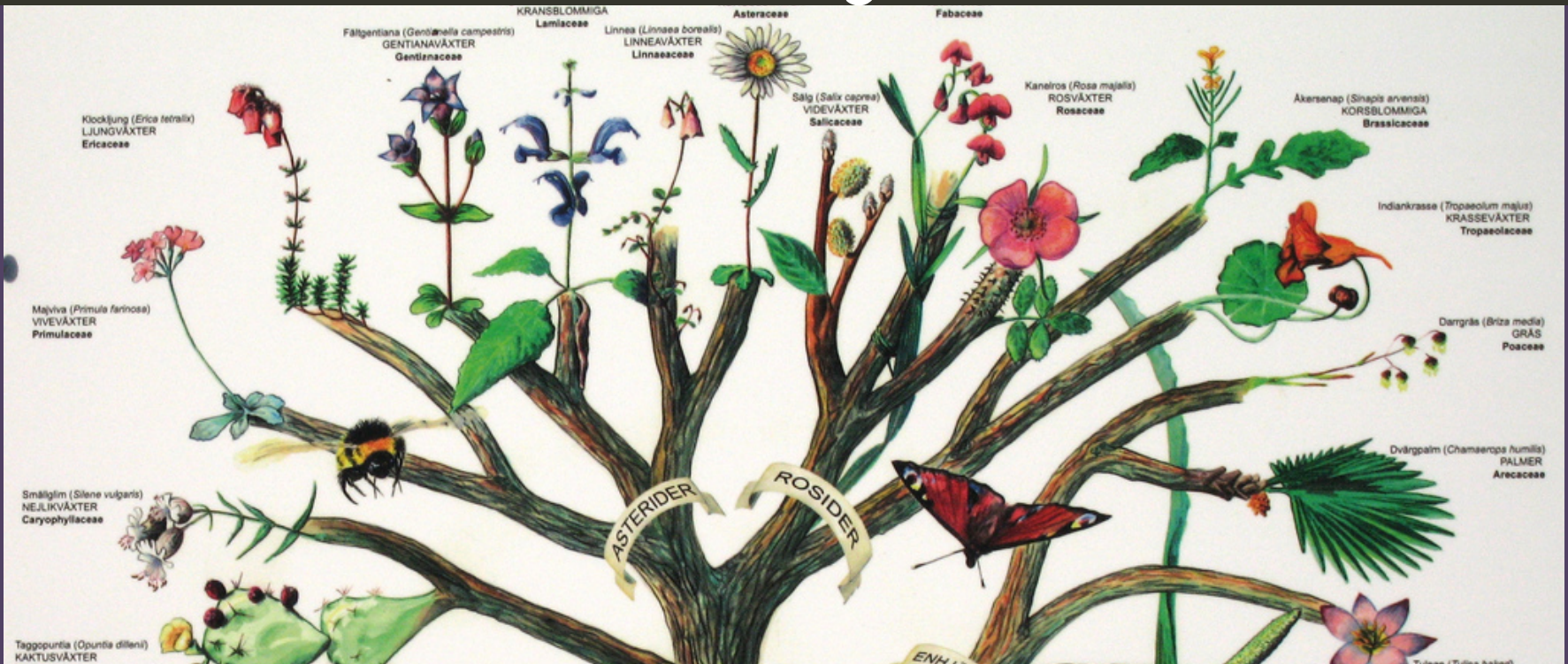**BLOMVÄXTERNAS EVOLUTION OCH MÅNGFALD**

Alla blomväxter härstammar från en växt som fanns på jorden för mer än 120 miljoner år sedan. Liksom hos nu levande blomväxter var dess frön inneslutna i en frukt och den hade blommor med ståndare och pistiller.

Revolutionary or Evolutionary?
Adapting Best Practices for Data Management

Aaron Albertson, Macalester College

Jonathan Carlson, College of St. Benedict/St. John's University

Kristin Partlo, Carleton College

Diana Symons, College of St. Benedict/St. John's University

What to Expect:

- Data Management Plans – Jonathan

- Talking to Faculty – Kristin

- Mapping Services to Need – Aaron

- Strategies for Getting Started – Diana

"In the context of research and scholarship, 'Data Management' refers to the storage, access and preservation of data produced from a given investigation. Data management practices cover the entire lifecycle of the data..."

Data Management Lib Guide
http://guides.library.tamu.edu/DataManagement

GRADUATE SCHOOL OF **LIBRARY AND INFORMATION SCIENCE**
The iSchool at Illinois

Specialization in Data Curation

certificate in community informatics

specialization in data curation

information in society specialization

socio-technical data analytics

special collections certificate

certificate in youth services

Search   Go

contact us   login

Data curation is the active and ongoing management of data through its lifecycle of interest and usefulness to scholarship, science, and education. Data curation enables data discovery and retrieval, maintains data quality, adds value, and provides for re-use over time through activities including authentication, archiving, management, preservation, and representation. The Specialization in Data Curation offers a focus on data collection, representation, and management, digital preservation and archiving, data standards, and policy. Our program provides a strong focus on the theory and skills necessary to work directly with academic, governmental, public, industrial, research center, etc. data to develop expertise in the curation of research data.

Program Overview (PDF, 1.51 MB)

The Graduate School of Library and Information Science (GSLIS) at the University of Illinois is the top ranked LIS school in the nation, situated within a world-class public research university with one of the world's largest library collections. Library and Information Science (LIS) as a discipline emphasizes the use of information technology to support new approaches to the organization of and access to information. As such, GSLIS is well suited to provide a specialization that prepares students to plan and manage data curation systems, create and maintain data collections, and to evaluation and apply data and metadata standards for varied uses across the sciences, humanities, and social sciences.

The Data Curation Specialization may be earned either as part of an ALA-accredited Master of Science in Library and Information Science (MSLIS) degree or, for students who have already completed their master's degree, as part of a Certificate of Advanced Study (CAS) degree. The MSLIS degree requires a total of 40 credit hours of course work that includes required core courses outlined below. Those in the specialization must complete two courses required of all MSLIS students as well as three courses required for the specialization. Students in the data curation specialization work with their faculty advisors to select electives for a more individualized program that will prepare them for either a general or specialized career path. At least two courses from the recommended data curation list are required for the specialization. We

GRADUATE SCHOOL OF LIBRARY AND
INFORMATION SCIENCE
The iSchool at Illinois

Specialization in Data Curation

Data curation is the active and ongoing management of data through its lifecycle of interest and usefulness to scholarship, science, and education. Data curation enables data discovery and retrieval, maintains data quality, adds value, and provides for re-use over time through activities including authentication, archiving, management, preservation, and representation.

The Data Curation Specialization may be earned either as part of an ALA-accredited Master of Science in Library and Information Science (MSLIS) degree or, for students who have already completed their master's degree, as part of a Certificate of Advanced Study (CAS) degree. The MSLIS degree requires a total of 40 credit hours of course work that includes required core courses outlined below. Those in the specialization must complete two courses required of all MSLIS students as well as three courses required for the specialization. Students in the data curation specialization work with their faculty advisors to select electives for a more individualized program that will prepare them for either a general or specialized career path. At least two courses from the recommended data curation list are required for the specialization. We

Basic Data Services Throughout the Library

- **Collections** of existing datasets

- **Cataloging and providing access** points to those collections

- **Reference**: helping students and faculty discover, access, and evaluate data sources in and beyond the library's collections

- **Instruction**: teaching students about the information landscape of data, and how to build good data management practices into their habits as researchers and as individuals

Data Management Plans

(and a little on Data Sharing)

Who requires Data Sharing or DMPs?

- **NIH** – data sharing mandate since October 1, 2003
  – Only applied to "final research data" on grants seeking > $500,000/year
- **NSF** – DMP required since January 18, 2011
  – Required Supplementary Document for ALL grants
- **Presidential Policy Memo**, February 22, 2013
  – "Increasing Access to the Results of Federally Funded Scientific Research"
- **Several Journals**
  – Ex: Dryad partners

Requirements by Directorate, Office, Division, Program, or other NSF Unit

Links to data management requirements and plans relevant to specific Directorates, Offices, Divisions, Programs, or other NSF units, are provided below. If guidance specific to the program is not provided, then the requirements established in Grant Proposal Guide, Chapter II.C.2.j apply.

Please note that if a specific program solicitation provides guidance on preparation of data management plans, such guidance must be followed.

- Biological Sciences Directorate (BIO)
    - Directorate-wide Guidance

- Computer & Information Sciences & Engineering (CISE)
    - Directorate-wide Guidance

- Education & Human Resources Directorate (EHR)
    - Directorate-wide Guidance

- Engineering Directorate (ENG)
    - Directorate-wide Guidance

- Geosciences Directorate (GEO)
    - Directorate-wide Guidance

- Mathematical and Physical Sciences Directorate (MPS)
    - Division of Astronomical Sciences
    - Division of Chemistry
    - Division of Materials Research
    - Division of Mathematical Sciences
    - Division of Physics

- Social, Behavioral and Economic Sciences Directorate (SBE)
    - Directorate-wide Guidance

NSF Data Sharing Policy

Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. See Award & Administration Guide (AAG) Chapter VI.D.4.

NSF DMP Requirements

Proposals submitted or due on or after January 18, 2011, must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results. See Grant Proposal Guide (GPG) Chapter II.C.2.j for full policy implementation.

Grant Proposal Guide, Chapter II.C.2.j

The DMP should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results and may include:

1) the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;

Grant Proposal Guide, Chapter II.C.2.j

The DMP may include:

2) the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);

Grant Proposal Guide, Chapter II.C.2.j

The DMP may include:

3) policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;

Grant Proposal Guide, Chapter II.C.2.j

The DMP may include:

4) policies and provisions for re-use, re-distribution, and the production of derivatives; and

Grant Proposal Guide, Chapter II.C.2.j

The DMP may include:

5) plans for archiving data, samples, and other research products, and for preservation of access to them.

DMP Exercise: The Scenario

As a librarian, you have been asked by a professor at your institution for feedback on his or her draft NSF Data Management Plan.

DMP Exercise

In your packet:

- A copy of the General NSF Data Management Plan Guidelines

- A real draft DMP

- In the colored paper (DO NOT OPEN YET) – my responses to the draft DMP

Tasks:

- Read through your Draft DMP

- Together, discuss and create a list of recommendations you would make to the researcher to improve their DMP

- Choose a spokesperson to report back to the large group

Exercise Two:

TALKING WITH FACULTY

"None of the researchers interviewed for this study have received formal training in data management practices, nor do they express satisfaction with their level of expertise. Researchers are learning on the job in an ad hoc fashion."

First Key Finding of the 2012 CLIR Study
"The Problem with Data: Data Management and Curation Practices Among University Researchers"
http://www.clir.org/pubs/reports/pub154/problem-of-data

Data Curation Profiles Toolkit

Welcome to the Data Curation Profiles Community!

This website is an environment where academic librarians of all kinds, special librarians at research facilities, archivists involved in the preservation of digital data and those who support digital repositories can find help, support and camaraderie in exploring avenues to learn more about working with research data and the use of the Data Curation Profiles Toolkit.

A Data Curation Profile is essentially an outline of the "story" of a data set or collection, describing its origin and lifecycle within a research project. The Profile and its associated Toolkit grew out of an inquiry into the changing environment of scholarly communication, especially the possibility of researchers providing access to data much further upstream than previously imagined. If researchers are interested in sharing or forced to provide access to data sets or collections, what does that mean for the data, for researchers, and for librarians?

Data Curation Profiles can:

- provide a guide for discussing data with researchers
- give insight into areas of attention in data management
- help assess information needs related to data collections
- give insight into differences between data in various disciplines
- help identify possible data services
- create a starting point for curating a data set for archiving and preservation

Goal of this exercise:
Participants, through discussion of the DCP Toolkit, will start to adapt an existing resource to their own setting and develop strategies for communicating with faculty at their own institutions.

Task:
Discuss the summary addressing the questions on the front of your handout.

Five Questions:

- Does your **research, creative work or teaching** require you to store a lot of/any information?
- If so, where do you **store** it? How do you store it? [Follow-up: backup]
- Do you ever **share** this information with anyone (e.g., other researchers, future students, current collaborators, etc.) beyond publishing an article or book?
- Do you hope to **archive** any of this work for the long term? If so, do you have plans for doing so? Are you hoping that Carleton will help you with this?
- What would "**help**" look like?

Questions to consider* when planning for the life of research materials, be they hand-written interview forms, digital images or quantitative data. How can faculty, librarians, and academic technologists work together to help prepare students to ask themselves these questions about their own work?

1. **Description**: What kinds of materials do you collect and create in the course of your research and how do those materials relate to the finished product of the research? What formats are they in and can you estimate their size?
2. **Stages**: What are the "lifecycle" stages of the materials as you move through the different stages of your research? Where do you store your research materials?
3. **Sharing**: At any point in your research lifecycle would you want or be willing to share your research materials with others?
4. **Access**: Would you consider putting the research materials into a collection or repository for use by other researchers or students?
5. **Transfer**: What would you need to do to prepare, describe and organize the materials in order to make them usable by someone else?
6. **Organization and Description**: How are your research materials organized and described? Are they organized and described sufficient for someone else to use them? Are they organized and described according to disciplinary norms or standards?
7. **Discovery**: Who is most likely to want to share your research materials and which groups are of the highest priority: students and researchers at your institution, students and researchers in your field, outside your field, or the general public?

• • •

Mapping Services to Need

Intro:
The Wilderness

What happened at Macalester:

A brief history of our data efforts from the past two years.

Initial Macalester Proposal

- "Advising on long term storage, preservation, and management."
- "Providing access that will enable continued use."
- "Linking to related scholarship as time and resources allow."
- "Providing metadata and standards that will facilitate discovery."

Since then:

Making connections across campus.

Where Macalester is now:

Honing the definitions of our services.

Envisioned categories:

- Planning For Your Data Needs
- Finding/ Gathering Data
- Working With And Organizing Data
- Preserving And Making Data Available For Sharing

Carleton's Approach

- **Finding and downloading files** - *Librarians*
- **Collecting and creating** - *Not identified*
- **Reformatting** - *Academic Technologist, GIS Specialist, Librarian*
- **Statistical Software Support** - *Academic Technologist, GIS Specialist*
- **Collection Management** - *Academic Technologist, GIS Specialist*

Carleton's Approach Cont.

- **Cleaning Data** - *Academic Technologist, GIS Specialist, Various Faculty*

- **Choosing Analysis Type** - *Academic Technologist, GIS Specialist*

- **Data Visualization** - *Academic Technologist, GIS Specialist, Librarian*

- **Analysis Process** - *Faculty, Classmates, Academic Technologist, GIS Specialist*

- **Writing It Up** - *Writing Center*

Discussion:

- How would you work to identify services the library should provide?
- What services do you think your library should provide?
- Who on your campus should you be partnering with to provide services?

Strategies for Getting Started

So Far We've Touched On…

- Data Management Plans
- Talking to Faculty
- Mapping Services to Need

What Steps Can We Take After LibTech?

Step 1: Keep Learning

- Look for useful reports, articles, upcoming workshops, conference presentations, etc.

- "Recommended Resources" at http://libguides.csbsju.edu/libtech2014

- LAREDAS listserv

Step 2:
Review Other Institutions

- Identify peer or aspirant institutions
  - Check their websites
  - Existing services
  - Organizational, service, & funding models
  - Who are their experts?

Step 3: Share What You Learn

- Contact us if you'd like us to add resources to http://libguides.csbsju.edu/libtech2014

- Library staff education
  - LibTech recap
  - Create an informal discussion group
  - Plan training sessions

Step 4: Needs Assessment

- **Explore your current data management needs**

  - **How is research data currently being created, stored, or used?**

  - **Institutional/departmental/individual strengths and weaknesses?**

Step 5: Create a Guide

- Develop a data management guide
  - Who will maintain this guide?
  - Who should faculty researchers contact for which services? Make it clear!
  - Further refine workflows, policies, & service models

Step 6:
Consider Institutional Culture

- "Institutional culture plays a significant role in determining what research data management services are needed, including how these services might need to evolve as new technologies emerge on campus."

~Jina Choi Wakimoto

Step 7:
Learn About Your Researchers

- Get a list of grant-funded projects from your institution's grants office

- Plan conversations with faculty

- Survey researchers' data management practices and needs

- Undergraduate or graduate student research activities?

Step 8: Identify Collaborators

- "In our small college environment, there was simply no way that one person or even one department could hope to build the infrastructure to support meaningful research data management." ~ Sarah Goldstein & Sarah Oelker (Mount Holyoke College)

- "Collaboration is essential in developing research data management services" ~Jina Choi Wakimoto

Step 8: Identify Collaborators

- Contact potential partners & stakeholders (IT, Office of External Grants, various research offices, etc.)
- How can you include researchers from various departments? They'll have different data management needs.
- Who else has relevant expertise?

Step 9: Storage & Maintenance Options

- Currently available data storage options?
  - Schedule a meeting with IT
  - Check on departmental storage options
- If you have a digital repository:
  - What are its storage limitations for data?
  - Are there clear advantages to other data management systems?

Step 10: Be Mindful of Differences

- "Libraries, data centers, academic departments—all organizations where data curation can be done—have varied, sometimes idiosyncratic, approaches and often entail different attitudes, cultures, and practices." ~The Problem of Data

- Brainstorm how best to bridge these organizational differences

Questions?
Comments?
Suggestions?