

Initializing and Advancing Data Management Conversations with Faculty, Researchers, & Scientists

Plato L. Smith II, University of Florida Libraries
Science Boot Camp for Librarians Southeast 2016:
Librarian Lightning Talks
University of Georgia

Identify Stakeholders

Communicate Stakeholders Interests in Data Management

Facilitate Collaborations among Stakeholders

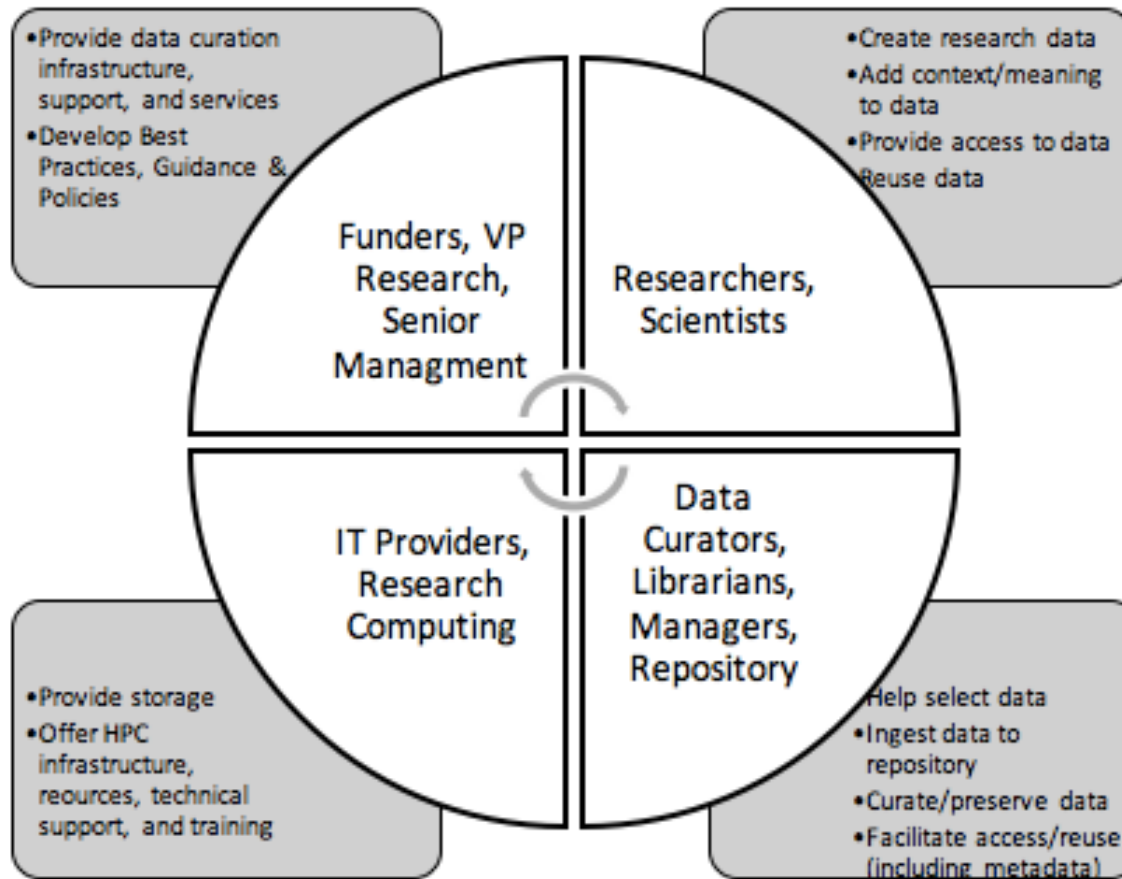


Fig. 1 Stakeholders' roles aligned with Data Curation Lifecycle Responsibilities

DAF Survey - Q26. How can the University (including your School, research lab cyberinfrastructure, and campus resources such as high performance computing (HPC)) make data management and storage easier for you?

• **Key Themes**

- Keep system stable & reliable
- Provide cost-effective backup
- Provide more storage & better transfer speed
- Relax policies for purchasing computer infrastructure to allow computers to be used across multiple projects
- Provide expertise in data management to provide advice when necessary
- Provide resources (financial and personnel) to document the data
- Provide more guidance on best practices



Fig. 2 Wordcloud of the 54 survey responses to DAF Survey Q26 (250 words)

<https://www.jasondavies.com/wordcloud/>

DAF Interview - Q41. On the scale of 1-3, please rate your level of confidence and awareness of the following research data management (RDM) and data management plan (DMP) matters.

#	Question	Not Very Confident	Confident	Very confident
1	Consideration of the RDM requirements/costs at the bidding stage	3	0	3
2	Expertise in developing DMP for bidding applications	2	0	4
3	Fulfilling research grant RDM obligations	1	0	5
4	Awareness of university RDM facilities and DMP policies	2	2	2
5	Ease of access to the research data during the life-time of the research project by yourself	0	1	5
6	Ease of access to the research data by collaborators	0	3	3
7	Re-accessing the data after the project life-time	1	3	2

Developing a Data Management Use Case at UF

From Bootcamp to Zenodo

- UF Digital Humanities Bootcamp meeting (2016)
- Research Computing Advisory Committee (RCAC)
- Department of Molecular Genetics & Microbiology Professor
 - Datasets (Variant Call Format .vcf; Common Separated Value .csv)

Dataset in Campus IR & Zenodo

The screenshot shows the Zenodo interface for a dataset. The header includes the Zenodo logo and the tagline "Research. Shared." with navigation links for Search, Communities, Browse, Upload, and Get started. There are also buttons for Sign In and Sign Up.

The dataset page displays the following information:

- Publication date:** 19 May 2016
- DOI:** 10.5281/zenodo.51693
- Keyword(s):** Genetics, Drosophila melanogaster, VCF file
- Related publications and datasets:** Supplement to: 10.1101/051286-015-1415-6
- Collections:** Communities > Large 'static' processed data files, Datasets, Open Access
- License (for files):** Creative Commons Attribution-NonCommercial-NoDerivatives
- Uploaded by:** plato (on 19 May 2016)

A note indicates: "Note: The DOI for this file is currently pending in Genetics <http://www.genetics.org/>."

Name	Date	Size	Download
SupplementalData1.tgz	19 May 2016	359.7 MB	Download

At the bottom right, there is a "New to Zenodo?" section with a "Sign Up" button and a link to "Read more about features and benefits."

Fig. 3 Large dataset deposited in an external repository(359.7 MB)

EarthCube

- What is EarthCube?
 - Started in 2011
 - Funded by NSF's Directorate for Geosciences (GEO) and Division for Advanced Cyberinfrastructure (ACI)
 - Improve data access, sharing, visualization, & analysis of geoscience data & related resources
- How to get involved?
 - Funding Resources
 - Governance
 - Connect with Funded Projects
 - Meetings & Conferences
 - Committees & Teams
 - Working Groups

<http://earthcube.org/info/get-involved>

<http://earthcube.org/info/about>

Thank you

Questions and comments

Plato Smith

plato.smith@ufl.edu