

# **Rights in the PREMIS Data Model**

**A Report for the Library of Congress**

**by**

**Karen Coyle**

**December 2006**

# Rights in the PREMIS Data Model

By Karen Coyle

December 2006

Preface .....	3
Executive Summary .....	4
I. Rights for Digital Preservation .....	5
Introduction .....	5
The Right to Preserve .....	5
Rights Metadata Today .....	6
The PREMIS Rights Entity .....	9
References .....	10
II. Digital Preservation Strategies and Rights Implications .....	11
Introduction .....	11
Digital preservation actions .....	12
Rights In the Law .....	15
Exceptions to Legal Rights .....	18
Mandatory or Legal Deposit .....	20
Digital Resources and Preservation .....	21
References .....	21
III. Rights Metadata for Preservation .....	25
Recommendations for the PREMIS Rights Entity .....	27
Conclusion .....	29
References .....	29
Appendix A: Preservation Actions and Rights .....	31

## Preface

The Preservation Metadata: Implementation Strategies (PREMIS) Working Group developed the *Data Dictionary for Preservation Metadata*, which is a specification containing a set of "core" preservation metadata elements that has broad applicability within the digital preservation community. It constructed a data model that defined entities involved in the preservation process and their relationships. One of the important entities in this data model is rights statements, which specify terms and conditions for using the objects in a preservation repository. The PREMIS Working Group chose to consider only rights required for preservation activities in scope for its work, rather than rights for access. Because of the ambiguity of the laws concerning intellectual property rights and the complexity in the roles that institutions play in digital preservation in relation to access, it was difficult for the Working Group to thoroughly cover all information needed about rights to preserve in the data dictionary.

In order to make progress, the Group included minimal metadata that a repository needs to know about the rights to preserve digital objects. Rights in PREMIS take the form of structured permission statements, which are defined in terms of preservation actions. The group felt that, as the laws were clarified in terms of preservation rights and permissions were better understood, that this section of the PREMIS data dictionary could be expanded.

The Library of Congress, as part of the PREMIS maintenance activity, commissioned Karen Coyle to provide this study to assist the newly established PREMIS Editorial Committee with its first revision of the data dictionary and schemas. The intention is to improve the specification so that institutions trying to assess their rights to preserve materials in digital formats will be able to provide enough information in their digital repositories to make such assessments about their materials over time. In this study Karen Coyle reviews the landscape of digital rights, analyzes various preservation rights scenarios and the sorts of preservation actions that digital repositories might take, relates copyright law to preservation actions, and provides recommendations for revision where the data dictionary needs expansion.

Sally H. McCallum

Network Development and MARC Standards Office

Library of Congress

December 2006

## Executive Summary

The PREMIS standard contains a *rights* entity that allows the association of rights with specific digital preservation actions. This paper looks at the various definitions of rights, the state of rights metadata, and surveys legislative actions taking place in many nations that will provide a legal standing for digital preservation activities.

The term "rights" is used in many contexts, in particular law and contracts. Rights that are relevant to digital preservation may be legal, such as a statutory right for certain institutions to preserve materials within their jurisdiction. They may also be contractual, in the form of a license agreement that allows all or some preservation actions.

The rights metadata as defined in the May 2005 version of the PREMIS metadata standard relies almost entirely on a view of preservation rights as explicit permissions. Such permissions can exist when there is a contract between the holders of intellectual property rights in the resource and the preservation agency. This is a common situation today in institutional repositories that are gathering, preserving, and providing access to works created by faculty and staff. However, many institutions are preserving, or intend to preserve, works for which the granting of permissions through a contract is not possible. The rights holders of these works are either not available to grant permissions, or the institution is engaging in a wide-ranging preservation activity for which gathering individual permissions is not feasible. These institutions therefore rely on law and policy as the foundation for their preservation actions.

The report suggests adding specific data elements to the PREMIS rights entity for the recording of preservation actions that have been interpreted to be permitted by law or statute. It also recommends that PREMIS metadata include at least basic data elements to record information about the copyright status of the work, when that is known. With these additions, the PREMIS rights entity will be able to support a variety of types of rights situations that are common in the digital preservation community.

# I. Rights for Digital Preservation

## Introduction

The PREMIS data model has five primary types of entity: intellectual entities, objects, events, agents, and rights. [PREMIS] The Rights entity takes the form of a structured permission statement that is linked to a digital object, presumably the object being preserved. Permissions are defined in terms of preservation actions that are described elsewhere in the PREMIS data dictionary. Each permission statement must have a granting agent.

The question posed in this document is how the PREMIS rights metadata, including the preservation actions, can be expanded and improved. To answer this question, we first must provide a picture of how law relates to preservation in general, and digital preservation specifically. We also have to look at a growing body of rights-related metadata efforts.

## The Right to Preserve

Certain cultural heritage institutions have a mandate encoded in national or local law to preserve the intellectual history of their community for future generations. In this situation, the right to preserve is defined by law and will often pertain to a wide range of materials. No specific permission will be required from rights owners because the law generally requires deposit by those owners and defines the role of the receiving institution in terms of its curatorial role over the deposited resources. For institutions working under this scenario, the key rights information is: the identification of the rights holder, the date of entry of the materials into the repository, and the version of the law that was in force at that time.

A more limited model for certain cultural institutions is found in the U.S. copyright law (USC 17, 108), "Limitations on exclusive rights: Reproduction by libraries and archives." This exception allows cultural heritage institutions to preserve resources they own, but only under certain circumstances. In this case rights metadata would need to record the date of the preservation action and the law under which the action took place. No express permissions are required of the rights holder.

Institutions often act as their own archives, preserving the documents they produce and own. This is often the case with institutional repositories, at least for a portion of their holdings. Because they are preserving their own intellectual property, no permissions are needed. In terms of rights metadata, a statement of ownership and reference to any contracts that discuss that ownership should suffice.

In absence of a binding law or institutional ownership, a direct permission model may be appropriate. For example, in some institutions of higher education, teaching staff own the intellectual property rights to their non-classroom activities such as writing for publication. Barring a work agreement with the instructors, an institutional repository cannot preserve those works without gaining permission. Such permission requires that a valid agent, either the rights holder or someone who can legally act for the rights holder, is party to the permission. The metadata must identify the agent and express the permissions that are granted. It also needs to record the date the permission was granted, and dates of the beginning and end of the agreement, if appropriate.

The permission model works relatively well for current works, but is less viable as the date of creation moves into the past and rights holders are harder to identify and locate. This is the case with orphan works, where the rights holder is unknown or cannot be located to obtain the needed permission. It is in this situation that one may be able to rely on certain legal exemptions, such as the Fair Use exceptions in U.S. copyright law (USC 17, 106). When such an exemption applies, preservation can take place without the actual permission of the rights holder. Preservation repositories in this situation need to record the exemption under which they are asserting their right to preserve, along with the date of the decision.

The last situation is that of the public domain. Works that are in the public domain can be preserved without obtaining permission, and there are no limits to the types of modifications or transformations that can be made of the item. There will be no agent and no permission statement for these works because none is needed. Yet the information that the work is in the public domain is very valuable to the preservation repository for decisions about the item, and the preservation metadata should record this fact.

This gives us the following preservation rights scenarios:

- Legal mandate to preserve, with preservation defined by the law
- Limited legal permission to preserve, with preservation defined by the law
- Preservation of ones own materials
- Preservation based on permission obtained from the rights holder
- Preservation based on fair use/fair dealing exception
- Preservation of items in the public domain

These are therefore the rights situations that PREMIS should address. Breaking these down into data elements, the metadata must be able to express:

- The law or regulation governing the preservation mandate or right, when preservation is done under those auspices.
- Ownership of the intellectual property rights in the resource at the time of entry into the preservation repository
- The public domain status of a work.
- When preservation is based on a specific permission, the permission and its grantor

## **Rights Metadata Today**

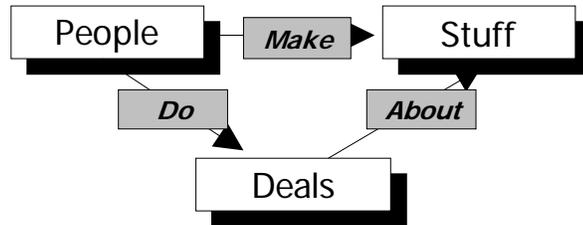
If we consider the rights work at Xerox PARC in the mid-1990's to be the beginning of modern rights metadata [COYLE], then this field is just over a decade old. Where other types of computing and coding have advanced greatly in similar time spans, rights metadata has not. There are some functioning uses of rights metadata, but for the most part this is an area where adoption has been slow.

One of the reasons for this slow uptake may be that the needs for rights metadata is more varied than was originally anticipated. There are at least four different types of rights metadata available to us or being developed today: transaction metadata, license metadata, metadata for technology protection measures, and copyright metadata.

## ***Transaction metadata***

Transaction metadata is based on the needs of e-commerce. Transaction metadata can be as conceptually simple as the exchange of a credit card number for merchandise. In general, transaction metadata is machine-actionable and transactions are machine-to-machine operations.

The basic elements of the transaction for intellectual property were expressed by Godfrey Rust as the triad: people, stuff, and deals [RUST]. This triad is combined with the verbs "make," "do" and "about" to create the following diagram:



This work comes out of the EC-funded INDECS (Interoperability of Data in E-Commerce Systems), an e-commerce based project that was tasked to investigate models for the trading of intellectual property in the networked environment. [INDECS] This INDECS model has influenced the creation of transaction metadata, such as the work on Electronic Rights Management (ERM). Other rights metadata schemes that cover transaction activities are the Open Digital Rights Language (ODRL), and the MPEG-21 (now ISO/IEC 21000-5) rights language that was derived from XrML. These are general-purpose languages for e-commerce that include metadata for technological protection measures (see below).

Transaction metadata is for materials being sold or licensed today. There is no need to refer to intellectual property law nor to take orphan works into account. The necessary assumption in transaction metadata is that an agent is available to represent the resource in the license.

## ***License Metadata***

Licenses prove to be more complex than mere transactions, as we can see in the recently issued draft of ONIX for Licensing Terms [EDITEUR]. This metadata model includes agent information (people), payment terms (deals), and three other categories not directly represented in Rust's model: supply terms, usage terms, and general terms. The supply terms describe the delivery mechanisms for the "stuff." The usage terms will cover "permitted and prohibited usages of the licensed materials." Details on usage are not included in the draft available at the time of this report. As for the general terms, this quote from the document provides a definition:

The "GeneralTerms" section covers general terms of the license that fall outside of supply, usage and payment, for example such things as liability, confidentiality or force majeure, or cannot usefully be expressed as structured XML since they are not machine-actionable. These will be covered by a combination of (a) a controlled value specifying the type of condition, and (b) a reference to and/or quotation from the text of the original license. [EDITEUR, p.2]

In this short definition we get a hint at how complex "deals" can be and how difficult it is to

express the deals in computer code. Although the ONIX work in this area is new, the model used is a common one in the area of computer software: a relatively simple transaction covers the purchase of software on disk or online, but a more complex license accompanies the purchase. This license is presented in a non-actionable, human-readable format, usually at the point of access or installation of the program or file.

In general, license metadata is *about* a license, it is not the license itself. The Creative Commons [CC] license is a clear and simple example of this: it uses simple icons to alert readers to the presence of a license, RDF-formatted code to name specific conditions of the license, with human-readable licenses that express the actual terms of the license in ways that could be used in a court of law. Although it is the goal of some technologists to develop a machine-actionable version of these licenses, in fact the licenses have a long-standing legal tradition behind them that does not lend itself to treatment as data.

### ***Metadata for Technology Protection Measures***

This is the metadata that supports what is generally known as "Digital Rights Management" or DRM. Metadata for Technology Protection Measures (hereafter referred to as MTPM) is designed to interact with programs and to express actions that can be executed within devices and over networks. MTPM must therefore be entirely machine-actionable. Two existing MTPM standards today are MPEG-21 (ISO/IEC 21000-5) and ODRL [ODRL]. They are also referred to as Rights Expression Languages (RELs) [COYLE 2004], and they can be used to support e-commerce from the transaction through the lifetime of the purchased or licensed resource. Although these are the best known of the REL standards, current products actually make use of a number of other proprietary solutions for technology protection. Probably the most common today is the Adobe PDF encryption function [ADOBE], followed closely by the iTunes DRM, Fairplay [WIKI/FAIR]. Both of these are specific to the types of files the software manages. The Adobe encryption allows the file creator to code rules for various actions such as printing the document or the filling in of forms. The iTunes DRM permits limited numbers of copies to be made to other devices. Where the standard MTPMs are used, they are used in limited implications, such as the version of MPEG-21 used in the Windows Media Format, and the Open Mobile Alliance's use of an ODRL-based TPM for its mobile phone products. [OMA]

These technologies define specific permissions relating to use of the digital resource. The actual uses found in working systems such as Adobe and iTunes are directly related to actions that can be performed on computer systems today, such as copying to the clipboard, modifying the file, printing or burning to CD. The MPEG model, using its REL and the Rights Data Dictionary that is under development, would allow the creation of highly complex permissions compared to those in use today. ODRL also allows very complex statements to be expressed, although using a much less abstract model for the development of new applications. In theory one could develop preservation-related permission statements using these rights languages, but to date the standards have focused on current e-commerce applications and do not have specific language for preservation actions.

## **Copyright Metadata**

None of the above rights metadata schemes record information about the copyright status of the work. This is not an oversight on their part; the use of licenses and technology protection supersedes copyright law in most interactions. Copyright law comes into play with any public use of materials with an intellectual property interest, but is especially important where no license governs the use of the materials. It also is key in arenas where the materials in question are not being currently managed by an agent. This is the case with many published materials that are out of print, as well as with a large number of unpublished materials. Orphan works, those for which the rights holder is unknown or cannot be found, are works that, by definition, have no agent to enter them into contracts or deals of any kind. As works age, their level of activity in the active commercial sphere tends to wane. The *stuff* remains, but the *people* and *deals* have faded away.

Some metadata has addressed copyright information, although the practice of providing this information is not widespread. The Cedars project produced metadata [CEDARS] that includes copyright fields (Copyright Statement, Name of Publisher, Date of Publication, Place of Publication, RightsWarning, Contacts or Rights Holder), license fields (Actions Permitted by License, License Text Pointer), and fields where laws and statutes could be noted (Actions Permitted by Statute, Legislation Text). The Cedars project ended in March of 2002, and it doesn't appear that the rights metadata was ever used in an active project. The draft rights schema for the Metadata Encoding and Transmission Standard (METS), METSRights [METSR], provides data elements for some copyright information (primarily the rights holder name and contract, but also a rights status field), although its main emphasis is on permissions and constraints. Copyright information is the focus of the metadata produced by the California Digital Library (CDL) as part of its rights framework project. This schema, called "copyrightMD [CDL-MD]," includes: creator name and dates; rights holder name and contact; a copyright status attribute; dates of creation, publication and copyright registration. The CDL schema is in test at this date and has not yet been applied to preservation resources.

## **The PREMIS Rights Entity**

In the PREMIS data model the *objects*, *events*, and *agents* are very close to Godfrey Rust's *stuff*, *deals*, and *people*. What PREMIS adds is the element of rights, something that is assumed in the e-commerce model because it is about rights-bound intellectual property. The e-commerce model does not have data elements to record situations in which *people* exchange and use *stuff* based entirely on the rights inherent in copyright law.

The PREMIS project is not alone in its need to express rights in terms of copyright law; this is a common issue in digital libraries today. Not only is there no standard set of metadata for this aspect of rights, there are no conventions for conveying this information to users. Part 2 of this report looks at the relationship between digital preservation and intellectual property law, and at some effort in progress to address the preservation rights needs through legislation. Part 3 will make recommendations for additional rights metadata for the PREMIS data dictionary.

## References

- [Adobe] Adobe Systems, Inc. *A primer on electronic document security. How document control and digital signatures protect electronic documents.* [http://www.adobe.com/security/pdfs/acrobat\\_security\\_wp.pdf](http://www.adobe.com/security/pdfs/acrobat_security_wp.pdf) (5/1/06)
- [CDL-MD] California Digital Library (CDL). *copyrightMD Schema, verision 0.9.* <http://www.cdlib.org/inside/projects/rights/schema> (4/24/06)
- [CEDARS] Cedars Project. *Metadata For Digital Preservation.* The Cedars Project Outline Specification Draft For Public Consultation. March, 2000. <http://www.leeds.ac.uk/cedars/documents/Metadata/cedars.html>
- [COYLE 2004] Coyle, Karen. *Rights Expression Languages.* February, 2004. [http://www.loc.gov/standards/Coylereport\\_final1single.pdf](http://www.loc.gov/standards/Coylereport_final1single.pdf)
- [COYLE] Coyle, Karen. *XrML – a Short History of Usage Rights.* August 19, 2004. <http://www.kcoyle.net/xrml.html>
- [CC] Creative Commons. <http://www.creativecommons.org>
- [EDITEUR] EDITEUR. *ONIX Publisher Licensing Terms, draft format.* February, 2006. <http://www.editeur.org/licensing/060202%20ONIX%20Licensing%20Terms%20draft.pdf>
- [INDECS] <the indecs project>. <http://www.indecs.org/project.htm>
- [ODRL] Open Digital Rights Language Initiative. <http://www.odrl.net>
- [OMA] Open Mobile Alliance. *OMA DRM 2.0 Enabler.* [http://www.openmobilealliance.org/release\\_program/drm\\_v2\\_0.html](http://www.openmobilealliance.org/release_program/drm_v2_0.html) (5/1/06)
- [PREMIS] PREMIS Working Group. *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group.* (May 2005) Available: <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>
- [RUST] Godfrey Rust. *Data Definitions.* Presentation at Dublin Core DC, November, 1998
- [WIKI/FAIR] Wikipedia. *FairPlay.* <http://en.wikipedia.org/wiki/FairPlay> (5/1/06)

## II. Digital Preservation Strategies and Rights Implications

### Introduction

The uneasy interaction between the technology of preservation and the environment of legal rights is a key problem that needs to be solved so that libraries, archives, and others can adequately preserve current digital materials for future generations. As is often the case, technology has evolved more rapidly than the legal environment, and this leaves some important questions unanswered. At the same time, digital preservation activities cannot wait for the resolution of the questions of intellectual property rights: if today's digital materials are not preserved now, they will likely be lost to future generations. To the extent that preservation of cultural materials is seen as an important social function, we can be confident that legal solutions will be developed to support that function. Indeed we are seeing considerable activity especially in the area of the development of legal deposit rules for digital materials. Meanwhile, institutions engaging in digital preservation are looking at current copyright law to seek answers to the question: is there a right to preserve materials in digital formats? [Ayre 2004] [Besek] [Cworth] [Hirtle] [NLA]

Discussion of preservation rights is complicated by the understandable desire of libraries and archives to provide access to the materials that they preserve. It is not difficult to find preservation repositories today whose contractual rights statement includes preservation copying and access to users in the same sentence. Preservation and access have been assumed to be simultaneous in the treatment of physical materials, even though some preservation copies, such as microfilm masters, are specifically designated as non-use. With digital materials, however, the equation of *a copy* with *a use* is harder to make. Digital preservation will require multiple copies to be made, both duplicate instances that serve as backups and possibly copies in different formats. Digital copies may or may not be accessible to users. For the discussion of digital preservation and rights in this paper we choose to focus on the preservation action apart from access. First, however, we need to define our most important terms: digital preservation, and rights.

### ***Digital Preservation***

In this paper, the term "digital" will refer to resources that are stored on general digital storage media, such as hard drives, that cannot be transferred in a physical manner. Resources that are digital in nature, such as digital forms of music recording, and are stored on physical media such as a compact disc will not be considered. The storage on general digital store media means that the digitally preserved resource must be copied into the preservation repository. It is in large part this act of copying that makes digital preservation a question of intellectual property rights.

In addition, the goal of digital preservation, at least in this paper, is more than the storage of the original resource's bits, although bit storage is an essential element of a repository's actions. The primary goal is to maintain a digital resource in a currently usable state. This is a complex goal because all digital materials require external mechanisms, in the form of hardware and software, to render them perceptible to human users. As we will see below, preserving resources in a usable state requires actions on the part of the repository, and these actions result in the creation of copies, and sometimes in the modification of the original resource.

A key aspect of digital preservation and a point of contrast to the preservation of analog materials is that digital preservation requires the repository to make a preservation copy of the resource from the first moment of acquisition, and to perform actions on that resource throughout the lifetime of the repository. Where analog preservation can, and often does, begin at the point that materials begin to show signs of deterioration, digital repositories must act before any deterioration can take place since even the smallest loss of integrity can render the resource entirely unusable. In addition, digital preservation must not only guard against deterioration of the resource itself, but against the obsolescence or unavailability of the technology needed to render the content for human interaction.

## ***Rights***

Intellectual property law defines author's rights; this includes the rights of exploitation (generally known as "copyrights") and moral rights. There are also recent additions to intellectual property law that address rights in databases, technical protection measures, and copyright management information. This article will refer to the Berne convention of 1979 [Berne] and the WIPO Copyright Treaty of 1996 [WIPO] for a broad concept of legal rights, and will highlight occasional significant variations from laws of some countries. This is in no way a thorough analysis of international law in this area.

In addition to the author's rights that are addressed by intellectual property law, resources can be bound by licenses and contracts. These agreements derive their authority from the legal rights but can specify a nearly unlimited set of requirements and prohibitions to be agreed by the signing parties.

The key questions before us regard the relationship of these rights and the actions required for digital preservation: can we preserve digital resources under the rights that we have in current law, can we define a specific set of permissions that should be granted for preservation, and what metadata is needed to document and manage these rights and permissions in a digital repository setting?

## ***Digital preservation actions***

Digital preservation efforts are beginning in countries all over the world and in a wide variety of institutions. In spite of the current interest, it remains more of a theory than a current practice, in part because the true preservation activities will require the passage of time before they can be put to the test. The problem set is well articulated, however. [CLIR, Gladney, Howell, Lorie 2000]

Digital preservation is not just a matter of preserving a resource, but may require preservation of the rendering environment. The methods used to preserve materials must be mindful of: issues of accuracy (possible loss of functionality over time); retention of the look and feel of the original; interaction of the resource with specific hardware; the overall software context, such as operating systems; environmental dependencies in general, such as screen capabilities; copy and access protection technologies. [Hedstrom, LC]

The actions that digital repositories may take on resources are given below. Note that these actions are the ones articulated today; it would be obviously incorrect to assume that no further development in preservation technologies will take place.

1. **Re-copying digital files onto new media.** This is a basic copy of the same

bit stream from one carrier to another. This includes the step that copies a digital resource initially into the repository, but it may take place periodically within the repository as well.

2. **Storage of multiple copies.** Unlike physical materials which can be rescued after deterioration is evident, a digital copy is most likely wholly unusable at the first sign of deterioration. Deterioration is also often not based on a slow passage of time but can instead be an unpredictable failure of the storage medium at any point in time. To protect against this kind of failure most repositories will store multiple copies and those copies may be kept in different locations and on different hardware. This increases the chance that a viable copy of the resource will always be available. [LOCKSS]
3. **Validation of digital files "in place."** Preservation repositories need to have a method of testing the validity of files in their care, usually through the calculation of a message digest for the resource. This is the digital equivalent of examining physical copies for signs of deterioration.
4. **Compression and de-compression of files for storage.** There is an inherent risk of change to the underlying file in many compression techniques.
5. **Migration of digital resources to current formats.** In addition to storing digital resources securely, the preservation repository has as its goal to maintain the resources in a *usable* condition. With files that are in program-specific formats, such as the output of word processing programs, subsequent versions of those programs may read and create files in new formats. In general, these newer program versions can read, and sometimes create, files in recent preceding formats. Conversion of older files into the newer format, using a later version of the generating program, is one way to maintain resource usability. In some cases, the successor format may be a wholly new technology, such as the migration of files in JPEG format to JPEG2000, or from a character set using the ISO 8859 standards to the Universal Character Set.
6. **Storing the content or information of a resource but not its format.** There are some digital formats that are considered to be better as preservation formats than others. Generally, the preferred preservation formats are non-proprietary and based on open standards. Preservation repositories may require that deposits be made in preferred formats, or they may create versions of the files in those formats as part of the preservation process. such as the reduction of language files to plain text, or the export of database data as delimited text. [CDL, p. 5]
7. **Emulation of the environment that supports a digital resource.** Emulation is used to access and use programs in non-preservation situations such as the emulation of the Microsoft Windows operating system on a Macintosh computer that will run programs designed for the Windows OS. Emulation has also been used by game enthusiasts to continue use of games from earlier platforms. Emulation for preservation would render the resource for use, leaving the original format of the file intact. Emulation of programs and of complex digital resources with interdependent components (e.g. network services) are significantly more difficult and may not be possible without modification of the target digital resource. [Grancer, Rothenberg]

8. **Universal Virtual Computer (UVC).** The UVC is based on the creation of a version of the data stream to be preserved in a standard format that can be used to recreate the file in an "understandable form" at a future time. An interpreter program will be written to allow the standard data stream to be rendered in the future computing environment. [Lorie 2001] Similar to the UVC in concept is the Universal Preservation Format, although this does not seem to be currently in development. [Shepard]

While these are the methods, the actual actions on digital resources inherent in the methods are more general.

1. **Copying.** Copying can entail the creation of a new, identical bit-level instance of the resource. However, all of the methods, with the possible exception of emulation, also make a new fixed copy of the digital resource, albeit with some differences.
2. **Access or use.** Access and use can act on the file for rendering but leave the original file unchanged. For example, testing of the validity of the file requires access. Methods that would transform the file by converting it to another format must access and render the file in some form. Emulation also accesses the file, but only at the time that the emulation environment is active.
3. **Migration.** A preservation strategy in which a version of a Digital Object is created in a different Format, where the new Format is compatible with contemporary software and hardware. [PREMIS] A migration is a particular type of transformation, as noted below.
4. **Transformation.** Process performed on a Digital Object that results in one or more new Digital objects that are not bit-wise identical to the source Digital Object. A format migration is a special case of transformation. Destructive transformations, such as deletion, can also be considered in this category.

The PREMIS event types can be categorized by these actions as well:

<u>Copying</u>	<u>Access or use</u>	<u>Migration</u>	<u>Transformation</u>
capture	decryption	migration	compression
ingestion	digital signature validation		deaccession
replication	fixity check		decompression
	message digest calculation		deletion
	validation		normalization
	virus check		

See Appendix A for a more detailed chart of preservation actions and rights.

## **Rights In the Law**

Intellectual property law confers certain exclusive rights on the authors of intellectual and creative works. Author's rights can be divided into two basic categories: the right to exploit the work, commercially or otherwise, and moral rights. The latter protect the integrity of the work and, by implication, the author's personal reputation. In recent years new rights have been added to protect digital files. These include protection for databases, prohibitions on circumventing technical protection, and protection for rights management information. [WIPO, US-DMCA]

In addition to these intellectual property rights, many countries have laws that require the deposit of intellectual and cultural works with one or more libraries in the country. Because the intention of these deposits is to create an archive of the nation's cultural heritage, preservation of these works is implied. These deposit laws are being modified to include digital resources and will be addressing the preservation issues.

### ***Copyright or Exploitation Rights***

Copyright law addresses the copying, distribution, and modification of works, and makes these the exclusive rights of the creator of the work. It is silent on the actions of users that do not infringe on these exclusive rights, such as the reading of a book or the viewing of a performance. The primary author's copyrights as described by Berne are the exclusive rights of:

- translation
- reproduction
- public performance
- broadcast
- public recitation

The particulars of the author's rights vary among the countries that are signatories to the Berne Convention. For example, the United States copyright law does not mention translation, but does reserve for the copyright holder the right to make "derivative works," which includes translations but also actions like fictionalization or dramatization [USC 17, 106]. US and other countries' law make reference also to "distribution" or "publication" of the work as an exclusive right. There are specific rights for non-textual materials that generally pertain to performance and broadcast, in particular for sound recordings and film. There are also rights that pertain particularly to the visual and performance arts, such as the right of display and performance.

Certain of the author's rights regard only public display or performance, so those can be excluded, since the scope of preservation as we have defined it does not include public access. (Public access can be provided concurrent with preservation but is not inherent in the preservation activity.) Of the basic copyrights, the right of reproduction is the one that is directly of interest for preservation. Because all digitally archived materials will, by their nature, be copies, and the act of entering an item into an archive is the making of a copy, copying a digital resource into a preservation repository may infringe on the author's rights. Whether or not copying itself is an infringement will depend on exceptions provided by copyright law, which are discussed below.

Most acts of the preservation process also make copies. Many of the acts will modify the work in some way, such as the migration of a digital resource to a more modern format. The degree to which these modifications are significant in relation to rights may depend on the nature of the work: conversion of a text to an alternate format (e.g. from Microsoft Word to Adobe PDF) may not affect the creativity expressed in the text, and therefore the modification would be considered a copy, but a similar conversion in digitally recorded music might result in an obviously modified work if the quality of the sound is changed in any way by the new format. In addition, there are transformations that can be considered to result in an entirely new work. Some of these are obvious, such as the creation of a film from a book, but others are more subtle, like the definition of a new edition of a book after making corrections to the text. The threshold for copying as opposed to creation of new works differs in the Berne treaty countries, as does the rights of a creator to control those actions.

Other rights, like translation or the preparation of derivative works, may have implications for preservation activities in some countries. For example, the UK copyright law extends "translation," one of the adaptations that are exclusive rights of the copyright holder, to computer programs: "a version of the program in which it is converted into or out of a computer language or code or into a different computer language or code." [UK, 21, section 4]. In the US Code, no such clarifications have been added for digital materials.

Note that not all works are given the same protection by copyright law. Berne excludes facts and "news of the day." U.S. law excludes Federal documents. Each country has the ability to make specific exclusions within some limits set by Berne, and any country can add protections beyond those in Berne.

There are also works whose copyright protection has expired due to the passage of time. These works cannot be "born digital" since they must date from the early 20<sup>th</sup> century, but when digitized they retain their public domain status. These works are free of the restrictions that copyright law would impose on digital preservation.

## ***Moral Rights***

Moral rights are recognized in many countries, although the exact meaning of these rights is nuanced in their laws. Berne defines moral rights as:

"... the author shall have the right to claim authorship of the work and to object to any distortion, mutilation or other modification of, or other derogatory action in relation to, the said work, which would be prejudicial to his honor or reputation."  
[Berne, 6bis]

Moral rights in some countries expire at the author's death, and in others are extended to the full length of the term of author's rights (life plus 50 or 70 years). Generally, moral rights cannot be transferred, but some countries do allow authors to waive these rights [Collins]. In the U.S., moral rights are only extended to works of visual art [USC 17, 106A].

Moral rights include the right of attribution, which means that authors have the right to have copies of their works or portions of their works carry the author's name. This right is particularly apt in the electronic environment where works and their metadata can become separated during copying or transmission. [Oppenheim]

In some countries, one of the moral rights of the author is the right to "retrait et de repentir," meaning that an author can retract a previously distributed work from the public

[France, Art. L. 121-4]. Spain has a nearly identical right in its law [Spain, Article 14, n. 6]. This ability to remove a work may conflict with the preservation philosophy of permanent storage.

The chief difficulty with the application of moral rights is that "derogatory" or "prejudicial" is in the eye of the beholder. Some acts may be obvious (painting a moustache on a famous portrait), while others will fall into a grey area (reproduction of visual art with a limited color palette). It is not known if preservation actions such as migration and transformation can be successfully challenged under moral rights.

### ***Specific Rights in Computer Files and Technology***

The World Intellectual Property Organization (WIPO) produced an amendment to the Berne Treaty in 1996 "Recognizing the profound impact of the development and convergence of information and communication technologies on the creation and use of literary and artistic works...." [WIPO, p. 1]. Here we find protection for computer programs and for databases, as well as prohibitions against the creation or trafficking in technology designed to defeat technical protections, and against the removal or modification of rights management information for the purpose of masking infringing uses of works. These measures have been implemented in varying ways in the countries participating in WIPO. For example, the UK has incorporated database protection into its copyright law, but the US has not. Yet the US produced the Digital Millennium Copyright Act [USC 17, Section 108] with detailed treatment of the prohibition on the circumvention of technical protection, which other countries have not.

Some of the key elements of this treaty, and of the laws that have implemented it, are:

- Computer programs are protected as literary works [WIPO, Article 4]
- Compilations of data (databases) are protected as such. This does not extend copyright to the underlying data, but to the database as a whole. [WIPO, Article 5] Database protection law prohibits the extraction and re-use of substantial portions of a database.
- Authors of literary and creative works have the exclusive right to communicate their works to the public, by any means. This extends the right of distribution and publication to the networked environment. [WIPO, Article 8]
- Laws in the countries that adhere to WIPO's treaty must provide protection against the circumvention of effective technological measures that are used by authors. [WIPO, Article 11.] This is the basis for the protections included in the US law [USC 17, 1201], and UK law [UK Part VII, 296].
- Laws in the countries that adhere to WIPO's treaty must provide protection against the knowing alteration of electronic rights management information. This information is defined very broadly as any information that identifies the work, the author of the work, terms and conditions on the use of the work, and that is attached to or is communicated with the work. [WIPO, article 12]. The law requires that the alteration be for the purpose of infringing on the intellectual property of the item. This is implemented in the US as Title 17, section 1202, and in the UK as section 296ZG.

## Exceptions to Legal Rights

Berne allows countries to develop exceptions to the rights of authors so long as those exceptions do not interfere with the normal exploitation of the work. The common exceptions are those of fair use and fair dealing, and exceptions for education activities, including libraries.

### *Fair Use and Fair Dealing*

Fair use and fair dealing exceptions are often cited in support of educational and non-profit copying of copyright-protected materials, including the creation of digital copies for preservation. In this sense, the act of preservation is itself a use of the material.

Subsequent access to those digital materials (which is generally understood to involve making a copy for the user) is also commonly cited as a fair use, but these would have to be made in separate determinations since they are two different actions. Here, we will only address the analysis of fair use and fair dealing for the creation of the preservation copy.

There is a significant difference between the exceptions of fair use and fair dealing. Fair dealing law lists specific exceptions to copyright, the most common being those derived from Berne [Berne, Article 10]: quotations from a work, for purposes of commentary, review, or news reporting; and utilization in an educational setting.

Fair use sets out criteria for use of a work, but leaves the question of what is fair open-ended. This provides a great deal of flexibility, especially as technology changes the uses that are made of materials, but also adds uncertainty for users who must make the determination of whether the use they wish to make is likely to be fair. Determination of fair use can only be made in a court of law, therefore occurs when rights holders and users are in dispute.

In some countries, fair dealing includes an extensive list of actions. For example, the UK copyright law has about 60 allowed uses of copyrighted works [UK, sections 28 through 75]. Australian law has about two dozen fair dealing exceptions, including ones relating specifically to computer programs [Australia, Sections 40-47]. Canadian law has about two dozen fair dealing exceptions [Canada 1999, Sections 29 & 30], many of which refer to educational activities and libraries. Although these exceptions name particular actions, they still require interpretation when applied to specific actions, so although the fair dealing environment is less flexible than that of fair use, it is not a matter of cut-and-dried fact.

Berne and other international agreements, such as the Agreement on Trade-Related Aspects of Intellectual Property Rights [WTO] include basic criteria for determining exceptions to copyright law. These criteria are applied in the creation of legal exceptions in and in the US, as criteria for the determination of fair use. The "three step test" reads:

exceptions and limitations to the rights of copyright owners must be confined:

- to certain special cases,
  - which do not conflict with a normal exploitation of the work, and
  - do not unreasonably prejudice the legitimate interests of the right holder.
- [Australia2, p. 11]

## ***First Sale***

The First Sale right exists in US law [USC 17, Section 109]. First sale allows the owner of a copy to dispose of that copy without the permission of the rights holder. It is the first sale exception that allows libraries to lend works and that is the basis for the used books market. First sale is one of the battlegrounds in the digital environment, with those representing rights holders arguing that because all transfers of the work require that a copy be made there is no "disposition" in the sense intended by section 109. [US 2000, Coyle] First sale might be cited as allowing a digital repository to discard or delete a work that it holds. First sale might also allow an institution to transfer items in its digital archive to another institution, much as one would donate a collection.

## ***Computer Programs***

Some countries have enacted specific exceptions to copyright law to allow owners of computer programs to make a single copy for the purposes of backup [USC 17, Section 117] [Canada, 30.6]. These laws also allow an adaptation to be made for the sole purpose of using the program as licensed. This adaptation in many cases will require de-compilation or reverse-engineering of the program. While this appears to allow some of the actions required by preservation, note that most computer programs purchased today will have license terms that govern these particular actions, so the owner may be prevented from taking advantage of the exception in the law.

## ***Exceptions for Libraries***

Not all Berne signatories have specific exceptions for libraries and Berne itself does not mention libraries in its text. It does, however, state that signatory countries can allow copies to be made "in certain special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author." [Berne, 9(2)] This seems to be a useful exception under which a variety of library activities, including preservation, can be defined.

Copyright law in the UK [UK, sections 32-44], Canada [Canada 30.1-30.5] and Australia [Australia, Part III, Division 5], include exceptions for libraries and archives that are similar in nature. These laws generally allow the selected copying of certain individual works, like journal articles or parts of longer works, under certain conditions by non-profit libraries.

These laws, as well as section 108 of US copyright law, include an exception for copying for the purposes of preservation, although the laws as written today are appropriate to the preservation of analog works and do not address the special needs of digital preservation. Copying for preservation only applies to damaged or deteriorating works for which a replacement copy is not otherwise available in the market, or for manuscripts where use of a service copy is necessary to substitute for handling the original. The use of deterioration as the trigger that permits preservation to take place is not suitable for digital resources because a file that has deteriorated, in many cases, cannot be used to create a usable copy of the original.

The US law's exceptions for libraries [USC 17, section 108] fall under this category of allowing preservation for deteriorating copies. Under section 108 libraries can make copies of works that are deteriorating and that are not available for purchase under normal market conditions. It also allows copying of unpublished works. This latter does not require

deterioration for the preservation copy to be made. US law specifies a particular number of preservation copies that can be made (three), a clear testament to its assumption of both analog originals and copies, since this doesn't address the fact that computer systems include many more copies in their system backups and even in their disk arrays. Preservation in this law is expressed solely as the making of copies and there is no mention of the type of transformations that digital preservation needs to perform to assure that the underlying content can be rendered in the future. It also limits use of those copies to the library premises, another assumption from the analog world. Analysis of section 108 to determine how it might be modified to include digital works is underway, but actual results from this study will likely take years to be realized. [LC 108]

## **Mandatory or Legal Deposit**

As shown in the IFLA-CDNL report [ICABS], most countries in Europe and the Western world are in the process of modifying their legal deposit laws to accommodate digital resources. Because it is the role of legal deposit to preserve the intellectual materials of a country's culture and to make those available to the citizenry over time, legal deposit laws must address the long-term preservation issues as well as ongoing access. A few countries, such as Canada and China, have already included digital materials in their mandatory deposit. Others have legislation that is coming into effect in 2005 or 2006 (these are: Denmark, France, Germany, New Zealand, UK) [ICABS, p. 13], while the remainder are studying the issue both from legal and technical viewpoints. Some national libraries allow voluntary deposit of digital materials today (the UK and US, among others). [Beagrie] [ICABS] These deposit laws, or their implementation directives, must define allowable preservation for digital materials. In addition, because their goal is both preservation and access, the relevant laws or regulations must also cover allowable access conditions.

The emphasis in these projects is on currently produced materials, and in general will involve resources that are "born digital." Because the materials are current, they will be under copyright, with the exception of materials that are always in the public domain such as some legislative works. The legal deposit repository may define its rights as related to the repository as a whole rather than on an item-by-item basis. For example, some countries intend to only allow access within the national library that receives the digital deposits; others will provide access within the context of educational institutions. In either case, the items within the repository are fairly uniform in terms of their legal status, which means that it may not be necessary for these countries to develop metadata that defines the rights status of an individual item.

The range of services that will be provided from legal deposit schemes will vary by country. The laws being developed will not apply to digital preservation outside of the legal deposit regime, but may promote greater acceptance of the role of libraries and other cultural institutions in the preservation of digital materials.

Because of the great variety of digital files that can be subject to deposit, some of the national libraries are taking a phased approach, either by format (e.g. html, text) or by publication type (e-journals). Few are currently planning to store software or entire databases. In the UK, the 2003 Copyright Act added section 44A that allows legal deposit libraries (of which there are 6 in the UK) to make copies of UK-related works from the Internet. This would appear to allow web crawling of at least the .uk namespace for the purposes of legal deposit.

## Digital Resources and Preservation

While the WIPO treaty specifically addresses digital resources, the particular prohibitions are primarily relevant to use, re-use, and economic exploitation of the digital materials. Their effect on preservation may be minimal. For example, the specific ruling on databases distinguishes between normal, allowed use of a database, such as end-user searching and display of retrieved items, but describes as infringing: "without the consent of the owner of the right, he extracts or re-utilizes all or a substantial part of the contents of the database." [UK, section 16(1)] This law protects the database from competitive exploitation by extraction of its data. The general copyright prohibition against copying or modification, which applies to the database as a whole, is more relevant to the act of preservation since presumably preservation would be applied to the database as a whole and not to extracted portions.

The anti-circumvention prohibitions could have a significant impact on the digital preservation of technology-protected files, mainly because this portion of the law will prevent the creation of technologies that could be used to unlock protected files. In the US, libraries have already been given an exception to the prohibition on circumvention for the purposes of reviewing resources prior to purchase. [USC 17, Section 1201 (d)] However, it is unlikely that libraries will have the means to circumvent technological protections since circumvention technologies themselves are prohibited by the same law.

## References

- [Adobe] Adobe Systems, Inc. *A primer on electronic document security. How document control and digital signatures protect electronic documents.* [http://www.adobe.com/security/pdfs/acrobat\\_security\\_wp.pdf](http://www.adobe.com/security/pdfs/acrobat_security_wp.pdf) (5/1/06)
- [Australia] Australia. *Copyright Act 1968, Part III Copyright in original literary, dramatic, musical and artistic works. Division 5 Copying of works in libraries or archives.* pp. 71-85. Available: <http://www.comlaw.gov.au/CopyrightAct1968>
- [Australia2] Australia. Attorney General's Department. *Fair Use and Other Copyright Exceptions. An examination of fair use, fair dealing and other exceptions in the Digital Age.* May, 2005
- [Ayre] Ayre, Catherine and Adrienne Muir. *Right to Preserve? Copyright and licensing for digital preservation project. Final Report.* Leicestershire, Loughborough University, March 2004. 319 p.
- [Beagrie] Beagrie, Neil. *National Digital Preservation Initiatives. An Overview of Developments in Australia, France, the Netherlands, and the United Kingdom and of Related International Activity.* Washington, D.C., Council on Library and Information Resources, and Library of Congress. April, 2003.
- [Berne] *Berne Convention for the Protection of Literary and Artistic Works.* September 28, 1979. Available: <http://www.wipo.int/clea/docs/en/wo/wo001en.htm>
- [Besek] Besek, June M., *Copyright Issues Relevant to the Creation of a Digital Archive: A Preliminary Assessment.* Washington, DC, Council on Library and Information Resources, January, 2003.

- [BL] British Library. *Code of Practice for the Voluntary Deposit of Non-Print Publications*. <http://www.bl.uk/about/policies/codeprac.html>
- [CDL-MD] California Digital Library (CDL). *copyrightMD Schema, version 0.9*. <http://www.cdlib.org/inside/projects/rights/schema> (4/24/06)
- [CDL] California Digital Library (CDL). *Overview of the UC Libraries Digital Preservation Repository (DPR)* Available: <http://www.cdlib.org/inside/projects/preservation/dpr/DPRoverview.pdf>
- [CDLa] California Digital Library (CDL). *UC Libraries Digital Assets Agreement*. [http://www.cdl.ucop.edu/inside/projects/preservation/dpr/Submission\\_Agreement.pdf](http://www.cdl.ucop.edu/inside/projects/preservation/dpr/Submission_Agreement.pdf)
- [Canada] Canada, *Copyright Act ( R.S., 1985, c. C-42 )*. 2005. Available: <http://laws.justice.gc.ca/en/C-42/>
- [Canada 1999] Canada. Department of Justice. *Copyright Act: Exception for Educational Institutions, Libraries, Archives and Museums Regulations*. Available: <http://laws.justice.gc.ca/en/C-42/SOR-99-325/7701.html>
- [CEDARS] Cedars Project. *Metadata For Digital Preservation*. The Cedars Project Outline Specification Draft For Public Consultation. March, 2000. <http://www.leeds.ac.uk/cedars/documents/Metadata/cedars.html>
- [Cworth] Charlesworth, Andrew. *The CAMiLEON Project: Legal issues arising from the work aiming to preserve elements of the interactive multimedia work entitle "The BBC Domesday Project."* University of Hull, 2002
- [Collins] Collins, Bowley Kerr, *Copyright Law in Canada*. 2002. Available: <http://www.trytel.com/~pbkerr/copyright.html>
- [CLIR] Council on Library and Information Resources (CLIR). *Access in the Future Tense*. Washington, D.C., CLIR, April 2004.. Available: [www.clir.org/pubs/reports/pub126/pub126.pdf](http://www.clir.org/pubs/reports/pub126/pub126.pdf)
- [Coyle] Coyle, Karen, *Comments to the Copyright Office*. 2000. Available: <http://www.copyright.gov/reports/studies/dmca/comments/Init007.pdf>
- [Flecker] Flecker, Dale, "Preserving Digital Periodicals." Appendix 3, *Preserving our Digital Heritage; Plan for the National Digital Information Infrastructure and Preservation*. Washington, DC, Library of Congress, October, 2002. Available: <http://www.digitalpreservation.gov/about/planning.html>
- [France] France. *Code de la Propriete Intellectuelle (Partie Législative)*. Available: <http://www.legifrance.gouv.fr/WAspad/RechercheSimplePartieCode?commun=&code=CPOINT>
- [Gladney] Gladney, H.M. and R.A. Lorie, "Trustworthy 100-Year Digital Objects: Durable Encoding for When it's Too Late to Ask." *Communications of the ACM*. V. 49, n. 2 (2006), pp. 111-116. 30 June 2003; rev. 14 March 2005.
- [Grancer] Grancer, Stewart, "Emulation as a Digital Preservation Strategy." *D-Lib Magazine*, v. 6, n. 10, October, 2000. Available: <http://dlib.org/dlib/october00/granger/10granger.html>

- [Hedstrom] Hedstrom, M. and S. Montgomery, *Digital preservation needs and requirements in RLG member institutions*. Mountain View, CA: Research Libraries Group. Available: <http://www.rlg.org/preserv/digpres.html>
- [Hirtle] Hirtle, Peter B. *Digital Preservation and Copyright*. Stanford University Libraries, Copyright and Fair Use. Available: [http://fairuse.stanford.edu/commentary\\_and\\_analysis/2003\\_11\\_hirtle.html](http://fairuse.stanford.edu/commentary_and_analysis/2003_11_hirtle.html)
- [Howell] Howell, Alan G., *Preserving Digital Information: Challenges and Solutions*. rev. 2004. Available: [http://www.alanhowell.com.au/papers/pdi\\_wkb.pdf](http://www.alanhowell.com.au/papers/pdi_wkb.pdf)
- [ICABS] IFLA-CDNL Alliance for Bibliographic Standards (ICABS), *Building Networks in Digital Preservation*. Den Haag, The Netherlands, June 2005. Review version.
- [LC] Library of Congress (LC), *Preserving our Digital Heritage; Plan for the National Digital Information Infrastructure and Preservation Program*. Washington, DC, Library of Congress, October, 2002. Available: <http://www.digitalpreservation.gov/about/planning.html>
- [LC 108] Library of Congress (LC). *Section 108 Study Group*. Available: <http://www.loc.gov/section108/>
- [LOCKSS] LOCKSS Alliance. *LOCKSS Program Home Page*. Available: <http://www.lockss.org>
- [Lorie 2001] Lorie, Raymond A., "A Project on Preservation of Digital Data." *RLG DigiNews*: 5(3), Jun3 15, 2001. Available: [www.rlg.org/preserv/diginews/diginews5-3.html](http://www.rlg.org/preserv/diginews/diginews5-3.html)
- [Lorie 2000] Lorie, Raymond A., *Long-Term Archiving of Digital Information*. IBM Research Report RJ 10185, 18 May, 2000. Available: [http://domino.watson.ibm.com/library/CyberDig.nsf/papers/BE2A2B188544DF2C8525690D00517082/\\$File/RJ10185.pdf](http://domino.watson.ibm.com/library/CyberDig.nsf/papers/BE2A2B188544DF2C8525690D00517082/$File/RJ10185.pdf)
- [NLA] National Library of Australia. *Preserving Accessing to Digital Information. Intellectual Property Rights Management*. Available: <http://www.nla.gov.au/padi/topics/28.html>
- [OCLC] OCLC, *OCLC Digital Archive Preservation Policy and Supporting Documentation*. Dublin, OH, January, 2005. Available: <http://www.oclc.org/support/documentation/digitalarchive/preservationpolicy.pdf>
- [Oppenheim] Oppenheim, Charles, "Moral Rights and the Electronic Library." in: *Ariadne*, Issue 4, 1996. Available: <http://www.ariadne.ac.uk/issue4/copyright/>
- [PREMIS] PREMIS Working Group. *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group*. (May 2005) Available: <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>
- [Rothenberg] Rothenberg, Jeff, *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. CLIR, January, 1999. Available: <http://www.clir.org/pubs/reports/rothenberg/pub77.pdf>

- [Shepard] Shepard, Thom and Dave MacCarn. *The Universal Preservation Format*. Boston, WGBH, 2001. Available: [http://info.wgbh.org/upf/pdfs/20010818-UPF\\_RP.pdf](http://info.wgbh.org/upf/pdfs/20010818-UPF_RP.pdf)
- [Spain] Spain, *Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el Texto Refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia*. Available: [http://www.mcu.es/propint/files/LeyProp\\_Intelectual\\_mod172.pdf](http://www.mcu.es/propint/files/LeyProp_Intelectual_mod172.pdf)
- [UKCS 2005] UK Copyright Service (UKCS), *Common questions about copyrights*. 2004-2005. Available: <http://www.copyrightservice.co.uk/copyright/questions/>
- [UK] UK Copyright Service (UKCS), *UK copyright law*. 11th August, 2004.
- [USC 17] United States Copyright Office. *Copyright Law of the United States, and Related Laws Contained in Title 17 of the United States Code*. Available: <http://www.copyright.gov/title17/>
- [US-DMCA] United States Copyright Office. *Digital Millennium Copyright Act Study*. 2000. Available: [http://www.copyright.gov/reports/studies/dmca/dmca\\_study.html](http://www.copyright.gov/reports/studies/dmca/dmca_study.html)
- [WIPO] World Intellectual Property Organization (WIPO). *WIPO Copyright Treaty*. December, 1996. Available: [http://www.wipo.int/treaties/en/ip/wct/trtdocs\\_wo033.html](http://www.wipo.int/treaties/en/ip/wct/trtdocs_wo033.html)
- [WTO] World Trade Organization. *Annex 1C. Trade-Related Aspects of Intellectual Property Rights (TRIPS)* [http://www.wto.org/english/docs\\_e/legal\\_e/27-trips\\_01\\_e.htm](http://www.wto.org/english/docs_e/legal_e/27-trips_01_e.htm)

### III. Rights Metadata for Preservation

We have seen in this analysis is that there are two primary types of rights that will need to be expressed in rights metadata for preservation:

**Legal rights, based on copyright law.** These will be the cases where the right to preserve is based on legal conditions such as fair use, preservation exceptions in the law, or the legal deposit rules for that country's national library. Legal rights also include the public domain.

**Permissions.** If the repository determines that preservation cannot take place under current law, permissions must be sought.

For either case we will need to know who holds the intellectual property rights in the object.

#### ***Rights Metadata Based on Law***

In many cases digital preservation will rely on statute or law for its rights. There are few instances today of law that directly addresses digital preservation (as opposed to digital deposit, which is becoming a subject of law), although there is a great deal of activity in progress that may result in such laws. The specific areas where law addresses preservation at the moment refer to the preservation of analog materials, and some digital preservation is taking place as interpretations of these laws.

An important legal designation is that an item is in the public domain. These materials are free of exclusive rights and therefore can be copied, modified, and displayed without restriction. (Although moral rights may impinge on some types of modifications, the general goal of preservation of the work in a form as close as possible to the original will usually be in keeping with moral rights.) It will be important for preservation repositories to identify materials that are in the public domain and to provide supporting information for that determination. The supporting information may be categorical, that is that the material in question is not covered by copyright in that jurisdiction (e.g. federal government documents in the United States), or it may be information documenting the expiration of rights, such as the date of creation and the death date of the creator.

Where materials are not in the public domain, it will be advisable to provide the citation to the specific area of copyright law that was used to allow the repository to include the material in the repository and to perform preservation actions. This may be an area of law that is not specific to preservation, such as law allowing Fair Use, or in the future it may be a future law supporting digital preservation. It remains to be seen, however, how specific digital preservation laws will be in terms of the technology of preservation. Past experience tells us that "less is more" when legislating technology-based areas of activity, and laws that are enacted may leave the particulars of digital preservation to the designated institutions and professions.

#### ***Rights Metadata Based on Permissions***

Where a repository has obtained permission from rights holders, these permissions need to be codified in the rights metadata. One approach to these permissions is to obtain a general permission to preserve. This general permission often includes all actions that will be necessary to preserve the resource in the future, although specific actions are not named.

"You are also permitting the Library to take the necessary steps to preserve your publication, and to make it accessible to the public via the Internet now and in perpetuity." [NLA]

Other contracts name specific actions, such as the storage of copies, the migration of file formats, and the transformation of files to different formats. Even these statements tend to be general in nature and do not elaborate on the meaning of the terms used.

"... electronically store, archive, copy and/or convert the Digital Assets for preservation purposes;" [CDLa]

They also tend to be a single statement of agreement, not a choice of individual actions, although some repositories offer a choice of bit-level or full preservation treatment. [OCLC 2005]. Note that many contracts include preservation and access in a single statement, and that in general there is greater emphasis on specific access rules than on specific permissions for preservation. This reflects the fact that the preservation has no visible impact on author's rights, while access is an act of distribution of the work. Access can affect the market potential of the work and the control over distribution that belongs to the rights holder.

Permissions may be obtained through written documents, online forms, or through "click-through" agreements. In general, these documents are maintained outside of the preservation repository and are not included in the metadata that accompanies the digital object. In some repositories, contracts are uniform either across the repository, or within a particular project, so item-level recording of the permissions are not deemed necessary. Permissions of a non-preservation nature are common, generally covering access and use. Some permissions found in metadata are instructions to end-users that display with the digital item.

Managers of digital preservation services are understandably reluctant to enumerate actions or technologies in their approach to rights, since these could change over time. Instead, they are taking one of these approaches to rights:

**Preserving without a specific contract for digital materials.** For example, this is occurring in some countries that have begun accepting digital materials in their legal deposit program, but that have not (yet) developed a specific law or ruling that will govern the terms of that deposit. It is also taking place in digital archives and libraries that are working with materials that are in the public domain or for which they have made a fair use assessment for the right to preserve.

**Accepting digital materials for preservation with a blanket contract.** These contracts vary in language, but what they have in common is that the act of preservation is presented as indivisible, and the contracting party agrees to "preservation" as defined.

**Providing levels of rights.** This was only seen in the OCLC contract, but it provides two options: a bit-level preservation, and a full preservation, the latter being a blanket contract.

Note that permission can be obtained only when a rights holder can be located and chooses to participate. In the case of legal deposit, a rights holder is always involved. In other cases, for example in the preservation of Internet resources, it frequently occurs that no rights holder is involved in the transaction. And it may be the case that, even with reasonable

effort, the rights holder can not be located. Depending on the source of deposits, some digital repository actions will necessarily have to rely on law rather than permissions.

## Recommendations for the PREMIS Rights Entity

The PREMIS Rights entity is primarily a statement of permissions. It includes a granting agent (embedding the PREMIS Agent entity) and an object identifier; essentially *deals*, *people*, and *stuff*. Permissions can be time-bound with a beginning and/or ending date. Granted permissions must be expressed as acts but can also include statements of restrictions. The acts are defined as acts that can be performed by the preservation repository.

One area that needs clarification is that of the granting Agent. It may be important to understand the relationship of the granting Agent to the intellectual property rights inherent in the object. I would recommend that a role designation be added to the Agent entity so that it can be made clear if the Agent granting the permission is the rights holder or a legal representative of the rights holder. This is important because only those two have the legal right to grant permissions.

There are other types of agents, such as archives that physically hold or own the material being preserved. These may not be rights holders themselves, and the preservation permissions may come from a donor agreement with the archive. In that case, the signer of the donor agreement is the one granting the permission, not the archive, unless the agreement specifically transfers the relevant rights to the archive at which point the archive is the rights holder. The language of the donor agreement then constitutes the permission, even if it is not worded in terms of specific preservation acts. In this case it may be useful to record that the archive or institution is the other party to the donor agreement.

The expression of licenses and statutory grants may require the addition or modification of the PREMIS Rights data elements. Laws and regulations that permit the preservation repository to take action need to be recorded as the source of those rights. These do not have a specific granting Agent in the sense of a person or corporate entity that is making an agreement with the preservation repository. Instead they have a jurisdiction that is responsible for the law, and the connection between the object and the rights granted is also based on jurisdiction. Because it can be important to recognize that preservation is covered by law rather than by an agreement between parties, it would be best to have a distinct data element for recording the particulars of the law, including the jurisdiction.

The right to preserve may not be granted explicitly by the law, but may be the subject of some interpretation. This is clearly the case where an institution is undertaking preservation based on a fair use or fair dealing assessment, but it can also be true for works that have been determined to be in the public domain. These assessments are made within a specific context and at a specific time. Using other criteria, the context, and therefore the assessment, can change. For this reason it is important to record the date of the decision.

The PREMIS element *act*, which carries the action or actions the preservation repository is permitted to perform, is mandatory. Laws and agreements may be purposely vague in terms of actions, giving broad permission to do whatever is necessary to preserve the digital object in a usable format. This means that the *act* element will need to be able to codify these general permissions, or the element must be defined as optional.

In addition to recording rights and permissions, the PREMIS Rights entity must be able to

record whether the object is or is not under copyright. If the object is in the public domain then specific permission to act is irrelevant, and the preservation repository is free to copy or modify the object. It should not be necessary to record a granting agent nor a permission statement for public domain works.

It must also be possible to record when the copyright status of the object is unknown. This will be the case for the large pool of orphan works that exist in archives today. In this case, there is no granting agent and no permissions granted, so the archive is required to base its preservation action on law or statute. In addition to recording the copyright status, I would also recommend considering the inclusion of detailed copyright information in the PREMIS standard. This could be accomplished by including the copyrightMD schema developed by the California Digital Library [CDL-MD], or by incorporating the copyright elements similar to those in the CEDARS project [CEDARS].

In summary, the PREMIS Rights entity as it exists today can be seen to be the data elements needed to express a license or agreement to preserve. To this must be added a section for rights conferred by law or statute, and a statement of copyright status. The Agent entity also needs to be expanded to express a role, specifically a set of rights roles that will be used in the grantingAgent element. Conceptually, the resulting Agent and Rights entities might look something like:

#### **Agent Entity**

- agentIdentifier
  - agentIdentifierType
  - agentIdentifierValue
- agentName
- agentType
- agentRole

#### **Rights Entity**

- permissionStatement
  - permissionStatementIdentifier
    - permissionStatementIdentifierType
    - permissionStatementIdentifierValue
  - linkingObject
  - copyrightInformation
    - copyrightStatus
  - permittedByLicense
    - grantingAgent
      - grantingAgentRole
    - grantingAgreement
      - grantingAgreementIdentification
      - grantingAgreementInformation
  - permittedByStatute
    - jurisdiction
    - statute
      - statuteIdentification
      - statuteCitation
    - determinationDate

- permissionGranted
  - act
  - restriction
  - termOfGrant
    - startDate
    - endDate
  - permissionNote

The above should be seen as only one possible revision of the PREMIS Rights entity. Further discussion of the role of copyright law in relation to preservation is needed, especially in light of the PREMIS event-driven approach to preservation.

## Conclusion

For those developing metadata for preservation today, there will be two rights areas that will need to be addressed: recording of the law or statute under which preservation has been undertaken, and recording of contracts or permissions for preservation.

In terms of contracts and permissions, the rights metadata that is developed to record this information must allow the encoding of the full variety of rights expressions that are in use in the community. This is difficult because the current (and perhaps future) expression of rights is textual in nature and not easily reducible to data elements.

In terms of intellectual property law, there is little in copyright law today that directly addresses digital preservation, although developments in the area of legal deposit and exceptions for libraries and archives will hopefully produce a body of law in this area. Other than the analog concept of copies, laws in the Berne treaty countries have no direct correlation with the preservation actions that are enumerated in PREMIS or named in repository contracts. However, copyright law is being used today to support the preservation of some resources, law and this fact needs to be captured in the preservation metadata.

## References

- [CDL-MD] California Digital Library (CDL). *copyrightMD Schema, version 0.9*. <http://www.cdlib.org/inside/projects/rights/schema> (4/24/06)
- [CDL] California Digital Library (CDL). *Overview of the UC Libraries Digital Preservation Repository (DPR)* Available: <http://www.cdlib.org/inside/projects/preservation/dpr/DPRoverview.pdf>
- [CDLa] California Digital Library (CDL). *UC Libraries Digital Assets Agreement*. [http://www.cdl.ucop.edu/inside/projects/preservation/dpr/Submission\\_Agreement.pdf](http://www.cdl.ucop.edu/inside/projects/preservation/dpr/Submission_Agreement.pdf)
- [CEDARS] Cedars Project. *Metadata For Digital Preservation*. The Cedars Project Outline Specification Draft For Public Consultation. March, 2000. <http://www.leeds.ac.uk/cedars/documents/Metadata/cedars.html>
- [NLA] National Library of Australia. *NLA Copyright License Frequently Asked Questions*. <http://pandora.nla.gov.au/licencefaq.html> (5/6/06)

[OCLC] OCLC, *OCLC Digital Archive Preservation Policy and Supporting Documentation*. Dublin, OH, January, 2005. Available:  
<http://www.oclc.org/support/documentation/digitalarchive/preservationpolicy.pdf>

[USC] United States Code, *Copyright Law of the United States and Related Laws Contained in Title 17 of the United States Code*. Available:  
<http://www.copyright.gov/title17/>

## Appendix A: Preservation Actions and Rights

The chart below is one interpretation of the intersection between copyright law and preservation actions. Because laws differ between countries, this should not be seen as representing any one country's law, but to be a composite of common laws in Europe and North America.

- **Copy** includes any act that results in a new, non-temporary copy
- **Adapt/transform** is marked to all actions that result in a new copy that differs from the original, that is anything that is not a bit-level copy. Not all transformations will violate this section of the copyright law, but we have no decisions that would give us a clear line between an insignificant change and a transformative one in the preservation context.
- **Moral rights** may be evoked for any changes to the content or format of a resource, although they are more likely to be of issue for creative rather than scientific works.
- The **database right** is included here even though infringement of it is highly unlikely in a preservation environment since it only relates to actions taken on portions of the database, and with possible market effects.
- **Anti-circumvention** prohibitions would be infringed any time a file is un-encrypted for the purposes of preservation. Actions that require rendering the file, such as migration, could result in this infringement. In reality, it is unlikely that preservation repositories will have the ability to render encrypted files with permission.
- The **rights management information** laws require that particular data that is included with a work not be lost. Any transformation does provide a risk of that loss.

	copy	adapt/transfor m	moral rights	database protectio n	anti- circumventio n	rights management information
Re-copy to new media	x					
Validate						
Store multiple copies	x					
Compress, de- compress	x	x				
Migrate to current formats	x	x	x		x	x
Store content but	x	x	x		x	x

not format						
Emulate environment			x		x	
UVC	x	x	x		x	x