

[Skip to content.](#)



- [More...](#)
- [Taporize](#)
- [Attach a file](#)
- [Edit this page](#)

## TADA Contents

[Main](#) > [Projects](#) > [ProblemOverview](#) > BestPractisesDigitalArchive

- [About TADA](#)
- [About Text Analysis](#)
- [Resources for Devs](#)
- [Projects](#)
- [News & Events](#)
- [Publications](#)

**Please Note:** This Page Has Been Archived

Archiving the Compendium was a project to figure out how to deposit an archive of the Globalization Compendium. These pages reflect our thinking in 2008, **not necessarily** how we finally implemented the archive. This and other pages were archived so changing them won't affect the project.

## Find topic

[Search text](#)

## Web tools

- [Recent changes](#)
- [Topic list](#)
- [Verbose topic list](#)
- [Access statistics](#)
- [Notify me of changes](#)
- [Web preferences](#)

- [Best Practices for Archiving a Digital Project](#)
  - [SSHRC Requirements](#)
  - [Existing Literature](#)
  - [Projects](#)
  - [Practises](#)
  - [Recommendations](#)

## Best Practices for Archiving a Digital Project

The **MCRI Globalization & Autonomy** project reaches its objectives and goals, it is necessary to address the means by which the online research product portion of the project can be maintained for posterity. The results of research were presented in three ways. The first is the **Globalization and Autonomy Online Compendium**. Results were also published through the Globalization and Autonomy Series and by individual team members as part of their work in journals and books. To ensure continued and future access to the online research, policy and procedures to provide long-term access must be developed. This raises questions as to best practises to maintain access to online, digital and dynamic research products, either through conversion to other media and/or maintenance of the existing media in the long term.

Ultimately, continued access itself is the crucial aspect of preservation...by ensuring access and maintaining currency the value of the repository and the research products is perpetuated and the value to the wider research community demands attention and re-evaluation of hosting and location of repository issues.

(Do we have an idea of how many people are consulting/using the compendium on an ongoing basis? Are there any metrics compiled?)

The project, specifically through the compendium, has maintained an open access policy. "The Globalization and Autonomy Online Compendium provides open access to all of its content on the principle that making research freely available to the public supports a greater global knowledge exchange. Such access is associated with

## Help

- [Text formatting rules](#)
- [TWiki documentation](#)
- [Sandbox \(test\) web](#)

## Tools

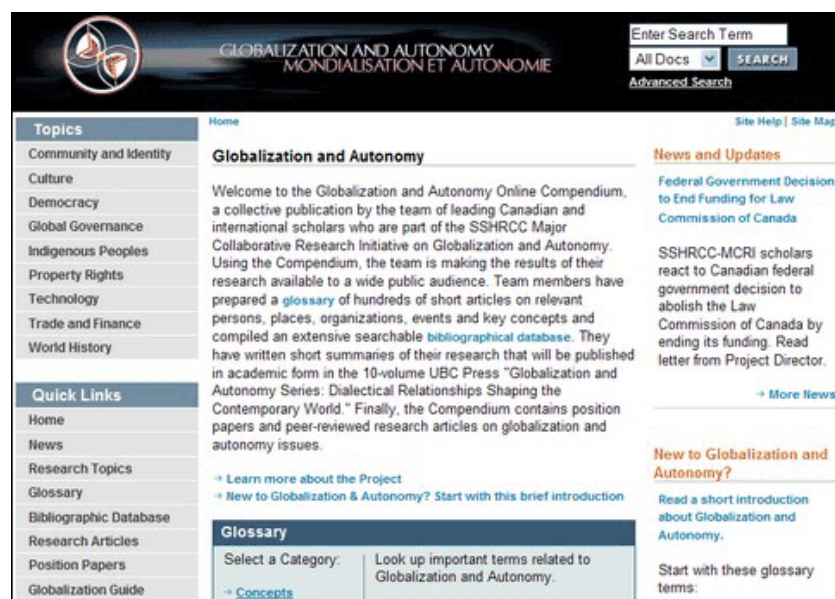
Analysis Tool Bar +

increased readership and citation levels. The Compendium uses open source software, adapted and extended by the [TAPoR](#) project, to help make open access economically viable, as well as to improve the scholarly and public quality of research."

(Question: this being the case, and although authors retain rights to the contributions, what are the legal implications of what we choose to do with the archived materials - what if par example we decided to go with a proliferation archiving project? Are there aspects of the signed agreements that relate to archiving or redistribution? How much control must be maintained to satisfy the terms and conditions of the compendium vis-a-vis its contributors?)

## SSHRC Requirements

Research funded under a SSHRC MCRI is subject to a [data archiving policy](#). This policy states that research data collected with public funds belong in the public domain and must be made available to other researchers. SSHRC policy stipulates that this must be accomplished within a reasonable amount of time – defined as being within two years of the conclusion of the project. Specific form that this data must take, nor the means of access are not addressed in the policy. The policy does state that research product should be placed in the library of the institution at which the research was carried out and failing that, it recommends libraries that they have certified to be able to host this data, should the library at the institution where the research has been carried out be unable to make this data publicly available. Costs associated with preparing this data for archiving are considered expenses under the terms of the SSHRC grant. The statement of SSHRC objectives is meant to be carried out in spirit of their intentions and to also conform to their privacy policy for research subjects.



## Existing Literature

Literature dealing with archival of digital projects is limited. Existing works pertain more specifically to digital objects as opposed to the codebases developed as part of a research project. Although these materials provide information relating to the media by

which digital items may best be preserved for perpetuity, they generally do not address issues such as code standards and the perpetuation of dynamic data. This would be one the crucial aspects of the digital archiving of the Globalisation Compendium. The compendium benefits from adherence to TEI standards for text markup. As a result, the text repository is itself inherently manipulable as a multifaceted digital object.

Available sources on digital project archiving include:

## Projects

[netpreserve.org](http://netpreserve.org)

In July 2003 the national libraries of Australia, Canada, Denmark, Finland, France, Iceland, Italy, Norway, Sweden, The British Library (UK), The Library of Congress (USA) and the Internet Archive (USA) acknowledged the importance of international collaboration for preserving Internet content for future generations. This group of 12 institutions chartered the IIPC to fund and participate in projects and working groups to accomplish the Consortium's goals. The initial agreement was in effect for three years, during which time the membership was limited to the charter institutions. Since then, membership has expanded to include additional libraries, archives, museums and cultural heritage institutions involved in Web archiving.

### NDIIPP National Digital Infrastructure and Information Preservation Program

The [NDIIP](#) is a collaborative effort of the U.S. Congress and the Library of Congress begun in 2000 to develop a national strategy for digital preservation and to draft policies relating to standards and technologies. [Web at Risk](#) [MetaArchive Strategy](#)

### JISC

[Joint Info Systems Committee \(JISC\)](#) was established in 1993 to provide a national strategy for the development of networking and specialist information services in the public sector for England, Scotland and Wales. (Interestingly - at least to me - is that JISC's website is very similar to the compendium in regard the type of materials served, the audience and in its overriding intention - I wonder how they approach the long-term maintainance and access???)

### LIFE

The [LIFE](#) project

### eSPIDA

the [eSPIDA](#) project

Specifically Canadian:

### [Canadian Digital Information Strategy](#)

- [Mapping the Current Situation in Canada](#)
- [Canadian Review of International Landscape](#)
- [Government of Canada WebArchive](#)
- [Alouette Open Digitisation Initiative](#)

- [SYNERGIES](#)

SYNERGIES seemed at first like an opportunity, but the initiatives seems to be single faceted...cataloging a collection and making it available merely by reference from its site. Hosting is the responsibility of the host institution.

- [CASPAR](#)

### CEDARS

The [CEDARS Project](#)

### AGORA

The [AGORA Project](#)

### Variable Media Network

[Variable Media Network](#) is

**encoded** "To say that a work is encoded implies that part or all of it is written in computer code or some other language that requires interpretation. In the case of works with nondigital components, this code can sometimes be archived separately from the work itself." [Purpose of VMN](#)

### Rhizome

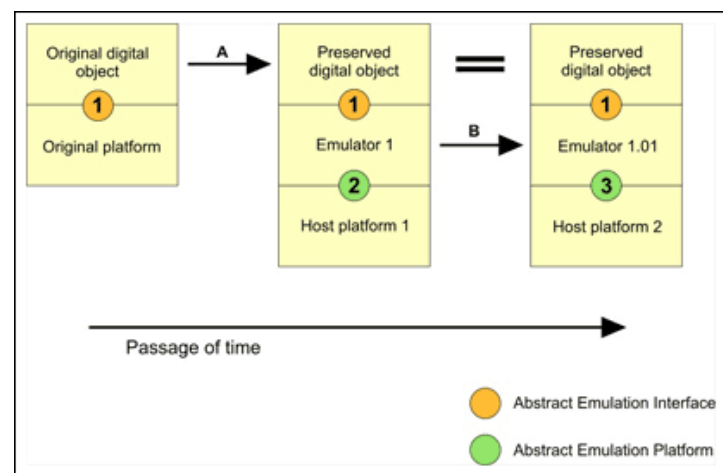
The [Rhizome project](#)

### Camileon

[Camileon Project](#)

### Citations

## Practises



## Technical Preservation

### OAIS [Open Archive Information Systems - ISO](#)

- [Resources](#)

### Code Repository

What is the long term value of code??

### Code Documentation

### Process Documentation

### Storage Location

### Storage Media

- Paper

### [Good General Archival Tips](#)

- - Value
  - Risk/Pro/Con
- DVD/CD/Dig Media

Basically, the expected longevity of dye-based DVD-R and DVD+R is anywhere from 20 to 100 years, the same as CD-R discs. Some dye formulations such as phthalocyanine and azo last longer, more than a 100 years compared to 15 to 40 years for less stable dyes. Re-writeable DVD formats like DVD-RAM, DVD-RW, and DVD+RW, which use metallic materials that change the phase of the light, rather than light-sensitive dyes, have an expected lifetime of anywhere between 25 to 100 years. It's important to remember that the quality of materials and production process can greatly affect the longevity of DVDs. Poor quality DVDs may deteriorate within a few years and produce errors or become unreadable within a year time.

There is no big difference between DVDs and CDs when it comes to materials. The dye chemicals used in write-once DVDs are similar to CD-R, though recording density and disk construction differ. So you are taking no greater risk when using DVDs instead of the old CDs.

When talking about "short" lifespan of optical media we need to take look at the alternatives as well: Magnetic media like tapes and diskettes last 10 to 30 years, acid-neutral paper can last about 100 years or longer, high quality microfilm is believed to last about 300 years or more. So it's obvious to see that CDs and DVDs alike have a relative long lifespan compared to other possible storage methods. It's also important to remember that digital storage media often becomes obsolete within 20 years, long before it physically deteriorates.

Recommendations from John [MacDonough?](#) and Tim Keefe at ther DHO Summer School: Early Writable DVDs are failing within 5 years. Must use Archival quality DVDs (5-10x price) never write on DVD surface...it will bleed through ... if you must, write on the clear spindle area Best recommendation is to get DVD's with printable surface...they have layers of protection Use major manufactureres and get certified and

tested. TDK, Verbatim, Matsui, - Unesco report on CDs

Source: <http://www.techmount.com/index.php/20051205/dvd-lifespan/>

Additional: [Life Span of DVDs](#) [Breaking Point](#)

- - Value
  - Risk/Pro/Con
- Server Imaging
  - Value
  - Risk/Pro/Con

## **Emulation**

[MAME](#) [MESS](#)

## **Maintaining Live**

A rudimentary interface was provided for editorial entry and population of the compendium. Once fully populated, this functionality no longer required support and hosting a live version of the compendium is less complex than is currently done.

## **Migration**

### **Ability to Resurrect**

## **Recommendations**

Most of this should be contingent upon the amount of cost and effort demanded by each alternative versus the potential research gains that each returns. (Has there been any measure of the value returned through the online presentation of this research?)

- Preservation by proliferation

Make everything widely available, not subject to rights management and

- Cooperative venture with Mills library

First choice. What sort of facilities exist within the McMaster library system to host the compendium in a live state?

-- [ShawnDay](#) - 27 Apr 2007

- [codePolicy.pdf](#)
- [BestPractisesGuidev0.3.pdf](#):
- [Christensen.pdf](#):
- [final-reportdigitalarchiving.pdf](#):
- [final-report.pdf](#):
- [news56.pdf](#):

- [ResponsibilityforDigitalArchivingandLongTermAccessto...pdf:](#)

---

Use this box to quickly add a comment to the page.

Add comment

Commenting is locked out by GeoffreyRockwell for at least 60 more minutes

[more options...](#)