



English:

<https://bit.ly/PIDMAGIC>

Français:

<https://bit.ly/PIDMAGIC>



# La magie des identifiants pérennes (PID) pour la recherche/les chercheur-euse-s



**Susan Brown**

Chaire de recherche du Canada sur la recherche  
numérique collaborative, Université de Guelph  
Directrice, CWRC | Cheffe de projet, LINCS

2023 Sommet des membres et l'AGA du RCDR  
4 Octobre 2023

# Magie administrative

Systemes avancés orientés vers les chercheur·euse·s  
qui tirent parti des PID

- e-cvs
- rapport de subvention
- candidatures aux subventions
- et plus

# Magie administrative

## Défis

- Plus de PID pour représenter un éventail plus large de produits de recherche
- Construire **des outils tournés vers les chercheur·euse·s!**
  - Centrés sur les chercheur·euse·s
  - Bien conçus

# Magie dans la recherche

## Revendications

- Les données générées par des chercheur·euse·s ont besoin de PID
- La recherche a besoin de données culturelles pour avoir des PID

## Infrastructures des sciences humaines

- Le collaboratoire scientifique des écrits du Canada (CWRC/CSÉC)
- L'Infrastructure Numérique pour les données Culturelles dans le web Sémantique (LINCS)

# Technologies sémantiques

**Extensible Markup Language (XML)**

**Web Ontology Language  
(OWL)**

**Resource Description Framework (RDF)  
and RDF Schema (RDF-S)**

**Web Annotation Data Model**

**CIDOC**

**Conceptual Reference Model (CIDOC-  
CRM)**



Source: [Ontotext](#)

L'objectif ultime de la technologie sémantique est d'aider les machines à comprendre les données.

— “Semantic technology”, *Wikipedia*

# Le Collaboratoire scientifique des écrits du Canada (CSÉC)

- Environnement de recherche virtuel subventionné par la FCI
- 35 projets; 400,000 objets numériques en formats XML, PDF, RDF, image, audio, et vidéo
- Construit avec Islandora extended pour prendre en charge le balisage sémantique et les métadonnées

The screenshot shows the CSÉC website interface. At the top, there is a navigation bar with the CSÉC logo and links for 'Collaboratoire', 'À propos', 'Conditions d'utilisation', and 'Actualités'. On the right, there are links for '? - FR -' and 'Connecter/Se joindre'. Below the navigation bar is a search bar with the placeholder text 'Rechercher...' and a dropdown menu set to 'Collaboratoire'. To the right of the search bar is a magnifying glass icon and a link for 'Recherche avancée'. The main content area features a large banner for LEAF (Linked Editing Academic Framework). The banner includes the CWRC/CSÉC logo and text: 'The Canadian Writing Research Collaboratory software turns over a new LEAF!'. To the right of this text is the LEAF logo and a paragraph: 'We've teamed up with Bucknell University (US) and Newcastle University (UK) to build the next generation of the CWRC Virtual Research Environment – and it's well underway!'. A blue button labeled 'Check out LEAF!' is positioned at the bottom right of the banner. Below the banner are three content boxes. The first box, titled 'Ce que nous écrivons', lists 'Italian Gardens' (10/3/23) by Dorothy Dean Duffield (9/12/23) and 'Grace Campbell' (9/12/23). The second box, titled 'Ce que nous collectons', lists 'Saving Frogs Transcript' (10/3/23) by Zavitz Pines Story Transcript (10/3/23) and 'Dorothy Dean Tate' (9/12/23). The third box, titled 'Ce que nous partageons', lists 'We Belong Here: Interview with author Frieda Wishinky and illustrator Ruth Ohi' (10/4/23), 'We Belong Here' (10/3/23), and 'Salma Writes a Book' (9/29/23).

# Les données générées par des chercheur·euse·s ont besoin de PID

Large éventail de formats et de types de données

- Historiques



Histoire de la  
Presse  
Francophone dans  
l'Ouest canadien



The People and the Text

# Les données générées par des chercheur·euse·s ont besoin de PID

Large éventail de formats et de types de données

- Historiques
- Contemporains



**Contextualizing  
COVID-19: An Archive  
for the Future**

**Canadian Women  
Playwrights Online**



# L'Infrastructure Numérique pour les données Culturelles dans le web Sémantique (LINCS)

➤ Cyber-  
infrastructure  
subventionnée par la  
FCI

➤ Suite d'outils open  
source (dont beaucoup  
sont réutilisés) pour la  
conversion, le  
stockage, et l'accès  
aux DOL

➤ Stocke uniquement  
les triplets de données  
liées (RDF): tous les  
biens culturels sont  
connectés par des PID



## Interagissez différemment avec les données.

Parcourez des millions de points de données et explorez leurs connexions. Effectuez des recherches dans ou dans des ensembles de données et des résultats de facette. Créez des données ouvertes liées pour répondre à des questions de recherche complexes.

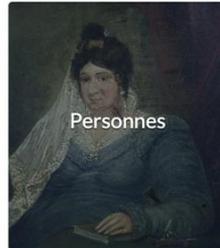
À propos

En savoir plus sur les données liées

Explorer

Contribuer aux données

### Jeux de données



Personnes



Des endroits



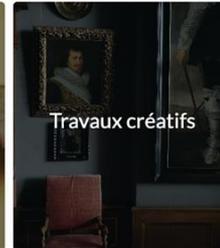
Organisations



Sources



Événements



Travaux créatifs

5,780,000 +

Triples in Progress

272,000 +

People, Places, Groups, Events, and  
Creative Works

8

Datasets in Progress

100 +

Researchers, Students, Staff and  
Partners

**L'Infrastructure Numérique pour les données Culturelles dans le web Sémantique (LINCS)** permet **aux chercheur-euse-s Canadien-ne-s** de créer des **données en lignes interreliées** et **contextualisées** sur la culture pour des bénéfices académiques et publiques.

<https://lincsproject.ca/fr/>

# Les données des sciences humaines avec des PID sont véritablement “FAIR”

F  
Findable



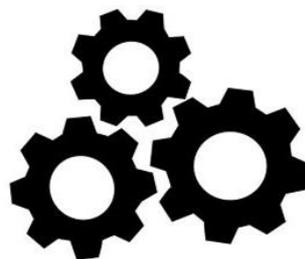
Trouvable

A  
Accessible



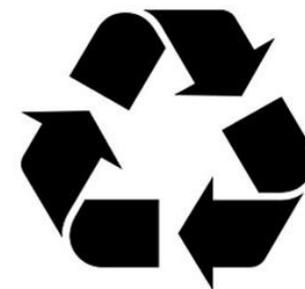
Accessible

I  
Interoperable



Interopérable

R  
Reusable



Réutilisable

Principes FAIR (2016): <https://www.go-fair.org/fair-principles/>

Image: [https://commons.wikimedia.org/wiki/File:FAIR\\_data\\_principles.jpg](https://commons.wikimedia.org/wiki/File:FAIR_data_principles.jpg)

# Les données générées par des chercheur·euse·s ont besoin de PID

## Défis

- Lier des données rassemblées par des chercheur·euse·s dans l'écosystème de recherche
- Élaborer une stratégie DOI (identificateur d'objet numérique) pour revenir à la version de dépôt

# La recherche a besoin de données culturelles pour avoir des PID

Early Canadian Cultural Journals Index (set de données sauvé et replatformé)

Près de 140,000 dossiers descriptifs au niveau éditorial du contenu de journaux anciens

Importante superposition avec les collections Canadiana/Héritage et d'autres contenus historiques numérisés



 Early Canadian Cultural Journals Index

[Home](#) [Journal Index](#) [Article Index](#)

## Illustrated News

This collection makes available the rich contents of Thomas B. Vincent's 1993 CD-ROM Index to Pre-1900 Canadian Cultural and Literary Magazines.

This ground-breaking index includes 137,943 records of prose, poetry, drama, fiction, and miscellaneous items (letters to the editor, sheet music, illustrations, advertisements, and obituaries) from 203 nineteenth-century periodicals published in Canada.

It includes *The Nova Scotia Magazine*, Canada's first literary magazine, longer-running titles such as the *Canadian Monthly and National Review*, and more ephemeral ones such as *Mayflower* or *Ladies' Acadian Newspaper*. Not all runs are complete; only what was found extant was included. Some major periodicals such as *Saturday Night* are not included. The Journals collection, which was not released with the original index on CD-ROM, records what was seen, where, what was entered in the collection, and what is missing. In addition to essential publication information, there are subject descriptors on all items.

**Related full-text content**

Much of the content described in these records has been digitized and is freely available online through the [Canadiana](#) Collections.

**Content warning**

As the Canadiana website notes: "The Canadiana Collections contain content created over five centuries that tell an incomplete, oftentimes distorted and, sometimes harmful, story of Canada. The content, metadata, and resource descriptions in the Canadiana Collections contain language that reflects the biases, norms, and perspectives of the time in which they were created. This includes harmful and offensive wording, cultural references, and stereotypes." This is true also of the metadata records, provided in this collection, which record the language of the time.

[Documentation](#)

Related Articles

["Opportunities and Pitfalls" - 1989 Article](#)

Related Resources

[Early Canadian Periodicals at Canadiana](#)

Creator

Thomas B. Vincent

# La recherche a besoin de données culturelles pour avoir des PID

## Défis

- Échelle
- Coût
- Maintenance
- Implications en matière d'infrastructure

# Le contenu des données culturelles a besoin de PID

- Entités nommées:
  - Figures historiques, endroits, Premières Nations, organisations, endroits historiques
  - Entités non nommées
- Vocabulaires pour des concepts

# Le contenu des données culturelles a besoin de PID

## Défis

- Ampleur et échelle
- La logistique et l'infrastructure pour permettre aux institutions d'établir des liens avec des chercheur·euse·s
- Encourager les chercheur·euse·s à devenir collaborateur·rice·s

# Des PID pour un écosystème intersectoriel

Les rôles doubles du RDCR pour donner accès aux:

- Publications académiques numériques
- Ressources de l'héritage culturel en ligne

**Une stratégie nationale de PID peut tirer parti des deux rôles: renforcer les connections entre chercheur·euse·s, et collaborer avec eux·elles dans la production d'un écosystème de connaissance meilleur et plus robuste**

“Alors que l’infrastructure des données dans les sciences humaines est sous-développée, (peut-être en grande partie en raison du manque de dépôts de domaines), **l’obstacle le plus important au partage de données** dans les sciences humaines serait **un manque de compréhension systématique sur comment réutiliser des données en tant que source principale.**”

[traduit par Alice Hinchliffe]

Dylan Ruediger and Ruby MacDougall  
“Are the Humanities Ready for Data Sharing?” *Ithaka S+R*. 6 March 2023.  
<https://doi.org/10.18665/sr.318526>

# Merci

**Contactez-nous si vous êtes intéressé·e·s**

[cwrc@ualberta.ca](mailto:cwrc@ualberta.ca)

[lincs@uoguelph.ca](mailto:lincs@uoguelph.ca)

## Avec gratitude

Les organisateur·rice·s du Sommet des  
membres du RCDR

L'équipe LINCS et CSEC et nos multiples  
collaborateur·rice·s

**Financeurs majeurs**

**INNOVATION**

Canada Foundation  
for Innovation

Fondation canadienne  
pour l'innovation

**canarie**



Mellon  
Foundation

