CANARIE project

Overview and Project background

Welcome to the CANARIE project page! Here you will find information related to the project "Dataverse for the Canadian Research Community: Developing reusable and scalable tools for data deposit, curation, and sharing."

Work has been completed on the 18-month development project (2019-2020) funded by CANARIE's call for Research Data Management software tools. Overall, the project has strengthened Scholars Portal's capacity to support Canadian researchers and demonstrated the value of cross-institutional collaboration.

As part of the CANARIE-funded project, the Scholars Portal Dataverse team developed the platform to better support the data deposit and sharing needs of Canadian researchers. This work focused on data curation tools, authentication mechanisms, system scalability, and large-file support.

Project Deliverables

Data Curation Tool

- The Data Curation Tool (DCT), launched in fall 2019, allows data owners and curators to create and edit metadata at the variable level for files uploaded through the tabular ingest process in Dataverse. Users of the DCT can view summary statistics and charts about their data. The DCT improves data curation workflows within Dataverse, improves the ability for data reuse, and supports the application of standards and best practices using the Data Documentation Initiative (DDI) metadata standard.

Shibboleth Authentication (CAF)

- Scholars Portal configured Dataverse to work with Shibboleth for institutional single sign-on through the Canadian Access Federation (CAF), an identity management service for Canadian research institutions run by CANARIE. This integration ensures secure and trustworthy exchange of identity information as well as provides a simpler log-in process for users with one less username and password to manage.

System Scalability

- Scholars Portal connected Dataverse to in-house cloud storage by hosting files in a test cluster of the Ontario Library Research Cloud (OLRC). This optimizes system architecture for scalable use and leverages an existing, distributed Canadian cloud data storage network.

Large File Support

- Scholars Portal developed proof-of-concept integration with Globus as a large-file transfer tool. Internal tests found that this integration can handle robust transfers up to 100 GB in size and up to 38,000 files. The Scholars Portal Dataverse team is continuing to collaborate and consult with Harvard's IQSS Dataverse team to bring this proof-of-concept development work into the core Dataverse code.
- Work continues into 2021

Project Team Members:

Direction	*Developers*	*Systems Support*
PI: Kate Davis *Technical Lead*: Amaz Taufique *Co-PI*: Amber Leahey *Project Manager*: Meghan Goodchild Kaitlin Newson	Jayanthy Chengan Sunil Manikonda Victoria Lubitch	Bikram Singh Sohaib Anwar Carlos McGregor Dawas Zaidi

For more information about this project, please visit the wrap-up blog post or contact us at dataverse@scholarsportal.info .