

Federated Geospatial Data Discovery for Canada - Geodisy

Eugene Barsky and Evan Thornberry, UBC
January 2019



Image - <https://www.flickr.com/photos/double-m2>

The good people of data...

- Amber Leahey, Data Services Metadata Librarian, Scholars Portal
- Marcel Fortin, Head, Map and Data Library, University of Toronto
- Jason Brodeur, Associate Director, Digital Scholarship Services, McMaster University Library
- Jason Hlady, Manager, Research Computing, University of Saskatchewan
- Eugene Barsky, Research Data Librarian, University of British Columbia
- Paul Lesack - GIS/Data Analyst, University of British Columbia Library
- Evan Thornberry, GIS Librarian, University of British Columbia Library
- Mark Goodwin, Metadata Coordinator, University of British Columbia Library
- Paul Dante, Software Developer, University of British Columbia Library
- Lee Wilson, Service Manager, Portage

Outline

- General overview of the project
- Including:
 - The problem
 - Our suggested solution
 - Steps
 - Timelines
 - Your feedback



Image - <https://www.flickr.com/photos/34547181@N00/>

The Problem

- How do I find, for example, the migration paths of humpback whales, the distribution of maple-syrup yields, infrared satellite imagery, distribution of artifacts in an archaeological site or the flow routes of water due to sea level rise?
- Text-based searches don't always work well with spatial data
- **Location** is a key



Image - <https://www.flickr.com/photos/blprnt/>

The Problem

- Most repositories lack a map-based interface
- How do I find data about mining in Northern BC?



The screenshot shows a search interface for Canadian research data. The header features the text "Find and Share **Canadian** Research Data" in white and red on a dark blue background with a network pattern. Below the header is a search bar with a grey background. The search query "mining AND North* AND ('British Columbia' OR BC)" is entered in the search bar and is highlighted with a red rectangular box. A large green arrow points to the left from the search bar. To the right of the search bar is a magnifying glass icon and the text "Deposit Data" with a cloud icon. Below the search bar, the text "Advanced search" is visible.

Find and Share **Canadian** Research Data

mining AND North* AND ('British Columbia' OR BC)

Deposit Data 

Advanced search

Potential Solution

Extend existing software to find and display research data in a search interface which is both map and text based, combining research data with the functionality of a product such as Google Maps.



Image by <https://www.flickr.com/photos/40032755@N06/>

Steps:

Step 1:

Software will query the Canadian Dataverse repositories (Scholar Portal, UBC, UofA, Dal, UNB, UofM) to determine if geospatial information is present within the digital object (e.g., a study, or a data deposit)



Image - <https://www.flickr.com/photos/idhancock/>

Steps:

Step 2:

- Software will harvest any geospatial metadata in the primary record (eg: main record page)
- More importantly, the software harvester will query and harvest any geospatially relevant file objects (satellite imagery, geospatial vector files, etc)



Image - <https://www.flickr.com/photos/a-g>

Steps:

Step 3:

Once the data have been harvested, the software will create and normalize relevant geospatial data from the (A) primary record and from (B) any associated digital objects, and extract all relevant metadata



Image -

<https://www.flickr.com/photos/wakingtiger>

Steps:

Step 4:

The extracted, cleaned and normalized (ISO 19115) data is deposited by the software pipeline into a geospatial data server, such as Geoserver, capable of distributing geospatial data in a wide variety of formats to various services

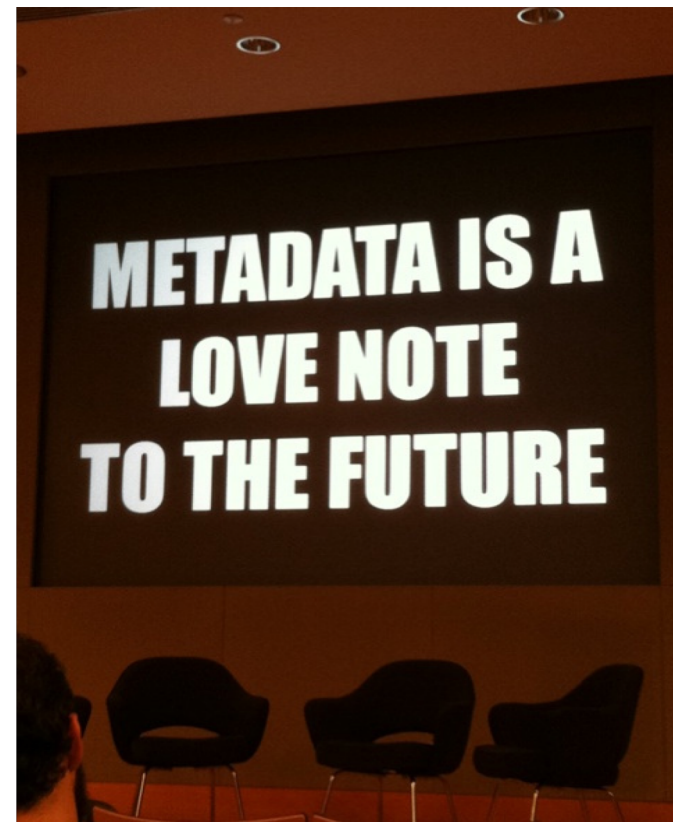


Image - <https://www.flickr.com/photos/centralasian/>

Steps:

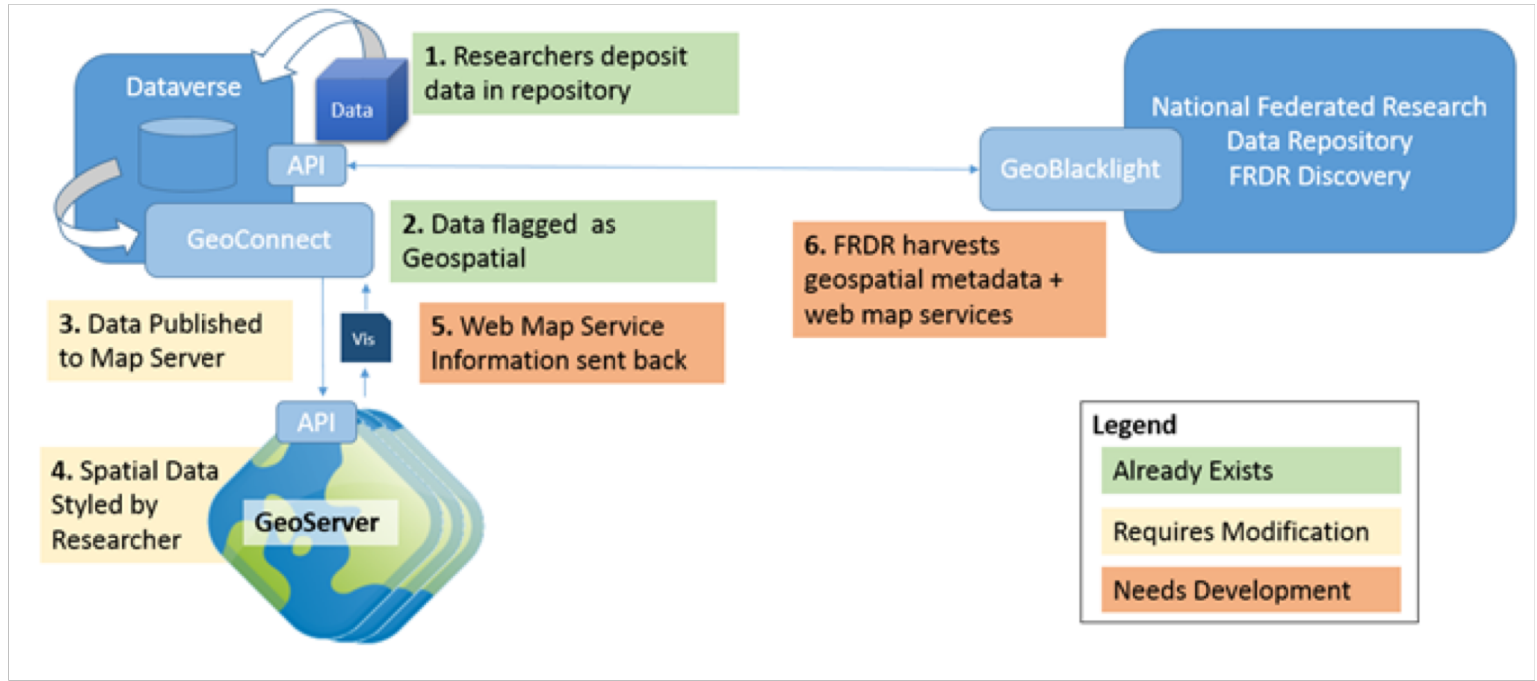
Step 5:

- Data will then be harvested by a geospatial search interface such as **GeoBlacklight**, an open source geospatial search tool
- The user interface will be customized to the needs of the Federated Research Data Repository project (**FRDR**), providing a unified map-based search interface for research data in Canada.



* Image - <https://www.flickr.com/photos/pamilne/>

Suggested Solution



Potential Partners and Synergies

- Big Ten Academic Alliance - Geospatial Data Discovery Project - <https://geo.btaa.org/>
- Canadian Historical Geographic Information Systems Partnership (CHGIS)- <http://geohist.ca/>
- And hopefully some of you!



Image - <https://www.flickr.com/photos/derpunk/>

FAIR Principles

- Enhance **Findability** and **Accessibility** of geospatial research data in Canada - metadata clean-up and crosswalks, guidance, and tools for the description of research data (e.g. Dublin Core, DDI, ISO 19115)
- **Interoperability** - open geospatial metadata exchange and open APIs
- **Re-usability** - open, fit-for-purpose interface for discovering and exploring geospatial data



* Image - <https://www.flickr.com/photos/wwworks/>

National services

- **Dataverse** - We are working with our Dataverse North partners, a potential national repository solution
- **FRDR** - We are working with our Portage and Compute Canada colleagues - incorporating an open source GeoBlacklight application into FRDR as another search option, a mapped based search



* Image - <https://www.flickr.com/photos/crysb/>

Questions?

