

Dataverse Fall 2021 Update Report

Report presented to representatives from Canadian academic library consortia that are institutional members of the Dataverse service.

Scholars Portal Dataverse (<https://dataverse.scholarsportal.info>), launched in 2012, is reaching its tenth year of service and has come a long way over the past decade. Rooted in the spirit of academic library collaboration, research experimentation, and stewardship of research data, the service has grown to include 60 institutional library partners from across Canada.

Fall 2021 Update Report

[Updates on Services](#)

[Staffing update](#)

[Migration to OLRC](#)

[Globus Integration](#)

[Dataverse Upgrades](#)

[Dataverse in Canada: Building a Community of Practice](#)

[University of Alberta Dataverse Collection Migration](#)

[Policy development and CoreTrustSeal Certification](#)

Prepared by: Scholars Portal/University of Toronto Libraries

Date: September 2021

Updates on Services

Staffing update

- The Dataverse team welcomes Jonathan Dorey from INRS/ENAP/TÉLUQ on a secondment at one day per week. He contributes expertise to support internationalization development projects of the Dataverse software, provides French translations of documentation, and will assist with French Canada Dataverse versions that will be contributed to the Global Dataverse Community Consortium [language packs repository](#), used by installations all over the world.
- The Dataverse team welcomes Meghan Goodchild back from her leave. Meghan works jointly with Scholars Portal and Queen's University as RDM Systems Librarian.

Migration to OLRC

- The Dataverse team successfully migrated the file storage for Dataverse to the [Ontario Library Research Cloud \(OLRC\)](#) in August 2021.
- Leveraging the robust cloud storage network, all Dataverse data is now replicated a minimum of 3 times across the OLRC's 5 storage nodes. This architecture will also enable storage to be more easily expanded for the service and will support ongoing digital preservation of research data in libraries.

Globus Integration

- The Dataverse team continues to investigate and develop infrastructure solutions for large file upload and download in Dataverse.
 - **What is it? Who is involved?** In collaboration with the Harvard IQSS team, UChicago/Globus team, and others in the broader Dataverse community, the Dataverse team is building a connection between Dataverse and Globus utilizing the Globus and Dataverse APIs. The team is developing a custom, bilingual, front-end integrated application that can be launched within Dataverse and is connected to users' personal Globus accounts for transfer of large files and volumes of data into Dataverse for long-term preservation and access.
 - **How can I use it? When will it be ready?** Researchers and libraries with data that exceeds the current 3GB file upload limit (using the browser via HTTP) will soon be able to test uploading large files and volumes in Dataverse using the Dataverse Globus Tool. The team anticipates there will be a test instance of the integration available to the community in late fall 2021.
 - **Want to learn more?**
 - A video demonstrating the current development's functionality and a prototype is available [here](#);
 - Additional information about large file support in Dataverse was showcased at the 2021 Global Dataverse Community Meeting (see video about Globus integration [here](#));
 - More information about the integration can be found on the Spotdocs [project space](#).

Dataverse Upgrades

- Planning for the next upgrade to Dataverse is well underway. We are currently reviewing options for the next release, including the upcoming release of Dataverse 5.7, with a focus on the latest features and functionality that are of interest to the community.
 - Some important features of interest to the community include:
 - Make Data Count full integration and options for enabling alongside existing dataset metrics;
 - Enhanced data licensing support and machine-readable licensing;
 - Institutional Metrics;
 - Globus API updates to support large file verification and publishing;
 - Updates will be provided by the team over the fall to inform the community of pending upgrades to Dataverse.

Dataverse in Canada: Building a Community of Practice

- The Dataverse team continues to host a monthly [Dataverse Community Call](#) for service partners to find out about new features and updates from staff.
- The team's hosted listserv, which includes over 100 individual library staff from across Canada, continues to be a source of support for end-users and library staff, to engage with Dataverse staff and one another about Dataverse use at their institution.
- The Dataverse team continues to work closely with partners at Portage/NDRIO to understand and engage with RDM experts and community members from across Canada.
 - Together with the NDRIO-[Portage Dataverse North Working Group](#), [Curation Expert Group](#), [Data Repositories Expert Group](#), and the [Preservation Expert Group](#), valuable Dataverse expertise is contributing to the development and adoption of RDM best practices in Canada.
 - Expert groups' input serves staff and the community at large by providing focused expertise for the development of best practices in Dataverse, as well as promoting knowledge sharing, training, and national collaboration.
 - Through these national initiatives, the Dataverse community is engaging with colleagues from across Canada about service models, data curation, data discovery, approaches for handling data from across different domains and disciplines, digital preservation, repository certification, and supporting broader efforts around outreach and communication in support of the Dataverse service.
- Several webinars are being planned for the fall by community members and coordinated by NDRIO-Portage and the Dataverse team, with the goal to increase adoption, build awareness and understanding, and engage with service partners and researchers about use of Dataverse in Canada.

University of Alberta Dataverse Collection Migration

- Over the summer 2021, the Dataverse team has been working closely with the University of Alberta Library to develop migration scripts and technical processing steps to facilitate the full migration of the [University of Alberta Dataverse](#) Collection to the Scholars Portal Dataverse.

- The Dataverse team anticipates the final production release of the UofA Dataverse Collection, which includes over 200 sub-Dataverse collections, 700 datasets, and over 14,000 files, to be in October 2021.

Policy development and CoreTrustSeal Certification

- Working with NDRIO-Portage's CoreTrustSeal Certification Support Cohort (launched in May 2021), the Dataverse team has led the [SP Dataverse CTS Project Cohort](#) of 11 institutions to receive training, engage in workshops, obtain expert feedback, and to guide the work necessary to achieve CoreTrustSeal requirements for certification of individual institutional Dataverse collections.
- This extensive work involves [gap analyses](#), Dataverse policy and documentation development, including rewriting the Dataverse Terms of Use and User Guides, information technology and security guidance, preservation planning, and curation functionality documentation.
- Several guidance materials will directly influence and support institutional policy development and Dataverse CTS applications moving forward.