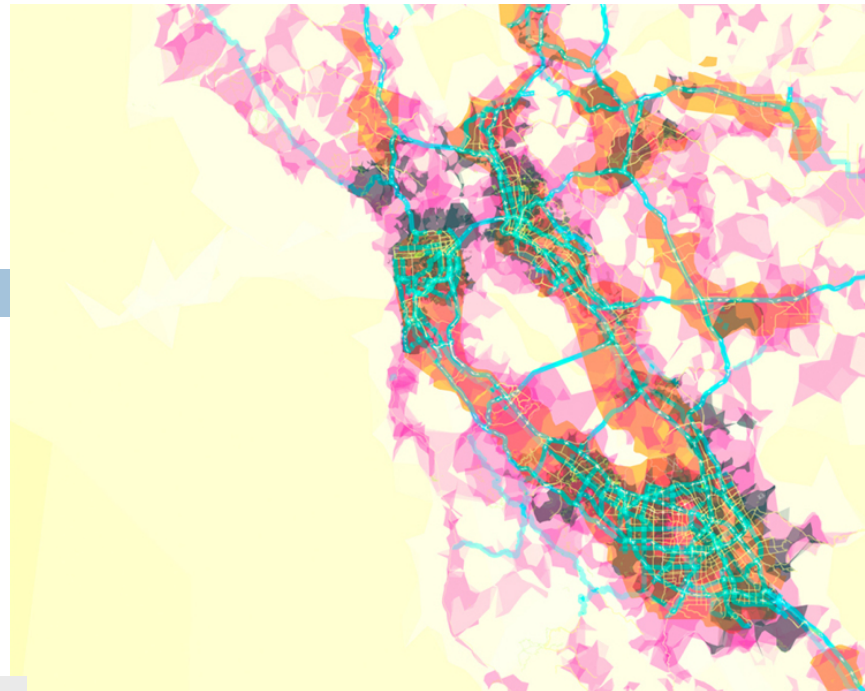
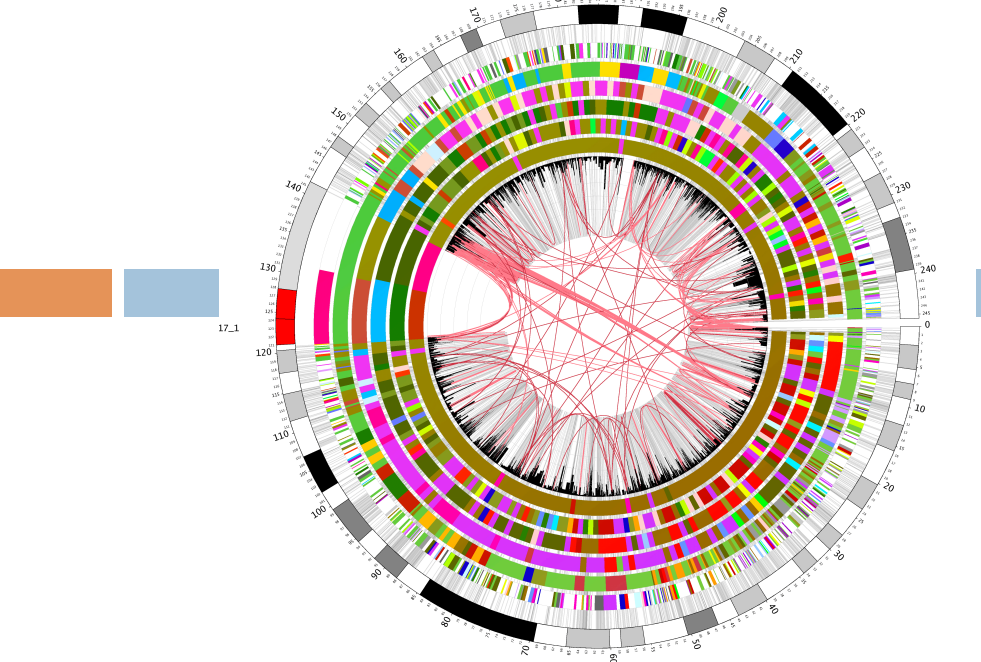
- What is the relationship between research data management and data curation as a set of scholarly practices?
- What does data curation in the digital humanities look like as library service?

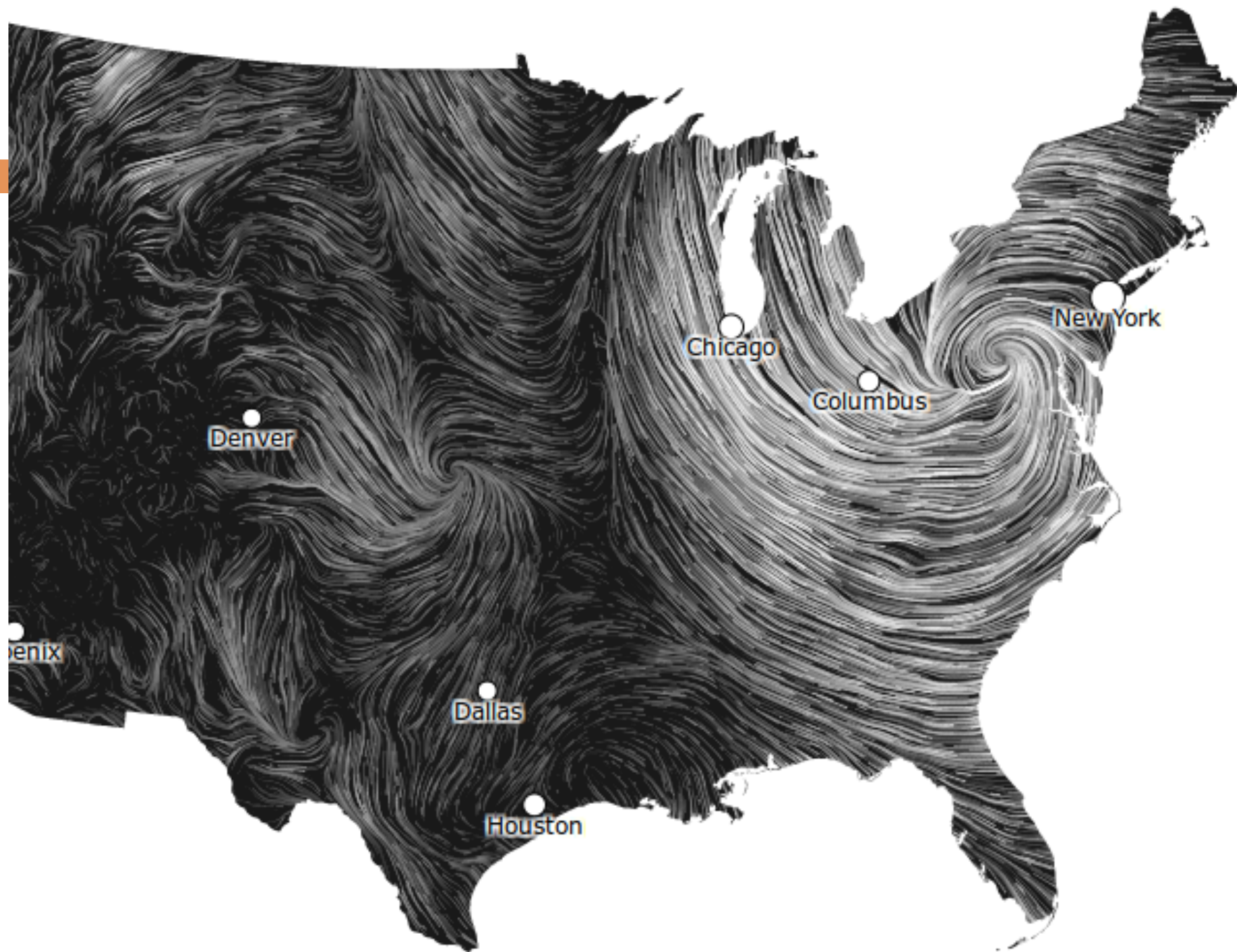
The Research Universe



Data: What is it?







Denver

Chicago

Columbus

New York

Dallas

Houston

Phoenix



Actually...



Design

- Literature review
- Website review
- Use cases

Definitions (JISC)

- Curation – managing and promoting the use of data; continuous enrichment and updating
- Archiving – ensuring that data is properly selected, stored, maintained over time
- Preservation – long-term storage and accessibility through successive change and obsolescence

Research Data Management



- Larger set of practices that includes institutional policy, planning, funding (local, regional, national, international), strategy development.
- Practical day-to-day of consultation, selection, acquisition, description, access, maintenance, preservation, storage, sharing, marketing, instruction.

Why humanities?



- ❑ Provide access to cultural information
- ❑ Enable manipulation of data
- ❑ Transform scholarly communication
- ❑ Enhance teaching and learning
- ❑ Make a public impact

In other words...



OPEN ACCESS

Use Cases (Applications)

- DataStaR @ Cornell
- ScholarSphere @ Penn State
- DRYAD
- Figshare
- Readux @ Emory



Cornell University



DATASTAR

[Home](#)

[People](#)

[Research Groups](#)

[Data Sets](#)

[Repositories](#)

[For Data Authors](#)

[Index](#)

Repositories

[Cornell University Geospatial Information Repository \(CUGIR\)](#) | GIS data repository | [CUGIR website](#)

[DataStaR \(Data Staging Repository\)](#) | Cornell institutional repository | [DataStaR](#)

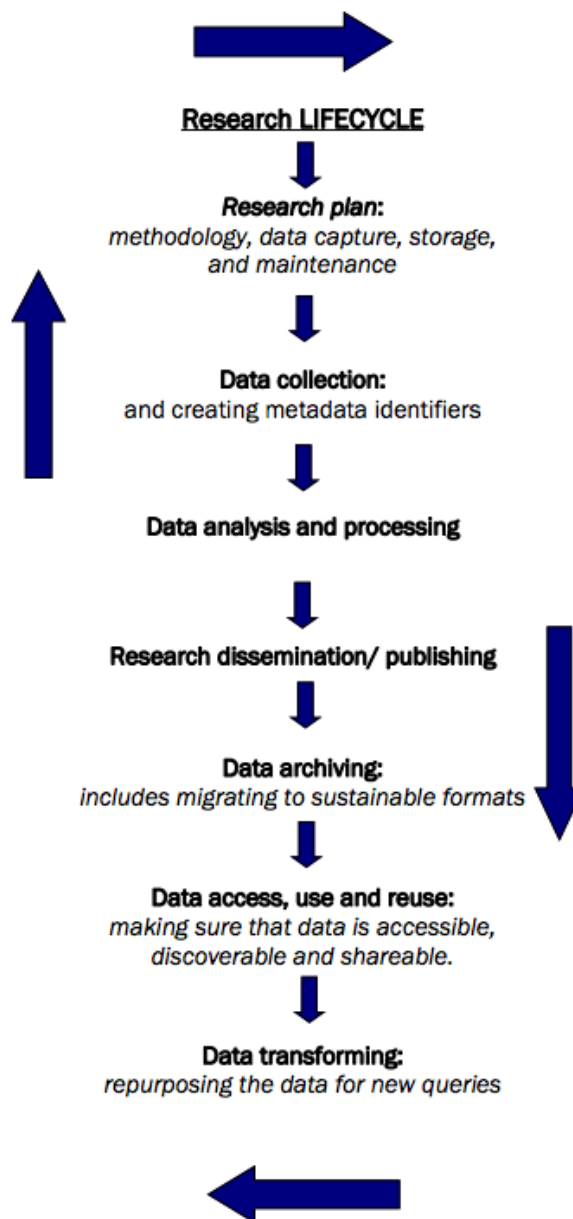
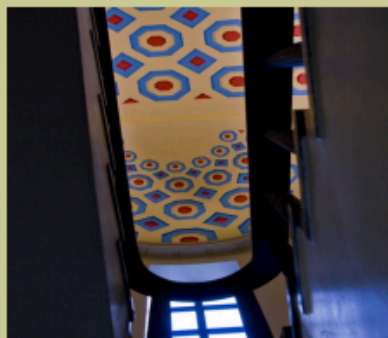
[eCommons@Cornell](#) | Cornell institutional repository | [eCommons@Cornell](#)

[Knowledge Network for Biocomplexity \(KNB\)](#) | Disciplinary data center | [Knowledge Network for Biocomplexity](#)

[New York State GIS Clearinghouse](#) | GIS data repository | [NYS GIS clearinghouse](#)

Data sharing is vital to ensuring the scholarly record for many reasons:

- **It is required** by NSF and NIH and likely to be adopted by other funders.
- **It fosters the reuse of data:** Allowing researchers access to your data opens up possibilities for new readings, new interpretations, and new discoveries.
- **It benefits research:** Being open about data can lead to sharing of process, workflows, and tools, all of which enhance the potential for replicating the research results.
- **It is good science:** Sharing data frames them as a public good, promoting community, collaboration, and conscientiousness.



Research Data Management Services Team (RDMST)

at the University Libraries



The Research Data Management Services Team (RDMST) at the University Libraries offers consultation services for data management planning to researchers applying for grant funding from agencies requiring a data management plan (DMP).

All data management inquiries should be addressed to RDMST at:
UL-RDMST@lists.psu.edu



About ▾

For researchers ▾

For organizations ▾

Contact us

Login Sign Up

1 DEPOSIT
DATA



2 GET PERMANENT
IDENTIFIER



<http://dx.doi.org/10.5061/dryad.20>

3 WATCH YOUR
CITATIONS GROW!



4 RELAX, YOUR DATA
ARE DISCOVERABLE
AND SECURE



Submit data now

[How and why?](#)

Search for data

Enter keyword, author, title, DOI, et

Go

[Advanced search](#)

Browse for data

Recently published

Popular

By Author

By Journal

Be part of Dryad



Publishers, societies,
universities, libraries,



discoverable

- Secure and accessible
- Easy to manage your research data
 - 1GB private space
 - Unlimited public space



shareable

- Publish your negative data
- Quick and simple upload
- All formats of research accepted



citable

- All published research is citable
 - Cloud based service
 - Always available

[Find out more](#)

Sign up for free

what's this?"/>[Sign up](#)

"Data which are determined to be a matter of public record and can therefore be made freely available, without restriction, are categorised as Public Access"

University of
Hertfordshire



"Data must be properly curated throughout its life-cycle and released with the appropriate high-quality metadata."

MRC

Medical
Research
Council

"NERC is committed to supporting long-term environmental data management to enable continuing access to these data."



NATURAL
ENVIRONMENT
RESEARCH COUNCIL



Start searching here...



browse media:

[My Dashboard](#)[My Profile](#)[My Groups](#)[My Notes](#)[My Bookmarks](#)[My Collections](#)[Need Help?](#)[Contact Us](#)

My Groups

Groups

Invitations

Groups I lead:

Group Title 1

Group Title 2

Group Title 3

[+ Create new group](#)

Groups I'm a member of:

Group Title 4

Group Title 5

Group Title 6

Group Title 1

[+ Invite New Member](#)[Email Group](#)

Members

Shared Collections

Group Settings

[Joebob Smith](#)[Susie Who](#)[Abraham Lincoln](#)[Bob Says-Hih](#)[+ Invite New Member](#)

BROWSE BY:

[Collection](#)
[Featured Collections](#)
[Public User Collections](#)
[Subject](#)
[Author](#)
[Date Range](#)
[Tags](#)

ABOUT US

[FAQ](#)
[CONTACT US](#)
[LOCATION](#)
[REGISTER NEW ACCOUNT](#)

CONNECT WITH US:



MY ACCOUNT:

[My Dashboard](#)
[My Notes](#)
[My Collections](#)
[My Bookmarks](#)
[View/Edit Profile](#)



Repository Statistics:

123,456 Total Repository Items

2,345,678 Items Downloaded

678,912 Downloads This Year

456 Total Members

120 Members Currently Online

Influence and impact



- Collection development
- Organization of knowledge (taxonomy, metadata)
- Systems and IT
- Instruction and outreach
- Reference and information seeking
- Leadership and organizational culture

Integrated services

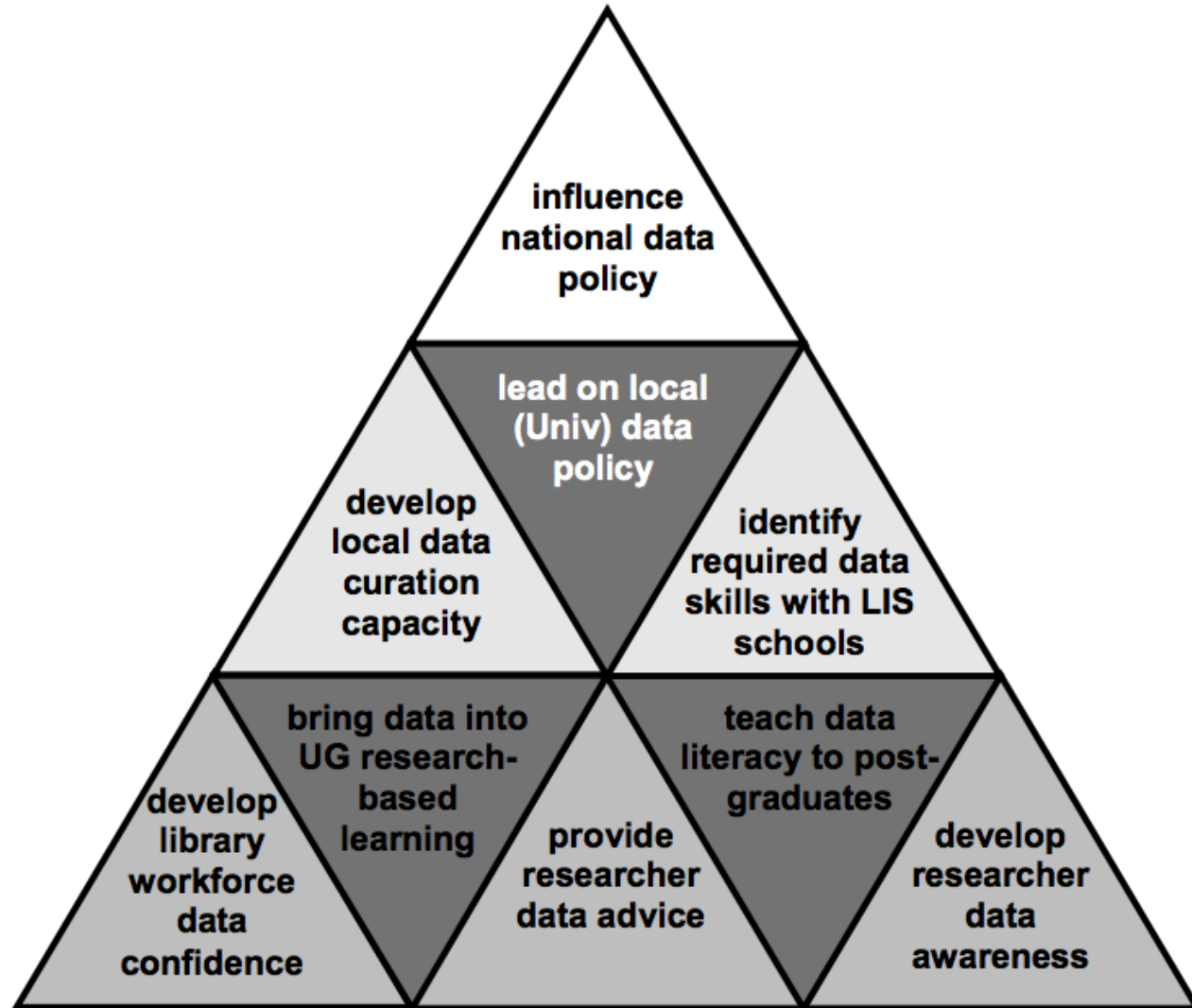


Figure 2 the research data management pyramid for libraries

Thanks!



- lydiaz@ualberta.ca
 - @lydia_zv