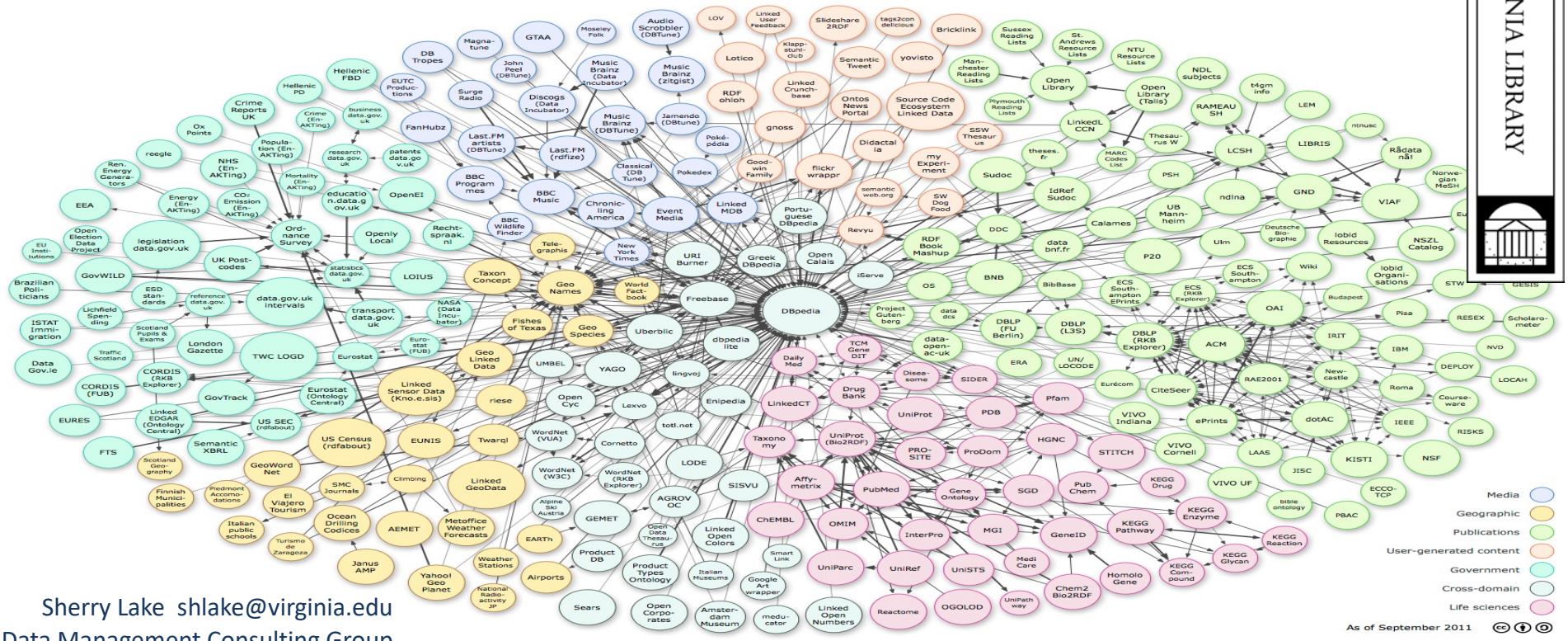
Choosing Between Data Sharing Repositories for Engineering

UNIVERSITY OF VIRGINIA LIBRARY



Sherry Lake shlake@virginia.edu
Data Management Consulting Group
University of Virginia Library

Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/

Today we will:

- Define data repository
- Discuss motivations for sharing data
- Describe issues around ownership and sharing of data
- Utilize tools to find data repositories
- Discuss considerations in choosing a data repository
- Visit some repositories in the life sciences

Data repositories:

- Storage location for datasets
  - Research data repositories
- Online
- Typically freely available, searchable
- Items (datasets) are permanent, linkable, citable

Reasons for sharing your data

- **Meeting expectations** of sponsors, funders, publishers and institutions
- **Linking** to research products like publications & presentations
- Enabling others to **replicate** and **verify results**
- Powering **future research and discovery**
- Receiving **credit for research** for career advancement

Can I share my data?

That depends on many factors

- Who owns the data?
- Requirements from sponsors, publishers, collaborators
- Institutional concerns such as IRB, data ownership
- Risks to sharing or not sharing the data
- Privacy and confidentiality issues with your data
- Commercial value of the data
- Intended uses of the data
- Method of sharing the data

Who owns my research data?

*"Data and notebooks resulting from sponsored research are the **property of the University of Virginia**. It is the responsibility of the principal investigator to retain all raw data in laboratory notebooks (or other appropriate format) for **at least five years** after completion of the research project (i.e., publication of a paper describing the work, or termination of the supporting research grant, whichever comes first) **unless required to be retained longer by contract, law, regulation**, or by some reasonable continuing need to refer to them."* – UVA Policy RES-002

Can I share my data?

Depends on:

- Requirements from sponsors, publishers, collaborators
- Institutional concerns such as IRB, data ownership
- Risks to sharing or not sharing the data
- Privacy and confidentiality issues with your data
- Commercial value of the data
- Intended uses of the data
- Method of sharing the data

Can I share my data?

At UVA you must request permission

The Associate VPR and your Dean will have to approve your request to share research data.

You will need to identify an appropriate repository or archive before seeking permission.

Why a data repository?

Why you shouldn't just put your data on a website

Probably no…

- Persistent identification
- Persistent access
- Provision for future preservation
- Professional backup

You will waste time…

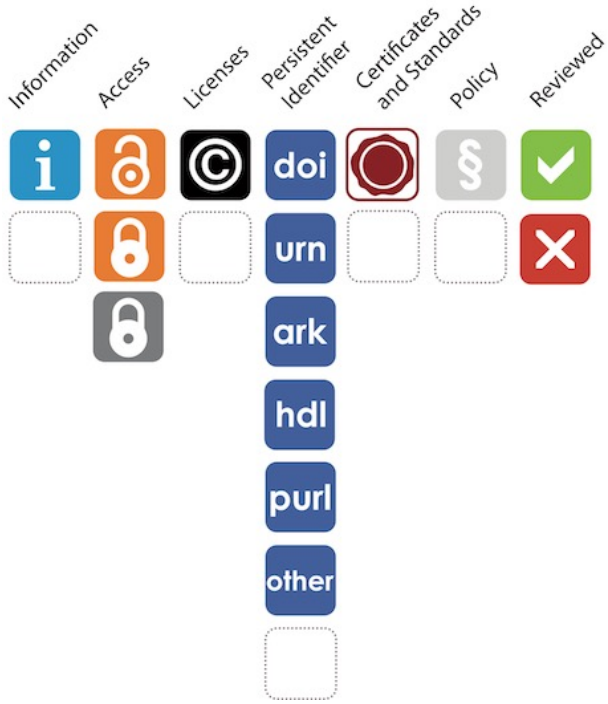
- Managing requests for access
- Preserving the data for reuse

An archive or repository can provide these, and more!

http://blogs.plos.org/mfenner/files/2013/06/figure2.png

Data Repository Advantages

- Persistent Identifiers -- unique and citable
- Access controls
- Terms of Use & Licenses
- Repository guidelines for deposit
- Data preservation -- migrating to new formats or emulating old formats
- Professional backup & documentation
- Repository Standards ensure commitment and quality



http://blogs.plos.org/mfenner/files/2013/06/figure2.png

Selecting a data repository

Questions to consider when selecting a repository or archive

- Does your **funder** specify a specific location or facility?
- Does your **discipline** recommend a specific repository or archive?
- Does your **publisher** require placement of data in support of an article in a specific location?
- Does your **institution** have specific requirements?

Data redundancy is important, so consider placing your data in at least two repositories or archives.

Selecting a data repository

- **Choose early:** There will be fewer surprises at the end of your research when you deposit your data.
- **Metadata:** Knowing the requirements at the start will enable you to design your data collection materials for easier metadata creation and facilitate your support documentation creation.
- **Persistent Identifiers:** Be sure the repository supplies one so your data is findable, citable, and can be linked to your publication(s).
- **Data embargo**: If you want to embargo your data be sure it is allowable, and learn about any restrictions before you submit.
- **Data access:** Identify any barriers that may limit or restrict data reuse.

Locating a data repository

International registries for data repositories

- **Databib** http://databib.org/
- **Re3data** http://re3data.org

You can start with these directories, or use them after determining if the funder, publisher, discipline, or institution have specific requirements.

Plus:
Simmons College hosts the Open Access Directory
http://oad.simmons.edu/oadwiki/Data_repositories

Exercise: Identifying a repository

- Search both **Databib** and **re3data**
- Locate a data repository for one of your projects.
- What types of data does the repository take?
- What additional information is needed for deposit (metadata)?

http://databib.org

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

http://re3data.org

UVa Institutional Repository: Libra

Libra

UVa Institutional Repository

- Online since 2011
- Thesis and dissertations
- Articles
- Conference paper, posters
- Article preprint
- Book
- Chapter in an edited collection
- Datasets

Other options

figshare: http://figshare.com/
"figshare is a repository where users can make all of their research outputs available in a citable, shareable and discoverable manner." Researchers can upload research in any format, and can include negative data.

bitbucket: http://bitbucket.org
Bitbuket is an online solution to hosting code, code management with distributed version control

GitHub: https://github.com/
Github is an online repository where you can share code and other files. It too offers version control.

We're available to help

- The Data Management Consulting Group provides consulting and training services to UVA researchers and graduate students in all aspects of data sharing.
- We can help you navigate and negotiate through the tricky issues and many approvals in order to responsibly share your research data.
- Contact us at dmconsult@virginia.edu

Photo credit
http://vprompt.com/wp-content/uploads/2013/10/data-mining-300x154.jpg

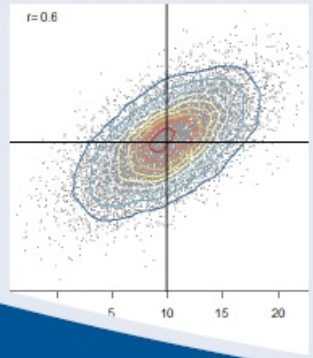
Additional links

- **Data Rights and Responsibilities Guidance 1.0** developed jointly by the Data Management Consulting Group, Office of General Counsel, and Office of the Vice President for Research - http://dmconsult.library.virginia.edu/data-rights-and-responsibilities-guidance-1-0/

- **Institutional Data Protection Standards** provided by the Information Security, Policy, and Records Office (ISPRO) - http://www.virginia.edu/informationsecurity/dataprotection/

RESEARCH DATA SERVICES

Offering expert data assistance at every stage of the research process.



1: PLANNING

We can assist you with developing a data management plan and designing your planned data analysis, including:

- Implementing plans, using tools, and creating workflows for managing research data
- Advising on study design, power analysis, and choice of statistical methods
- Helping to meet increasingly stringent criteria from funding agencies

2: FINDING & COLLECTING

We have access to thousands of sources of data and experts who will help you:

- Locate, evaluate and format data
- Create metadata and data documentation protocols for new data collection
- Capture data using best practices and appropriate technology

3: ANALYZING

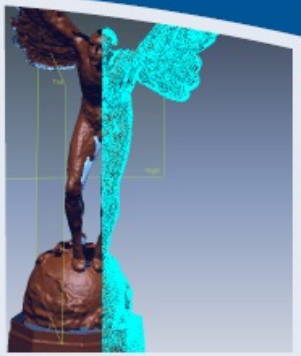
Get expert assistance from statistical, spatial, or media specialists to analyze your data and present your research:

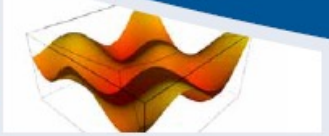
- Learn to use cutting-edge tools and methods
- Experiment with high-resolution visualization technologies
- Develop graphical representations that bring impact to your analysis

4: SHARING & ARCHIVING

We can consult with you on strategies to help others discover or access your research by:

- Adhering to data sharing policies and norms
- Selecting a data-sharing repository
- Making your data easier to discover and reuse





University of Virginia Library

Workshops • 1:1 Consultations • Class Presentations

Contact us at researchdataservices@virginia.edu