# **Assessing Scientific Data Management Practices and Needs**

Sherry Lake, Bill Corey, Andrew Sallans Scientific Data Consulting Group, University of Virginia Library, Charlottesville, Virginia, U.S.A.

## Abstract

SciDac

The University of Virginia Library is supporting the new data management requirements in science and engineering by developing a model that draws upon the close collaboration between data experts and subject librarians, and culminates in policy and infrastructure recommendations to the University's Office of the Vice President for Research (VPR) and the Office of the Vice President/Chief Information **Officer** (VP/CIO). The basis of our model is the Data Interview which assesses the researcher's data management practices and needs and establishes a baseline awareness of current practice. The second part of the model is the Data Management Plan Support and Review. Information collected from the Data Interviews is compiled, file format and file size results are furnished to the Institutional Repository (IR) team for planning purposes, and information on data management practices and needs are furnished to the subject librarians to enhance their knowledge of the data needs of their departments. In aggregate form, all this information is provided to the VPR and VP/CIO as policy and infrastructure recommendations. The entire process cycles back to the researcher with specific recommendations and solutions that will help improve the research process.

	No formalized document that outlines how re- searchers in the lab manage data.	See the NSF Data Management Pla
	Lack of clear file organization/structure - files are loosely organized	SciDaC recommends reviewing comvention standards, developing and method, and documenting the struct
	Lack of consistent file naming conventions - need Version Control	In order to reduce complications an control of the data, SciDaC will eval manage data version control, and the tion into the research process.
	Ownership and rights responsibilities of data is unclear	In coordination with the University produced a 2-page memo: <i>Data Rig</i> to clarify IP rights around research o
	Lack of Backup Strategy; need better Security & Storage Solutions	Involvement with data management to see new opportunities for scalable vide redundancy, access, and long-te explore appropriate options, reviewes storage solutions and help to integre
	No documentation or adopted metadata stan- dards currently exist	All research processes need to be do project and the datasets. For a proj text of data collection, methods use of data files, sources used, data vali techniques, stages of analysis from policies for confidentiality, access a the variable names and describe the cation schemes used, identify any a data, and identify the file formats a sion numbers.
	No discipline specific repository to deposit & share their data	SciDaC will review appropriate com this into the research process.

University **VP** Research



Virginia Library

Office of **Sponsored Programs** 



**Consulting Group** 

# **Subject** Librarian

an Reveal table (to the right).

nmunity-developed naming conl implementing a file organization Icture of all of the data.

nd inefficiency around version aluate appropriate software to then will help integrate the solu-

Library's Legal Counsel, SciDaC ghts and Responsibility that aims data.

nt at a broader level is allowing us ble storage options that will pro--term preservation. SciDaC will w available central/networked rate this into research processes.

locumented; both for the research ject this should include the consed to collect data, organization lidation and quality assurance raw data to output data, and and use. For the datasets, identify em, define the codes and classifialgorithms used to transform your and software used, including ver-

nmunity options and integrate

### **Consulting Services**

data storage options.

UVA's Institutional Repository (IR) "Libra".

Data Interview



THE

UNIVERSITY

OF

VIRGINIA

LIBRARY

The goal of our data interviews is to develop an understanding of how our science and en-gineering researchers manage their research data and initiate a discussion about how to simplify processes and improve practices. Each interview is scheduled for 60 minutes and includes the Scientific Data Consultants, the Subject Librarian and the researcher. A final interview report is then distributed to all Subject Librarians helping to give them a better understanding of research data processes beyond their own fields.

- Identify common research data practices and needs,
- Identify the types of digital "data" that being created,
- Identify communities and individuals who are under the most pressure from emerging grant regulations,
- Identify potential partnerships for institutional repository data
- Establish baseline awareness of data management,
- Develop opportunities to provide data management recommendations and training, and
- Compile data management recommendations and offer solu-

A Data Management Plan is a critical tool in protecting the continuity of the research process. SciDaC works with researchers to develop appropriate plans for managing data.

SciDaC works with researchers to develop an appropriate plan for managing their data. Once in place, a DMP can continually be updated, provided to new members of the lab as guidelines, and easily be applied to future grant proposals.

SciDaC reviews appropriate community repositories and helps to develop plans and processes for data sharing that will meet or exceed requirements.

Institutional policies assure that certain scholarly content is preserved (VPR). If UVa required deposit of data then there would be a need for increase capacity for storage (CIO/ITC).

NIVERSITY

IRGINIA

BRARY