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

Mark Goodwin - Metadata Coordinator, UBC

June 15, 2022

Research Data Management Summer Lunch and Learn Series

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

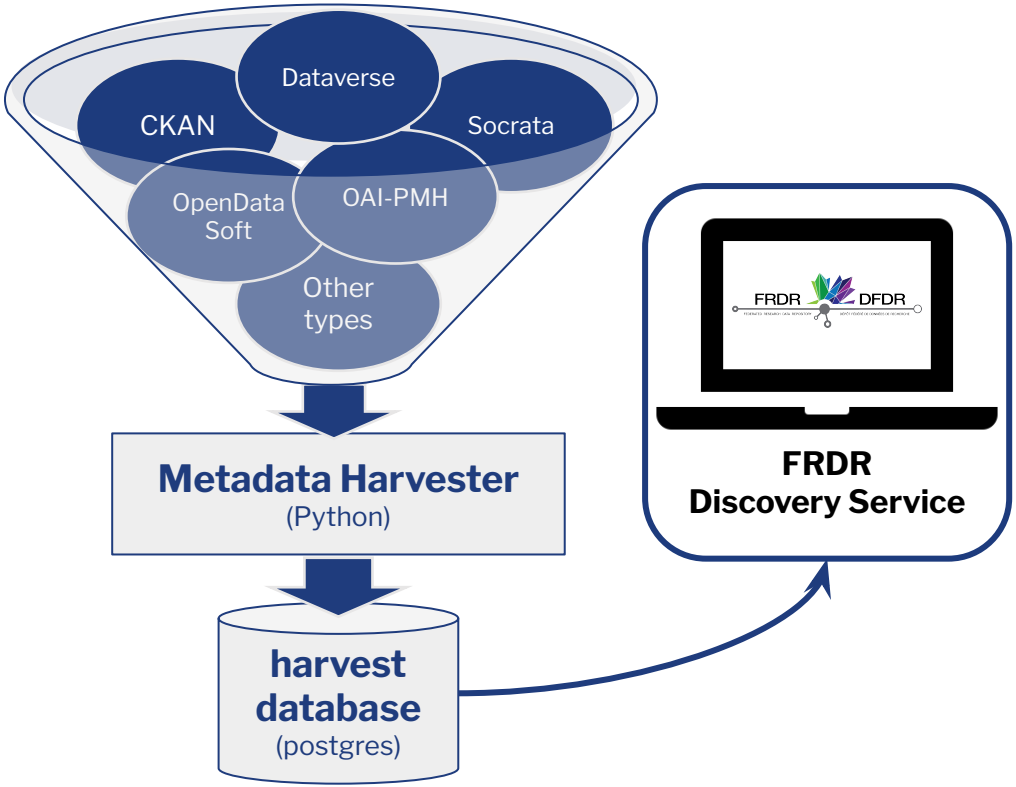
In this presentation:

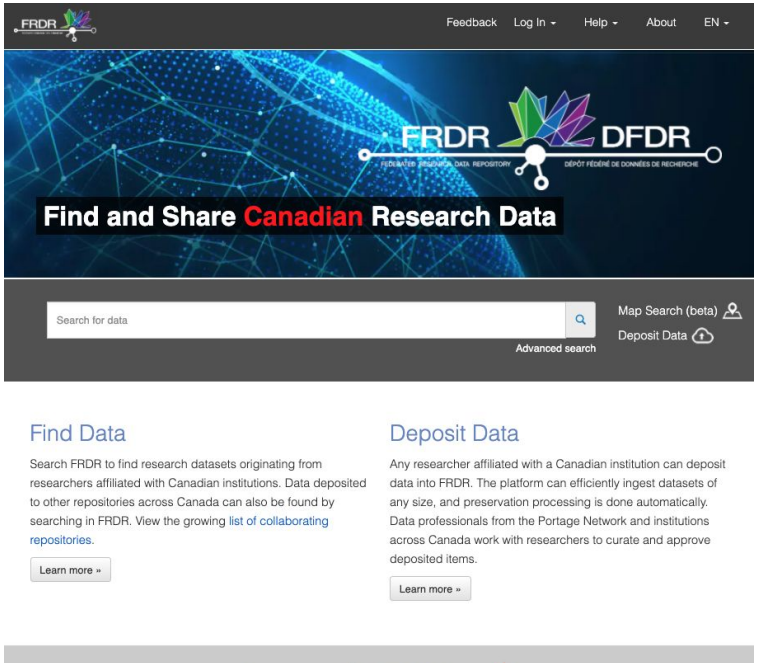
1. Introduction to the Federated Research Data Repository (FRDR), and the FRDR Discovery Service
2. The FRDR Discovery Redesign Project

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

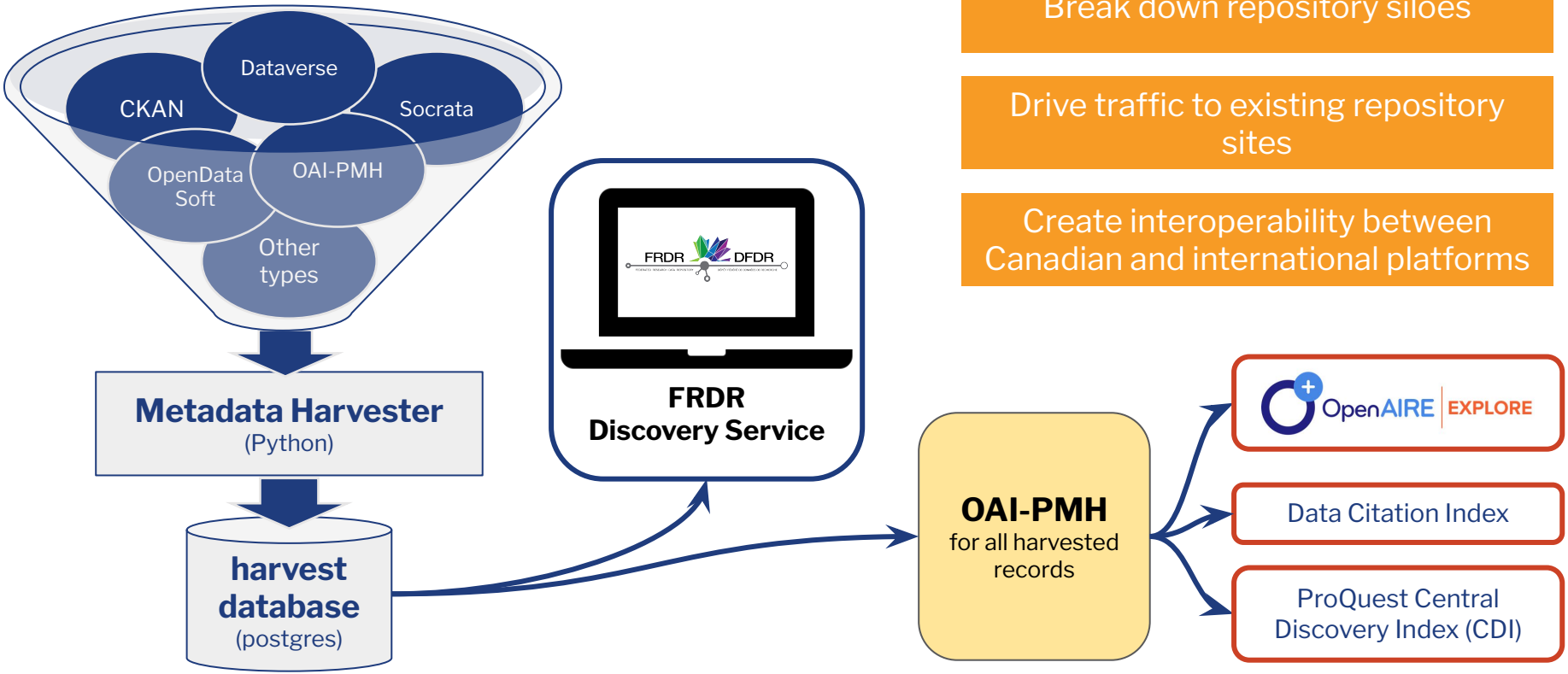




https://www.frdr-dfdr.ca/

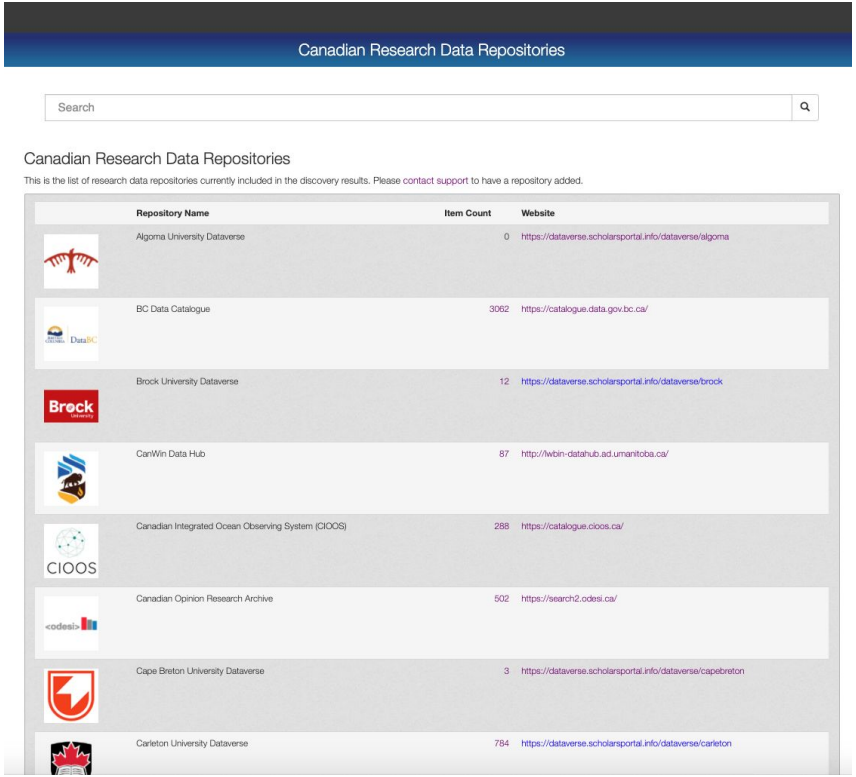
What are the benefits?

- Improve discovery of Canadian research (meta)data
- Break down repository siloes
- Drive traffic to existing repository sites
- Create interoperability between Canadian and international platforms



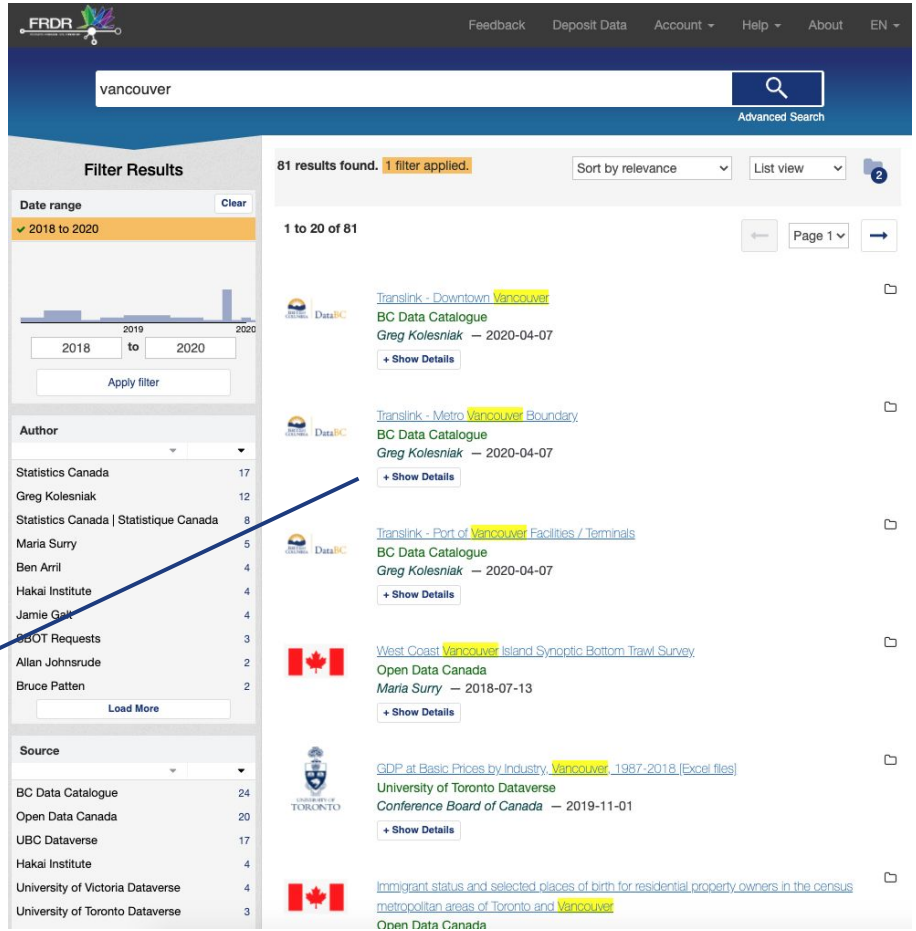
Collaborating repositories

- Over 100 discoverable research data repositories including:
  - **University** repositories, including Scholars Portal Dataverses
  - **Government** repositories at the federal, provincial, and local levels
  - **Domain-specific** repositories
  - Datasets deposited in **FRDR**



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 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

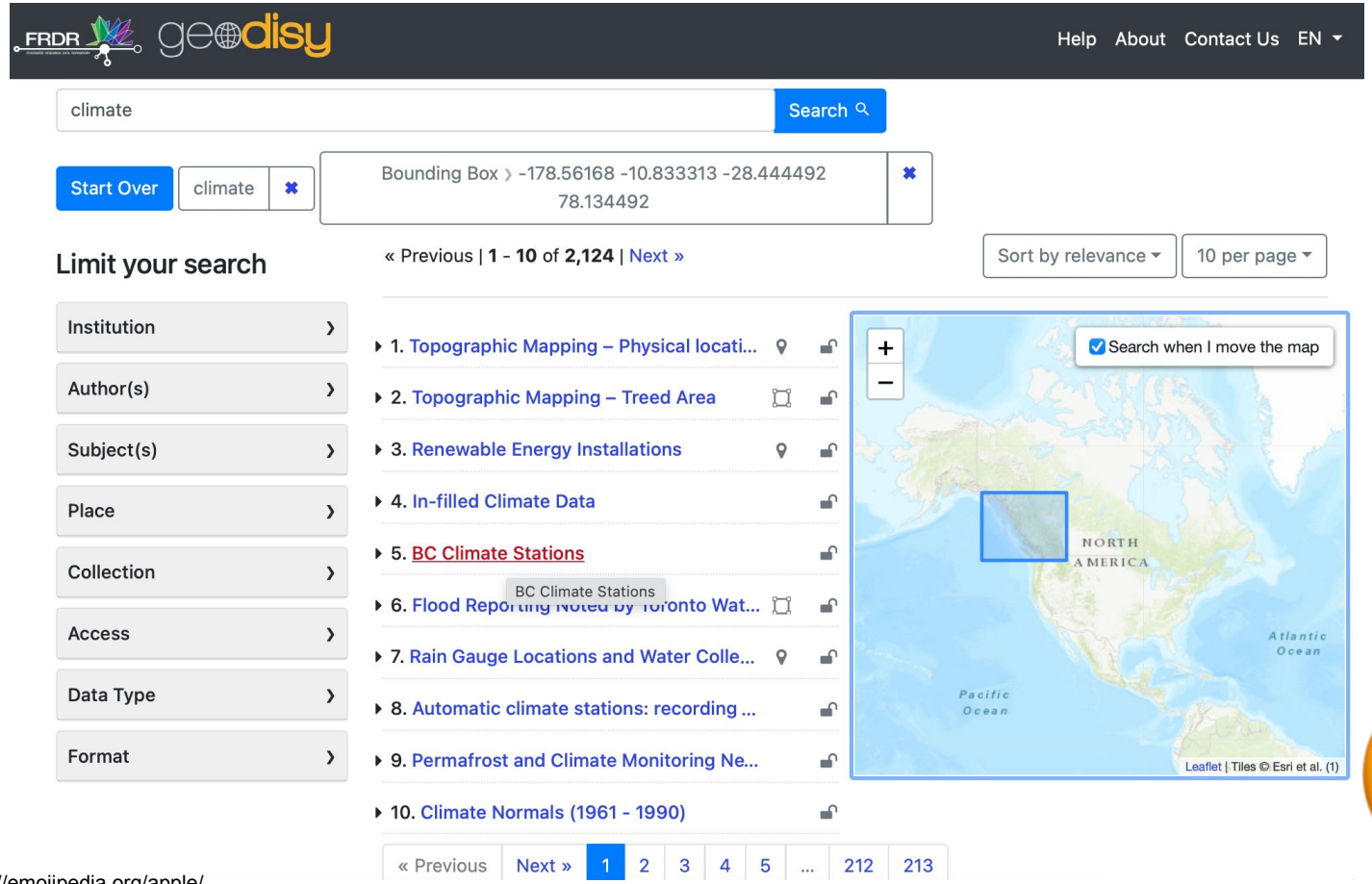
Why is geospatial search useful?

- Data can be difficult to find! When searching for data about a particular place, keywords can be hit or miss. A text search might look something like this:

((British Columbia OR BC OR B.C.) N2 (north*)) OR (Alaska N2 south*) OR (Yukon N2 south*)
OR (Tatshenshini-Alsek ADJ2 park) OR (Glacier Bay ADJ2 (park or preserve))
OR (Kluane ADJ2 (park OR reserve)) OR (Atlin ADJ2 (park OR recreation area) OR …

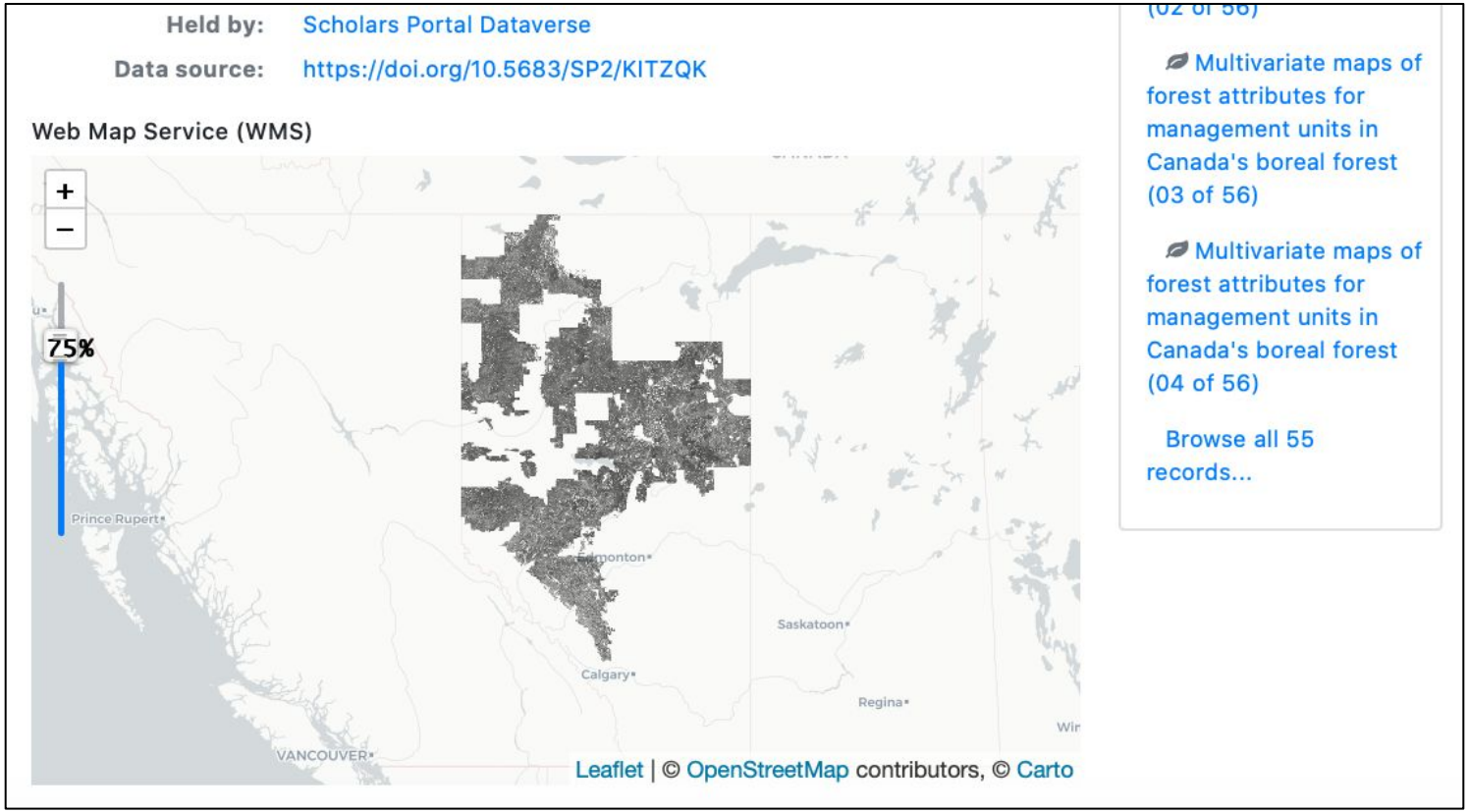
Emoji image from https://emojipedia.org/apple/

FRDR Map Search: search results page

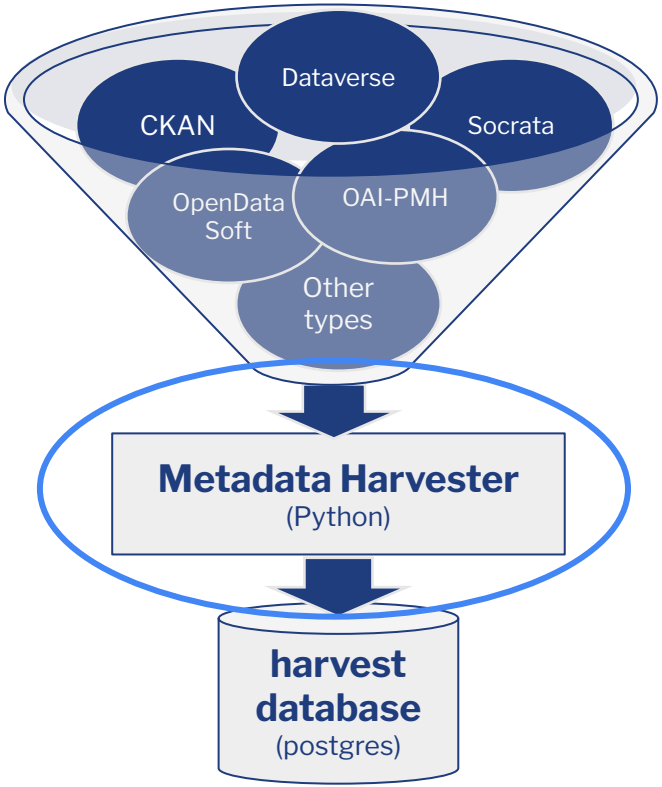


Emoji image from https://emojipedia.org/apple/

FRDR Map Search: record page - geospatial preview



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

The FRDR metadata profile

- Primarily **Dublin Core** elements
  - dc.contributor.author
  - dc.title
  - etc.
- Some elements from the **DataCite** schema
  - creatorAffiliation (“Author Affiliation”)
  - resourceTypeGeneral (“Type”)
- Some **custom** elements
  - access
  - item_url

Fields
Title (EN)
Title (FR)
Date
Author
Author Affiliation
Contributor
Description (EN)
Description (FR)
Geolocation (place)
Geolocation (box)
Geolocation (point)
Item URL
Keyword (EN)
Keyword (FR)
Subject (EN)
Subject (FR)
Access
Rights
Series
Publisher
Type

Metadata crosswalking

- Crosswalking involves mapping from one set or “schema” of metadata elements to another
- This is how we transform the source repository metadata to fit FRDR’s metadata profile
- Multiple metadata schemas are reconciled to a set of general elements

  → FRDR users can search multiple repositories with one metadata profile
- Not all of the source metadata is stored in FRDR records

  → FRDR users can view complete metadata on the source repository’s landing page

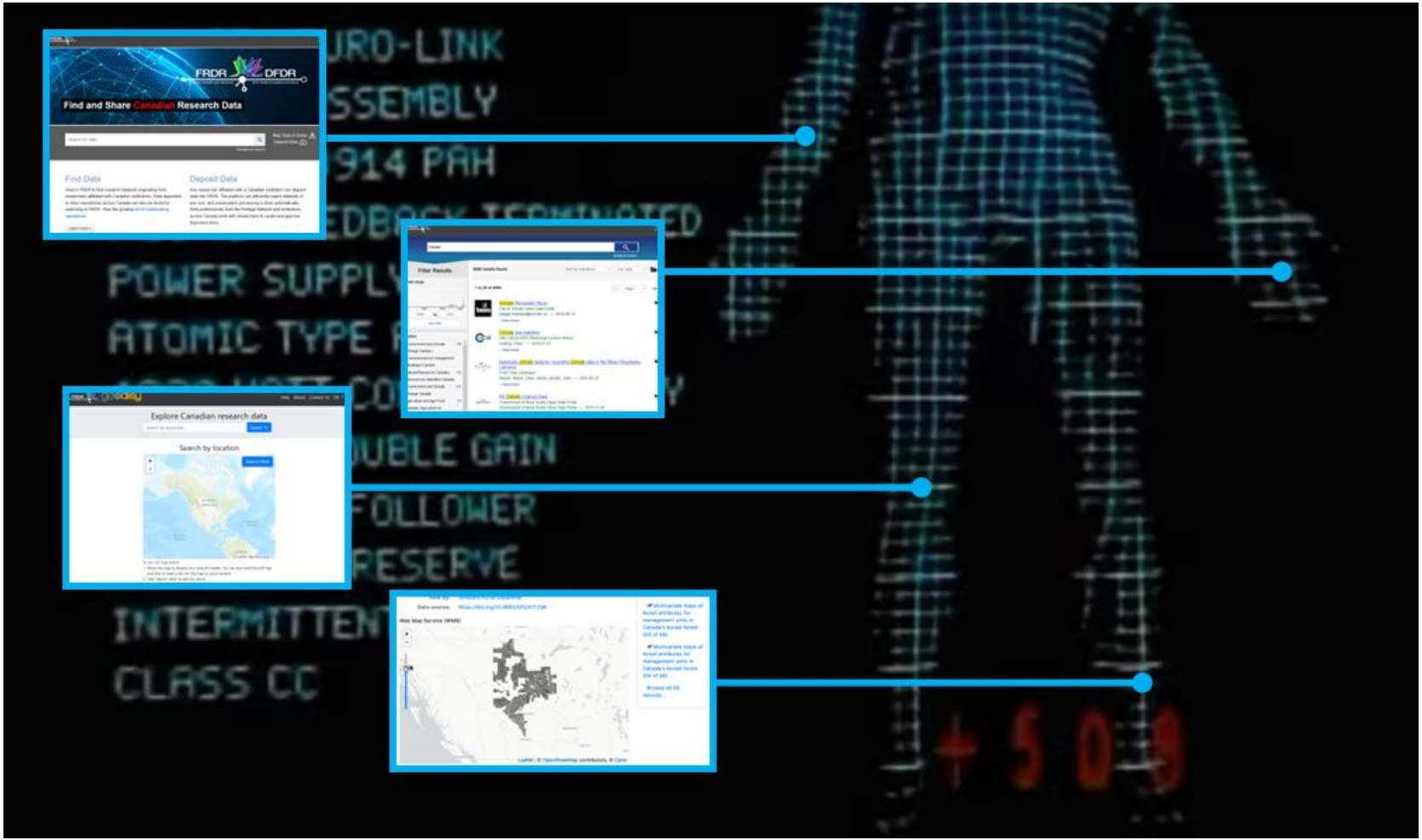
We can rebuild it

We have the technology

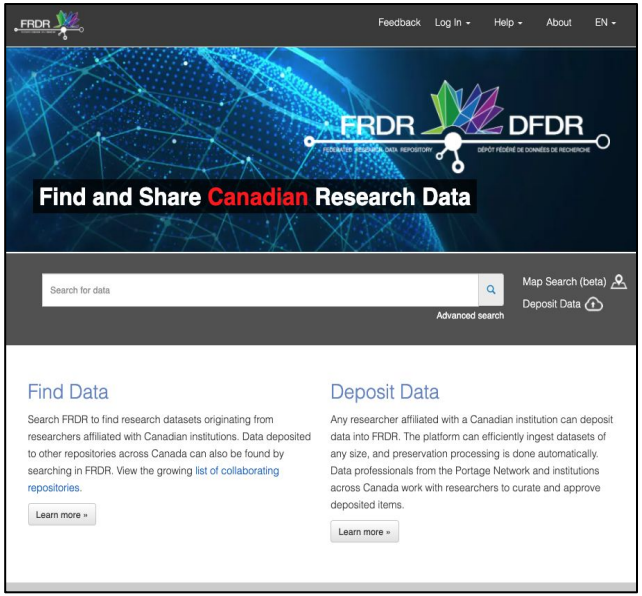
We can make it better than it was

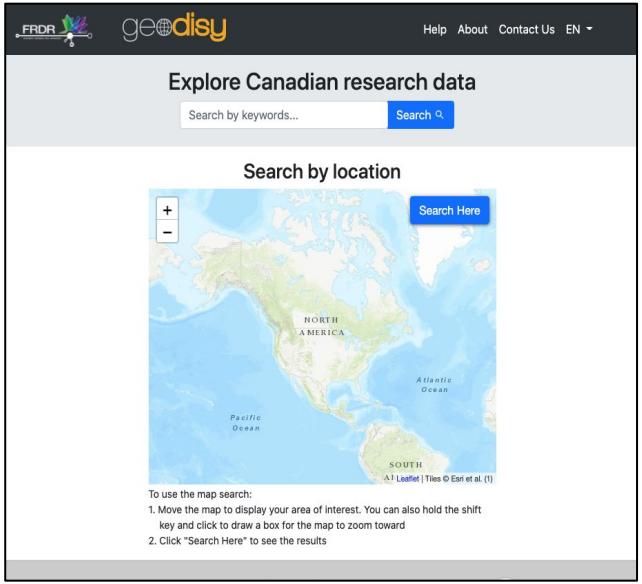
Better, stronger, ~~faster~~ geospatial-er

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

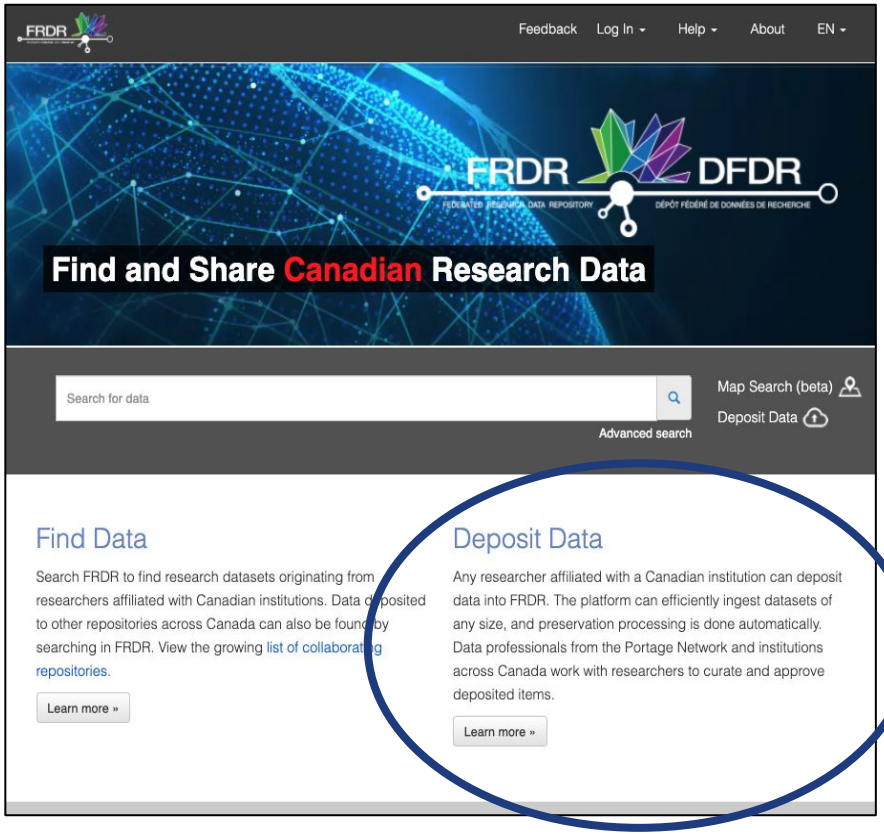


Screenshot image from https://www.youtube.com/watch?v=cRgqouS1O6E


FRDR
Feedback
Log In
Help
About
EN
FRDR DFDR
Find and Share Canadian Research Data
Search for data
Advanced search
Map Search (beta)
Deposit Data
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.
Learn more »
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.
Learn more »


FRDR
geodisy
Help
About
Contact Us
EN
Explore Canadian research data
Search by keywords...
Search
Search by location
Search Here
NORTH AMERICA
Atlantic Ocean
Pacific Ocean
SOUTH
Leaflet | Tiles © Esri et al. (1)
To use the map search:
1. Move the map to display your area of interest. You can also hold the shift key and click to draw a box for the map to zoom toward
2. Click "Search Here" to see the results

lunaris


FRDR
Feedback
Log In
Help
About
EN
FRDR
DFDR
Find and Share Canadian Research Data
Search for data
Advanced search
Map Search (beta)
Deposit Data
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.
Learn more »
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.
Learn more »

FRDR
Federated Research
Data Repository

DFDR
Dépôt fédéré de
données de recherche

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.

Discovery Redesign Project: Process

Completed:

- Environmental scan and comparative analysis
- Established working group
- User research - visioning exercise and user survey
- Confirmed high-level website structure (discovery will be separate from repository)
- Creation of semi-functional figma mockup
- Initial usability testing with figma mockup
- Worked with design firms to establish logo, branding, and wireframe design

Discovery Redesign Project: Process

Completed (continued):

- Design modules with working group
  - Module 1: Record pages
  - Module 2: Search results page
  - Module 3: Home search, browse page, & repo list
  - Module 4: Keyword and map search functionality/algorithm
  - Module 5: Filters and metadata display
  - Module 6: Linking and enriching research data
  - Module 7: API functionality and analytics

In progress:

- Development of alpha prototype

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

+

GDAL: Geospatial file processing

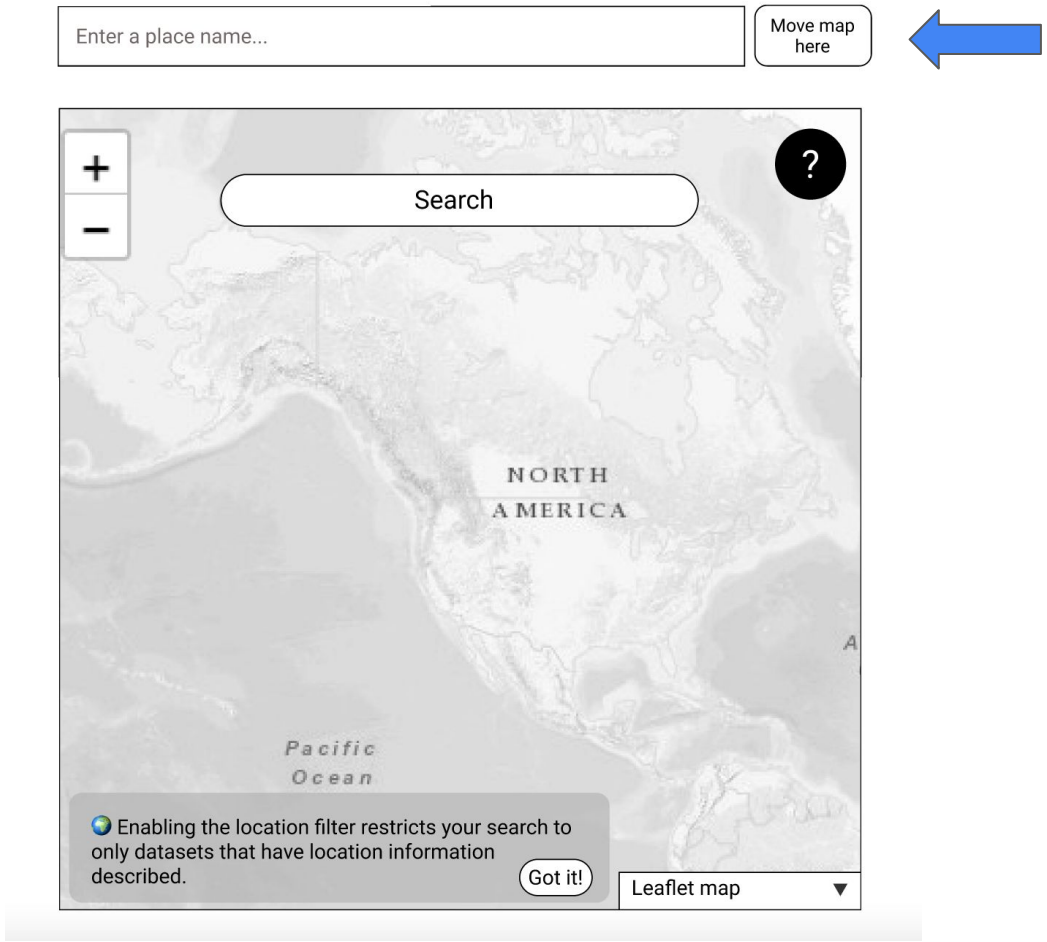
+

GeoServer: Server for publishing and distributing geospatial data

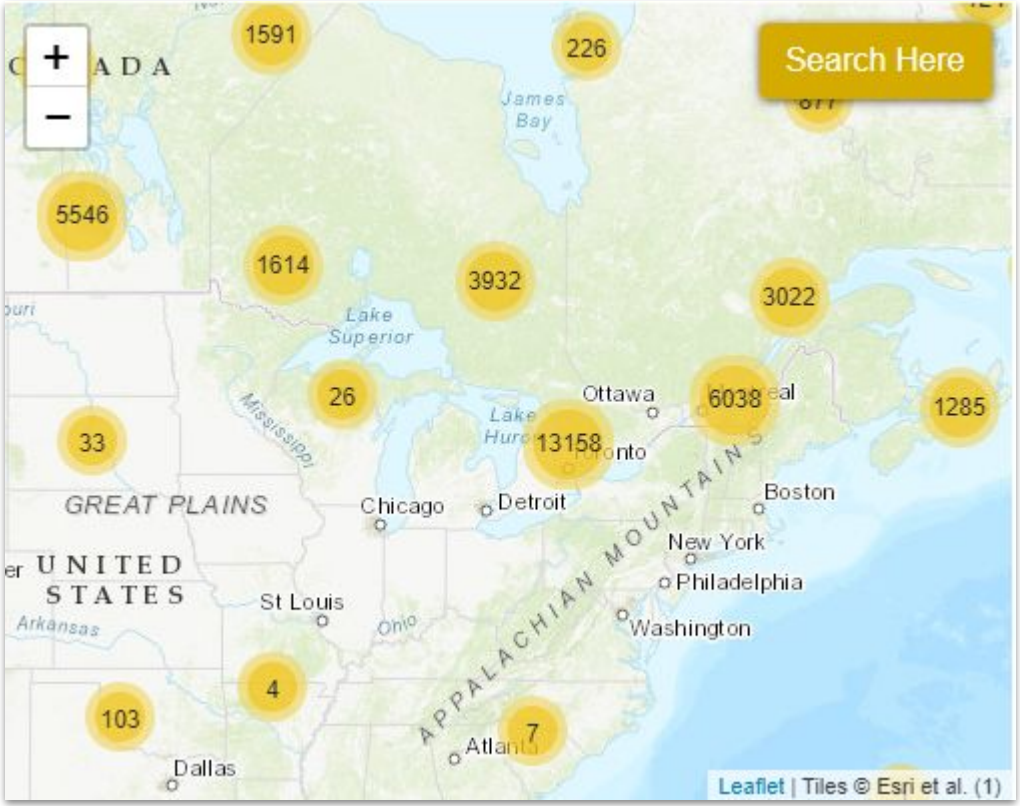
+

GeoBlacklight: Discovery layer

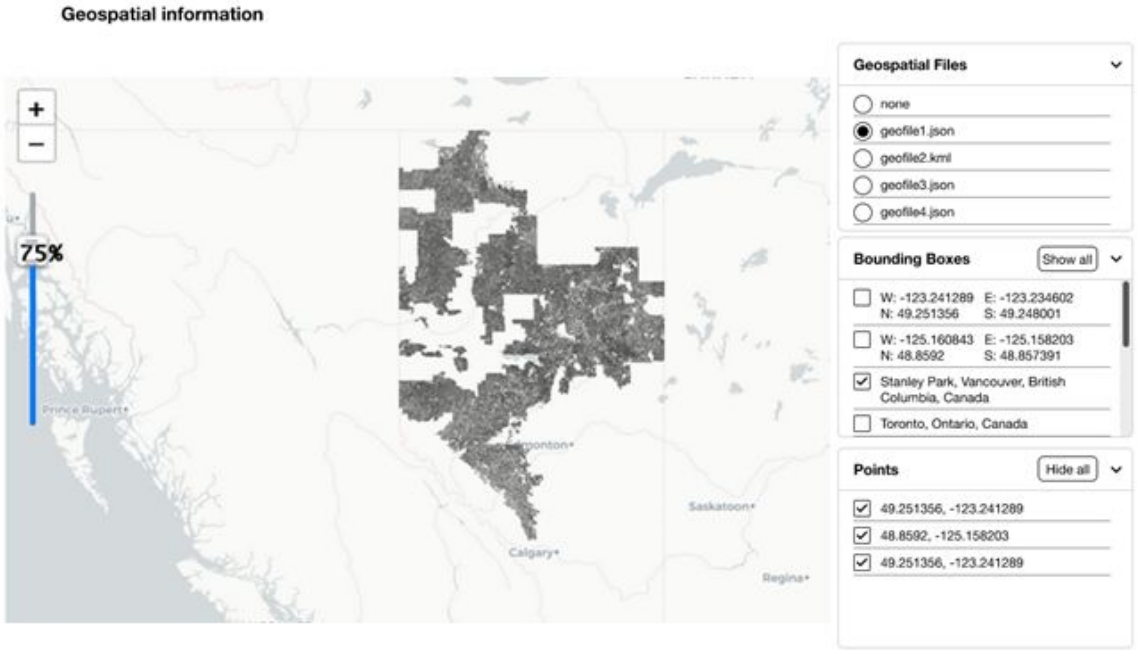
Planned enhancements:
Search by place name



Planned enhancement:
Clustering to preview dataset locations



Planned enhancements:
Ability to preview multiple geospatial files on a single dataset record page



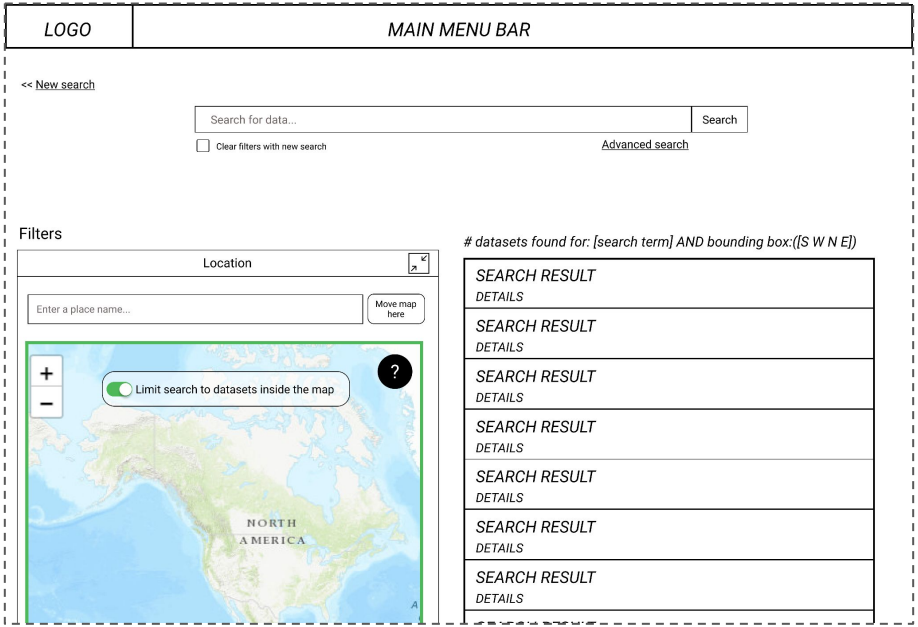
Planned enhancements

- Links to related publications in other repositories
- Metadata export (eg. ISO 19139, DDI, DC)
- Improved advanced search field options
- Stay tuned for more!

Discovery Redesign Project: Challenges

Challenges

- User research - existing users and potential users
- Testing prototype without functional map search (needed to do this to confirm overall flow before further development)
- Incorporating geospatial search into a database in which not all records are geospatial
    - There are many models for dedicated geospatial search, but few for hybrid. Research Data Australia offers a hybrid and we have collaborated with their team.



Discovery Redesign Project - next steps

- Continued development of alpha prototype
- User testing with alpha prototype
- Continued working group engagement
- Promotion
- Launch 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

Special thanks to Kelly Stathis

Thanks! Questions?

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

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada