Redesigning the FRDR Discovery Interface: Better, Stronger, Geospatial-er

Mark Goodwin - Metadata Coordinator, UBC
Paul Dante - Software Developer, UBC
Eugene Barsky - Research Data Management Librarian, UBC

June 2, 2022

Canadian Research Software Conference
In this presentation:

1. Introduction to the Federated Research Data Repository (FRDR), and the FRDR Discovery Service
2. The FRDR Discovery Redesign Project
What is FRDR?

- Federated Research Data Repository (FRDR) / Dépôt fédéré de données de recherche (DFDR)
- Scalable federated platform for digital research data management and discovery
- A service provided by the Digital Research Alliance of Canada (The Alliance)
- Three components:
  - **Discovery**: National discovery layer indexing Canadian research data repositories
  - **Deposit**: Data repository with dedicated curation support
  - **Preservation**: Archivematica integration and preservation pipeline
The FRDR Discovery Service

Metadata Harvester (Python)

harvest database (postgres)

FRDR Discovery Service

Find Data
Search FRDR to find research datasets originating from researchers affiliated with Canadian institutions. Data deposited to other repositories across Canada can also be found by searching in FRDR. View the growing list of collaborating repositories.

Deposit Data
Any researcher affiliated with a Canadian institution can deposit data into FRDR. The platform can efficiently ingest datasets of any size, and preservation processing is done automatically. Data professionals from the Portage Network and institutions across Canada work with researchers to curate and approve deposited items.

https://www.frdr-dfdr.ca/
Text-based discovery

- **Keyword** search, with support for English and French language analysis
- **Filters** for date range, author, and source repository
Map-based discovery

FRDR map search:
- An adaptation of the open-source CANARIE-funded Geodisy project
- Uses the open-source software GeoBlacklight
- Search results are driven by an interactive map
- Directly tied to FRDR’s research data indexing
- Contains datasets that are geospatial in nature or simply associated with a location
- Available in English and French
Metadata harvesting

Different for each repository type
(OAI-PMH, Dataverse, CKAN, OpenDataSoft, Socrata, etc.)

Steps:

1. Query the repository API to get list of records
2. For each record:
   a. query to get full metadata from repository
   b. “crosswalk”—or map—as many metadata fields as we can to the FRDR schema
   c. write the crosswalked metadata to database
3. Periodically refresh each record
We can rebuild it
We have the technology
We can make it better than it was
Better, stronger, faster, geospatial-er
The Discovery Redesign Project will result in a standalone discovery platform that integrates map-based searching with the existing FRDR discovery platform and enhances overall search capabilities.
Discovery Redesign Project: Goals

- Integrate the existing FRDR Discovery Service platform with FRDR Map Search, ensuring a seamless experience that combines text- and map-based searching for users
- Expand the capabilities of FRDR’s search, through mechanisms including filters and advanced search
- Build a foundation upon which FRDR can better leverage persistent identifiers and controlled vocabularies, including ORCID, ROR, and FAST
- Respond to diverse community needs for discovery of Canadian research data

The project is supported by a core project team and the Discovery Redesign Working Group, ensuring the project is community-driven and incorporates usability best practices.
lunaris will (re-)use a variety of open-source software components

**FRDR Harvester:** Repository metadata harvester
  +

**Globus Search:** Text and geospatial search and filtering
  +

**Geodisy:** Metadata and data retrieval and processing (created with CANARIE funding)
  +

**GDAL:** Geospatial file processing
  +

**GeoServer:** Server for publishing and distributing geospatial data
  +

**GeoBlacklight:** Discovery layer
Planned enhancement: Search by place name
Planned enhancement:
Clustering to preview dataset locations
Additional planned enhancements

- Ability to preview multiple geospatial files on a single dataset record page
- Links to related publications in other repositories
- Metadata export (eg. ISO 19139, DDI, DC)
- Stay tuned for more!
Discovery Redesign Project - next steps

- Continued development of alpha prototype
- User testing with alpha prototype
- Continued working group engagement
- Promotion
- Release of beta version: early 2023
Contributors

**Geodisy Team**
Eugene Barsky
Paul Dante
Mark Goodwin

**Discovery Redesign Project Team**
Paul Dante
Mark Goodwin
Adam McKenzie
Neha Milan
Todd Trann
Lee Wilson

**Discovery Redesign Project Working Group**
Paul Dante
Krista Godfrey
Mark Goodwin (Chair)
David Kemper
Amber Leahey
Winnie Li
Kathleen Matthews
Neha Milan
Nicholas Rochlin
Kristi Thompson
Lee Wilson
Thanks! Questions?

Paul Dante, Software Developer, UBC
paul.dante@ubc.ca

Mark Goodwin, Metadata Coordinator, UBC
mark.goodwin@ubc.ca