



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

14 avril 2023 // 14 April 2023



Digital Research
Alliance of Canada

Alliance de recherche
numérique du Canada





Reconnaissance du territoire // Land Acknowledgement





Code de conduite // Code of Conduct

- 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 évènements 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

- 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





Lunaris

Découvrez les données de recherche canadiennes

Mark Goodwin – Responsable de produits, Services de découverte



lunaris

Discover **Canadian Research Data**

Découvrir les **données de recherche canadiennes**



Ordre du jour

1. Introduction à Lunarix
2. Aperçu du logiciel
3. Historique du projet
4. Démonstration de la plateforme
5. Prochaines étapes – Mise en production complète

Lunaris est un service national de découverte des données de recherche offert par l'Alliance de recherche numérique du Canada, et il compte parmi les principaux services de gestion des données de recherche de l'organisation

Introduction à Lunaris

- La version bêta de Lunaris (www.lunaris.ca/fr) a été lancée le 29 mars 2023 – les commentaires sont les bienvenus à support@lunaris.ca
- Lunaris faisait autrefois partie du service de découverte du Dépôt fédéré de données de recherche (DFDR) – le DFDR et Lunaris sont désormais des plateformes et services distincts
- Plateforme évolutive bilingue misant sur des fonctions intégrées de recherche textuelles et cartographiques
- Les renseignements sur les ensembles de données sous forme d'enregistrements de métadonnées sont moissonnés à partir des dépôts et deviennent accessibles aux fins de découverte
- Lunaris a indexé plus de 80 000 ensembles de données provenant de plus de 100 dépôts et collections de données du Canada
 - Dépôts universitaires, dont le Dataverse Borealis
 - Dépôts des gouvernements fédéral, provinciaux et locaux
 - Dépôts propres aux domaines
 - Initiatives de dépôt nationales, dont le DFDR

Introduction à Lunaris

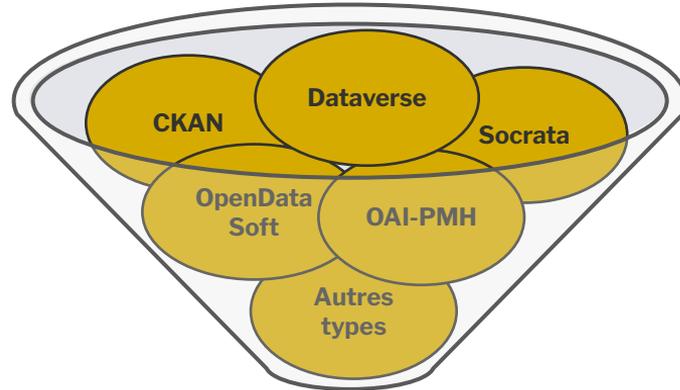
Fonctionnalités de Lunaris

- Recherche textuelle propulsée par Globus Search
- Recherche cartographique propulsée par Geodisy, une initiative basée à l'Université de la Colombie-Britannique
- Options de recherche par filtre et de recherche avancée
- Générateur de citation d'ensembles de données
- Aperçu des données géospatiales (à venir sous peu)

Avantages de Lunaris

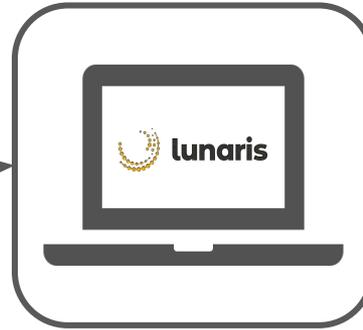
- Outil et service spécialisés qui améliorent la découverte des données de recherche canadiennes
- Décloisonnement des dépôts
- Renvoi du trafic vers les sites de dépôt existants
- Création d'une interopérabilité entre les plateformes canadiennes et internationales

Aperçu du logiciel

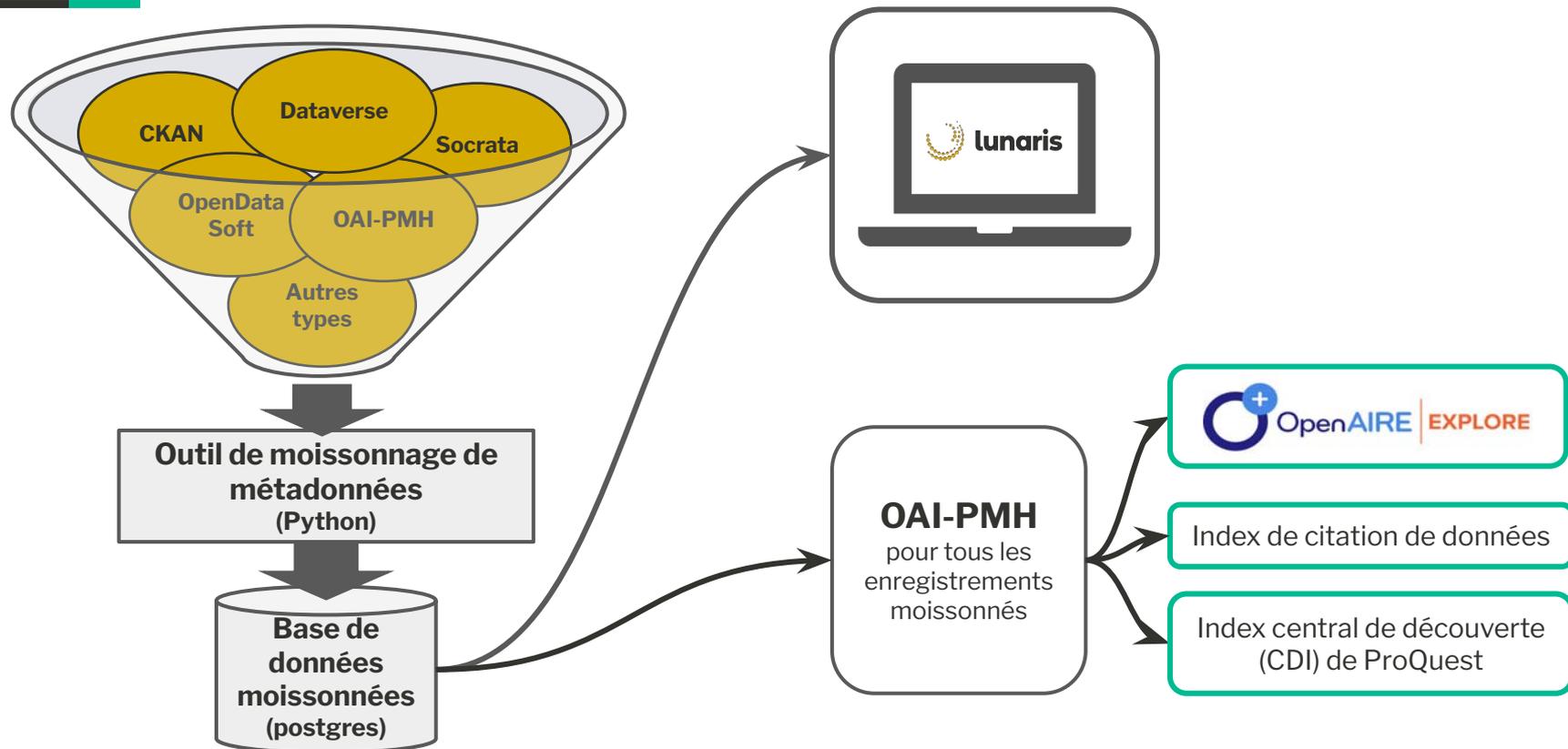


**Outil de moissonnage de métadonnées
(Python)**

**Base de données moissonnées
(postgres)**



Aperçu du logiciel



Aperçu du logiciel

Lunaris (ré)utilise diverses composantes logicielles :

- **Outil de moissonnage du DFDR** : outil de moissonnage du dépôt
- **Globus Search** : recherche textuelle et géospatiale et filtre
- **Geodisy** : recherche cartographique
- **GDAL** : traitement de fichiers géospatiaux
- **GeoServer** : serveur de publication et de répartition des données géospatiales
- **GeoBlacklight** : interface de découverte

Historique du projet Lunarisis

Le **projet de refonte de l'interface de découverte du DFDR** a débuté en 2021

Objectifs du projet :

- Intégrer la plateforme du service de découverte du DFDR dans la fonction de recherche cartographique du DFDR, offrant ainsi aux utilisatrices et utilisateurs une expérience de recherche harmonieuse combinant la recherche textuelle et la recherche cartographique
- Accroître les capacités de la fonction de recherche du DFDR en mettant en place différents mécanismes, dont des fonctions de filtre et de recherche avancée
- Améliorer les fonctionnalités de la plateforme et l'expérience d'utilisation pour répondre à divers besoins de la communauté de recherche du Canada

Les travaux sont soutenus par une équipe de projet interne et par le **groupe de travail sur la refonte de l'interface de découverte du DFDR**, garantissant ainsi un projet axé sur la communauté et tenant compte des pratiques exemplaires en matière d'utilisabilité

Historique du projet Lunaris

Activités liées au projet

- Évaluation du contexte et analyse comparative
- Recherche sur l'utilisation – exercice de visualisation et sondage auprès des utilisatrices et utilisateurs
- Élaboration de plusieurs modèles, maquettes et prototypes
- Conception de modules avec le groupes de travail
 - Module 1 : Pages des enregistrements
 - Module 2 : Page des résultats de recherche
 - Module 3 : Recherche locale, page de navigation et liste de dépôts
 - Module 4 : Fonction/algorithme de recherche par mot-clé et par carte
 - Module 5 : Filtres et affichage de métadonnées
 - Module 6 : Mise en corrélation et enrichissement des données de recherche
 - Module 7 : Fonctionnalités de l'API et analytique
- Collaboration avec des firmes de conception pour trouver un nom, créer un logo et concevoir la maquette
- Deux rondes d'essais d'utilisabilité
- Actualisation du contenu des pages de soutien (À propos, Guide d'utilisation, etc.)
- Traduction et révision du contenu en français
- Développement et mise à l'essai de la version bêta de Lunaris

Démonstration

www.lunaris.ca/fr

Prochaines étapes

La mise en production complète est prévue en mai. Mises à jour anticipées :

- Aperçu des ensembles de données géospatiaux sur la page des enregistrements
- Vérification indépendante du français
- Options de couches cartographiques
- Mises à jour en fonction des commentaires de la communauté

Contributions

Groupe de travail sur la refonte de l'interface de découverte du DFDR

- Krista Godfrey, Université de Waterloo
- David Kemper, Université McMaster
- Amber Leahey, Scholars Portal
- Winnie Li, Université Canada West
- Kathleen Matthews, Université de Victoria
- Nicholas Rochlin, Université de la Colombie-Britannique
- Kristi Thompson, Université Western

Partenaires

- Globus
- Université de la Saskatchewan
- Université Simon-Fraser

Équipe de développement et des opérations à l'interne

- Mark Goodwin, responsable de produits, Service de découverte
- Neha Milan, responsable de produits, Dépôt fédéré de données de recherche
- Todd Trann, directeur technique
- Mike Winter, développeur principal
- Adam McKenzie, analyste en gestion des données de recherche
- Joel Farthing, développeur
- Victoria Smith, coordonnatrice des politiques, de la confidentialité et des données sensibles
- Lee Wilson, directeur de la gestion des données de recherche

Un merci particulier aux anciens membres de l'équipe
Kelly Stathis et Paul Dante

Envoi de commentaires

Tous vos commentaires sont les bienvenus pendant la phase bêta de Lunarisis! Veuillez nous écrire à support@lunaris.ca.

Merci!

Le Dépôt fédéré de données de recherche

Publier et partager des données de recherche canadiennes

Erin Clary, coordonnatrice de la curation, Alliance de recherche numérique du Canada

Neha Milan, responsable de produits, DFDR, Université de la Saskatchewan

DFDR
Dépôt fédéré de
données de recherche



FRDR
Federated Research
Data Repository



Ordre du jour

1. Qu'est-ce le DFDR?
2. Pourquoi utiliser le DFDR?
3. Qui peut utiliser le DFDR?
4. Processus de dépôt et de publication des données
5. Démonstration
6. Fonctionnalités à venir

Qu'est-ce le DFDR ?

- Le Dépôt fédéré de données de recherche (DFDR) est une plateforme de publication bilingue qui permet de partager et de préserver les données de recherche canadiennes.
- Le service a initialement été mis sur pied dans le cadre d'une collaboration entre la Fédération Calcul Canada, l'Association des bibliothèques de recherche du Canada et un certain nombre d'établissements postsecondaires canadiens.
- La mise en production complète du service a eu lieu en février 2021.
- À l'heure actuelle, le service est offert à l'échelle nationale et son financement est administré par l'intermédiaire de [l'Alliance de recherche numérique du Canada](#).

Pourquoi utiliser le DFDR ?

- Publier librement des données de recherche dans une option de dépôt nationale **bilingue et canadienne**.
- Soutenir les **ensembles de données de moyenne et grande taille** au moyen de Globus File Transfer.
- Permettre aux utilisatrices et utilisateurs d'accéder à une option de **stockage dans un dépôt structuré** (sans frais, jusqu'à un certain seuil), avec des mécanismes en place pour obtenir du stockage additionnel, au besoin.
- Le DFDR attribue des **DOI** pour faciliter la citation des données.
- Option de stockage dans un dépôt sécurisé **réparti géographiquement** dans plusieurs sites d'hébergement nationaux de calcul informatique de pointe.
- Possibilité de collaborer avec **plusieurs personnes** sur une seule soumission.
- Possibilité de définir des périodes d'**embargos facultatifs** sur les enregistrements de données et de métadonnées.

Qui peut utiliser le DFDR ?

- N'importe qui peut chercher et télécharger des données.
 - Les données peuvent temporairement sous embargo.
 - Les licences attribuées aux chercheuses et chercheurs précisent la façon dont chaque ensemble de données peut être utilisé.
- Qui peut déposer des données?
 - Les membres du corps professoral et leurs personnes désignées (gestionnaires de données, étudiantes et étudiants, titulaires de bourse postdoctorale, etc.) d'établissements ayant droit au financement des trois conseils.
 - Allocation par défaut de 1 To d'espace de stockage de données sans frais.
 - Possibilité d'obtenir du stockage additionnel.

Processus de publication des données



Tri



Processus de dépôt
des données



Processus de curation



Approbation et indexation
aux fins de découverte



Processus de préservation

- ✓ Consultation des chercheur(-se)s (facultatif)
- ✓ Évaluation des données

- Taille des données
- Types/formats des fichiers
- Considérations spéciales concernant les licences, sensibilités ou droits d'auteur
- Unité minimale publiable
- Documentation
- Solution de stockage actuelle
- Soutien pour gérer les données, etc.

Processus de publication des données



Tri

- ✓ Consultation des chercheur(-se)s (facultatif)
- ✓ Évaluation des données



Processus de dépôt des données

- ✓ Choix d'une collection
- ✓ Métadonnées requises et recommandées
- ✓ Télécharger des fichiers de données

Petits ensembles de données via un navigateur Web

Globus File Transfer
>5 Go ou 300 fichiers



Processus de curation



Approbation et indexation aux fins de découverte



Processus de préservation

- Taille des données
- Types/formats des fichiers
- Considérations spéciales concernant les licences, sensibilités ou droits d'auteur
- Unité minimale publiable
- Documentation
- Solution de stockage actuelle
- Soutien pour gérer les données, etc.

Processus de publication des données



Tri

- ✓ Consultation des chercheur(-se)s (facultatif)
- ✓ Évaluation des données



Processus de dépôt des données

- ✓ Choix d'une collection
- ✓ Métadonnées requises et recommandées
- ✓ Télécharger des fichiers de données

Petits ensembles de données via un navigateur Web

Globus File Transfer
>5 Go ou 300 fichiers



Processus de curation

- ✓ Examen interne
- ✓ Les chercheur(-se)s sont une partie intégrante du processus de curation

C	Consulter
U	Un peu plus en profondeur
R	Recommander
A	Améliorer
T	Transposer
I	Inclure ID pérennes et les licences/ententes de réutilisation
O	Optimiser selon les principes FAIR
N	Noter les actions réalisées



Approbation et indexation

aux fins de découverte Processus de préservation



- Taille des données
- Types/formats des fichiers
- Considérations spéciales concernant les licences, sensibilités ou droits d'auteur
- Unité minimale publiable
- Documentation
- Solution de stockage actuelle
- Soutien pour gérer les données, etc.

Icônes de the Noun Project : croix médicale par TMD; compte par Adrien Coquet; chargement infonuagique par Hassan ali; curation par SANB; approbation par Aneeque Ahmed

Processus de publication des données



Tri

- ✓ Consultation des chercheur(-se)s (facultatif)
- ✓ Évaluation des données



Processus de dépôt des données

- ✓ Choix d'une collection
- ✓ Métadonnées requises et recommandées
- ✓ Télécharger des fichiers de données

Petits ensembles de données via un navigateur Web

Globus File Transfer
>5 Go ou 300 fichiers



Processus de curation

- ✓ Examen interne
- ✓ Les chercheur(-se)s sont une partie intégrante du processus de curation

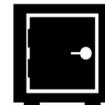
C	Consulter
U	Un peu plus en profondeur
R	Recommander
A	Améliorer
T	Transposer
I	Inclure ID pérennes et les licences/ententes de réutilisation
O	Optimiser selon les principes FAIR
N	Noter les actions réalisées



Approbation et indexation aux fins de découverte

Processus de préservation

- ✓ Publication
- ✓ Enregistrement des DOI (DataCite)
- ✓ Indexation des données dans Lunarix, OpenAIRE, Google Dataset, etc.



- Taille des données
- Types/formats des fichiers
- Considérations spéciales concernant les licences, sensibilités ou droits d'auteur
- Unité minimale publiable
- Documentation
- Solution de stockage actuelle
- Soutien pour gérer les données, etc.

Icônes de the Noun Project : croix médicale par TMD; compte par Adrien Coquet; chargement infonuagique par Hassan ali; curation par SANB; approbation par Aneeque Ahmed

Processus de publication des données



Tri

- ✓ Consultation des chercheur(-se)s (facultatif)
- ✓ Évaluation des données

- Taille des données
- Types/formats des fichiers
- Considérations spéciales concernant les licences, sensibilités ou droits d'auteur
- Unité minimale publiable
- Documentation
- Solution de stockage actuelle
- Soutien pour gérer les données, etc.



Processus de dépôt des données

- ✓ Choix d'une collection
- ✓ Métadonnées requises et recommandées
- ✓ Télécharger des fichiers de données

Petits ensembles de données via un navigateur Web

Globus File Transfer
>5 Go ou 300 fichiers



Processus de curation

- ✓ Examen interne
- ✓ Les chercheur(-se)s sont une partie intégrante du processus de curation

C	Consulter
U	Un peu plus en profondeur
R	Recommander
A	Améliorer
T	Transposer
I	Inclure ID pérennes et les licences/ententes de réutilisation
O	Optimiser selon les principes FAIR
N	Noter les actions réalisées



Approbation et indexation aux fins de découverte

- ✓ Publication
- ✓ Enregistrement des DOI (DataCite)
- ✓ Indexation des données dans Lunaris, OpenAIRE, Google Dataset, etc.



Processus de préservation

- ✓ Préservation des bits
- ✓ Stockage réparti géographiquement
- ✓ Évaluation et sélection pour un traitement de préservation additionnel à long terme par l'intermédiaire d'Archivematica

Icônes de the Noun Project : croix médicale par TMD; compte par Adrien Coquet; chargement infonuagique par Hassan ali; curation par SANB; approbation par Aneeque Ahmed

Le DFDR réinventé!!

www.frdr-dfdr.ca



À l'horizon...



Le DFDR dans le futur

- Intégration d'ORCiD
- Fonction « examen privé par des pairs »
 - Partage de certaines données avant la publication publique
 - Conception peaufinée pendant les essais d'acceptation des utilisateur(-trice)s
- Mécanismes de gestion des accès contrôlés aux données
 - Application de chiffrement à connaissance nulle
 - Cadre stratégique des données sensibles
 - Collaboratifs, dirigés par les établissements, interorganisationnels
 - Ressources et soutien pour les données sensibles

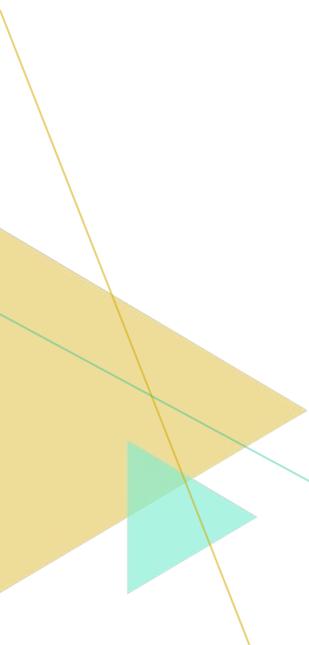


Ressources utiles



Le Dépôt fédéré de données de recherche

- Service d'assistance du DFDR <support@frdr-dfdr.ca>
- [Conditions d'utilisation du DFDR](#) (et autres politiques)
- [Obtenir un compte](#)
- Serveur [démon](#) du DFDR
- [Préparation de vos données](#) pour les déposer
- Modèle de fichier Lisez-moi ([en français](#) et [in English](#))



Merci! Des questions?

Restez en contact ! <support@frdr-dfdr.ca>