chin.gc.ca

MORE THAN JUST BACKUPS: PROTECTING DIGITAL ASSETS IN YOUR MUSEUM

AMNB Conference – November 7, 2013

Madeleine Lafaille
Canadian Heritage Information Network

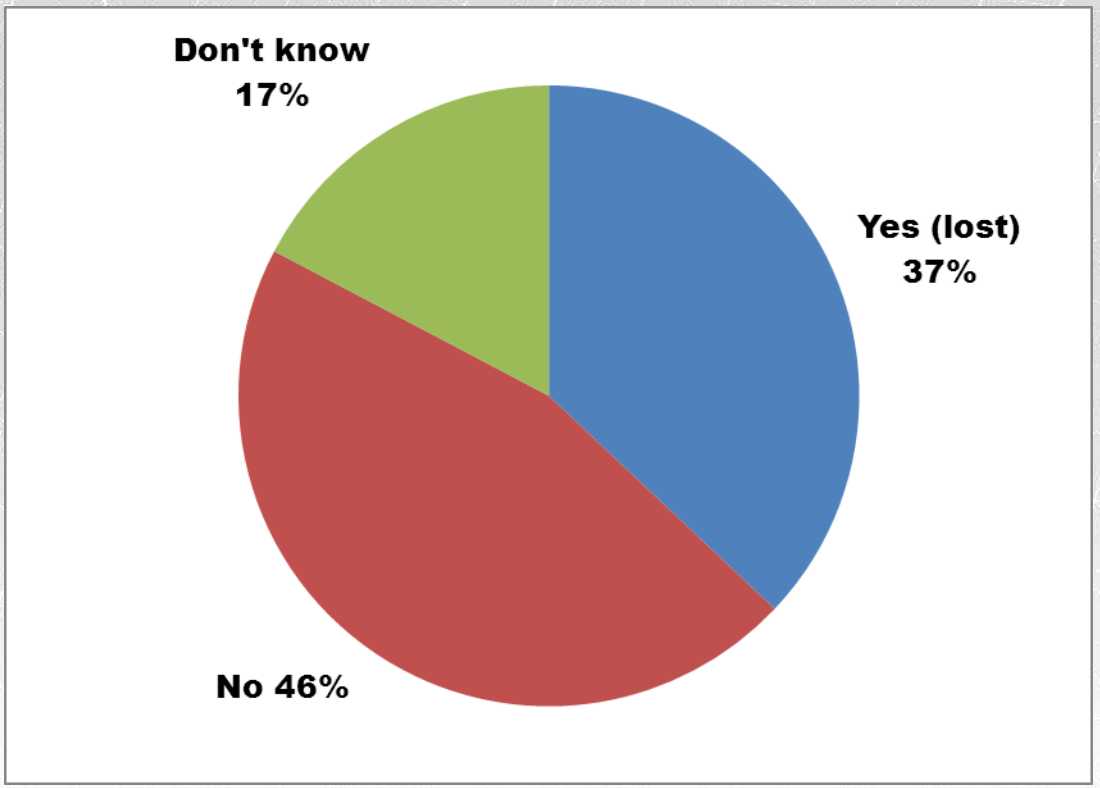
Overview

Digital Preservation in Museums

- CHIN Research and Survey
- What is Digital Preservation?
- Why should museums care?
- A Digital Preservation Toolkit

CHIN's Survey, 2011

Has your organization ever physically lost, or lost access to, any digital assets?



Question 11, CHIN's Survey, 2011

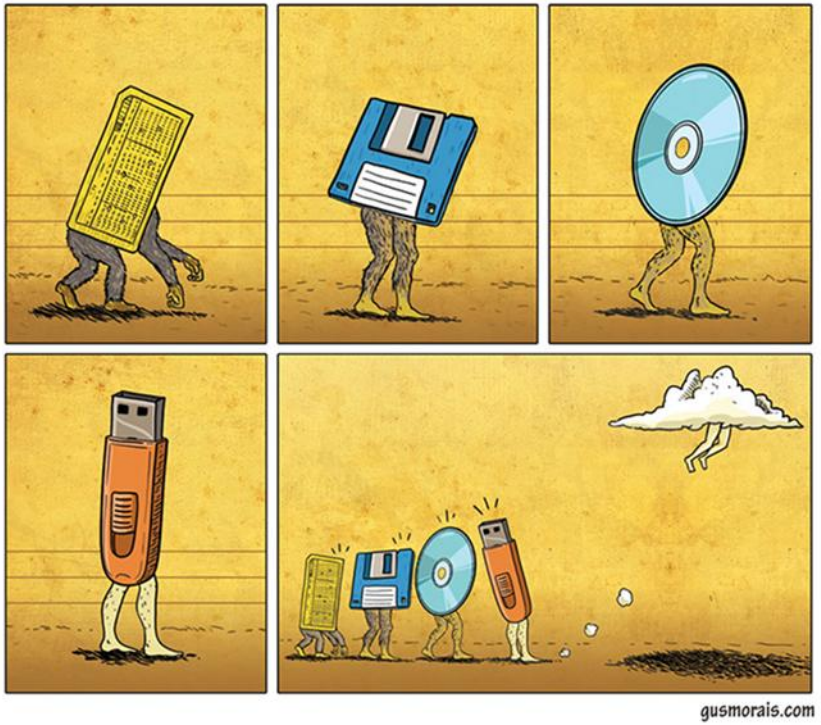
What is it?

- Digital Asset :
  - A broad term for any resource that is in a digital format
- Digital Preservation :
  - Term applied to all guidelines and actions required to insure long term preservation and access to existing trustworthy digital resources

Where is my stuff?

Information created and stored in digital form has advantages…

… And disadvantages!



Digital Assets Commonly Found in Museums

- Administrative Materials
- Records of a museum's physical holdings
- Born-Digital Materials

Why is Digital Preservation Important to Museums

Digital Preservation mitigates problems such as:

- Changes to File Format
- Operating System Obsolescence
- Software Obsolescence
- Hardware Obsolescence
- Media Degradation
- Threats to Storage Location
- Timely Access to Stored Assets
- Control of Access to Stored Assets
- Provenance & Copyright
- Usage Policies

Where should you start?

Asking questions such as:

- How is digital material likely to be used in the future?
- Who should have access to preserved material?
- Who is expected to preserve and provide access to the material in the long run?
- What are the costs of properly preserving your museum's digital assets?

A Holistic Approach

Digital Preservation is :

A holistic approach for long term life expectancy and access to digital resources, with respect to their integrity and applicable rights

CHIN Digital Preservation Toolkit

A suite of documents that offer concrete steps to identify:

- digital material found in your institution
- the potential risk and impact of lost material, and
- how to get started in the development of preservation policies, plans and procedures

Designed for museum professionals who may be aware of digital preservation, but who are not experts in the field.

1. Inventory Template

Inventory Template Section A: Summary of Digital Asset Groups **Identify all groups of digital assets held by your museum. Include a brief description for each, as well as the approximate number of assets in each group, and the approximate amount of file space required to store the entire group.**				
Name of Digital Asset Group	Brief Description of Group (i.e. what is the group used for, how does it differ from similar identified groups).	Approximate Number of Digital Assets in the Group	Approximate Amount of File Space Required to Store Group	Minimum Number of Copies of Assets in this Group (if multiple copies are kept)
(add rows as required)				

2. Policy Framework Guidelines

These guidelines help museums and archives develop a policy on what digital materials will be preserved and under what conditions.

- Outlines all the components of a digital preservation Policy.
- Ensures that it will conform with existing standards for digital preservation.

3. Decision Trees

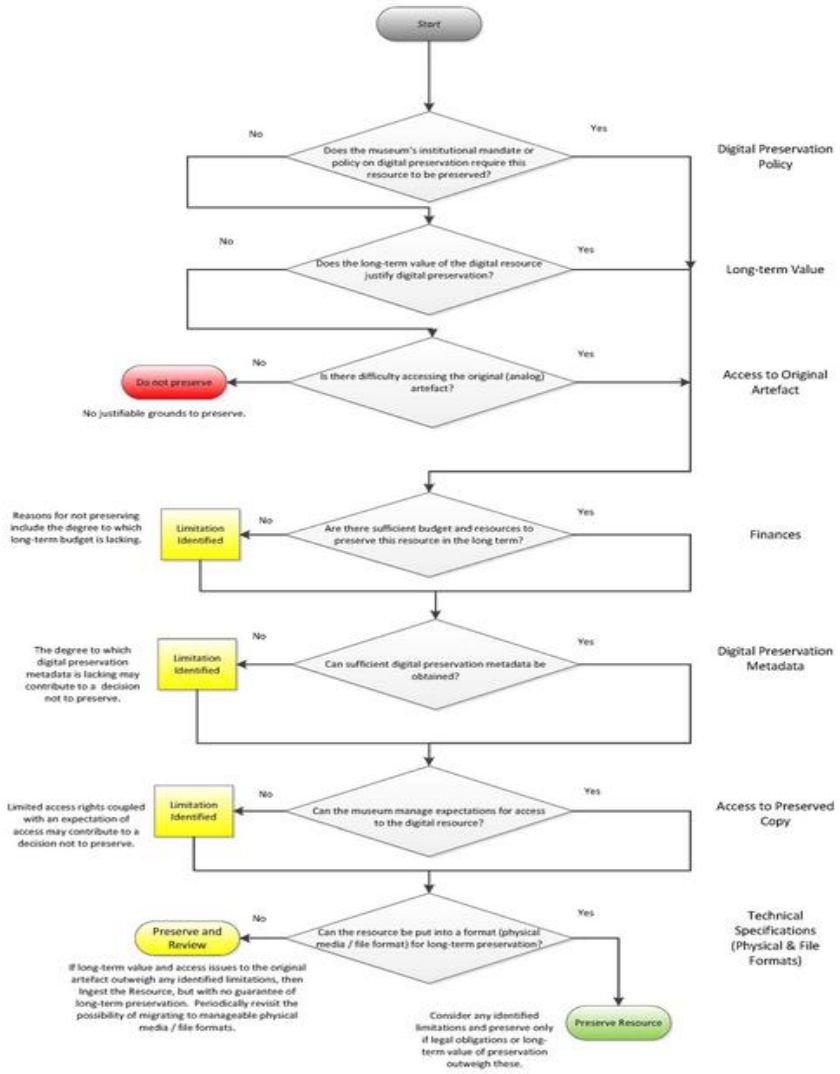
Help you **decide**, at a glance, whether a digital resource should be preserved, and under what conditions.

- A simple, high-level map of all issues that should be considered
- Can be used at policy development stage

Three Decision Trees are available:

- Digital Copies of Physical Objects
- Born Digital Artefacts
- Administrative Digital Resources

Decision Tree for Digital Copies / Records of Existing Physical Artefacts
Should an Existing Digital Resource be Preserved?



4. Digital Preservation Plan

A framework that helps museums develop an **action plan** to implement their policy.

Examples of framework items include:

- Organizational Preparedness
- Cost Considerations
- Preservation Standards
- Technical Considerations

5. InterPARES Creator Guidelines

As Accessibility

1. Select hardware, software and file formats that offer the best hope for ensuring that digital materials will remain easily accessible over time.

Accessing digital materials depends on having the appropriate software. Software that is not compatible with previous versions (backward compatibility) or with future versions (forward compatibility) makes it difficult to access records over time. Software for one application also needs to work well with that of other applications and systems (interoperability). Paying attention to the following six factors can help ensure that your software and hardware maintain accessibility.

A. Choose software that presents materials as they originally appeared. Ideally, materials should keep the same look over time to be fully intelligible and accessible. Be sure that new software will be able to read your older materials in the software format in which you kept it and display it on the screen in the same documentary form in which it was originally displayed. In other words, new software should be backward compatible with older software.

B. Choose software and hardware that allow you to share digital materials easily. Software should be able to accept and output files in a number of different formats. The ability to interact easily with other technology is called **interoperability**. It will make it easier to access your materials and also to move them to other systems.

C. Use software that adheres to standards. This is one of the best things you can do to ensure your material will last. Standards endorsed by national and international organizations are best. These are called *de jure* standards. If these do not exist for your material, you can help ensure longevity by adopting software that is very widely used. In the absence of an official standard, such software is often referred to as a *de facto* standard. Open source software, that is, freely available non-proprietary software, is preferable (see subsection G on the next page).

<< De Jure Standard >>

Standard adopted by official standards-setting body, whether national (e.g., ANSI), multi-national (e.g., CEN) or international (e.g., ISO). For computer file formats, two recent de jure standards are PDF/A (PDF standard for archiving) and ODF (OASIS OpenDocument Format).

<< De Facto Standard >>

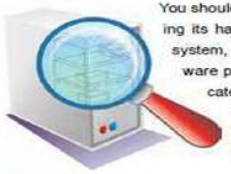
Standard not adopted by any official standards-setting body, but nevertheless widely used and recognized by its users as a standard. Well known and widely used computer file formats that are considered de jure standards include PDF, TIFF, DOC and ZIP.

As Accessibility

D. Keep the specifications of software. This kind of documentation (e.g. the owner's manuals or any other more detailed description of the software you might have) will be essential in the future to access the materials or to migrate them to a new computer environment as technology advances. It is particularly important to fully document any software that you build yourself.

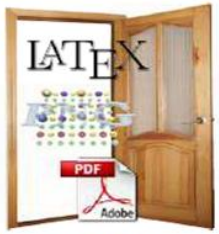
E. If you customize software, make sure you document the changes you make. Give detailed information about the changes and describe clearly the characteristics and features of the material these changes produce, as well as the outcomes you are trying to achieve by customizing the software. A good way to do this is to include the information as comments in the software code. The information will not get lost, as it is part of the file, and it will be very helpful to those who need to make adjustments later, as technology advances.

F. Document the construction of your system as a whole to help ensure its accessibility.



You should document your system's structure and functions. This means identifying its hardware and software components, including peripherals, its operating system, and software packages. Such documentation will identify how the software packages represent information, and how they process it and communicate it to each other and to users. These basic specifications will ensure that those who come after you understand the context in which you are working now. They will provide the information necessary to update the system as hardware and software evolve.

G. Choose widely-used, non-proprietary, platform-independent, uncompressed formats with freely available specifications where possible. These are often called "open formats," which means that their specification is published and freely available. However, it may also mean that the format is free of patent or royalty fees or the possibility of such fees being applied in the future, and/or that it is widely adopted. It should be noted that "open" formats are not necessarily the same as formats produced by open source software, as the latter term describes software for which the code is made freely available and can be modified. Open source software does not always produce non-proprietary formats. Distinguish between file formats, wrapper (or container) formats, and tagged formats such as XML-tagged files, and ensure that version, encoding and other characteristics are clear and fully specified. For XML files, make sure that the files are well-formed and valid and accompanied by the relevant DTDs or schemas. If it is not convenient for you to follow this recommendation, consult with an archives that accepts digital materials and choose among the formats that it recommends for long-term preservation. You should not compress your digital materials, if at all possible, since this can lead to problems for their long-term preservation. If you need to compress them, choose lossless compression techniques that conform to accepted international standards.



6. InterPARES Preserver Guidelines

3.3. Keep the oldest available logical format. The logical format in which the records were originally created, or in which they are held by the creator at the time of transfer should, whenever feasible, be maintained by the preserver, in addition to any preservation or reference copies generated after the transfer. Should selected preservation strategies, such as a specific conversion path, fail over time, continued custody of the initial logical format will allow the preserver to essentially re-start the preservation process with the most authoritative copy of the records, by applying a different preservation strategy to the records. Over the long periods during which preservers hold records, experience may show that other preservation strategies are more stable over time, or can more easily be carried forward over the long-term. Alternately, new methods of preservation may have been developed following the acquisition and initial processing of the records.

<< Logical Format >>

The organized arrangement of data on electronic media that ensures file and data control structures are recognizable and recoverable by the host computer operating system. Two common logical formats for files and directories are ISO 9660 for CD-ROMs, and Universal Disk Format (UDF) for DVDs.

3.4. Avoid duplicates. Because of the ease of replication of digital records, the preserver must put in place procedures to ensure that digital records from a specific series are transferred by a specific creator to the preserver only once. Accurate identity information is an important first step in avoiding duplication of effort by the creator and the preserver. Also, if reference copies are provided by the preserver to the creator after the transfer of the records, they should be clearly identified and marked as such to prevent accidental re-transfer.

3.5. Document all processing. Initial processes applied during and immediately after transfer may or may not be related to preservation per se. Confirming the identity of the transferred material, checking for viruses, and confirming completeness of files tend to leave the transferred file unchanged. File conversion, renaming digital entities, and encapsulating files are more intrusive activities. In both cases, preservers must document all processing of digital records, and its effects, while they are in their custody (see Appendix B, Requirement B.2). This documentation should include information such as:

- why certain processes were applied to the records;
- what records were processed;
- the date when the process was performed;
- the names of persons performing and documenting the various steps of the process(es);
- the impact of the process performed on the records' form, content, accessibility and use; and
- the description of any damage, loss or other problems encountered as a result of the processing, including any effect on the elements expressing the records' identity and integrity.

Should the preserver produce copies of the acquired records, it is important to remember that, as discussed in Section 1.5, these copies should be produced in an environment that satisfies the relevant requirements[12] from the InterPARES 1 Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records.

4. Preserve Accessioned Records (A4.4)

Pr
Preserving

The designated records preserver is the entity responsible for taking physical and legal custody of, and preserving (i.e., protecting and ensuring continuous access to) a creator's records. Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a trusted custodian for a creator's records. The authentic copies of the creator's records are kept by the trusted custodian in a **trusted preservation system** (see Appendix C), which should include in its design a description and a retrieval system. This trusted preservation system must also have in place rules and procedures for the ongoing production of authentic copies as the existing system becomes obsolete and the technology is upgraded.

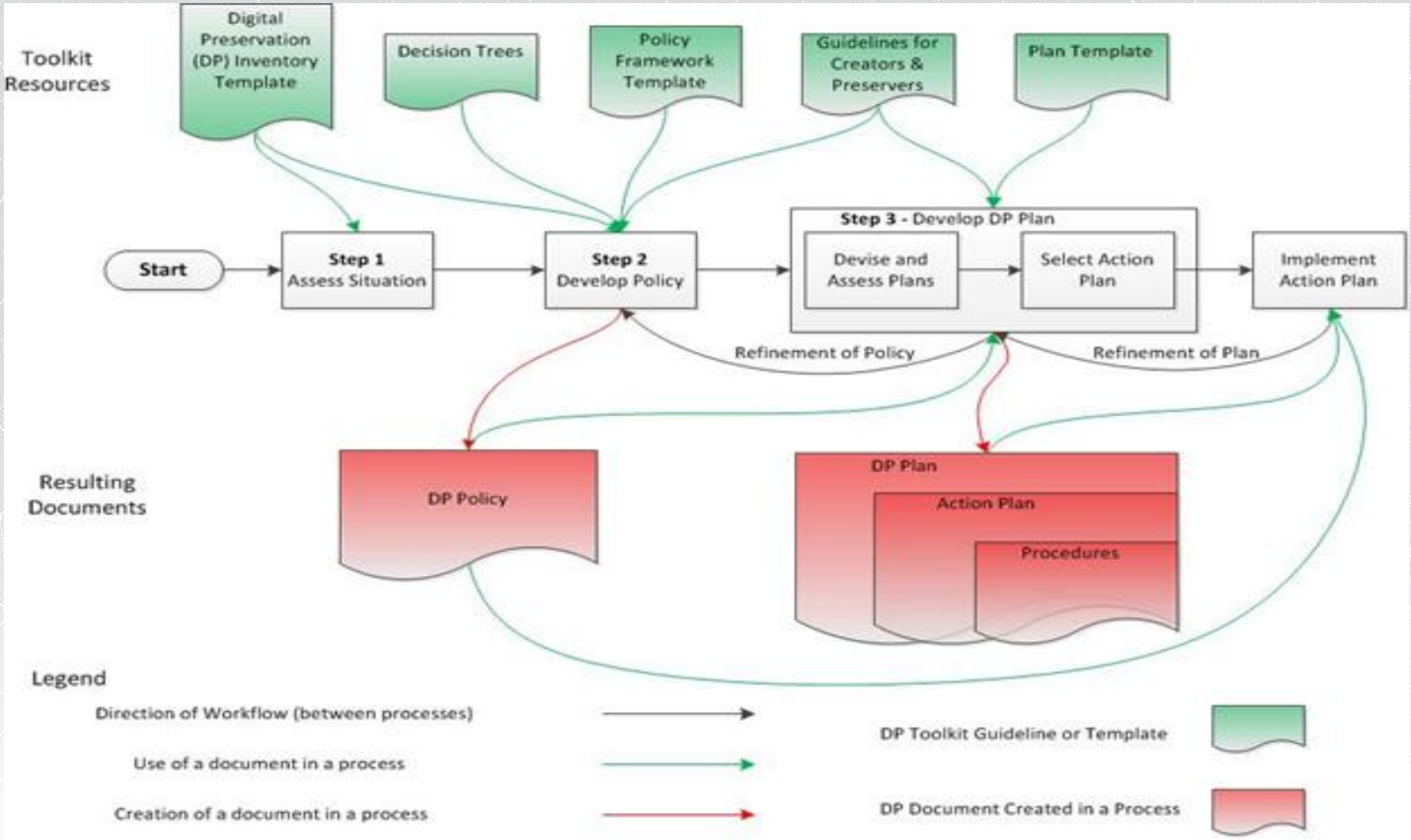
4.1. Describe the records. The information about the records and their contexts collected during the appraisal and processing stages should form part of the archival description of the fonds or series in which the records belong (see Appendix B, Requirement B.3). This should also include information about intellectual property rights or privacy concerns.

The archival description of the fonds or series containing the digital records should include—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the digital records of the creator have undergone since they were first created. The description should also include an overview of the transfer and preservation processes based on the documentation discussed in Section 3.5 and the explanation of the relationships among digital components discussed in Section 2.7.

<< Trusted Custodian >>

A preserver who can demonstrate that it has no reason to alter the preserved records or allow others to alter them and is capable of implementing all of the requirements for the authentic preservation of records.

4.2. Identify legal ramifications of preservation actions. When a preservation strategy is selected, its legal implications should be reviewed. For example, format conversion out of a proprietary environment could involve the preserver in illegal actions. In the United States, the Digital Millennium Copyright Act has made it a criminal offence to produce tools that can circumvent copyright protection measures. Internationally, the World Intellectual Property Organization Copyright Treaty (WIPO WCT) contains provisions that include copyright protection for software as well as digital works, and that introduce criminal penalties for infringement, which ranges from unauthorized copying of material placed on a Web site to the removal or alteration of rights management controls from digital works. Most software packages also include some type of similar restrictions, which users must agree to during the installation process.



Note: other data (listed in the templates) is used in the development of the Policy and Plan.

Few Other Resources

- SPECTRUM Digital Asset Management - Collections Trust
  http://www.collectionslink.org.uk/spectrum-resources/1688-spectrum-digital-asset-management

- Canadian Conservation Institute, "Electronic Media Collections Care for Small Museums and Archives"
  http://www.cci-icc.gc.ca/caringfor-prendresoindes/articles/elecmediacare/index-eng.aspx

- Trusted Digital Repositories : Attributes and Responsibilities, an RLG-OCLC Report, *www.oclc.org/programs/ourwork/past/**trusted**rep/**repositories**.pdf*

- The Open Archival Information System Reference Model Introductory Guide, Brian Lavoie, OCLC, *www.dpconline.org/docs/lavoie_**OAIS**.pdf*

CHIN's Digital Preservation Toolkit :
www.pro.rcip-chin.gc.ca/

Madeleine.Lafaille@pch.gc.ca

Questions ?

Comments ?