Support Your Data

A Data Management Guide for Researchers

John Borghi, Stephen Abrams, Daniella Lowenberg, Stephanie Simms, John Chodacki

Introduction

Researchers are faced with an evolving array of expectations related to the management and sharing of data. Drawing upon models from the broader digital curation community as well as surveys, focus groups, and informal conversations with active researchers, we have developed a suite of tools designed to help researchers assess and advance their data management practices throughout the lifecycle of a project.

Elements of the Guide

At present, our tools include a rubric that describes RDM-related activities using researcher-friendly language and a series of customizable guides that provide actionable information about how to advance RDM practices as necessary or desired.

Our tools are designed to be user friendly, free of jargon, and customizable to meet the needs of researchers in different disciplines and with different levels of expertise.

RDM Rubric

	Ad hoc	One-time	Active and informative	Optimized for re-use
Planning your project	When it comes to my data, I have a "way of doing things" but no standard or documented plans.	I create some formal plans about how I will manage my data, but I generally don't refer back to them.	I develop detailed plans about how I will manage my data that I actively revisit and revise over the course of a project.	I design my plans for managing data to streamline future use by myself or others.
Organizing your data	I don't follow a consistent approach for keeping my data organized, so it often takes time to find things.	I have an approach for organizing my data, but I only put it into action after my project is complete.	I have an approach for organizing my data that I implement prospectively, but it not necessarily standardized.	I organize my data to the so that others can navigate, understand, and use it without me being present.
Saving your data	I decide what data is important while I am working on it and typically save it in a single location.	I know what data needs to be saved and I back it up after I'm done working on it to reduce the risk of loss.	I have a system for regularly saving important data while I am working on it. I have multiple backups.	I save my data in a manner and location designed maximize opportunities for re-use by myself and others.
Getting your data ready	I don't have a standardized or well documented process for preparing my data for analysis.	I have thought about how I will need to prepare my data, but I handle each case in a different manner.	My process for preparing data is standardized and well documented.	I prepare my data in such a way as to facilitate use by both myself and others in the future.
Analyzing your data	I often have to redo my analyses or examine their products to determine what procedures or parameters were applied.	After I finish my analysis, I document the specific parameters, procedures, and protocols applied.	I regularly report the specifics of both my analysis workflow and decision making process while I am analyzing my data.	I have ensured that the specifics of my analysis workflow and decison making process can be put into action by others.
Sharing your data	I share the results of my research, but generally I do not share the underlying data.	I share my data only when I'm required to do so or in response to direct requests from other researchers.	I regularly share the data that underlies my results and conclusions in a form that enables use by others.	Because of my excellent data management practices, I am able to share my data efficiently whenever I need to with whomever I need to.

The RDM rubric is designed to emphasize that the management of data is an active and iterative process that occurs throughout the different phases of a project. Activities in different phases are described without jargon and are represented in separate rows. Each row includes a series of declarative statements which describe activities at one of four levels depending on the extent to which they are described, documented, and integrated with those implemented at other phases.

The rubric is intended to be descriptive rather than prescriptive. It is designed to help researchers understand where they are.

Ad hoc - Every time a researcher has to manage their data they have to design new practices and procedures from scratch.

One time - Practices or procedures implemented at one phase of a project are not designed with later phases in mind.

Active and informative - Practices and procedures are standardized, documented, and well integrated with those implemented at other phases.

Optimized for re-use - Data management activities are designed to facilitate the re-use of data in the future

One Page Guides

Each row of the rubric is complemented by a one page guide designed to provide researchers with information about RDM-related requirements and advice about advancing practices as required or desired. We developed the guides so that they can be easily customized by RDM specialists to meet the needs of different disciplinary or institutional communities.

Our aim is to help researchers get where they want or need to be.

Next Steps

Stay tuned as we work on the presentation of the Support Your Data material in both physical media and online. We welcome any and all feedback about the content and application of our tools.

University of California
CDL
California Digital Library

UC3