Integrating Research Data into Digital Preservation Workflows for Libraries & Archives: Mapping the Landscape

Amber Leahey and Grant Hurley, Scholars Portal, Ontario Council of University Libraries

Summary: This poster surveys the suite of research data management services and workflows currently in development and production at Scholars Portal, the information technology service provider for the Ontario Council of University Libraries (OCUL) in Canada. OCUL’s efforts in this domain have been prompted by consortium member requests for infrastructure and tools to assist researchers in effectively managing research assets during their active phases through to deposit and long-term preservation. Using shared cloud-based storage and open-source technologies, OCUL intends to support researchers with a suite of accessible management and preservation services for research data. We will do so by enabling early interventions in the long-term preservation and reuse of research data through small-scale contributions during active phase data creation and management, such as capturing metadata and documentation and ensuring it matches frameworks for access and preservation. The following workflows illustrate the current models in development and testing and the systems and tools in place, alongside related functions and activities and roles and policies.

Workflows

Active phase management in OwnCloud with researcher-driven deposit in Dataverse and preservation using Archivematica with storage in the OLRC.

Active phase management in the Open Science Framework with researcher-driven deposit to a Dataverse and preservation using Archivematica with storage in the OLRC.

Curation phase deposit by researcher to an institutional repository and automated preservation using Archivematica with storage in the OLRC.

Curation phase deposit by researcher to Dataverse and automated preservation using Archivematica with storage in the OLRC.

Optional Workflow Connectors

Researchers may choose to manage or move files between OwnCloud and the Open Science framework depending on needs.

The Open Science Framework allows researchers to upload data to, and pull from, Dataverses.

Data may enter Permafrost directly via legacy media stored in archives and libraries. Access copies can be made available via institutional repositories or Dataverse.

Tools

DMP Assistant
Tool for preparing research data management plans describing metadata, formats, storage, etc., planning as part of grant process.

OwnCloud
Dropbox-like application using the OLRC for storing and managing active-phase data.

Open Science Framework
Third party hosted platform for research data management and collaboration.

Dataverse
Multi-institutional hosted research data repository for data deposit, discovery and reuse.

Private settings enable active-phase file management, file permissions, encryption.

Functions & Activities

Assess data management needs and requirements.

Manage research data creation, including filing versioning, syncing, etc.

Coordinate and communicate data creation/management functions across individuals involved in project.

Collect, document, organize, clean and analyze data for publication and presentation.

Generate basic documentation and related metadata.

Appraise data and select final versions for deposit, including appropriate final formats and required documentation.

Roles & Policies

Funder-driven policies.

Researcher-driven policies, procedures and workflows.

Guidance on appraisal and file formats from RDM liaison librarians at institutions.

Institutional policies for sensitive and private data.

12th International Digital Curation Conference, Edinburgh - 20-23 February, 2017