

Innovative Solutions for Data
Management Challenges: An
Introduction to FRDR and Lunarix // Des
solutions novatrices aux enjeux de
gestion des données : Une introduction
au DFDR et à Lunarix

14 April 2023 // 14 avril 2023



Digital Research
Alliance of Canada

Alliance de recherche
numérique du Canada





Land Acknowledgement // Reconnaissance du territoire





Code of Conduct // Code de conduite

- The Alliance seeks to provide a welcoming, engaging, and safe community for everyone.
- Discriminatory language is not appropriate in any venue or platform.
- The Alliance does not tolerate harassment in any form. Harassment is understood as any behaviour that threatens another person or group or produces an unsafe environment.

Adapted from Code4Lib Creative Commons Conference Code

- L'Alliance s'engage à offrir à tous les participants et toutes les participantes une expérience accueillante, stimulante et sécuritaire.
- Le langage et les images discriminatoires ne sont aucunement tolérés pendant les événements ou sur les plateformes.
- L'Alliance n'accepte par ailleurs aucune forme de harcèlement, qui est défini comme étant tout comportement qui menace une autre personne ou un groupe, ou qui produit un environnement non sécuritaire.

Adapté du code Code4Lib Creative Commons Conference





Lunaris

Discover Canadian Research Data

Mark Goodwin - Product Lead, Discovery Services



lunaris

Discover **Canadian Research Data**

Découvrir les **données de recherche canadiennes**



Agenda

1. Introduction to Lunaris
2. Software overview
3. Project background
4. Platform demonstration
5. Next steps - full production launch

Lunaris is a national research data discovery service provided by the Digital Research Alliance of Canada as one of its core Research Data Management services

Introduction to Lunaris

- The beta version of Lunaris (www.lunaris.ca) launched on March 29, 2023 - feedback welcome at support@lunaris.ca
- Formerly part of the Federated Research Data Repository (FRDR) as the FRDR Discovery Service - FRDR and Lunaris are now separate platforms/services
- Scalable and bilingual platform with integrated text- and map-based search functionality
- Information about datasets in the form of metadata records are harvested from repositories and made available for discovery
- Over 80,000 datasets from over 100 Canadian repositories and data collections are indexed by Lunaris
 - University repositories, including Borealis Dataverses
 - Government repositories at the federal, provincial, and local levels
 - Domain-specific repositories
 - National repository initiatives, including FRDR

Introduction to Lunaris

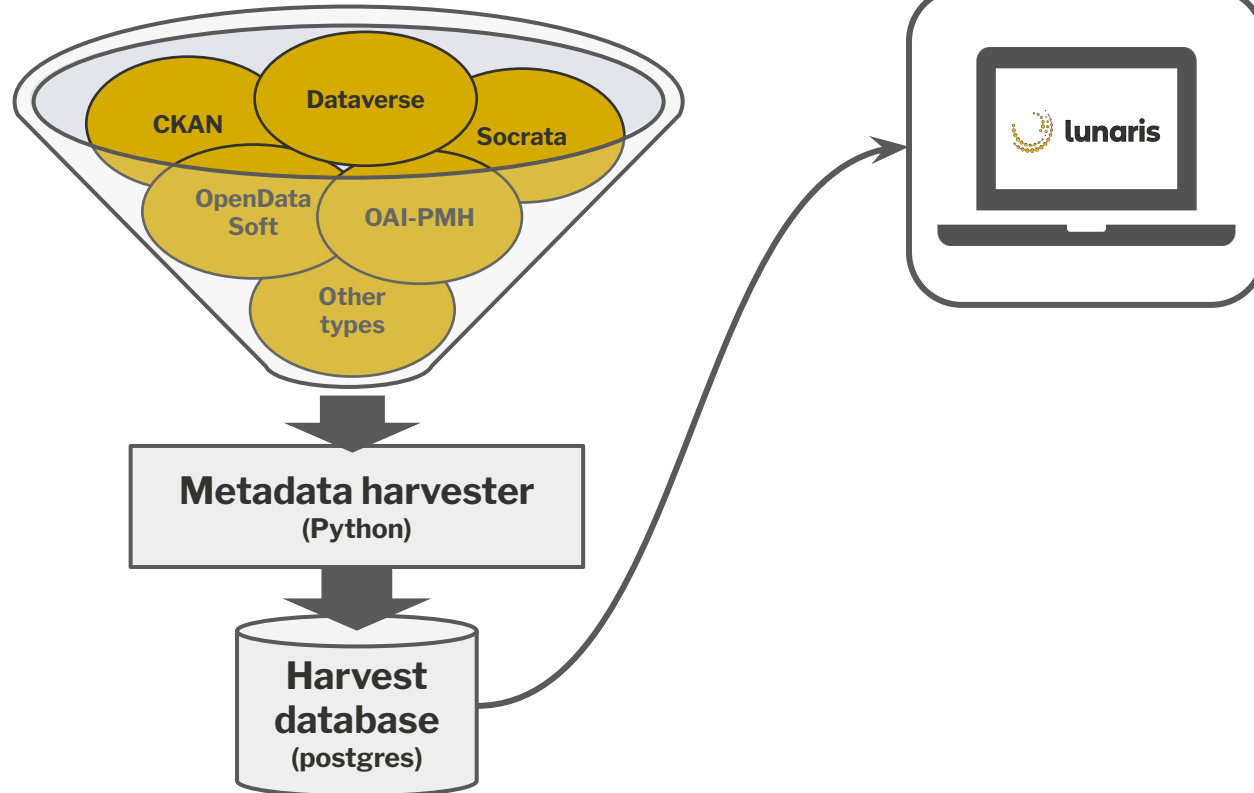
Lunaris features

- Text search powered by Globus Search
- Map search powered by UBC-based initiative Geodisy
- Filtering and advanced search options
- Dataset citation generator
- Geospatial data preview (coming soon)

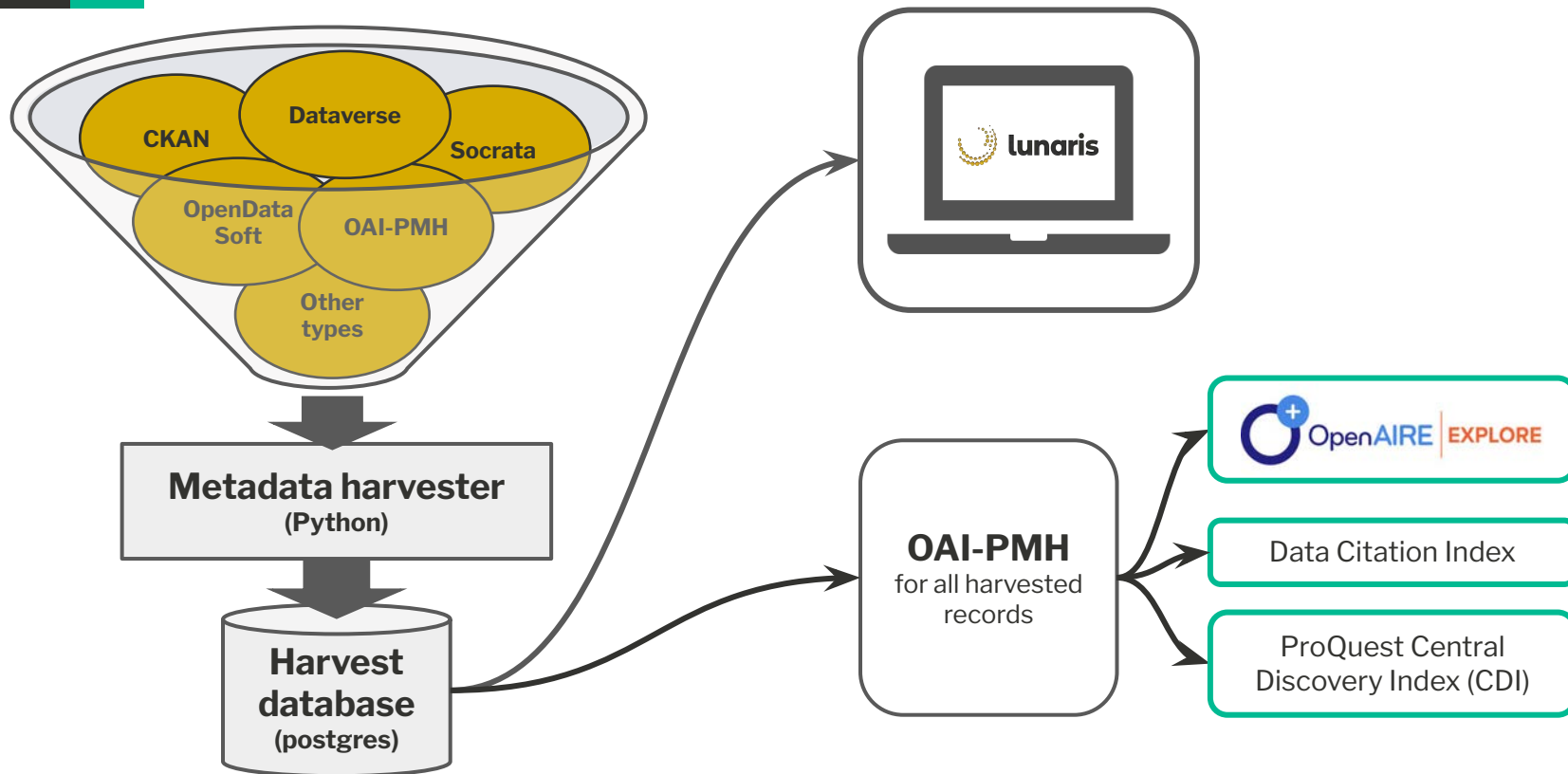
Benefits of Lunaris

- Specialized tool and service that improves discovery of Canadian research data
- Breaks down repository silos
- Drives traffic to existing repository sites
- Creates interoperability between Canadian and international platforms

Software overview



Software overview



Software overview

Lunaris (re)uses a variety of software components:

- **FRDR Harvester**: repository metadata harvester
- **Globus Search**: text and geospatial search and filtering
- **Geodisy**: map search
- **GDAL**: geospatial file processing
- **GeoServer**: server for publishing and distributing geospatial data
- **GeoBlacklight**: discovery layer

Lunaris project background

The **FRDR Discovery Redesign Project** began in 2021

Project goals:

- Integrate the 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
- Improve platform functionality and user experience to meet diverse needs of the Canadian research community

Work is supported by an internal project team and the **FRDR Discovery Redesign Working Group**, ensuring a community-driven process that incorporates usability best practices

Lunaris project background

Project activities

- Environmental scan and comparative analysis
- User research - visioning exercise and user survey
- Development of multiple mock-ups, wireframes, and prototypes
- 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
- Worked with design firms to establish name, logo, and wireframe design
- Two rounds of usability testing
- Support pages (About, User Guide, etc) content refresh
- French language translation and review
- Development and testing of Lunaris beta version

Demonstration

www.lunaris.ca

Next steps



Full production launch in May. Planned updates:

- Geospatial dataset preview on record page
- French language independent audit
- Updates based on community feedback

Contributors



FRDR Discovery Redesign Working Group

- Krista Godfrey, University of Waterloo
- David Kemper, McMaster University
- Amber Leahey, Scholars Portal
- Winnie Li, University Canada West
- Kathleen Matthews, University of Victoria
- Nicholas Rochlin, University of British Columbia
- Kristi Thompson, University of Western Ontario

Partners

- Globus
- University of Saskatchewan
- Simon Fraser University

Internal Development and Operations Team

- Mark Goodwin, Product Lead, Discovery Services
- Neha Milan, Product Lead, Federated Research Data Repository
- Todd Trann, Technical Lead
- Mike Winter, Senior Developer
- Adam McKenzie, Research Data Management Analyst
- Joel Farthing, Developer
- Victoria Smith, Policy, Privacy, and Sensitive Data Coordinator
- Lee Wilson, Director of Research Data Management

Special thanks to past team members Kelly Stathis and Paul Dante

Providing feedback



During the beta phase of Lunarix, feedback is welcome! Please contact us at support@lunarix.ca.

Thank you!

Federated Research Data Repository

Publish and Share Canadian Research Data

Erin Clary, Curation Coordinator, Digital Research Alliance of Canada

Neha Milan, FRDR Product Lead, University of Saskatchewan

FRDR
Federated **Research**
Data Repository



DFDR
Dépôt fédéré de
données de recherche



Agenda

1. What is FRDR?
2. Why use FRDR?
3. Who can use FRDR?
4. Deposit process / Data publication workflow
5. Demonstration
6. Upcoming features

What is FRDR?

- Bilingual publishing platform for sharing and preserving Canadian research data
- Developed as a collaboration between Compute Canada Federation, Canadian Association of Research Libraries, and a number of Canadian post-secondary institutions
- Service fully launched in Feb 2021
- Currently a national-level service with funding administered through the [Digital Research Alliance of Canada](#)

Why use FRDR?

- Publish research data openly in a **Canadian, bilingual** national repository option
- Support for **medium-to-large sized datasets** via Globus file transfer
- Users have access to **curated repository storage** without cost up to a certain threshold, with mechanisms in place to obtain additional storage as needed
- FRDR assigns **DOIs** to facilitate data citation
- Secure repository storage, **distributed geographically** across multiple National Advanced Research Computing Host sites
- You can work with **multiple collaborators** on a single submission
- Allows for **optional embargoes** on data and metadata records

Who can use FRDR?

- Anyone can search for and download data
 - Data may be temporarily embargoed
 - Researcher-assigned licenses specify how each dataset can be used
- Who can deposit data?
 - Faculty or their designates (data managers, students, post docs, etc.) from institutions eligible to receive Tri-Agency funding
 - Default allocation of 1TB of repository storage at no cost
 - Additional storage may be provisioned

Data Publishing Workflow



Triage



Data Deposit Workflow



Curation Workflow



Indexing for Discovery



Preservation Workflow

- ✓ Researcher consultation (optional)
- ✓ Data Appraisal

- ❑ Size of data
- ❑ File types/formats
- ❑ Special considerations around licensing, sensitivities, or copyright
- ❑ Minimum publishable unit
- ❑ Documentation
- ❑ Current storage solution
- ❑ Data management support, etc.

Data Publishing Workflow



Triage

- ✓ Researcher consultation (optional)
- ✓ Data Appraisal



Data Deposit Workflow

- ✓ Choose Collection
- ✓ Required and recommended metadata
- ✓ Upload Data Files



Curation Workflow



Indexing for Discovery



Preservation Workflow

- ❑ Size of data
- ❑ File types/formats
- ❑ Special considerations around licensing, sensitivities, or copyright
- ❑ Minimum publishable unit
- ❑ Documentation
- ❑ Current storage solution
- ❑ Data management support, etc.

Small datasets via web browser

Globus File Transfer
>5 GB or 300 files



Data Publishing Workflow



Triage

- ✓ Researcher consultation (optional)
- ✓ Data Appraisal



Data Deposit Workflow

- ✓ Choose Collection
- ✓ Required and recommended metadata
- ✓ Upload Data Files

Small datasets via web browser

Globus File Transfer
>5 GB or 300 files



Curation Workflow

- ✓ Internal review
- ✓ Researcher is integral part of curation process

C	Check
U	Understand
R	Recommend improvements
A	Augment
T	Transform
I	Include persistent IDs and a reuse licence/agreement
O	Optimize for FAIRness
N	Note down curation activities



Indexing for Discovery



Preservation Workflow

- ❑ Size of data
- ❑ File types/formats
- ❑ Special considerations around licensing, sensitivities, or copyright
- ❑ Minimum publishable unit
- ❑ Documentation
- ❑ Current storage solution
- ❑ Data management support, etc.

Icons from the Noun Project: Medical cross by TMD; account by Adrien Coquet; Cloud Upload by Hassan ali; curation by SANB; approved by Aneeqe Ahmed

Data Publishing Workflow



Triage

- ✓ Researcher consultation (optional)
- ✓ Data Appraisal



Data Deposit Workflow

- ✓ Choose Collection
- ✓ Required and recommended metadata
- ✓ Upload Data Files

Small datasets via web browser

Globus File Transfer
>5 GB or 300 files



Curation Workflow

- ✓ Internal review
- ✓ Researcher is integral part of curation process

C	Check
U	Understand
R	Recommend improvements
A	Augment
T	Transform
I	Include persistent IDs and a reuse licence/agreement
O	Optimize for FAIRness
N	Note down curation activities



Indexing for Discovery

- ✓ Publication
- ✓ DOI registration (DataCite)
- ✓ Indexing in Lunaris, OpenAIRE, Google Dataset, etc.



Preservation Workflow

- ❑ Size of data
- ❑ File types/formats
- ❑ Special considerations around licensing, sensitivities, or copyright
- ❑ Minimum publishable unit
- ❑ Documentation
- ❑ Current storage solution
- ❑ Data management support, etc.

Icons from the Noun Project: Medical cross by TMD; account by Adrien Coquet; Cloud Upload by Hassan ali; curation by SANB; approved by Aneeqe Ahmed

Data Publishing Workflow



Triage

- ✓ Researcher consultation (optional)
- ✓ Data Appraisal



Data Deposit Workflow

- ✓ Choose Collection
- ✓ Required and recommended metadata
- ✓ Upload Data Files

Small datasets via web browser

Globus File Transfer
>5 GB or 300 files



Curation Workflow

- ✓ Internal review
- ✓ Researcher is integral part of curation process

C	Check
U	Understand
R	Recommend improvements
A	Augment
T	Transform
I	Include persistent IDs and a reuse licence/agreement
O	Optimize for FAIRness
N	Note down curation activities



Indexing for Discovery

- ✓ Publication
- ✓ DOI registration (DataCite)
- ✓ Indexing in Lunaris, OpenAIRE, Google Dataset, etc.



Preservation Workflow

- ✓ Bit-level preservation
- ✓ Geographically distributed storage
- ✓ Appraisal and selection for additional long-term preservation processing via Archivematica

- ❑ Size of data
- ❑ File types/formats
- ❑ Special considerations around licensing, sensitivities, or copyright
- ❑ Minimum publishable unit
- ❑ Documentation
- ❑ Current storage solution
- ❑ Data management support, etc.

Icons from the Noun Project: Medical cross by TMD; account by Adrien Coquet; Cloud Upload by Hassan ali; curation by SANB; approved by Aneeqe Ahmed

Reinvented FRDR!!

www.frdr-dfdr.ca



On the Horizon...



FRDR's future



- ORCID integration
- Private for peer-review feature
 - Selectively share data ahead of public release
 - Design refined during user acceptance testing
- Developing a mechanism for controlled access
 - Zero-knowledge encryption app
 - Sensitive data policy framework
 - Collaborative, Institution-led, Cross-jurisdictional
 - Sensitive data resources & supports



Helpful Resources



Federated Research Data Repository

- FRDR Support Desk <support@frdr-dfdr.ca>
- [FRDR Terms of Use](#) (and other policies)
- [Getting an account](#)
- FRDR [Demo](#) server
- [Preparing Your Data](#) for Deposit
- README template ([in English](#) et [en français](#))

Thank you! Questions?

Please keep in touch! <support@frdr-dfdr.ca>

