



# Repenser la préservation numérique

## Au-delà du dépôt

Conférence du RCDR 2020 | 15 octobre 2020

Corey Davis

Digital Preservation Librarian, University of Victoria

Visiting Program Officer, Canadian Association of Research Libraries

# Reconnaissance territoriale



# Aperçu de la séance

- Qu'est-ce que la préservation numérique ?
- Pourquoi les bibliothèques ont-elles des difficultés?
- Quelles sont les solutions ?
- Discussion

# Qu'est-ce que la préservation numérique ?

- «La préservation numérique consiste en un certain nombre de méthodes qui visent à assurer l'accessibilité permanente aux matériaux numériques. Faire cela implique qu'on ait trouvé les moyens de représenter ce qui a été présenté à l'origine aux utilisateurs moyennant une combinaison d'outils logiciels et matériels agissant sur des données ».
  - UNESCO  
<https://fr.unesco.org/themes/information-preservation/digital-heritage/concept-digital-preservation>

# Enquête de l'ABRC sur la préservation numérique

- « 96% des répondants ont indiqué qu'ils donnent accès aux documents numériques qu'ils préservent [et] 79% ... utilisent un dépôt d'objets numériques ou une plateforme de découverte ».
- Hurley, G., & Shearer, K. *Rapport final de l'enquête sur la capacité et les besoins en matière de conservation numérique des établissements canadiens de la mémoire, 2017-18*. Association des bibliothèques de recherche du Canada.

# Enquête de l'ABRC sur la préservation numérique

- « Seuls quelques répondants estiment avoir la capacité de préserver leurs fonds... »  
– *Ibid.*

# Une question de priorités et d'attention

- « Pour les établissements de mémoire, la préservation numérique n'est pas encore une activité prioritaire. Ils n'investissent pas beaucoup de ressources dans ce domaine ».  
– *Ibid.*

# Action collective pour accroître la capacité

- Groupe d'experts sur la préservation (Portage)
  - « L'objectif du groupe est de conseiller au sujet de l'élaboration de l'infrastructure de gestion des données de recherche et de pratiques exemplaires en préservation des données et métadonnées de recherche ».
- Groupe de travail sur la conservation numérique (ABRC)
  - « Son travail comprend l'évaluation de la capacité, les ressources disponibles et les possibilités de financement présents dans le milieu canadien de la recherche en matière de conservation numérique, et l'identification et la promotion des approches, des normes, des pratiques et des technologies pertinentes ».

# Aperçu de la présentation

1. Il est compliqué d'établir un lien entre les dépôts individuels et les services de préservation partagés.
2. L'écosystème actuel privilégie l'accès à moyen terme (par des dépôts) plutôt que l'accès à long terme (par la préservation).

COPPUL's WestVault

# ÉTUDE DE CAS

# Étude de cas : WestVault

- Le réseau privé de COPPUL LOCKSS pour le stockage de préservation distribué était sous utilisé
  - Difficile de créer des plug-ins de dépôt pour LOCKSS
  - Chaque boîte LOCKSS a dû être agrandie pour répondre aux besoins des plus gros dépôts
  - Manque d'ententes de service

# Étude de cas : WestVault

- Réinitialisation !
  - Dépôt « glisser-déposer »
  - Machines virtuelles dans d'importants centres de données
  - Politiques robustes, guides d'utilisateurs et formations, personnel attiré de COPPUL

# Étude de cas : WestVault

- Résultats
  - L'extraction du matériel à déposer exige encore beaucoup d'attention et d'intervention manuelle

# Étude de cas : WestVault

- Impératifs :
  - Accroître les capacités au niveau local et par l'entremise de consortiums régionaux et d'organisations nationales
  - **Mettre les dépôts à l'échelle comme nous l'avons fait pour les services de préservation; établir un lien entre ces deux éléments**
  - **Repenser potentiellement nos méthodes d'accès et de préservation !**

Étudier des systèmes exemplaires pour trouver des solutions

# METTRE LES DÉPÔTS À L'ÉCHELLE

BC ELN et COPPUL

# INFRASTRUCTURE EXEMPLAIRE 1

# BC ELN et COPPUL

- Arca
  - « La plateforme partagée offre aux organisations des dépôts à libre accès ».

<https://bceln.ca/services/shared-services/arca>
- WestVault
  - « Réseau de stockage à redondance élevée, administré de façon indépendante par les pairs qui est déployé dans les 4 provinces de l'ouest ».

<https://bceln.ca/resources/westvault>

# Cadre réglementaire simple...

## BC ELN and COPPUL working together: Arca Collaborative Digital Repository and COPPUL Digital Preservation Network

 [Printer-friendly version](#)

**VANCOUVER, B.C., November 30, 2016** — The [British Columbia Electronic Library Network](#) (BC ELN) and the [Council of Prairie and Pacific University Libraries](#) (COPPUL) announced today the signing of an agreement that will enable COPPUL members to join the [Arca Collaborative Digital Repository](#) while providing existing Arca members access to COPPUL's emerging suite of digital preservation services.

Both BC ELN and COPPUL have a common interest in ensuring that the high-value digital materials of their member institutions are accessible to their communities over the long term. By leveraging their respective expertise and service infrastructures in the areas of institutional repositories and digital preservation, both BC ELN and COPPUL members will benefit from this agreement.

Under this agreement, COPPUL members will be able to participate in the [Arca Collaborative Digital Repository](#), which is based on a shared instance of the open source [Islandora Digital Repository Framework](#). At the same time, Arca participants will have access to an emerging set of digital preservation services being developed as part of the [COPPUL Digital Preservation Network \(COPPUL DPN\)](#), including preservation storage in [COPPUL's Private LOCKSS Network](#), and preservation processing services such as [Archivematica-as-a-Service](#).

# E·COLLECTIONS

AT EMILY CARR

[Home](#) | [Browse](#) | [About + Contact](#) | [Submit](#) | [ECU Library](#) | [Login](#)

Explore Emily Carr University's eCollections

Search Emily Carr University's Institutional Repository. Find scholarly research materials and other digital assets from our archives and the Wosk Masterworks Print Collection.

[Search All Collections](#)

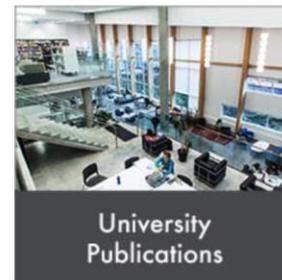
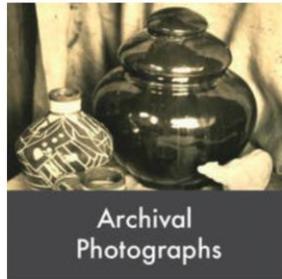


[Advanced Search](#)

[Browse Collections](#)  
[Browse All Materials](#)

[About](#)

Heinz Laffin in ceramics studio, circa late 1950's - early 1960's.



[Home](#) > [Browse Repository](#) > [Graduate Theses](#)

# Aurora

View

IP Embargo

Manage

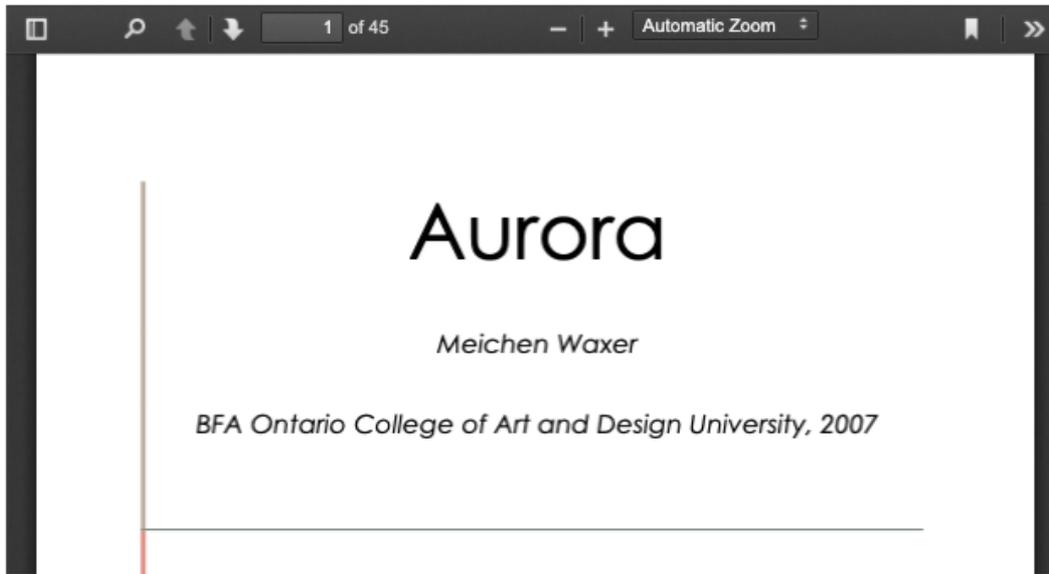
Edit metadata

## Style

APA

Choose the citation style.

Waxer, M. (2016). Aurora. doi:10.35010/ecuad:8455



Part of: [Aurora \(10 objects\)](#)  
manage parent

Next

Aurora



Untitled (Evenlode)



Bird of paradise, tree of life

Preserve this object?

Checking this box creates a PRESERVATION datastream, which will flag this object for preservation on the next cron run. Un-checking this box will delete the PRESERVATION datastream, which will not delete the preserved objects, but stop the preservation of updates to the object.

Save Configuration

Preserve this object?

Checking this box creates a PRESERVATION datastream, which will flag this object for preservation on the next cron run. Un-checking this box will delete the PRESERVATION datastream, which will not delete the preserved objects, but stop the preservation of updates to the object.

New children of this collection (excludes current children)

All children of this collection (existing and new)

Current children only - new objects will not be preserved

Child collections. Child collections must be configured individually.

Save Configuration

# Intégration de la préservation

- COPPUL a investi 3 000 \$ pour aider BC ELN à développer Islandora WestVault
- Arca = 5 100 \$/année (allocation de 150Gb)
- WestVault = 200 \$/ année par 100Gb

# Accroître l'accès au secteur « GLAM »...



## Hosting Initiative Expands Access to Province's Historical Treasures

**Burnaby, B.C., August 24, 2020** – Past, present, and future BC History Digitization Program (BCHDP) grant recipients have access to free, sustainable, standards-based hosting for their online collections into 2025 thanks to an extension of the BCHDP-Arca Hosting Initiative. A renewed 5-year commitment by the British Columbia Electronic Library Network (BC ELN) and the Irving K. Barber Learning Centre (IKBLC) at the University of British Columbia Library builds on an established 3-year partnership that has allowed BCHDP grant recipients to host their digitized material on Arca, BC's collaborative digital repository, at no cost.

Ten GLAM sector (Galleries, Libraries, Archives, and Museums) organizations – groups like the Potters Guild of BC and grunt gallery – have taken part in this initiative over the preliminary 3-year pilot period, and several more are in discussions to add their collections to Arca. They join a growing number of post-secondary institutions and GLAM organizations that are making their local digital assets openly available via the Arca website ([arcabc.ca](http://arcabc.ca)).

IKBLC's continued funding, along with BC ELN's dedicated administration of the initiative and of the award-winning Arca platform, ensures that the BCHDP-Arca Hosting Initiative remains an ongoing and sustainable option for grant recipients. The successful partnership advances both BCHDP's goal of promoting access to British Columbia's unique historical materials and Arca's vision of sharing digital scholarship and heritage and building communities in British Columbia and beyond.



# Extensible à d'autres services de dépôt

 **bondjimbond** Merge pull request [#46](#) from bondjimbond...  ✓ 2a5c2cb on Aug 26, 2019  **117** commits

 includes	Add maximum filesize check	14 months ago
 .travis.yml	Change travis php dist	13 months ago
 LICENSE.txt	Get started	2 years ago
 README.md	Update Readme with config rules	16 months ago
 islandora_westvault.drush.inc	Remove debug line	13 months ago
 islandora_westvault.info	Setup paged and compound for use with islandora_bagit_c...	16 months ago
 islandora_westvault.module	Change compound child preservationto foreach loop	16 months ago

README.md

## Islandora Westvault

---

### Introduction

---

Islandora Westvault lets you flag individual objects, newspapers, books, or whole collections to be preserved in COPPUL's Westvault LOCKSS service or another OwnCloud-based service.

Scholars Portal et COPPUL

# **INFRASTRUCTURE EXEMPLAIRE 2**

# Scholars Portal et COPPUL

- Dataverse
  - « Plateforme de dépôt de données qui est accessible publiquement. Elle est entièrement hébergée sur des serveurs canadiens à University of Toronto Libraries ». <https://dataverse.scholarsportal.info/>
- Archivemata-as-a-Service
  - « Système de préservation numérique de source ouverte [qui peut être utilisé] sans installer ou gérer localement Archivemata ». <https://coppul.ca/archivemata>

# Travailler ensemble en Ontario...



[News](#) [Contact](#)

[About](#)

[Community](#)

[Services](#)

[Resources](#)

[Vendors](#)

[Log In](#)

[Home](#) » [Dataverse-Archivematica Integration Now Available for Testing](#)

## Dataverse-Archivematica Integration Now Available for Testing

**Posted on:** 5 March 2019 **By:** Kaitlin

The Ontario Council of University Libraries (OCUL) and Scholars Portal are pleased to announce that the first version of the Dataverse-Archivematica integration for long-term dataset preservation is now available for community testing and feedback.

For the past several years the Scholars Portal team has been working with Artefactual Systems Inc. to create an integration between the preservation processing tool Archivematica and Dataverse. This technical integration enables users of Archivematica to select and copy datasets from Dataverse instances and process them for long-term access and digital preservation. The Archival Information Packages (AIPs) include metadata and all files stored by Dataverse, and are packaged for long-term storage and management using Archivematica's standards-compliant workflow. Integration between Archivematica and Dataverse expands capacity to preserve research data, opening up new avenues for librarians and archivists to make valuable research outputs accessible into the future.





**University  
of Victoria**

UVic Research Data  
Collection



Ocean Networks Canada



Canada Health **Infoway**  
Canada Health Infoway

[Advanced Search](#) **Dataverses (4)** **Datasets (40)** **Files (291)**

### Dataverse Category

Organization or Institution (4)

### Publication Year

2019 (21)

2018 (13)

2020 (10)

### Author Name

1 to 10 of 44 Results

Sort

Data from: How far have we come? A review of MPA network performance indicators in reaching qualitative elements of Aichi Target 11



Sep 15, 2020 - CHONe

Miller, Mairi; Ban, Natalie; Devillers, Rodolphe; Singh, Gerald; Claudet, Joachim, 2020, "Data from: How far have we come? A review of MPA network performance indicators in reaching qualitative elements of Aichi Target 11", <https://doi.org/10.5683/SP2/FPYIVC>, Scholars Portal Dataverse, V1, UNF:6:RpVEXjI4EVf/w858NYQ3Gw== [fileUNF]

This dataset is based on a systematic literature review to identify indicators used to assess MPA network effectiveness in achieving the qualitative elements. Web of Science core collection database (1900 to April 2019) and Elsevier's Scopus database (1995 to April 2019) has been...

Staging path:

/var/archivematica/storage\_servi  
ce

Absolute path to a staging area. Must be UNIX filesystem compatible, preferably on the same filesystem as the path.

Host:

dataverse.scholarsportal.info

Hostname of the Dataverse instance. Eg. apitest.dataverse.org

API key:

2a4c8992-e710-4c56-ba0c-df12!

API key for Dataverse instance. Eg. b84d6b87-7b1e-4a30-a374-87191dbbbe2d

Agent name:

University of Victoria Dataverse

Agent name for premis:agentName in Archivematica

Agent type:

Dataverse

Agent type for premis:agentType in Archivematica

Agent identifier:

https://dataverse.scholarsportal.i

URI agent identifier for premis:agentIdentifierValue in Archivematica

Edit Space

Standard



Transfer type

Transfer name

Accession no.

Access system ID

Browse

Start transfer

Approve automatically

Transfer

UUID

Transfer start time



Dataverse



Transfer type

Transfer name

Accession no.

Access system ID

Browse

Start transfer

Approve automatically

The University of Victoria Dataverse



- 72612 (2014 National Survey of Canadian Community Pharmacists: Use of Digital Health Technologies in Practice)
- 72603 (2014 National Survey of Canadian Nurses)
- 72614 (2014 National Survey of Caregivers)
- 72621 (2016 National Survey of Community-Based Pharmacists: Use of Digital Health Technology in Practice)
- 72617 (2017 National Survey of Canadian Nurses: Use of Digital Health Technology in Practice)
- 75038 (2018 Canadian Physician Survey)
- 105311 (Algal lipids in the Canadian arctic)
- 74061 (B.C. Ministry of Environment Environmental Quality Information System (EQUIS) Database)
- 71772 (Compiling and mapping global permeability of the unconsolidated and consolidated Earth: GLobal HYdrogeology MaPS 2.0 (GLHYMPS 2.0). [Supporting Data])
- 77377 (Data from: A climate-associated multi-species cryptic genetic cline in the northwest Atlantic")
- 93020 (Data from: An hypoxia-tolerant flatfish: Consequences of sustained stress on the slender sole *Lyopsetta exilis* (Pleuronectidae) in the context of a changing ocean)
- 96763 (Data from: Biodiversity and habitat assessment of coastal benthic communities in a sub-Arctic industrial harbour area)
- 72090 (Data from: Composition and functional diversity of macrofaunal assemblages on vertical walls of a deep northeast Pacific fjord)
- 111204 (Data from: Continental margin sediments underlying the NE Pacific oxygen minimum zone are a source of nitrous oxide to the water column)
- 93083 (Data from: Food and initial size influence overwinter survival and condition of a juvenile marine fish (age-0 Atlantic cod))

Add

Transfer

UUID

Transfer start time



# Test\_Dataset\_for\_UVic\_Dataverse Archival Information Package

<b>UUID</b>	cd64bee3-0fe8-47c1-bd95-ae36956f1f59	
<b>Size</b>	12.03 MB	
<b>Date stored</b>	2020-05-12 17:57	
<b>Status</b>	Stored	
<b>Encrypted</b>	False	
<b>Location</b>	<input type="button" value="Download"/>	[...] / Test_Dataset_for_UVic_Dataverse-cd64bee3-0fe8-47c1-bd95-ae36956f1f59.7z
<b>Pointer file</b>	<input type="button" value="View"/>	

## Actions

### Metadata-only DIP upload to Atom

Insert slug \*

Only compatible with Atom 2.4 or newer.

# Groupe d'experts de la préservation – Portage.

- Recommandations pour une stratégie nationale de préservation

Ce n'est pas juste une question de préservation !

# AUTRES AVANTAGES

# Encourager les investissements dans l'infrastructure

- À quand remonte la dernière fois que vous avez obtenu une subvention pour améliorer votre dépôt ?

## **CANARIE Awards up to \$3.2M in Funding to 9 Research Teams to Develop Research Data Management Software Tools**

Posted on [November 27, 2018](#) in [Press Releases](#)

The ability to reuse research data helps accelerate discovery, allows for reproducibility of scientific results, and maximizes return on investment of research funding



**Funded software tools aim to strengthen  
Canada's capacity to manage research data**

# Bâtir la confiance avec les communautés de recherche

- « Les dépôts doivent gagner la confiance des communautés qu'ils veulent desservir et prouver qu'ils sont en mesure de gérer adéquatement les données qui leurs sont confiées ».
  - Lin, D., Crabtree, J., Dillo, I. *et al.* The TRUST Principles for digital repositories. *Sci Data* **7**, 144 (2020).  
<https://doi.org/10.1038/s41597-020-0486-7>



## Core Certified Repositories

[Home](#) > [Why certification](#) > [Core Certified Repositories](#)

Applications are made public only once certification of a data repository has been approved by the CoreTrustSeal Board. Certification is against the version of the Core Trustworthy Data Repositories Requirements named in the link to the public application (e.g., 2017–2019). The CoreTrustSeal for Data Repositories is valid for three years from the certification date listed within the public application.



# Avoir plus d'agilité en combinant les ressources de développeurs

- « Il faut se préparer pour l'évolution rapide de la technologie pour répondre aux besoins des communications savantes. L'agilité est essentielle ».
  - Arlitsch, K., & Grant, C. (2018). Why so many repositories? Examining the limitations and possibilities of the institutional repositories landscape. *Journal of Library Administration*, 58(3), 264-281.

## Et encore....

- Amélioration de la découverte avec des métadonnées plus uniformes
- Intégration avec d'autres systèmes (p.ex. graphique PID, ORCHID, etc.)
- Amplification de l'impact par l'effet réseau

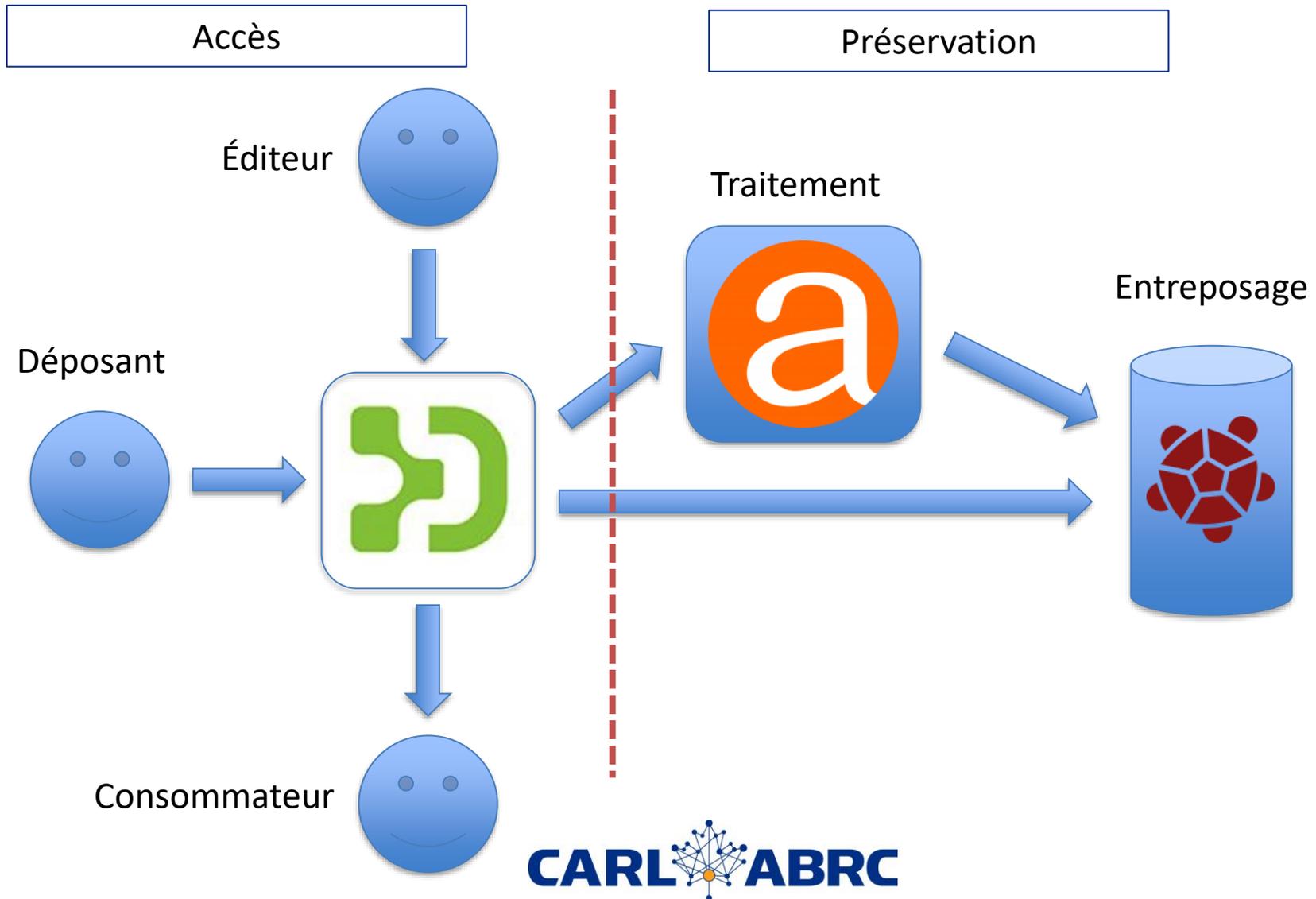
Réfléchir au-delà des dépôts

# REPENSER L'ACCÈS ET LA PRÉSERVATION

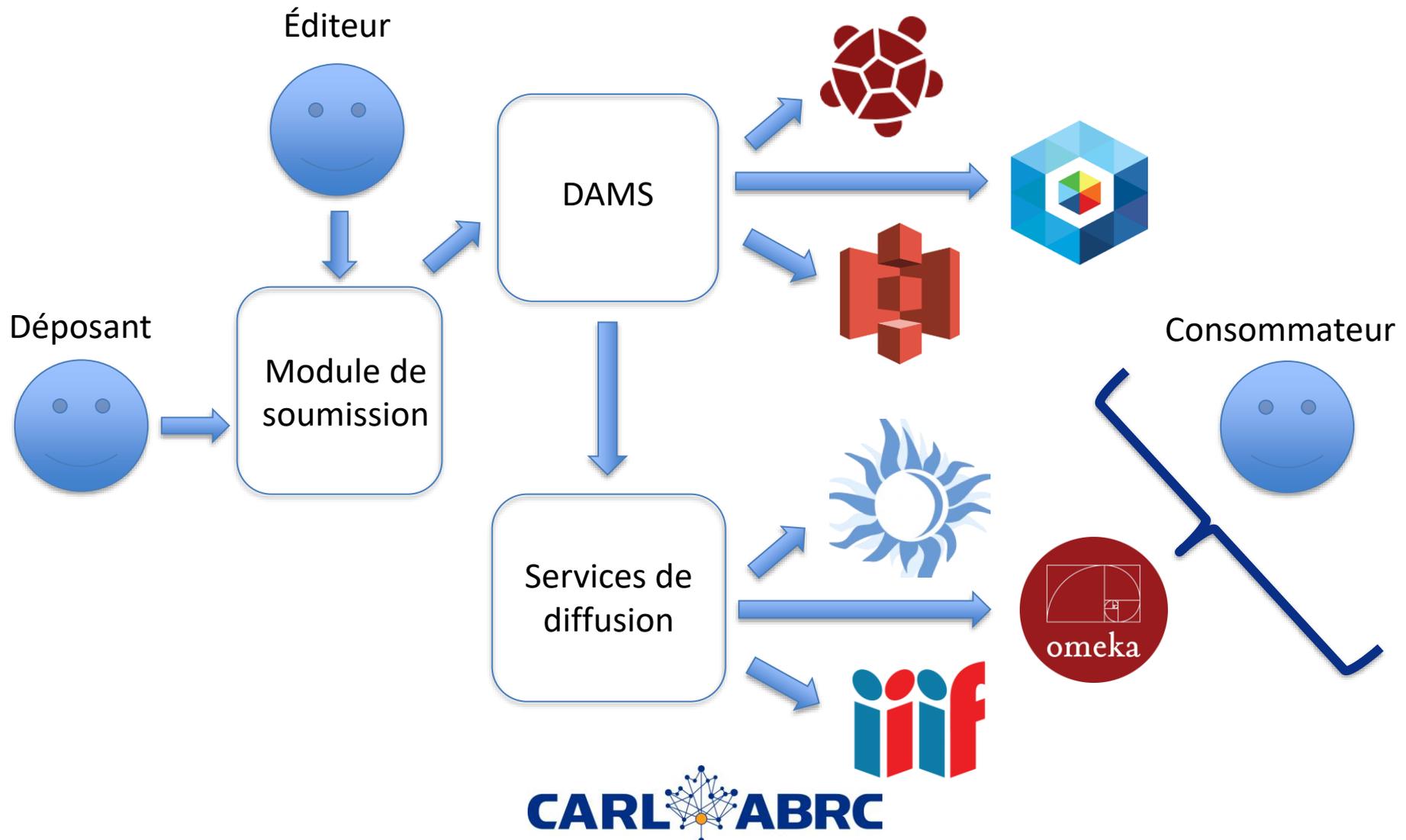
# À l'instar de U of T, Scholars Portal, UBC et d'autres intervenants...

- Et si nous repensions nos processus pour donner la priorité à la préservation !

# Processus d'un dépôt traditionnel de préservation



# Réfléchir au-delà du dépôt



# Open Archival Information System (Système ouvert d'archivage d'information)

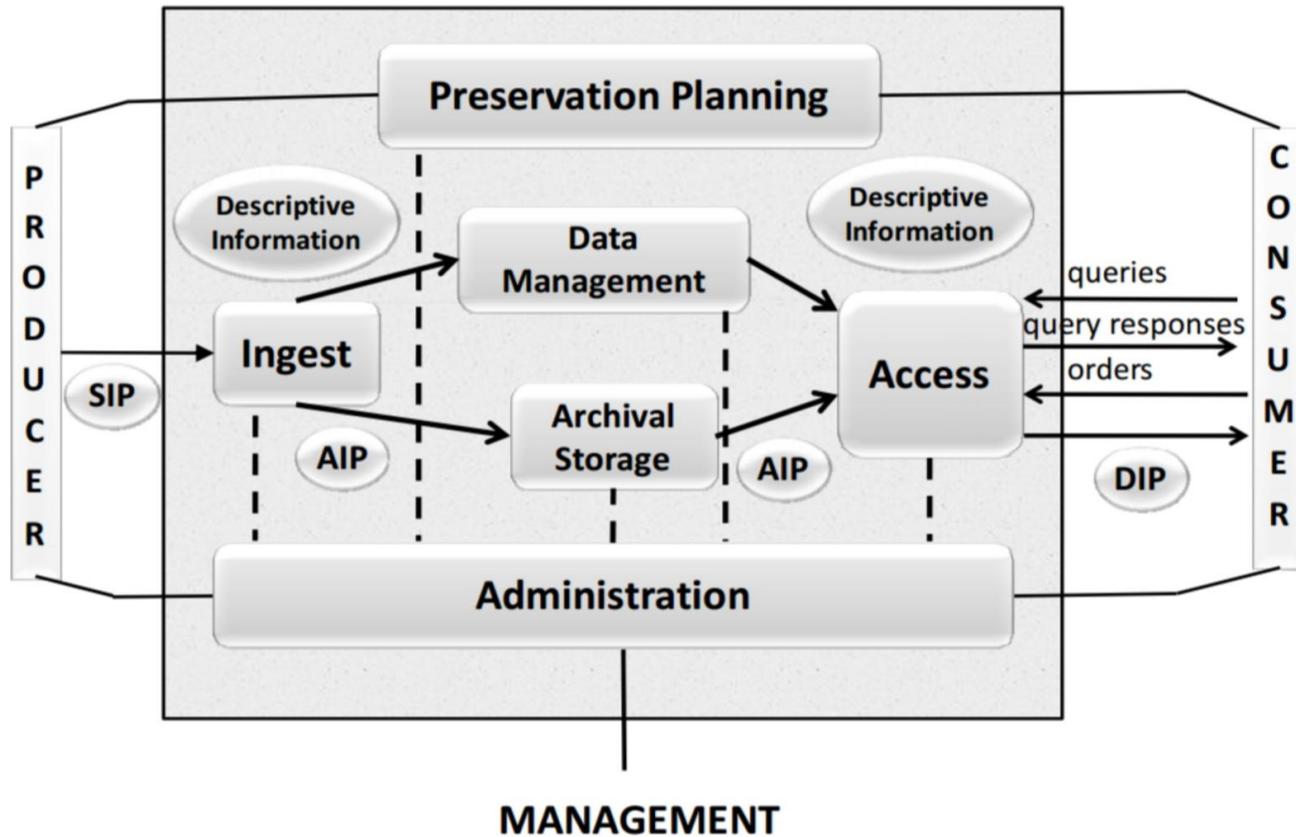


Figure 4-1: OAIS Functional Entities

# L'avantage archivistique

- Les bibliothèques ont toujours privilégié l'accès
- Les archives ont intégré la préservation à l'accès (elles ont beaucoup d'avance sur les bibliothèques en matière de préservation numérique)
- L'état d'esprit des bibliothèques privilégie les dépôts aux dépens de la préservation

# OAIS

- OAIS offre une approche à l'accès qui est davantage axée sur les services, les chercheurs et la préservation
- Importance pour les bibliothécaires et les archivistes de travailler ensemble

# DISCUSSION

# Questions et commentaires

- Diversité ou mise à l'échelle : quel est le bon niveau ? Établissement, région, nation ?
- Dans quelle mesure les besoins des dépôts locaux sont-ils uniques ?
- Quels sont les coûts de transfert ?

# Merci !

- [coreyd@uvic.ca](mailto:coreyd@uvic.ca)