

Digital Rights Management for Research and Education

December 2, 2004

Geoff Collier
Robby Robson

This paper examines digital rights management from the perspective of the research and education community. It proposes a set of digital rights management requirements for this community and looks at the technologies, standards and community initiatives that are evolving to support these requirements.

Sponsored by MELCOE (Macquarie E-Learning Centre of Excellence)
www.melcoe.mq.edu.au

About the Authors

Geoff Collier (gcollier@eduworks.com) is a senior partner at Eduworks Corporation with many years of involvement in designing and developing student and learning administration systems.

Robby Robson (rrobson@eduworks.com) is president of Eduworks Corporation, Chair of the IEEE Learning Technology Standards Committee and director of the National Digital Science Library Reusable Learning project.

1	Introduction.....	1
1.1	Context	1
1.2	Definition of Digital Rights Management.....	2
1.3	The Need for DRM in Research and Education.....	3
1.3.1	What has changed?	3
1.3.2	Institutional Repositories.....	4
1.3.3	A Barrier to Use	4
1.4	Scope of DRM for Research and Education	5
1.4.1	Objectives.....	5
1.4.2	Rights Management Models	5
1.4.3	The Scope of Automation	7
2	Legal Context and Evolution.....	8
2.1	Copyright Treaties and Laws	8
2.2	Fair Use / Fair Dealing	8
2.3	Copyright Treaties and Laws for the Digital World	9
3	DRM Requirements	11
3.1	Use Cases.....	11
3.1.1	Wide Distribution	11
3.1.2	Limited Distribution	12
3.1.3	Secure Distribution.....	12
3.1.4	Commercially Licensed Research.....	13
3.1.5	Commercially Licensed "Course Packs"	13
3.1.6	Managing Institutional and Federated Repositories	14
3.1.7	Rights Management for Catalog Records	15
3.2	Stakeholder Requirements	16
3.2.1	Author / Researcher	16
3.2.2	User / Reuser	17
3.2.3	Research Institution Digital Repositories / Libraries	18
3.2.4	Publisher.....	19
3.3	Digital Rights Management Processes	19
3.3.1	Define	20
3.3.2	Assign	22
3.3.3	Enforce.....	23
3.3.4	Track.....	23
3.4	General Requirements	24
3.4.1	Standardization	24
3.4.2	Limited Set of Licenses	24
3.4.3	Privacy	25
3.4.4	Persistence.....	25
3.4.5	Support for Diverse Technologies.....	26
3.4.6	Support for Aggregation / Disaggregation	26
3.5	Requirements Matrix	26
4	DRM Technologies / Solutions	28
4.1	Persistent Identifiers	28
4.1.1	The Role of Identifiers	28
4.1.2	ISSN, ISBN	29
4.1.3	URL, URN, URC	29
4.1.4	Handles and Registries	29
4.1.5	Digital Object Identifiers (DOI)	30
4.1.6	Persistent Uniform Resource Locator (PURL)	31
4.1.7	Archival Resource Key (ARK)	31

4.2	Authentication and Authorization	32
4.2.1	Overview	32
4.2.2	Shibboleth.....	33
4.2.3	XACML - eXtensible Access Control Markup Language.....	35
4.2.4	PERMIS.....	35
4.3	Rights Expression	36
4.3.1	The Role of a Rights Expression Language.....	36
4.3.2	Access Control versus Rights Expression	37
4.3.3	A Review of Rights Expression Languages	37
4.3.4	Relation to Other Metadata Standardization Initiatives.....	40
4.4	Tracking and Reporting.....	41
4.5	Protection and Enforcement	41
4.6	Trusted Computing	42
5	DRM Implementations in Research and Education	43
5.1	JORUM	43
5.2	Celebrate.....	43
5.3	eduSource	43
5.4	COLIS	44
5.5	Guth, Neumann and Strembeck.....	44
5.6	RoMEO	45
5.7	Federated Digital Rights Management (FDRM)	45
5.8	Creative Commons.....	45
5.9	AE Sharenet.....	46
6	Technology Patents.....	47
6.1	Patents Affecting DRM	47
6.2	Patent Pools.....	48
6.3	Observations.....	48
	Glossary of terms and abbreviations	50
	References	53

1 Introduction

1.1 Context

Digital content repositories are steadily expanding in number and size across the research and education community. Efforts are also underway to federate groups of repositories, to provide 'one stop shopping' for digital content for researchers and educators. To be successful, federated repositories should agree on compatible approaches to issues such as access management and intellectual property rights.

The Meta-Access Management System (MAMS) project is developing a middleware component that supports the "*integration of multiple solutions to managing authentication, authorisation and identities, together with common services for digital rights, search services and metadata management*" for Higher Education in Australia. (MELCOE, 2004)

This paper is prepared for Macquarie University, the lead university on the MAMS project. Its purpose is to examine digital rights management from the perspective of the research and education community. It proposes a set of digital rights management requirements for this community and looks at the technologies, standards and community initiatives that are evolving to support these requirements.

The current state of technology and practice can be summarized as follows:

- There are established access enforcement technologies for authenticating users and for authorizing or denying access to digital collections, down to the individual item level, based on group membership and other user attributes.
- Newer technologies are beginning to roll out that allow authentication from one institution to be recognized at another, supporting inter-institutional access control for digital repositories.
- Technology exists to enforce restrictions on viewing, copying, printing or forwarding for a piece of digital content in any environment. However, this technology is simplistic from a policy perspective in that it applies the same restrictions or password to anyone accessing the content.
- Languages and supporting technology are beginning to emerge that allow rights management policies to be expressed in machine readable form, and that allow rights to vary by user. These technologies are only beginning to be adopted.
- Intellectual property law underlies digital rights management policies and technologies, and this law is contentious and changing.

The analysis contained in this paper leads to the following conclusions about the actions that should be taken by any organization implementing rights management for digital repositories in a research and education environment:

- Implement robust authentication and authorization technology, including support for inter-institutional access control.

- Implement a persistent unique identifier solution to support the identification and management of digital content within and between repositories.
- Implement a rights management solution that can associate rights policies with content.
- Select technologies that are based on open standards.

1.2 Definition of Digital Rights Management¹

Digital Rights Management can be simply defined as:

The management of intellectual property rights for digital content through digital means

Where:

- **Management** refers to the definition, assignment, enforcement and tracking of rights.
- **Intellectual property rights** refers to permissions to use digital content and the conditions for that use (for example: the right for students to read and copy a research paper from an electronic journal as long as they are enrolled at a subscribing university.)
- **Digital content** refers to content represented electronically (for example: e-books, pdf files, software, video files, and MS word documents.)
- **Digital means** refers to the use of software and electronic data to perform the management functions listed above.

It is also instructive to look at other definitions for digital rights management, which vary depending on the perspective of the organization or person doing the defining.

- From the publishing perspective, Thomson Publishing says that: *"Digital rights management (DRM) systems help protect the copyright of materials by defining how the content can be used. These rights are determined by the publishers."* (Thomson Corporation, 2004)
- B. LaMacchia, an analyst working for Microsoft, says that a digital rights management system must *"project policy, with confidence that the policy will be respected, from the content owner to the remote environment where the content will be used."* (LaMacchia, 2002)
- From the Higher Education perspective D. Norris says that *"digital rights management is about enabling people to both share knowledge and share its control."* (Norris et al, 2003)

The definitions vary depending both on the role of the definer in the content life cycle (producer, consumer or technology provider), and on the nature of the content.

One clear distinction can immediately be seen between DRM in a commercial context and DRM in a research and education context. This is summarized succinctly in a report on DRM prepared for JISC:

¹ Sections 1.2, 1.3 & 1.4 draw heavily on two ECAR research papers (Collier, Piccariello & Robson, 2004a & 2004b)

"DRM in an academic environment should be an 'enabler' not a 'preventer'." (Duncan et al, 2004).

All the discussions in this paper revolve around four basic processes that make up rights management (digital or otherwise):

Define – Determine and express the rights associated with digital content.

Assign – Assign rights permissions and conditions to content users.

Enforce – Enforce the rights assigned to users.

Track – Track and report on the usage of digital content.

This paper is intended to provide a high level understanding of where DRM has been, where it appears to be going, where there is clarity, where there is confusion, and what makes sense to do today. No single paper can cover all technologies, legal considerations or options for managing digital rights. The focus here is on rights management and related technologies that are relevant to research and education.

1.3 The Need for DRM in Research and Education

1.3.1 What has changed?

The digital world puts control over content directly into the hands of the researcher and other end users, with less dependence on intermediaries such as publishers and librarians. The Internet allows researchers to find and retrieve content from many sources. More digital repositories are being developed and deployed, and repositories are harvesting metadata from one another, with the aim of providing powerful federated search capabilities that make this search and retrieval process faster, easier and more effective. Once digital content is brought into the end user's personal computing environment, they can easily copy this content, and share it with others directly or by posting it on the web, either in its original form or in a modified form.

In the world of paper-based publication, there are effective business models and physical processes that support rights management for research and education. In the non-digital world, the physical manifestation of content (a book or an article in a journal) acts both as a rendering of the content and as a physical token of exchange for the rights associated with the content. The content is a 'work' and the physical book is an 'expression' of the work (IFLA, 1998). Two fundamental changes affecting rights management occur when the work is expressed in digital rather than physical form.

First, a book or journal article can only be read by one person at a time, therefore effectively enforcing a 'single user' license. Physical copies could be made, but the quality of those copies will not be as good as the original, and businesses that provide copy services will refuse to make multiple copies of a book or journal article in the absence of documented permission. Digital content, on the other hand, can be copied with perfect quality and these copies can be easily transmitted to hundreds or thousands of others through email lists and web download sites.

Second, the physical existence of a book or journal allows it to be tracked as it moves from publishers through bookstores and libraries to the individual user. The physical nature of

the book supports this tracking and control even in the absence of automated purchase and inventory systems. This same level of tracking of individual copies of content is not available yet in the digital world, and there are no easy ways to control the 'number of digital books on the shelf'.

In the digital world, wide distribution of digital assets can be carried out by one person sitting at their desk. There are virtually no barriers to participation in a network of digital content distribution, and therefore a rights holder cannot count on trusted business relationships to prevent the unauthorized use and distribution of their content.

1.3.2 Institutional Repositories

Another key trend for research and education is the establishment of digital repositories supported by research institutions or by consortiums.

Digital repositories can provide a simple and fast method for scholars to share information with other scholars, without requiring personal inquiry and passing of information to occur from individual to individual. This allows scholars to have faster, more complete access to a wider range of work for use in both research and teaching. Refereed scholarly journals are an important, high quality venue for academic communication, but they are by necessity a channel of limited size, and they add delay and cost to the process of sharing information. Institutional repositories support a networked method of academic communication.

Coupled with the expansion of digital repositories is a move towards self-publishing. Some organizations and individuals believe that the dysfunctions in formal scholarly publishing have reduced dissemination of scholarship and constrained libraries. This view of the world sees scholarly journals as having a monopoly on the publishing of academic research, leading to ever increasing prices, forcing universities and colleges to 'buy back' their own research, and putting a chokehold on idea sharing. The adoption of institutional repositories is seen as a required first step in the development of a new and 'frictionless' academic communication model that eliminates knowledge cartels, and enables knowledge sharing directly between institutions, faculty and students.

Digital repositories fill the role of distributor and publisher. As such they need to be able to allow authors to define licenses for the material they publish, and they must make users aware of these license terms. They may need to provide methods for enforcement of license terms and conditions. To support self publishing in place of existing academic journals, they also need to support varying subscription models that may or may not be fee-based.

"The management of rights for digital materials will be essential. ... we need methods of recording and documenting the rights and permissions associated with works that facilitate the goals of the research and education community." (Lynch, 2003)

1.3.3 A Barrier to Use

Digital content and repositories hold great potential for supporting a 'revolution in knowledge sharing' (Norris et al, 2003), but the lack of agreed upon and effective approaches to digital rights management act as a barrier to the effective sharing of that knowledge (Robson, Muramatsu & Collier, 2004).

Some authors and other rights holders are hesitant to make all their intellectual property available in digital form if they cannot effectively define their rights to the content and

enforce conditions for its use. Users of digital content need to be aware of the conditions for using, reusing and modifying content. The lack of easily accessible expressions of these rights and conditions makes it difficult for researchers and educators to determine what is allowed, and what is not, outside of general legal constructs such as copyright laws and the right of fair use.

The research and education community can take a large step to removing this barrier by agreeing to use a small number of standard licenses to grant usage rights to researchers and educators. The key is for the community to agree on this small number of standard sets of terms and conditions. Once these have been defined, it is a simpler problem to determine what tools and techniques should be used to express, communicate and exchange rights data.

1.4 Scope of DRM for Research and Education

1.4.1 Objectives

The objective of effective digital rights management solutions for research and education is to reach a state where:²

- Copyright holders can express the rights they hold on digital content, and can express permitted uses for that material and conditions for that use.
- Copyright holders can expect that these rights and conditions for use will be respected.
- Researchers, educators and students know who holds the rights to digital content, what the terms of use are, and what they can and cannot do with the content.
- Where there are restrictions on use, or licensing requirements prior to use, these restrictions and requirements are respected.
- Digital repositories are able to provide easy access to large quantities of content with confidence that they are supporting the rights of copyright holders and the needs of their users with regard to rights.
- Content can be shared among researchers, educators, repositories, institutions and countries in a way that effectively supports these objectives.
- The use of content can be tracked for the purposes of attribution, evaluation and collaboration.

1.4.2 Rights Management Models

Before moving ahead with implementation of rights management processes and technologies, the research and education community needs to reach agreement on what rights it intends to manage, and what operational models it intends to support.

"Before a community can implement [rights management] it must identify and agree upon the underlying market and intellectual property management models. Market models might include retail and wholesale models, public funding models, free distributions models, and

² These objectives are derived from a number of sources, including the CETIS report on DRM (Duncan et al, 2004), a paper on the Federated Digital Rights Management initiative (Martin et al, 2002a), and several others.

federations and cartels. Property management models might include centralized and decentralized control and client / server, distributed networks, and peer-to-peer networks. Each model and management approach has commensurate rights management and tracking requirements." (Collier, Piccariello & Robson, 2004b)

Establishing a rights management policy requires identification of the type of digital content that the community intends to manage and what rights models the community will support for different types of content. Once agreed upon, these policies need to be clearly defined in procedural, technological and legal terms.

Commercial Models

The purpose of creating intellectual property is not the same in the research and education community as in the commercial world, and therefore the methods and technologies required to support these goals are not the same.

In commercial implementations, the default assumption is that there is no right to access or use a resource, a license must be negotiated first. In research and education, the default assumption for most content is that anyone can access the content and use it within the legal parameters of copyright law and fair use, unless there are explicit additional permissions or constraints specifically attached to the content.

In the commercial world, digital rights management focuses on revenue-based licensing. The assumption of much commercial DRM work is that without enforcement through legal and technological means, revenue will not be collected because content will be stolen and shared indiscriminately through the Web. Therefore, commercial DRM systems typically focus on tightly controlled packaging, encryption and enforcement that prevents unauthorized use, copying or redistribution. In research and education, DRM solutions focus on making content widely accessible in distributed, collaborative environments.

The commercial model also assumes a formal, sequential content life cycle. The roles of author, publisher, distributor and consumer are performed by different people in separate, distinct organizations. It is also assumed that the consumer uses, but does not alter, the content. In the research and education community the situation is different. There, digital content is typically part of an ongoing, collaborative process where the people and organizations involved are not defined in advance, and content is constantly being reworked and incorporated into new intellectual creations.

The focus of digital rights management in a commercial context is therefore on licensing, distribution, and the protection of content from unlicensed use. Commercial licensing requirements cannot be ignored by the research and education community. They are needed to support the requirements of revenue-based publishers, who are a part of the community. However, the commercial approach is not the only model of intellectual property management, and it is not the one that is most central to the needs the community, particularly in the area of research and collaboration.

Research and Education Models

Research and education activities require policies, standards and technologies that support the sharing and reuse of digital resources, rather than limiting that sharing and reuse. Effective rights management is needed to enable sharing while respecting the conditions and requirements that rights holders have associated with their content.

The goal of most researchers in academia is to share their intellectual property as widely as possible while receiving appropriate attribution whenever and wherever their works are used. Attribution is the lifeblood of the academic and research world. What they want to prevent with digital rights management is plagiarism and non-attribution.

Besides these differences in objectives, there are also major differences in the commercial and academic content life cycles. Firstly, content in research and education is often modified by the 'consumer', and is incorporated into new content that contains intellectual property from multiple sources. Secondly, researchers and their organizations act as producers, distributors and consumers of digital content. These roles are not separate and distinct.

1.4.3 The Scope of Automation

The automation of rights management for research and education can only succeed if the scope of that automation is closely controlled and carefully justified. Overly complex models do not work in early implementations.

Rosenblatt, Trippe & Mooney (2002) point out that not everything is suitable for automation:

"DRM will be of the most value in an environment where there are simple, small and frequent transactions involving the use or exchange of intellectual property. If the transactions are infrequent then automation is not cost effective. If the transactions are overly complex, then automation may not be feasible. If the transactions are more suitably handled via traditional negotiations and contracts, then automation is not called for. Not all types of rights transactions are appropriate for automation."

The IT infrastructures of research and education institutions can provide access control and security within their own environment. This technology can be used to manage access to digital content stored and used within this IT environment. However, content sharing in academia crosses the boundaries of technology environments and organizations to form highly distributed systems that can span the globe. Content may be distributed to any number of delivery media, platforms or environments. Therefore, the terms and conditions associated with digital content should be 'persistent' in the sense that they should be available in human and machine readable form whenever and wherever the content is used or distributed for use.

In the commercial world, the automation of rights enforcement is important, and needs to persist all the way down to the user environment, to protect content from unpaid use or redistribution.

In the research and education community this type of automated enforcement technology is seldom needed once content is distributed to a user. Enforcement more typically occurs at the point of distribution, where a decision is made on whether or not to grant the researcher or educator the right to access the content.

2 Legal Context and Evolution

2.1 Copyright Treaties and Laws

Providing the underlying international context for copyright laws, the Berne Convention of 1886 developed an international agreement for the Protection of Literary and Artistic Works. (Berne Convention, 1979) This convention evolved partly out of the refusal of foreign exhibitors to attend the International Exhibition of Inventions in Vienna in 1873, because they were concerned that their ideas would be stolen and used in other countries without any recognition or return to the creator of the work. Under this international agreement, a copyright comes into existence upon creation of a work and is recognized by all countries that ratify the treaty, regardless of the country of origin. The copyright may be held by individuals or corporations, or there may be joint ownership.³

The World Intellectual Property Organization (WIPO) is the part of the United Nations system responsible for administering various treaties on copyright, patents, trademarks, and designs. (WIPO, 2004) The countries who ratify these international treaties agree to use them as the basis for laws developed and enforced in their own countries, in order to provide some international consistency of intellectual property law, and to provide internationally recognized protection for intellectual property, no matter where it is created. Once a treaty is developed, each country still has to go through the process of ratifying a treaty, and of developing statutes in their own legal systems which reflect the conditions of the treaty as they interpret them.

Individual countries also address the issue of copyrights and patents in their legal systems. For example, in succinct prose, the U.S. Constitution lays out the underlying authority for the Copyright Laws of the United States:

The Congress shall have Power... To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries (United States Constitution, Article I, Section 8). Subsequent laws have not again achieved this level of brevity.

In Australia, the Copyright Act of 1968 and subsequent amendments form the legal context for digital rights management (Australian Copyright Council, 2004). This act deems that copyright exists at the moment of creation, and that copyrighted material cannot be reproduced or redistributed without permission of the copyright holder. Most countries have similar acts controlling copyright in their legal systems.

2.2 Fair Use / Fair Dealing

For research and education, an important exception to these general rules is the 'fair use' or 'fair dealing' exemption. The Berne Convention provides general guidelines for these exemptions, and requires attribution, but leaves it to individual countries to enact the specifics. (Berne Convention, 1979)

"Article 9 (2)- It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works 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."

³ For background on international intellectual property law, see (LexMercatoria, 2004).

Article 10 - Certain Free Uses of Works: 1. Quotations; 2. Illustrations for teaching; 3. Indication of source and author

(1) It shall be permissible to make quotations from a work which has already been lawfully made available to the public, provided that their making is compatible with fair practice, and their extent does not exceed that justified by the purpose, including quotations from newspaper articles and periodicals in the form of press summaries.

(2) It shall be a matter for legislation in the countries of the Union, and for special agreements existing or to be concluded between them, to permit the utilization, to the extent justified by the purpose, of literary or artistic works by way of illustration in publications, broadcasts or sound or visual recordings for teaching, provided such utilization is compatible with fair practice.

(3) Where use is made of works in accordance with the preceding paragraphs of this Article, mention shall be made of the source, and of the name of the author if it appears thereon."

Again using the U.S. as an example, the fair use exemption to the U.S. Copyright Act allows for the use of copyrighted works for teaching and scholarship without permission of the rights holder. However, this exemption is more restrictive than many people understand. For example, reuse of the 'heart of a work' is not allowed without permission, and fair use does not apply if it would affect the economic value of the work (prevent sales of a book or journal article for instance). Fair use allows the reuse without permission of only portions of a digital resource and does not likely apply in a case where a researcher or educator wants to incorporate an existing resource into their own content.⁴

2.3 Copyright Treaties and Laws for the Digital World

In the years since 1886 there have been a series of international treaties that address specific areas of concern or respond to developments in commerce and technology.

The Berne Convention Treaty itself has been amended several times (most recently in 1979) to extend the terms of the treaty to new technologies, as well as to extend the term of copyright protection.

WCT and WPPT

In 1996 the WIPO Copyright Treaty (WCT)⁵ and the WIPO Performances and Phonogram Treaty (WPPT)⁶ were adopted, and have resulted in the update of copyright laws in many countries. These two treaties update provisions for the international protection of intellectual property to reflect the capabilities of technologies such as the internet and personal computing. Major (and controversial) components of these treaties include:

- Agreement that circumventing electronic protection schemes is illegal.
- A ban on removing or altering rights information.

TRIPS

The Agreement on Trade-Related Aspects of Intellectual Property Rights was created by the World Trade Organization in 1994 (WTO, 2004). It covers much of the same 'territory' as

⁴ For background on fair use in U.S. law, see (Stanford, 2004) and (NCSU, 2004).

⁵ WIPO Copyright Treaty, <http://www.wipo.int/treaties/en/ip/wct/index.html>

⁶ WIPO Performances and Phonograms Treaty, <http://www.wipo.int/treaties/en/ip/wppt/index.html>

the WCT, and its copyright provisions are based on the Berne Convention. A report from the Foundation for Information Policy Research (FIPR, 2001) characterizes TRIPS as a strategy led by the U.S. and Europe to move control of intellectual property treaties away from WIPO. They indicate that this agreement is binding on all WTP members and wields the threat of trade sanctions for non-compliance.

Digital Millennium Copyright Act (DMCA)

A U.S. law that is particularly important to digital content is the Digital Millennium Copyright Act of 1998 (US Copyright Office, 1998) which supports the anti-circumvention agreements in the WIPO treaties. Passage of the DMCA was backed most strongly by publishers and the recording and movie industry, whose businesses depend on generating revenue from the sale of intellectual property. Other groups opposed many parts of the DMCA, arguing that it goes too far in favor of protection, and does not allow enough exceptions for the legitimate use of intellectual property and technologies for distributing intellectual property. For example, circumvention of protection technologies is not allowed, even if the use itself would be allowed (under fair use exceptions for example).

European Union Copyright Directive

The EU intended the Copyright Directive (EU, 2001) to clear up differences in copyright laws between its members, and to respond to changes in technology and international markets. Some extracts from the act highlight its impact on digital rights management for research and education:

"Any harmonisation of copyright and related rights must take as a basis a high level of protection, since such rights are crucial to intellectual creation.

A common search for, and consistent application at European level of, technical measures to protect works and other subject-matter and to provide the necessary information on rights are essential

... exception in the case [of]... use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author's name, is indicated, unless this turns out to be impossible and to the extent justified by the non-commercial purpose to be achieved"

Criticisms of the EU directive are summarized by this quote from the Foundation for Information Policy Research report on its implementation (FIPR, 2001) *"As with the development of the Digital Millennium Copyright Act in the US, rights holder organizations saw the opportunity for a second chance to push proposals that had been rejected by the diplomatic conference that led to the WIPO treaties. In particular, they were keen to reintroduce the detailed anti-circumvention rules previously proposed by the US but rejected in favour of the simpler and more flexible language of Articles 11 and 18 of the final WIPO Copyright and Performance and Phonogram treaties. The developing nations who forced the WIPO compromise were missing from this second round. Copyright user organizations were also grossly underrepresented compared to the number of industry lobbyists in Washington DC and Brussels. Unsurprisingly therefore, both the EU and US ended up with legislation specifically outlawing acts of circumvention and circumvention devices, rather than concentrating on acts of copyright infringement."*

3 DRM Requirements⁷

Before considering what DRM solutions should be put in place, a research and education community must determine which DRM requirements they intend to support. This section takes a broad view of potential requirements, and then collects them together at the end as a brief list of high-level requirements, along with a prioritization of these requirements. Conclusions about these priorities are based on the considered opinions of the authors of this paper, and relate to the priorities of these requirements for the MAMS (Meta-Access Management System) project being led by MELCOE (MELCOE, 2004).

3.1 Use Cases

The following are some use cases that are representative of the research and education environment. For each use case, we also describe the 'ideal' DRM functionality that would support the use case for each major stakeholder. (Note that the implementation of all 'ideal' capabilities may not be feasible or desirable for all communities.)

3.1.1 Wide Distribution

A researcher authors a paper on geology and also creates Java™ applets that visually demonstrate some of the principles discussed in the paper. The researcher makes the paper and the applets available to anyone who wants to use them, and is willing to allow others to create derivative works from her work as long as she is given proper attribution and as long as these derivative works are shared with others under the same terms. The researcher posts the paper and the applets to her personal web site.

Ideal DRM Functionality:

Researcher:

- Define terms of use that allow anyone to read, reuse and repurpose the content, as long as they properly attribute the source of the work.
- Provide an attribution statement.
- Require that derivative works are made available to others under these same terms of use.
- Track the use of her work by others.

User:

- Access rights information and terms of use for the paper and the applets.
- Accept license conditions.
- Provide proper attribution when reusing or modifying the paper or applets.
- For derivative works, define the same terms of use as the original creator.

⁷ The requirements described in this section are distilled from several sources, most notably - (Byrne, 2002), (Collier, Piccariello & Robson, 2004a & 2004b), (Duncan et al, 2004), (Martin et al, 2002b), (Rosenblatt, 1997), (Rosenblatt, Trippe & Mooney, 2002), and unpublished documents from the IEEE LTSC DRM working group.

3.1.2 Limited Distribution

A PhD student is finishing her research. She would like to make her thesis publicly available as a PDF document on the university's open repository. However, she also would like to store some other artifacts there including the word document used to create the PDF, some java code that she created, plus the source data that she used. She only wants to make these other items available to her colleagues in the department.

Ideal DRM Functionality:

Researcher:

- Define terms of use for the PDF document that allow anyone to download, read, print or forward the document, as long as they properly attribute the source of the work.
- Define terms of use for the other items that restrict access to faculty members and PhD students in a specific department at her school.
- Provide an attribution statement.
- Track the use of her work by others.

Users Outside the Department:

- Access rights information and terms of use for the PDF.
- Accept license conditions.
- Provide proper attribution when reusing all or part of the thesis.
- Do not see the existence of the other artifacts (not returned by a search of the repository)

Users Inside the Department:

- See the existence and terms of user for the other artifacts, based on authentication of their position as students of faculty in the authorized department
- All the same capabilities as outside users

3.1.3 Secure Distribution

A psychology professor conducts a human subject study on the personality effects of long term addiction to a prescription drug. The professor creates a database of results and a paper describing the study and the results. The research is sponsored by a pharmaceutical company that owns the rights to the research data. The company wants to release the data to other researchers in the field as long as they agree to keep the report and database content confidential.

Ideal DRM Functionality:

Researcher:

- Provide an attribution statement.
- Receive attribution for his work.
- Track the use of his work by others.

Company:

- Allow the data to be accessed only by legitimate research psychologists, and only if they agree to the terms of confidentiality.
- Protect the data and the report from being forwarded to anyone else.
- Protect the data and the report so that only authorized researchers who accept the terms of use can access it.

- Track the use of the work by others.

Other Researchers:

- Access rights information and terms of use for the paper and the database.
- Accept (or decline) terms of use (note that individuals not recognized as legitimate psychology researchers cannot access at all).
- Once the terms of use are accepted, be able to access and view the report and database.
- Report usage back to original researcher.

3.1.4 Commercially Licensed Research

A university subscribes to a commercial research service, and buys a license that makes the content available to all faculty members and research assistants, but not to students or the administration.

Ideal DRM Functionality:

Researchers (employees of the research service):

- Provide an attribution statement.
- Receive attribution for their work.

Research Service:

- Define a license that allows the research reports to be accessed only by faculty and research assistants at the subscribing university.
- Protect the reports from being forwarded to anyone else.
- Encrypt the data and the report so that only researchers with an authorized license can access it.
- Track the use of research reports.

University:

- Track the use of research reports (to evaluate activity and benefit to the university.)

Faculty and Research Assistants:

- Access license information for the research reports service.
- Access and view (but not forward or copy), reports in the research database.
- Report usage to university and research service.

3.1.5 Commercially Licensed "Course Packs"

A college bookstore acquires online "course packs" from a textbook publisher. The course packs complement and support the textbook being used for a particular course. Course packs are loaded directly into a campus course management system through which students access them. (This scenario is from Collier, Piccariello and Robson, 2004a).

Ideal DRM Functionality:

Publisher:

- Define a distribution license that allows the university to redistribute the course packs to a predefined number of students.
- Protect the course packs from being copied and shared with other students.

- Track the number of uses of the course packs, and sell additional licenses to the university when the limit is reached.

University:

- Define a usage license that allows students to access and use the course pack materials as part of their class.
- Protect the course packs from being copied and shared with other students.
- Track the number and type of uses of the course packs.

Student:

- Access license information for the course pack and accept (or decline) the terms of use.
- Access and use (but not forward or copy), the course pack.

3.1.6 Managing Institutional and Federated Repositories

A university's institutional digital repository gathers digital content from educators and researchers at the university, and supports the cataloging of these materials. The institutional repository houses some digital content itself, and also points to content housed on various departmental and special purpose servers. The repository provides an interface that permits users to find content based on keywords, subject classifications, technical requirements and other characteristics.

In addition, the repository allows federated digital libraries to "harvest" some of its catalog records, although some are designated as 'internal use only'. These harvested records are made available in the external federated digital libraries where they can be searched by researchers, educators and the general public through the open web.

Ideal DRM Functionality:

Author / Researcher:

- Select from a standard set of licenses that lay out different terms of use for the content they contribute to the repository.
- Specify the audience that is allowed to access the content (for example - departmental only, university wide, or unlimited)
- Provide an attribution statement.
- Receive attribution for their work.
- Track the use of their work.

Repository:

- Define a set of standard licensing options.
- Restrict access to posted content as determined by the Author.
- Support searching based on different terms of use (for example – show only material that allows derivative works to be developed)
- Display rights statements and terms of use, and require acceptance of the terms prior to granting access to the content.
- Expose rights information as part of the content metadata, for that content where the author granted the right to share the content outside the university.
- Track the use of repository content.

Researchers, Educators and Students Accessing the Repository:

- Access license information for repository content and accept (or decline) the terms of use.
- Download content to their personal computing environment.
- Display license restrictions every time the content is opened, and have personal computing technology that automatically enforces those restrictions.
- Provide proper attribution.

Other Repositories Federating with This Repository:

- Harvest metadata, including rights metadata.
- Support rights management capabilities (requirements are similar to those for institutional repository)
- Report usage back to the source repository.

3.1.7 Rights Management for Catalog Records ⁸

A digital library reviews and maintains catalog records of material for K-12 mathematics and science teaching. The materials come from a variety of sources including

- Individual authors or schools
- Commercial publishers
- Educational outreach programs sponsored by agencies such as NASA
- Educational programs produced by public and commercial television
- Catalog records from other educationally oriented digital libraries

The records maintained by the digital library point to original sources from the material; with rare exceptions the digital library does not “house” any content itself.

Teachers, students and school districts use the digital library to find professional development material for teachers and content for students. The content is used in both purely online and traditional classroom settings. In addition, the digital library allows other digital libraries to “harvest” its catalog records, just as it harvests records from them.

The digital library employs professional catalogers to review all materials in the library, to create summary descriptions of the material, and to generate the information needed for searches. This information is called *metadata* and it constitutes valuable intellectual property that the digital library has generated through the application of their resources. If teachers, parents and students know that a review or classification came from this particular digital library, they can assume that it is authoritative and accurate.

The digital library would like to appropriately protect its own intellectual property (the metadata records) in order to maintain “branding” and prevent misuse or misrepresentation of the reviews and metadata it has created.

Ideal DRM Functionality:

Repository:

- Define a simple license for the use of metadata records that allows use under the conditions of:
 - proper attribution

⁸ This is a real-life scenario based on issues faced by the Eisenhower National Clearinghouse (www.enc.org). For an example of an ENC record, see <http://www.enc.org/resources/records/full/0,1240,025708,00.shtm>.

- retention of branding
- no modification
- reporting back on usage
- Track the use of metadata records

Researchers, Educators and Students Accessing the Metadata record:

- See the attribution and branding of the metadata record

Other Repositories Harvesting Metadata Records:

- Harvest rights statement for the metadata along with the metadata record
- Provide proper attribution and retain branding
- Report usage of the metadata record back to the digital library

3.2 Stakeholder Requirements

This section summarizes digital rights management requirements and expectations from the perspective of different stakeholders in the research and education community.

3.2.1 Author / Researcher

The 'Rights Metadata for Open Archiving' project, funded by JISC, undertook a survey which gathered responses from over 500 academics in 57 countries. A summary of the key findings of the RoMEO Project (Gadd, Oppenheim & Proberts (2003b) reports that *"The majority of respondents (60% or more) were happy for others to display, print, save, excerpt from, and give away their research papers as long as the respondents were attributed as the authors, and that all copies were exact (verbatim) copies of the original work. Most respondents wanted to prohibit sales of their works and 55% wanted to limit usage of their works to certain purposes, e.g., educational or non-commercial. A comparison between these usage limits and those provided by UK copyright law and many electronic journal licences showed that the academics' conditions were far more liberal."*

Another study carried out by the Authors' Licensing and Collection Society (ALCS) found that writers' primary concerns were *"moral rights, accreditation and retaining integrity of the work. Modification of work was clearly the area which authors showed most concern about. ... In particular very few were in favour of editing of words and phrases within text. ... the majority were not concerned about altering those aspects which have less effect on the integrity of the work such as fonts, formats and colours."* (Duncan et al, 2004)

Based on these studies it is possible to summarize research and education author's DRM top requirements are as follows:

- Recognition as the creator of a work
- Proper attribution when the work is cited or reused
- Protection of the integrity of the work
- Specify whether or not derivative works can be created from the original work
- Require that derivative works be distributed on the same terms as the original work

- Restrict the way their work can be used and redistributed (commercial vs. non-commercial)
- Tracking the use of their work

If the work includes software, authors can choose to make source code available or not. If reusability is the goal, then the code should be made available. For reusability, licenses that allow modification are essential. It is a poor strategy to rely on exemptions to copyright laws. A better alternative is to explicitly attach rights and conditions to a resource that allow the desired type of reuse.

3.2.2 User / Reuser

Just about anyone is a potential user of research and education content. The term 'reuser' (from Robson, Muramatsu & Collier, 2004) refers to people who want to either adopt the content for use in other contexts (such as a class), or who want to modify (adapt) the content in some way.

These users need to know the answers to the following questions:

- What is in the content, and am I interested in using it? Is the content authentic?
- Am I allowed to access the content, and if so what are the terms of the license that grants me that access?
- Can I use the content for the purpose I have in mind?
- Am I allowed to modify the content to create derivative works?
- How do I license the content?
- How do I access the content?
- What are the attribution requirements, and what attribution should I use?
- Can I access the content anonymously, so that my privacy is protected?

An effective DRM solution must enable a user of content to quickly and easily resolve these questions.

The most interesting intellectual property rights issues arise when someone wants to modify and reuse a digital resource. Anyone in that position hopes that the work is licensed in a way that grants permission to modify and redistribute. If it is not, they must make judgments about fair use and possibly seek permission. If payments or conditions apply and the resource is being combined with other resources, there may well be unanswered questions concerning what rights and fees apply to the new aggregation and as to whether a new (or derived) work has been created. Derived works have their own copyright and terms of use.

The ability to modify content also requires that the work be made available in a modifiable format, perhaps as a rich text file rather than just as a pdf, or that the source code be available as well as the executable.

3.2.3 Research Institution Digital Repositories / Libraries

Digital repositories and libraries run by research institutions collect resources from multiple sources, catalog them with relevant metadata that supports effective search and use of the collection, and provide services intended to make it possible for contributors and users to interact efficiently with the repository.

As part of these responsibilities, digital repositories have to address the management of digital rights. These functions can be summarized as:

- Define the rights models that will be supported by the repository (many repositories will not support fee based license models, for instance)
- Provide contributors with a method of defining rights and terms of use for the content they deposit in the repository
- Expose the terms of use and rights information to repository users, and provide them with a method of accepting or rejecting usage terms
- Protect content from unauthorized use, including protecting paid subscription databases from access by unlicensed, external users
- Expose rights metadata if the collection is federated with other collections
- Import or harvest rights metadata from other sources
- Export rights expressions along with content when the content is downloaded by a user
- Define rights and permissions for the use of metadata (catalog records)
- Protect paid subscription databases from unauthorized access

Note that this list does not specify the method by which any of these requirements should be achieved. The solutions may be procedural, moral, legal, technological, or a combination of all four.

Digital Libraries have a responsibility to provide an environment that supports appropriate rights management. As an intermediary in the digital content life cycle an organization that makes copyright-infringing work accessible online may be held responsible for supporting copyright infringement, even if they have a general statement on their site indicating that depositors must have the right to deposit the content they put on the site.

3.2.4 Publisher

Publishers of electronic journals and other digital content are businesses that require fee based licensing or purchase of content to support their business models. Their basic requirements are to:

- Identify the individual or groups accessing their content, either to confirm that they have an existing license or to assign them a license for use
- Uniquely identify each instantiation of digital content for licensing and tracking purposes
- Define access rights that determine whether a user may view, print, copy or distribute an item, whether he or she is buying, borrowing, or accepting transference of rights, and under what conditions those rights are conveyed (Rosenblatt, 1997)
- Protect content from unauthorized use
- Track the use of digital content
- Process financial transactions related to licensing

3.3 Digital Rights Management Processes

Digital rights management is an integral part of the larger context of content management. Decisions about rights management must be made within this larger context, considering how the life cycle is managed for any particular body of content.

The table below (from Collier, Piccariello & Robson, 2004b) provides a basic model of a content management life cycle and some general rights issues associated with different parts of the life cycle.

Create Content		
Author	Create new content, from scratch or by modifying existing material	Licenses should be defined when the content is created. Supporting technology should be considered. Rights need to be consolidated when content is assembled from multiple sources.
Assemble	Bring together content from multiple sources and assemble them into a more comprehensive work.	
Offer Content		
Publish	Prepare and issue content for public (or institutional) distribution.	Rights should be published, cataloged and distributed together with content. Rights data should be preserved throughout publishing and distribution. Use licenses are different than distribution licenses.
Catalog	Classify and record attributes of content.	
Distribute	Distribute content to the targeted users.	
Acquire Content		
Find	Search for content and discover content that meets the search	Rights information can be used as a search criterion. Rights must be

	criteria.	acquired before content can be used.
Acquire	Acquire access to the content in the format needed to support the desired use.	
Use Content		
Use	Display, interact with and otherwise use content	Rights determine if and how content may be used. Rights may be enforced when content is used.

Content Life Cycle and Rights Issues

This model suggests four groups of rights management processes that support the content management life cycle.

Define – Determine and express the rights associated with digital content.

Assign – Assign rights permissions and conditions to content users.

Enforce – Enforce the rights assigned to users.

Track – Track and report on the usage of digital content.

3.3.1 Define

Determine and express the rights associated with digital content. This process includes the following requirements:

Unique Identification – Each instance of the content is assigned a unique identifier that continues to be associated with the content instance as it moves through the content life cycle in a distributed network environment.

Attribution - The author or other rights holder defines the attribution and copyright statements to be associated with the content.

Offers - The author or other rights holder determines what rights he wants to grant to potential users of the content, and what the terms the users must agree to and/or what conditions they must meet before these rights are granted to them.

- **Standardized Offer Templates** – An effective DRM environment will agree upon and define a small set of standardized contract offers, while retaining the ability to create customized offers for unique types of content.
- **Multiple Offers** - There may be more than one offer defined for the content, although in most research and education repositories it is anticipated that there will be only one offer associated with content.

Rights - Offers may include the following rights:

- **General usage** – The rights to read, listen to, view, or otherwise personally use content as it was intended for an 'end user'.
- **Print** – The right to print one or more hardcopies of an electronic document.

- **Copy** – The right to make one or more digital copies of the content.
- **Modify** – The right to embed the content in a different content item entirely, and the right to edit / modify the content.
- **Distribute** – The right to distribute the content to others.
- **Commercial use** – The right to use the content in a way that generates revenue or the potential of revenue for the user.
- **Usage Volume Limits** - Specification of constraints limiting the number of times digital objects can be used, accessed, distributed, and/or replicated.

Conditions – Offers may require that certain conditions be met before a user is eligible to be granted a usage license. Those conditions identified as relevant to the research community are:

- **Group Membership** – Authors and digital libraries often wish to differentiate the rights they will grant to different classes or groups of users. For example, they may wish to grant modification rights to other researchers in their field, usage rights to any member of the higher education community, and no access rights to anyone outside of that community.
- **User Attributes** – Rights may be granted based on a user’s attributes.
- **Share Alike** – A common condition in the research and education community is that the right to create derivative works is granted on the condition that those derivative works are made available to others under the same terms and conditions defined by the creator of the original work. This approach is widely used in the Open Source Software community.
- **Financial Conditions** – Some content may have a cost, which will need to be reflected in financial requirements as part of the license offer. Different types of financial conditions can include.
 - **One time purchase**
 - **Subscription**
 - **Site license**
 - **Pay per use**
 - **Pay per unit of time**
- **Machine or Address Based Access** – Restricting access to particular IP addresses or physical machines.

Expression Method – The rights, attribution, offers and licenses should be defined in a way that can be shared in a distributed network, and that can adapt to different technologies and standards over time. This requires a structured syntax and a standardized set of codes (semantics) for expressing these rights and offers. This approach supports machine readability, and ‘future proofs’ the expressions because they can automatically reformatted into a different syntax and semantics.

- **Extensible** - Ideally, the coding structure chosen would be able to express all the different licensing options that are currently available, or are reasonably anticipated for the near future. The expression language should be extensible to allow alternative schemas wherever sensible to meet the future needs of the research and education community.

Distribution Licenses – Every time content changes hands, a new license is needed to express the rights of the recipient. As the content is passed along, new sets of conditions and permissions may come into play. For example, a publisher might attach a license to a course pack for use on a PC that grants a school district the right to distribute the course pack, provided they pay a license fee and that no more than one thousand copies are made in total. This is called a *distribution* license. When the district distributes the course packs to individual schools or classes, it might create new licenses that allow a fixed number of copies to be made but that does not have a payment condition. When a copy is downloaded to a student's computer, the right to make a copy might be removed completely. This is called a *usage* license. (Collier, Piccariello and Robson, 2004b)

3.3.2 Assign

Assign rights to participants in the content life cycle. This process includes the following requirements:

Authenticate – Establish the individual identity of the user, or determine that the user has certain attributes or is a member of a specified group.

Authorize – Determine if the user is authorized to access the content collection (or which parts of the collection he is authorized to access).

Search – Expose rights metadata to the search process, to allow the user to search based on rights as well as on other content metadata.

Expose – Expose the terms and conditions of use (the offers) to the user.

Accept License – Require or allow the user to accept the terms and conditions associated with a license.

Financial Transaction – Support interaction with a financial processing service to allow the user to pay for a content license.

Assign License – Assign a license to an individual user.

Package and export license – Package the license and export it to the user's computing environment.

Separate Content from Definitions of Rights, Offers and Licenses - Note that rights and licenses can be defined and acquired separately from the actual content. If a license provides a key for accessing the content, then the license must be able to be separated from the content.

3.3.3 Enforce

Enforce the rights assigned to participants. This process includes the following requirements:

Re-Authenticate – If the user is accessing content that is restricted in any way, then it is necessary to re-authenticate the user each time they access the content. Ideally this will occur at any place or time where the user requests access to the content. To be completely effective, this authentication process will extend to the user's personal environment, allowing the enforcement logic to restrict access to the appropriate user regardless of who else may have access to the user's computing environment.

Access and Interpret License – Access the usage license each time the user accesses the content, and interpret the terms and conditions of use.

Grant Access – Grant or deny access to the content based on the license. Physical protection can be provided through encryption technology that requires a de-encryption key provided by a license. Persistent protection solutions can be provided by always distributing content in secure files that require a license to open.

Enforce License Restrictions – Enforce restrictions on the use of the content once access is granted (such as restrictions on copying, printing or distributing).

For much of the content created by the research and education community there is little need for automated enforcement of detailed rights such as copying, printing or distributing. Some content is restricted to users who belong to particular groups or who have specified attributes, but these enforcement decisions are made at the point of initial user authentication and content distribution. Therefore, once the content is released to the user, 'low-tech' means can be used to express these rights. For example, the user can be presented with a clear statement of licensing terms on the first page of the content, or on a pop-up window that requires them to acknowledge and accept the terms of use before they access the content.

3.3.4 Track

Track and report on the usage of digital content.

Activity Tracking for License Restriction Purposes - Some licenses impose limits on the number of times content is used, the number of hours of use allowed, or the number of people to whom content can be re-distributed. Given that this tracking is strictly for license enforcement purposes, it is possible for this tracking to be contained within the content itself, without the need to access a central tracking service.

Activity Tracking for Financial Purposes - If the DRM system supports licenses that require tracking and charging for 'per use' transactions, then some sort of centralized usage tracking service becomes essential in a widely distributed ecosystem.

In the absence of such a service, licenses that charge per use can be supported in two ways:

- Purchase a license for each individual usage of the content, and have that license restrict the user to a single use.
- Access content through a portal that provides the necessary tracking and financial

processing, and do not allow local access.

Activity Tracking for Reporting and Evaluation Purposes –Information what content is being used, by whom and for what purposes is valuable to content providers and repository managers, even if they do not charge for access to content. Researchers want to know where their work is being used and cited, for a number of different reasons. As an example of this tracking, some electronic journals will send emails to their contributors, indicating how often their work is being accessed.

3.4 General Requirements

3.4.1 Standardization

In a fully implemented DRM ecosystem, all tools that author, assemble, publish, catalog and distribute content would provide their users with the ability to interpret, display and define rights, offers and licenses. This can only occur if there is agreement on standardized approaches to defining, assigning, enforcing and tracking these digital rights management attributes.

Standardization refers both to technology and to business models and business practices. To quote the Gartner web site "*[the] fundamental challenges holding back DRM are a lack of interoperability and business models.*"

Interoperability is the key benefit that standards are intended to provide.

3.4.2 Limited Set of Licenses

As noted in section 3.3.1 above, an effective DRM environment will agree upon and define a small set of standardized contract offers, while retaining the ability to create customized offers for unique types of content. This is a social issue, not a technical issue. Based on the research conducted for this paper, the authors would recommend that a research and education community consider limiting the terms and conditions of license offers to the following options (and less than this set if possible):

Restrict license offers to:

- Members of specified groups
- Users with specified attributes

Allow modification / creation of derivative works:

- Allow
- Allow with the condition that derivative works be offered on the same license terms as the original
- Do not allow

Allow the user to copy and redistribute the work under the same conditions as the original license (respecting limitations on audience):

- Yes, redistribute all or part of the work
- Yes, but only if the work remains intact
- No

Allow commercial use of the work:

- Yes
- Yes, but only for specified commercial purposes
- No

Require users to report back to the rights holder on their use of the work:

- Yes, and specify reporting mechanism
- No

Individual licensing fee

- Yes, with specifics
- No, no individual cost for the specified audience

Note that there is no option for attribution. It is assumed that attribution is always required in the research and education community.

3.4.3 Privacy

Tracking is very important both to researchers and to organizations that fund research. Researchers want to easily track who is accessing their work, and how their work is being used. However, the existence of centralized services that track individuals and their access to digital content runs directly counter to privacy requirements, both moral and legal.

The development of 'trusted' or 'secure' computing environments that control access to content all the way down to individual personal computing devices is also a concern. This technology raises issues related to privacy and the legal right to do what you want with your own property in the privacy of your home and office. An article by Richard Stallman called "The Right to Read" provides an entertaining, but troubling scenario of one possible future for rights management and trusted computing. (Stallman, 2002)

DRM systems must provide privacy safeguards, to prevent against the unauthorized use of personal data. Legal requirements in many countries are that personal data may be collected and used only for the purposes authorized by the person about whom the data is being collected.

Protecting privacy includes designing procedures and systems that do NOT collect data about individuals if it is not needed. DRM systems can support this by authenticating users and authorizing their access to content based on specified attributes or group membership, without gathering personal information. The Shibboleth architecture is designed to support authentication while protecting individual privacy.

3.4.4 Persistence

Persistence refers to maintaining the integrity of rights data and rights processes as content moves from one system to another. Persistence is an important underlying concept that is important when restricted content is supported in a distributed network environment.

3.4.5 Support for Diverse Technologies

We cannot expect researchers, educators and students to all use standardized, controlled and trusted platforms. This group will always operate with a heterogeneous mix of workstations, servers and operating systems. Some participants will use the 'cutting edge' of evolving technology, and others will run on old platforms with 'ancient' versions of tools and operating systems.

The research and education community is increasingly moving towards the use of open source software, for both operating system and application system software. Therefore the DRM solutions chosen for use by research and education must be supported by the open source software community. Also, the adoption of DRM solutions provided by commercial software vendors will be possible only if those vendors include support for the open source software initiatives that are crucial to the research and education community.

3.4.6 Support for Aggregation / Disaggregation

Some researchers and educators may wish to create new digital content that is composed of an aggregation of smaller content objects from multiple sources. Similarly, some digital content may be structured in such a way that makes it suitable for disaggregation into multiple independent parts which can then be used individually or reassembled with content from other sources.

To fully support digital rights management for aggregated content objects, it would be necessary to assign identifiers and associate different rights holders, terms and conditions with multiple parts of a single digital asset.

3.5 Requirements Matrix

The following matrix presents a summarized set of generic requirements for DRM solutions in the research and education community. The matrix is intended as a template for organizations to assign priorities to these requirements for their own initiatives, and to evaluate potential solutions against these priorities.

The priorities filled in on the matrix in this document relate to the priorities of these requirements for the MAMS (Meta-Access Management System) project being led by MELCOE. For a description of this project, see (MELCOE, 2004). The authors assume that these priorities will be different for other initiatives.

This prioritization places requirements into one of three categories for MAMS:

Required - a high priority item

Nice to have - but not high priority

No – Not needed, more negatives than positives if implemented

Requirement	Required	Nice to Have	No
DEFINE			
Unique Identification	X		
Attribution	X		
Offers	X		
Standardized Offer Templates		X	
Multiple Offers		X	
Rights – General Usage	X		
Rights – Print		X	
Rights – Copy		X	
Rights – Modify	X		
Rights – Distribute	X		
Rights – Commercial Use	X		
Rights – Usage / Volume limits		X	
Conditions – Group membership	X		
Conditions - Attributes	X		
Conditions – Share Alike	X		
Conditions – Financial		X	
Conditions – Financial – One time purchase		X	
Conditions – Financial – Subscription	X		
Conditions – Financial – Site License	X		
Conditions – Financial – Pay per use			X
Conditions – Financial – Pay per unit of time			X
Conditions – Machine or Address Access	X		
Expression Method	X		
Expression Method – Extensible		X	
Distribution License		X	
ASSIGN			
Authenticate	X		
Authorize	X		
Search on rights data		X	
Expose to user	X		
Accept license	X		
Financial Transaction		X	
Assign License	X		
Package and export license		X	
Separate content from rights data		X	
ENFORCE			
Re-authenticate		X	
Grant access		X	
Enforce license restrictions		X	
TRACK			
Track for license restrictions		X	
Track for financial purposes		X	
Track for reporting and evaluation		X	
GENERAL REQUIREMENTS			
Standardization	X		
Limited Set of Licenses	X		
Privacy	X		
Persistence		X	
Diverse Technologies	X		
Aggregation / Disaggregation		X	

4 DRM Technologies / Solutions

Many of the necessary components and services that make up a complete DRM ecosystem have not yet evolved into standard, accepted approaches that cross enterprises. The deployment of the technology is in its early stages and operational and business models for DRM are still being worked out in different sectors of the economy. The technology needed and the standards which the technology will be based on are still being developed, and there are unresolved questions about patents that may apply to the use of some technologies. (Barlas, Cunard & Hill, 2003) (Duncan, et al, 2004)

However, there are significant implementations underway in some areas, particularly content identification, authentication, and metadata, which are vital components of effective digital rights management.

Complete 'end to end' DRM solutions are not available outside a few cases in the commercial sector, most notably in the publishing and music industries. For the most part, these are proprietary, integrated solutions. Commercial DRM initiatives for text content include the Digital Asset Server from Microsoft which provided DRM support for customized electronic publications, the Adobe Content Server, and the Palm Digital Media DRM system for distributing eBooks (used by most major online eBook stores).

This section focuses on the technologies where considerable work is underway and that are also directly relevant to the immediate needs of research and education: Persistent Identifiers; Authentication and Authorization; and Rights Expression. It also touches briefly on: Tracking and Reporting; Content Protection; and Trusted Computing

4.1 Persistent Identifiers

4.1.1 The Role of Identifiers

Persistent, unique identifiers for digital content are a part of a distributed digital rights management solution, where content needs to be dependably identified across multiple environments. These identifiers allow different components of the DRM environment to reliably locate and exchange both the content itself and other artifacts related to that content, including digital rights management information. A common identification design and set of services is particularly important when managing digital rights in a highly distributed research and education environment.

A number of initiatives are underway to design and roll out a persistent unique identifier solution. Two of the most well known in research and education are the Digital Object Identifier system (DOI, 2004) and the Persistent Uniform Resource Locator (OCLC, 2004). For a good overview and a collection of general references about identifiers, see (PADI, 2002).

Persistent, unique identifiers may also be used when it is important to identify the individuals and organizations involved in DRM processes. For example, a license may need to be associated with a person or organization, and may need to link to contact information. A distributed control system has to have a unique identifier for each party that can be resolved by different components of the distributed infrastructure. For privacy purposes, these identifiers and the related may not be shared or have any meaning outside of the context of a specific license.

It is important to note that there are many situations in research and education where it is not necessary to identify specific individuals. Attributes about a user are often sufficient to determine what rights they have without the need to determine their specific identity.

4.1.2 ISSN, ISBN

The need for persistent unique identifiers is not new to the publishing world. The ISBN (International Standard Book Number) is a “unique machine-readable identification number, which marks any book unmistakably.” It is administered in a coordinated fashion by a number of different organizations in 159 countries. The ISBN is only used to identify complete books, so it is not a sufficient solution for research and education’s digital needs.

The ISSN (International Standard Serial Number) is an eight-digit number which identifies periodical publications as such, including electronic serials.

Although neither of these number standards are a complete solution, it is important to be able to incorporate them into digital solutions.

4.1.3 URL, URN, URC

A Uniform Resource Locators (URL) specifies the *location* of a resource on the web by including a protocol, domain name and the actual name of the file within which the resource resides. This is not a satisfactory means of persistently identifying a digital resource. The URL simply points to the current location of the resource. If a resource is moved to a new location, the previous URL is no longer useful and links to the resource that are embedded in other documents or databases also become redundant. Also, a URL does not exist for all digital objects because not all digital objects reside on the web.

A Uniform Resource Name (URN) is intended to serve as a persistent, location-independent, resource identifier. To find the location (URL) of a digital resource based on a URN, a service is required to look up the URL based on the URN. All URNs will include a Namespace Identifier (NID) code and a Namespace Specific String (NSS). The NID indicates the identification system being used for the URN and facilitates the interpretation of the NSS. The NSS is the local code that identifies the individual document.

Uniform Resource Characteristics (URCs) refer to encoded information about digital resources. The DOI and ARK initiatives are developing systems to include this information.

(Scharf, 2002) (W3C, 2004)

4.1.4 Handles and Registries

Systems designed to support the URN concept, typically involve ‘handles’ and ‘registries.’ A registry is a service that registers digital content, assigns a number and stores information on the content, including the location from which it can be accessed. A handle is a pointer to a registry that includes the content identifier provided by the registry. Registries can potentially associate rights metadata with content.

Descriptive and location information is recorded in the registry by the creator or publisher of content. If the location or status of the content changes there is only one place where the information needs to be updated, and distributed links to the content through the handle will not be ‘broken’ when the location of the content changes.

The CORDRA initiative within Advanced Distributed Learning has plans to create registry systems for learning objects (LSAL, 2004).

The Corporation for National Research Initiatives (CNRI) has developed a Handle System which conforms with the URN framework. The Handle System is "a comprehensive system for assigning, managing, and resolving persistent identifiers, known as 'handles', for digital objects and other resources on the Internet" (CNRI, 2004) The Handle System is the technology which underlies the DOI and PURL initiatives.

4.1.5 Digital Object Identifiers (DOI)

The most widely known initiative in the research and education community is the Digital Object Identifier (DOI) system, which is composed of three parts, the DOI itself, a DOI Registration Agency, and a DOI Registry Database.

A Digital Object Identifier has two parts. The prefix is an assigned number that identifies the organization that registers identifiers. The suffix is an identifying key assigned by registration organization. This approach allows the DOI initiative to support the decentralized assignment of IDs and maintenance of registries, while retaining global uniqueness for the complete identifier. Existing identification systems such as the ISBN or ISSN can therefore be incorporated into a DOI scheme.

The DOI is also not limited to representing an entire work. DOI's may be assigned to individual articles, chapters and other sub-components of content such as individual images.

The Digital Object Identifier (DOI) system was initiated by the Association of American Publishers, and it is now administered by the International DOI Foundation (DOI, 2004).

Paskin explains that the Digital Object Identifier (DOI) infrastructure includes:

- an open standard describing a syntax for unique identifiers (ANSI/NISO)
- a technical implementation infrastructure, based on persistent identification and resolution, using the Handle system
- a "social" implementation infrastructure covering governance and policy
- the ability to associate metadata and services with identifiers using the extensible indecs ontology

Paskin (2003) reports that there is considerable adoption of the DOI approach with "*ten million DOIs assigned from several hundred organizations through a number of Registration Agencies in USA, Europe, and Australasia, supporting large scale business uses.*"

This article also reports that a number of communities beyond English language publishing are adopting the service, including national libraries, government agencies, and non-English language markets in France, Germany, Spain, Italy and Korea. Most of these adoptions are supporting text-based content, but other sectors are discussing implementations.

These developments greatly increase DOI's relevance and applicability to the research and education community.

However, Dalziel (2003b) explains that there is a significant cost barrier to adopting DOI in the research and education community. "*DOIs are currently provided on a variable pricing model (ie, it costs money for each DOI) rather than representing a fixed cost for an*

unlimited number of identifiers. This is important in two areas: (a) it seems likely that the number of identifiers needed for e-learning could be very large ... and (b) the growing open content movement which seeks to make content freely available to others for academic purposes will not accept any infrastructure required to support open content which contains an embedded variable pricing component."

4.1.6 Persistent Uniform Resource Locator (PURL)

The Online Computer Library Center (OCLC) has developed and deployed PURL as a naming and resolution services for internet resources. Instead of pointing directly to the URL location of a resource, a PURL points to a PURL Resolution Service which associates the PURL with the actual URL and returns that URL to the client. The client can then complete the URL transaction in the normal fashion. In Web parlance, this is a standard HTTP redirect. (OCLC, 2004)

A PURL has three parts, a protocol a resolver address and a name. For example, in the PURL <http://purl.oclc.org/melcoe>, http is the protocol, purl.oclc.org is the resolver address, and melcoe is the name. The URL www.melcoe.mq.au.edu could be registered under this name with the resolver service, and if the URL ever changes in the future, it only has to be updated once, with that service. All references to the PURL will continue to work, even though the underlying URL has changed.

The use of the PURL approach has several advantages:

- It requires no browser modifications and makes use of existing, adopted protocols
- It adds no cost
- It can be easily converted to a URN structure if and when this protocol becomes a standard accepted by the Web community and widely incorporated into browsers

Given these advantages, PURL may be a better solution for research and education than the DOI initiative. However, the success of the PURL initiative depends on voluntary support from organizations such as OCLC. *"This system relies more on voluntary participation than international standards and hasn't really taken off."* (Scharf, 2002)

More information about PURLs can be found at the OCLC PURL Home Page. (OCLC, 2004)

4.1.7 Archival Resource Key (ARK)

ARK is a more recent proposal for persistent identification scheme. The approach follows a similar construct to the DOI:

"An ARK has four components:

[<http://NMAH/>] ark:/NAAN/Name

an optional and mutable Name Mapping Authority Hostport part (NMAH, where "hostport" is a hostname followed optionally by a colon and port number), the "ark:" label, the Name Assigning Authority Number (NAAN), and the assigned Name. The NAAN and Name together form the immutable persistent identifier for the object." (Kunze & Rogers, 2004)

As Kunze and Rogers point out further, the issue isn't so much the scheme itself as it is the long term sustainability of the services and organizations that support any scheme. What the ARK approach focuses on is a "promise for stewardship", and the ARK naming scheme includes a link from an object to a promise for stewardship of that object.

4.2 Authentication and Authorization

4.2.1 Overview

Authentication is the process of establishing the identity of a user, or of establishing that they have specified attributes or membership in a group. Increasingly, institutions are implementing 'single sign on,' which means that users authenticate once, and do not have to provide a login ID and password (i.e. do not have to re-authenticate) each time they access a new system. The ability to authenticate centrally, and to have all subsystems recognize this authentication, is an important part of a distributed DRM solution.

Single sign-on among a controlled set of applications within an enterprise is possible and is something that most research and education institutions are moving towards using technologies such as LDAP and Kerberos or architectures such as those proposed by the OKI initiative. However, there is no single standard approach for supporting single sign-on, and this is a major challenge to initiatives trying to 'prove out' approaches to DRM in research and education. For example, in a summary of the COLIS project, James Dalziel commented that "*no widely accepted standard for Single-Sign-On exists at present, which proved a significant challenge for the project.*" (Dalziel, 2002)

Authorization is the process of determining what users are permitted to do once they have been authenticated. In enterprise portals and content management systems, authorization functions often determine what content a user can access and what functions they can perform on that content. However, once the user moves content outside of the enterprise IT infrastructure boundaries (e.g. - emails the content out, or copies it to a personal computer), then the enterprise authentication and authorization technologies no longer apply, and DRM must then rely on rights information and restrictions contained in the content itself, and on the user's understanding of, and respect for, those rights.

Digital rights and security have an important relationship. They are not the same thing, but an effective distributed rights management ecosystem has a critical dependence on the existence of an effective distributed security ecosystem. Authentication on all platforms is needed to identify the 'entity' that has been licensed to use the content, whether an individual, an organization, or a particular physical device. Identifying this entity as an individual or as an entity with specific attributes allows the completion of the link to the appropriate license.

Role based authorization - In research and education rights management frameworks, users are typically authorized to perform certain actions with content based on their roles within an institution or course (student, auditor, professor, teaching assistant, etc.). A user's role is determined on the basis of established identity (authentication) and the role is then used to determine what the user may do (authorization).

Fine grained access control allows the access to specific content items and actions to be controlled based on a user's personal characteristics, roles and group memberships.

As an example of fine-grained access control technology, Oracle 8i introduced the notion of a Virtual Private Database (VPD). "*A VPD offers Fine-Grained Access Control (FGAC) for secure separation of data. This ensures that users only have access to data that pertains to them. Using this option, one could even store multiple companies' data within the same schema, without them knowing about it.*" (Oracle, 2004) This type of capability has always been possible to implement with a relational database, but to do so used to require the

developer to define the security and access control tables themselves, and to create and maintain complex views created from joins to security tables. Companies such as Oracle now provide this security embedded into their tools.

4.2.2 Shibboleth

Shibboleth is an architecture designed to support the sharing of digital content between organizations by providing cross-institution attribute exchange which can be used to assist with authentication and access authorization. It is an initiative to develop an open, standards-based solution to the needs for organizations to exchange information about their users in a secure manner, while preserving the privacy of the user's actual identity. A Shibboleth exchange determines if a person using a web browser has the permission needed to access a resource at another institution. Rather than requiring personal identity information, the Shibboleth exchange shares information about the relevant attributes, roles and group memberships that the person has, which may be sufficient to allow access to the resource. This could be based, for example, on enrollment in a class, or employment as a faculty member in the math department.

If the institution being asked for access requires more personal information before granting access, the Shibboleth architecture allows a user to choose whether or not to release that additional information.

Origin sites are responsible for authenticating their users, but can use any reliable means to do this. Clearly, Shibboleth requires that a trusting relationship first be established between computing environments, which means it is best suited to use within communities of well-known agencies.

Shibboleth is being widely experimented with and implemented at research and education institutions worldwide. For example:

- Over the next four years JISC is planning the implementation of an authentication system based on Shibboleth. (Duncan et al, 2004)
- Penn State has implemented Shibboleth to enable 1,200 students to access resources at North Carolina State. (Internet2, 2004)
- The National Science Digital Library (NSDL) uses Shibboleth to support access to its collections. (Internet2, 2004)
- The LionShare peer to peer project is implementing a version of Shibboleth to control access to LionShare peers through LionShare servers. (Penn State, 2003)

Shibboleth is designed to eliminate the need for multiple passwords and signons at many institutions, along with the resulting improvements in security, while preserving individual privacy by eliminating the unnecessary disclosure of personal attributes

The following table from the Shibboleth Web site provides some use cases. (Internet2, 2004)

Consider some activities for a typical graduate student:

Activity	How it is done today	Problems with current approach	What Shibboleth could do
Accessing digital library resources from off-campus	Proxy servers, shared passwords or no service	<ul style="list-style-type: none"> - Proxy servers hard to maintain - No access from home - IP address-based restrictions easily compromised - Privacy can be compromised if identity is inappropriately passed to library 	<ul style="list-style-type: none"> - Permits access directly to content without campus proxy server - Requires campus authentication, though identity is not passed to library - Be used by libraries for new licensing approaches to content
Using distance education courseware or using external grading services	Additional username/passwords	<ul style="list-style-type: none"> - New accounts - Users frequently set external passwords to be same as internal; significant security exposures - External agencies are limited in verification options 	<ul style="list-style-type: none"> - Use local campus authentication and have campus pass appropriate identifier passed to courseware or service - Requires remote resources to trust campus enrollment / authentication
Accessing a research web site at another university or managing a shared polar instrument	Group class accounts or new remote individual accounts	<ul style="list-style-type: none"> - New accounts for users - Shared passwords represent security and audit concerns 	<ul style="list-style-type: none"> - Enables use of local campus account - Permits role-based access - Requires active privacy management by user
Accessing a co-taught class web site at another university	Group class accounts or new remote individual accounts	<ul style="list-style-type: none"> - For users, too many accounts - Individual accounts could compromise privacy - High management overhead from account management 	<ul style="list-style-type: none"> - Permits use of campus account - Preserves privacy - Target management may be done by content owners - Users may be required to approve attribute release

4.2.3 XACML - eXtensible Access Control Markup Language

The OASIS Extensible Access Control Markup Language (XACML) is intended to express authorization policies in XML. It is a general access control policy language, designed to support policy enforcement (a permit / deny decision) for any application requesting access to a resource. (Cover Pages, 2004b).

Adoption of XACML would allow applications to use a common access control language, and eliminate the use of multiple proprietary or application-specific access control policy languages, thereby allowing access policies to be shared across different applications. This is also intended to provide an incentive for the development of common tools for managing security policies, and to provide a language that is extensible and able to express a wide range of access policies, which allows new requirements to be met as they arise. (Sun Microsystems, 2003a)

XACML resolves a part of the complex interplay of authentication and authorization between computing environments, and it assumes that both sides of the exchange contain sophisticated security processing capabilities, and are known and trusted by each other.

Sun Microsystems and Jiffy Software have development efforts underway. (Coyle, 2004)

XACML and Shibboleth

"There is interaction between the OASIS committee and ... Shibboleth. The combination of Shibboleth authentication of users and the XACML access language could end up looking not unlike the Federated Digital Rights Management System outlined by Martin, et al. for library-licensed materials." (Coyle, 2004)

Researchers are considering XACML to replace current access control functionality in Shibboleth. *"XACML addresses problems with htaccess files and scales well to a distributed, decentralized environment. XACML's ability to work with policies and attributes managed in arbitrary locations greatly supports the distributed nature of Shibboleth."* (Lorch et al, 2003)

XACML and SAML

The Security Assertions Markup Language (SAML) is an XML-based framework for communicating user authentication, entitlements and attribute information. MELCOE is investigating the possibility of combining XACML and SAML to support the authentication of users and user attributes, and for authorizing access to digital resources based on that authentication.

4.2.4 PERMIS

The European Community PERMIS project (PriviEge and Role Management Infrastructure Standards Validation) is intended to provide a 'single sign on' option that can be used by any web application to identify and authenticate an individual, and to determine what authorizations they should be given once signed on based on their current roles, attributes and group memberships. Conceptually it is an extension of the digital certificate standard that provides only identification of an individual, to a system that handles roles and privileges by offering trusted "attribute certificates".

From a user perspective, this attribute certificate allows them to only have to remember one Web ID and password. An application making use of the PERMIS attribute certificates would

have to internally match the roles and privileges on the attribute to their own policies or licenses internally.

PERMIS and XACML

"PERMIS defines a hierarchical role based access control (RBAC) policy language in terms of those roles and permissions. The RBAC policy (in XML format) is used to control access to all the targets within the policy domain and is composed of a number of sub-policies. The PERMIS project is currently investigating the use of XACML as a core language to replace parts of their proprietary policy language." (Lorch et al, 2003)

4.3 Rights Expression

4.3.1 The Role of a Rights Expression Language

As described above, authentication and authorization technologies allow access to content to be controlled when researchers, educators and others search for and retrieve content from digital repositories. However, these technologies are not designed to provide rights expressions that can be associated with the content and content metadata anywhere in a distributed environment.

Rights Expression Languages (RELs) 'extend the reach' of digital rights management beyond the boundaries of the digital repository environments. RELs allow expressions of rights and permissions to be directly linked with content and content metadata, even when that content is passed to environments outside the control of digital repositories, ideally all the way down to the desktop environment. They also provide a way to express digital rights and permissions in a language specifically designed for that purpose. However, whereas access control supported by authorization and authentication technologies are well accepted and widely adopted, digital rights management languages and related technologies are in their infancy, and are only beginning to be adopted.

A Rights Expression Language provides a grammar in which permissions and conditions can be expressed in a machine (and human) readable form. A Rights Data Dictionary is a standardized vocabulary for expressing rights and conditions in a rights expression language. The rights expression language defines a structure for expressing permissions and conditions, while the rights data dictionary precisely defines the meaning of the permissions and conditions expressed.

The vision and promise of rights expression languages is that, once fully implemented, they will provide cross-repository interoperability for expressing and interpreting rights, offers and licenses. The rights expressions would follow digital content wherever it goes, to allow full digital rights management capabilities to be realized throughout the full distributed network, all the way down to the user's personal computing environments.

The attraction of these capabilities to the commercial world is obvious, but in research and education there is great ambivalence about these goals, with concerns for privacy and the potential for stifling fair use rights and innovative collaboration being major concerns.

Within this context, the digital rights management community in research and education is struggling with some basic questions about rights expression languages:

- What rights, permissions and conditions need to be expressed for research and education?
 - This topic is discussed at length in the requirements section of this document.
- Which rights expression language best supports these goals?
 - This is discussed in section 4.3.3 below.
- If we have cross domain authentication and authorization capabilities for accessing resources such as those embodied in Shibboleth, XACML and PERMIS, do we really need a rights expression language at all?
 - This is discussed in section 4.3.2 below.
- How do rights expressions fit in with other content metadata?
 - This is discussed in section 4.3.4 below.

4.3.2 Access Control versus Rights Expression

Rights expression languages are designed to express the policies that a rights holder has established for the use of specific content. These can then be associated with the content or the content metadata, expressed to users, used to enforce rights, and shared with other repositories. Rights expression languages are intended to provide an expression of rights that can 'persist' as part of the content metadata, and that can be accessed anywhere the content is acquired or used. This persistence is important to communities (such as the commercial digital content community) that need rights enforcement and expression to occur everywhere, in order to protect the content from unauthorized use.

Access control technologies are part of the enforcement process, but they do not provide a method of sharing rights with all components of a distributed content management environment. As was discussed above in section 4.2 'Authentication and Authorization', architectures such as Shibboleth can be combined with languages and protocols such as XACML and SAML to provide fine grained control over access to digital content. These capabilities may be sufficient to support the level of access control required for much of the content in the research and education community. In this community, most of the 'enforcement' of digital rights that occurs after content is accessed does not need to be automated. There is an expectation that researchers and educators will adhere to professional ethics and legal constraints and will 'self enforce' their use of digital content as long as they are aware of the rights and conditions associated with that content. Therefore the requirement becomes one of associating a clear, human-readable statement of permissions and conditions with the content, and making sure that this expression is visible to the users of digital content whenever they access it.

4.3.3 A Review of Rights Expression Languages

The recent white paper by Karen Coyle prepared for the Library of Congress (Coyle, 2003) provides a useful analysis of rights expression languages, particularly ODRL and MPEG-21 REL.

Rights expressions should have the following capabilities:

- Express the rights that an individual or group holds for a piece of intellectual property (license, agreement, copyright, etc), and the terms, conditions and permissions associated with those rights.

- Express the terms and conditions under which a rights holder is willing to grant permissions to potential users, and what those permissions are (sometimes referred to as an offer)
- Identify the parties involved (rights holders, licensees, etc)
- Be machine readable and interpretable for expression and enforcement purposes
- Allow a controlled vocabulary to be defined for a community
- Support the extension and evolution of the language structures and vocabulary
- Be designed to 'persist' independent of any particular environment (ie – continue to be associated with the content and/or the content metadata even when that content or metadata moves from one digital repository to another or moves to a user's personal computing environment)

Rights expression languages are relatively new developments and none of the languages listed below have yet been used extensively in any large, widely used implementations. (Coyle, 2003) However, several repository initiatives in research and education are implementing, or plan to implement, a rights expression language for the sharing of digital rights data (see the section on DRM Implementations).

Digital rights should be expressed in a structured language using a standardized set of codes, at the appropriate level of granularity. This will 'future proof' whatever choice is made, because with such an implementation it will always be possible to translate these definitions into another schema, and therefore to adapt to new standards as they evolve, which they surely will.

The Rights Expression Languages considered below can be grouped into three general categories:

- REL's intended to support rights expression and enforcement
- An REL intended to support rights expression
- A Proprietary REL intended to support rights management in specific environments

General REL's Intended to Support Rights Expression and Enforcement

These general purpose rights expression languages are designed to express a comprehensive set of digital rights in a structured way with a controlled vocabulary. This makes them suitable for supporting both the expression of rights for presentation to humans, and for automated interpretation by software that can use the rights expressions to provide automated enforcement of those rights.

This category of languages is being promoted and adopted by commercial digital content interests because it supports their desire to extend rights management *enforcement* to all points in a distributed content management economy. The full set of DRM capabilities supported by these languages also makes them technologically appropriate to meet the access control and rights expression requirements of the research and education community.

However, there are issues that argue against the immediate adoption of REL as part of the DRM solution for research and education:

- The standards are still in development.
- These RELs are not widely adopted by operating systems, content management applications, browsers, etc
- There are patent issues that may affect the open adoption of these languages (discussed in a later section)

- RELs have been complex to implement in the pilot projects that are underway

Three languages in this category are described below: ODRL, MPEG-21 and XrML

ODRL – Open Digital Rights Language

ODRL was developed by IPR Systems (Australia), led by the work of Renato Iannella. It is being used in several demonstrator projects and repository development efforts in research and education communities in Australia, the UK and Canada. ODRL was selected along with MPEG REL as one of the languages supported by OpenIPMP, and IPR systems has also submitted ODRL to the World Wide Web Consortium (W3C).

ODRL has been selected by the Open Mobile Alliance (an industry group for mobile communications such as cell phones) as the basis for its rights expression language (OMA, 2003). OMA found that ODRL meets its requirements of a light-weight and simple language for expressing rights, easy to implement, optimized expression for delivery over constrained bearers and suitability for specifying rights independently of the content type and transport mechanism.

MPEG-21 REL and Rights Data Dictionary (RDD)

MPEG (Motion Picture Experts Group) selected XrML as its REL of choice based on a head to head comparison that included ODRL. As Coyle points out, MPEG has very specific and structured requirements for the commercial entertainment industry, and XrML's precision and complexity was more suited to their needs. The International Standards Organization approved MPEG-21 REL as an international standard in 2004. (ISO/IEC, 2004) MPEG also developed a Rights Data Dictionary (RDD) specific to the needs of their industry.

The Open eBook Forum also selected MPEG REL as its rights expression language.

ContentGuard also submitted XrML to the OASIS e-business standards consortium.

XrML – eXtensible Rights Markup Language

XrML was developed by ContentGuard, and was submitted to a number of standards bodies for consideration. It was chosen by MPEG as its REL, and was modified to meet the needs of that group. ContentGuard turned over XrML (version 2.1) to the OASIS Rights Language Technical Committee in 2002, but as of July 29, 2004 this committee was deactivated. (Cover Pages, 2004c)

An REL intended to support rights expression

Rights enforcement in the research community typically occurs at the point where a decision is made on whether or not to give someone access to digital content. Once that decision is made, the primary need beyond that point is to clearly express those rights to the user of the content. Therefore, an REL that is designed to support rights expression, but not automated rights enforcement, may be a viable solution for some communities. METSRights is an example of this type of language.

METSRights Language

The METSRights language is targeted to the library community. The language is designed to handle more detailed copyright information than is currently supported by MARC, and the

Dublin Core. No automated control over the materials is anticipated from the rights expressed in the METSRights elements. (Duncan et al, 2004) (Coyle, 2004)

A Proprietary REL

Commercial software vendors are well aware of the need to provide rights management that persists with content no matter where the content goes. In response, they have developed proprietary solutions where content is developed by their tools and can be accessed only by their 'content viewers'. This allows them provide an end to end digital rights management solution. These proprietary solutions are valuable to the research and education community because they work and they are available today. They allow a researcher or educator to control who can access their content, and what functions (print, copy, modify) they can perform on that content. However, these capabilities depend on the use of a common set of technology from a single vendor. The Adobe Content Manager provides an example of a proprietary rights expression language supporting proprietary digital rights management.

Adobe Content Manager (ACM)

Adobe Content Manager is an example of a product that incorporates a proprietary rights expression language. However, it can be used only in the Adobe product environment to control access and permissions for protected PDF files. Rights are enforced through the use of encryption, with authentication / authorization provided through document passwords or individual digital certificates and public key infrastructure (PKI). Permissions can be specified to control printing, copying and several different revision options for a protected PDF document.

4.3.4 Relation to Other Metadata Standardization Initiatives

Metadata creation, management and use are key concepts throughout the content life cycle. Rights management metadata is a part of this metadata management process, and without standards for overall metadata management, DRM processes cannot work effectively in a distributed ecosystem. Rights metadata, as defined in a rights expression language is one type of metadata. RELs provide structure for that metadata and rights dictionaries provide codes for that structure. They must be consistent and work together.

In his paper, "*Do we need a Rights Expression Language*", D. Rehak (2003) wonders how we integrate rights expressions into learning technology, and suggests that REL's may be redundant with other learning technology standards and capabilities (such as metadata and sequencing). There certainly are areas of overlap, and it is important to minimize redundancy between these specifications. However, it is probably more important in the long run to ask 'how do we integrate learning technology standards into the much larger world of content management, including rights management?' The world of learning technology is only one voice in the debate about the form that DRM standards should take.

The IEEE Learning Technology Standards Committee (<http://ltsc.ieee.org>) is working on the specification of Rights Expression Language requirements for the learning and education technology community. Once these are completed, they intend to submit these requirements to other bodies to try to ensure that the needs of this community are met by the broader standards as they develop.

4.4 Tracking and Reporting

Tracking of the usage of content is very important to the research community. Researchers and the organizations supporting their research want to know who is using their work, and how it is being applied. This information is important both to assess the value of a researcher's work and to provide feedback that helps to inform future research activity.

In the education and learning community, tracking also refers to the ability of a content delivery system to record data on the use of digital content (time spent with a piece of content and results from assessments are examples of this data). Standards such as SCORM enable data of this form to be communicated between content and a delivery system in a way that is initiated and controlled by the content rather than by the system. This is necessary if the same content is to run and exhibit the same behavior on multiple systems. In the future, this type of content interaction tracking may form the basis for reporting content usage back to rights holders and repositories, which would make these standards relevant to researchers as well as to educators.

However, there are no standardized content tracking and reporting services that are in wide operation in the research and education communities. In the entertainment industry, MPEG 21 is in the early stages of developing a standard for reporting. However, the purpose of tracking systems in the commercial world is to support and enforce revenue generation models, which is a very different purpose than that of tracking in the research and education community.

4.5 Protection and Enforcement

An examination of the content protection technologies that can be used to enforce digital rights is beyond the scope of this paper. There are many options and technologies available for these purposes, and massive implementations of these technologies in the security and content management industries.

In general, there are three general categories of enforcement and protection that are relevant to research and education.

- **Industry Practice and Ethical Standards** – Proper attribution, requesting permission for reuse, and severe professional penalties for plagiarism are practices that are well embedded into the culture of research and education.
- **Legal Statutes, Penalties and Remedies** – Copyright and related laws in most countries of the world lay out the rules under which society is expected to operate, and provide penalties and remedies for violating these laws. The legal environment is discussed elsewhere in this document.
- **Technological protection** – This refers to the protection of content and the enforcement of digital rights terms and conditions through technological means. It includes things such as:
 - encryption
 - watermarking
 - digital signatures
 - fingerprinting

- privacy technologies
- payment systems

4.6 Trusted Computing

Trusted computing initiatives intend to "... create a networked environment that is 'trusted', based on the secure identification of users, devices and software modules, ensuring that content can only be exploited in line with rules set by the owners of the material." (Barlas et al, 2003)

The general purpose personal computer is the main barrier to the implementation of trusted computing in a distributed environment. The conflict between control of content and the capabilities of a computer that can perform many functions is one that will be difficult to resolve. The thrust of much Trusted Computing work is on making personal computers more secure environments both for their users and for companies who target content to the personal computer platform. According to LaMacchia *"owners of digital content will not distribute their works to platforms they consider 'hostile' (or potentially so) and the same is true of individual users requested to reveal private information to remote systems. Every content owner needs some way to be convinced that the remote system receiving his or her valuable information will behave as the owner expects."* LaMacchia (2002)

An open standards approach is being taken by the Trusted Computing Platform Alliance which *"develops and promotes open industry standard specifications for trusted computing hardware building blocks and software interfaces across multiple platforms."* (Trusted Computing Group, 2004)

Microsoft's Next-Generation Secure Computing Base is a proprietary approach that appears to be targeted to the Longhorn release. However, there is little updated information available about this initiative. (Microsoft, 2004)

Digital rights management in the research and education community is unlikely to ever be able to assume the existence of this type of trusted computing. Researchers, educators and students do not tend to operate on standardized platforms. They work with a heterogeneous mix of workstations, servers and operating systems, including an increasing use of open source operating system software. Open access to all software source code is a fundamental principle of the open source software approach. This directly contradicts the concept of protecting that code, which is an underlying requirement for trusted computing platforms.

In addition, there is considerable hostility towards the concept of software and hardware companies having full control over the capabilities of personal computers, and we can expect the research and education community to be among the most implacable opponents to the adoption of this technology.

5 DRM Implementations in Research and Education

This section briefly describes some of the digital rights management initiatives occurring in the research and education community.

5.1 JORUM

The Joint Information Systems Committee (JISC) is developing JISC Online Repository for [Learning and Teaching] Materials (JORUM). JORUM is scheduled to be made available to Further and Higher Education Institutions in the UK beginning in August 2005. The repository will be used by faculty and contains learning materials they can provide to students and resources such as lesson plans for faculty.

Part of this initiative is the implementation of digital rights management capabilities, which in the longer term include a number of models for providing materials that are not free. In 2004, JORUM's plans called for the implementation of a simple set of licenses without any technical enforcement. They planned to use the work done by the RoMEO project (see below, and Gadd, Oppenheim & Proberts, 2003b) which expressed Creative Commons licenses using ODRL. In 2004 the project planned to allow depositors to specify one of two licenses, one which would allow repurposing of deposited content, and one which would not. (Halliday, 2004)

5.2 Celebrate

Celebrate (Context eLearning with Broadband Technologies) is a project funded by the European Commission intended to "*implement and explore ideas around the federation of various LMS's, LCMS's and repositories.*" (Simon & Colin, 2003) Celebrate implements a federated search that allows partners to search through each other's repositories and share learning objects. They also created a system which allows digital rights to be assigned to learning objects and managed. The initial implementation of Celebrate seems fairly trivial from the user perspective (three options, no view, view only, download), but underlying this is a fairly sophisticated DRM technology implementation, including cross-repository sharing of rights statements using a rights expression language.

To achieve these goals, Celebrate implemented a rights expression language and a rights management protocol. They chose to use ODRL over MPEG21 and XrML because "*it better suits the requirements of our application domain.*" An 'offer' is included in the object's metadata, and once a requestor accepts the offer the Celebrate brokerage creates and 'agreement' which is stored by the brokerage. When the user accesses the learning object, the agreement is retrieved and the brokerage checks to see if all conditions and constraints are met. If so, the brokerage returns a 'handle' that supports running or downloading the content, depending on what was requested and approved. (Simon & Colin, 2003)

5.3 eduSource

The eduSource project is a test bed network of Canadian learning object repositories providing federated access to educational resources. The intent of eduSource is to "create a test bed of linked and interoperable learning object repositories."

Rights are described using ODRL, contained in files managed by the provider broker, and accessed by means of pointers in the learning object metadata exchanged within the eduSource network. eduSource selected ODRL as part of their open source, open standards approach, because it can be used under an open source license.

eduSource is designed to support multiple digital rights models, including free access, cooperative sharing, fee-based and subscription-based.

The eduSource project uses a web services architecture, and has built a wizard for specifying ODRL rights and conditions. They have also developed an ODRL wizard for creating licenses, and they have developed standard license templates in ODRL, including several Creative Commons licenses. (Downes et al, 2004)

5.4 COLIS

The COLIS (Collaborative Online Learning and Information Systems) demonstrator project is funded by the Australian Federal Government Department of Education, Science and Training (DEST). Its primary goal in 2002 was:

"to develop a distributed systems framework for online learning and information services which would demonstrate [among other things] incorporation of specified levels of digital rights management [and] seamless movement between applications using single sign-on, incorporating directory services." (Dalziel, 2002)

COLIS used ODRL as the rights expression language for expressing and controlling the digital rights associated with Learning Objects, packaged with other metadata. The project was successful in implementing end to end digital rights management for a limited set of licenses. It uncovered significant challenges to the implementation of distributed rights management, particularly the need for unified access and identity management throughout the environment. This is necessary to *"implement the license requirements of an ODRL agreement at all stages of the lifecycle of creation, trading, downloading, arranging and student use of Learning Objects. The project also resulted in the definition of 'education market specific' ODRL license templates."* (Dalziel, 2002)

5.5 Guth, Neumann and Strembeck

Guth, Neumann & Strembeck (2003) undertook a project to enforce access rights extracted from contracts expressed in ODRL.

This project is particularly interesting in part because they built and experimented with rights interpreter that builds on top of the rights expression language layer. This resulted in a generic run-time model that is independent of any particular rights expression language and from the applications that use contract data. They demonstrate components that act as a mediator between a contract in a rights expression language and access control technology, which actually enforces the rules embedded in the rights expression language. Their approach *"uses ODRL-based contracts as a means to disseminate access control information in distributed systems."* (Guth, Neumann & Strembeck, 2003)

5.6 RoMEO

One of the results from Project RoMEO (Rights Metadata for Open Archiving) is an XML-based system for the expression of rights and permissions governing metadata and resources in institutional repositories. The project team developed an XML metadata notation using the Open Digital Rights Language (ODRL) and Creative Commons licenses for disclosure of the rights expressions under the OAI-PMH.

They rejected the use of XrML because they saw it as "a patented product with unclear licensing terms, and at the time of the project XrML also did not have a data dictionary component. By contrast, ODRL was an open source language with a form of Data Dictionary. RoMEO project worked with the Creative Commons to develop an XML schema for their RDF, and the project would also developed ODRL versions (XML instances) of the CC licenses that would conform to the ODRL XML schema." (Gadd, Oppenheim & Proberts, 2003b)

5.7 Federated Digital Rights Management (FDRM)

"The goals of the Federated Digital Rights Management (FDRM) project are to support local and inter-institutional sharing of resources in a discretionary, secure and private manner, while endeavoring to maintain a balance between the rights of the end-user and those of the owner." (Martin et al, 2002a)

In this approach there is shared administration of access controls between the origin site and the resource provider: the origin site is responsible for providing attributes about the user to the resource provider. FDRM applies and extends the federated access control mechanisms of Shibboleth.

"The ability to track usage, protect the integrity of a resource, and manage version control is possible in FDRM, without compromising the privacy of the individual." (Martin et al, 2002a)

5.8 Creative Commons

Creative Commons is a non-profit organization supported by grants and the Stanford Law School. It was founded in 2001 with support from the Center for the Public Domain. (Creative Commons, 2004)

The Creative Commons provides an environment where licenses are easy to create, easy to link to content, and easy to understand, while still providing the endless legalistic verbiage that is needed in an enforceable legal document. Its licenses are based on the open source movement. The Creative Commons web application allows people to dedicate their work to the public domain, or to retain copyright while licensing content as free for some uses, under certain conditions. Unlike the GNU GPL, Creative Commons licenses are not designed for software, but rather for other kinds of creative works.

The CC licenses are expressed in three forms: human, machine and lawyer readable. Users of the system select from a list of licenses or public domain dedications and receive support to express these declarations.

The Creative Commons licenses are beginning to be quite widely used in the research and education space. One thing to note is that Creative Commons licenses are designed for situations where individuals have ownership, so they may not be applicable to situations where an organization owns the rights. (Duncan et al, 2004)

5.9 AE Sharenet

AEShareNet is a collaborative system to streamline the licensing of intellectual property so that Australian learning materials are developed, shared and adapted efficiently. (AEShareNet, 2004) Like the Creative Commons, AESharenet has created a number of predefined licenses. The rights holders decide which license looks right to them, place a license mark on the material, and upload a description of the material to the AEShareNet catalogue.

6 Technology Patents

Disclaimer: The authors of this White Paper are not lawyers and the opinions expressed in this section are not legal advice. Anyone wishing legal advice should consult with a licensed attorney or solicitor.

6.1 Patents Affecting DRM

The research and education industry has expressed considerable concern about patents held by ContentGuard for digital rights expression. ContentGuard holds 12 US patents, 5 European patents, 1 patent in Japan and 1 patent in Korea related to the distribution of digital works and to any rights language. These can all be found using the US Patent and Trade Office Patent Search (USPTO, 2004), and the worldwide search in Europe's Network of Patent Databases (ESP@cenet, 2004). Use the search terms 'digital' and 'expression' in the title field.

These patents contain broad fundamental claims covering:

- Association of usage rights to content.
- A grammar to define rights or conditions.
- Persistent protection.
- Distribution of composite digital content.
- Fee accounting and reporting associated with the distribution or use of content.

ContentGuard is offering their patents to companies and organizations implementing rights expressions on 'reasonable and non-discriminatory' (RAND) terms. However, specifics about these terms are not available on the ContentGuard web site. ContentGuard's stated position in regard to these patents is that they are not specific to XrML but apply to the use of any digital rights language. A DRM product or system may infringe their patents without using any rights expression language.

Further research into patents leads to the discovery that there are many, many patents besides ContentGuard's that may apply to the development or implementation of digital rights management technology.

- Intertrust holds 31 U.S. patents and has over 100 patent applications pending worldwide in the areas of digital rights management, digital policy management, and trusted computing. (Intertrust, 2004)
- Searching for the words 'digital' and 'rights' in the patent title field in the USPTO database turns up 15 other granted patents, and 110 patent applications. (USPTO, 2004) These patents and applications are held or filed by a wide range of companies and individuals including General Instruments, Digimarc, Microsoft (quite a number), Sun Microsystems, Sony Ericsson Mobile, ContentGuard, Nokia, Xerox, Pitney Bowes, IBM, Case Western University, and many others.
- Searching Europe's Network of Patent Databases using the worldwide search and the terms 'digital' and 'rights' and 'management' in the title yields 228 results. These include the USPTO patents and applications, but also includes many additional foreign patents and patent applications. (ESP@cenet, 2004)

- A search on the terms 'digital' 'object' and 'identifier' in the title or abstract turns up 20 patents in Europe's Network of Patent Databases using the worldwide search. (ESP@cenet, 2004)

6.2 Patent Pools

MPEG LA (a limited liability corporation) defines reference models based on MPEG standard sets and provides pools of patents to protect all those implementing against the reference models. Their stated business model is to gather all essential patents needed to support development of DRM as envisioned by MPEG, to license them to "industry participants", and to indemnify those participants against patent claims. They state that they provide *"convenient, fair, reasonable, nondiscriminatory access to a portfolio of essential worldwide patent rights under a single license."* They are in the process of defining a Digital Rights Management Reference Model, and have not yet developed the DRM patent pool associated with this reference model (MPEG LA, 2004a & 2004b)

Carl Shapiro points out that *"As we move from pure "R" to applied "R" and ultimately to "D," one can fairly ask whether our legal and commercial institutions are in fact properly designed to promote rather than discourage the creation of products and services that draw on many strands of innovation and thus potentially require licenses from multiple patent holders. To complete the analogy, blocking patents play the role of the pyramid's building blocks."* (Shapiro, 2001) Shapiro takes the position in this article that although we may question the wisdom of how patents are being granted and enforced, it is still necessary to find ways through this "patent thicket". His proposition is that this is best done by an intelligently designed patent pools that indemnifies organizations making use of a defined set of standards. *"...standard-setting organizations like the ITU or the American National Standards Institute (ANSI) typically require that participants agree to license all patents essential to compliance with any standard on 'fair, reasonable, and non-discriminatory' terms. Rules such as this are explicitly intended to reduce or eliminate any 'hold-up' problems. However, it is well to note that many standard-setting organizations are wary of sanctioning any specific agreement regarding the magnitude of licensing terms for fear of antitrust liability, as such agreements might be construed as price fixing."* (Shapiro, 2001)

Indeed, licensees may well welcome such a pool, both for the convenience of "one-stop shopping" and because a subset of the required patents may be of little or no value by themselves. Thus, from the licensee's perspective, licensing the entire package is simpler and avoids the danger of paying for some patent rights that turn out to be useless without other complementary rights.

6.3 Observations

There is the potential for patent claims against just about any technology choice that might be made for digital rights management. A few observations and considerations:

- Stick with tools and standards that are being widely adopted in your community, and work with other members of the community to resolve the patent challenges if they arise.
- Licenses to use DRM patents should not be acquired piece-meal. The various patent holders should form a patent pool that gathers in all applicable patents, and then

make them available under a single license that indemnifies organizations against patent claims if they follow the specified model.

- Research and education is a small content market, at least from a revenue perspective, so it is not clear that it would make financial sense for any organization to assemble and support a content pool that meets the needs of this market. Also, if the research and education market assembles an open source infrastructure to meet their needs, then there is no software license fee revenue stream available to support the cost of 'fair and reasonable' patent licensing fees.
- Commercial software vendors are usually required by their clients to provide an indemnification clause in their purchase contract that indemnifies the client against any patent claims arising out of their use of the vendor's technology.

Glossary of terms and abbreviations

Access Control: Limiting or granting access to a file system, Web site, or other digital environment, usually via some sort of *authentication*.

Attribution: Assigning credit to the creator of an original work when the work is referenced, copied, distributed or performed.

Authentication: Establishing the individual identity of a user, or determining that the user has certain attributes or is a member of a specified group.

Authorization: Establishing what an individual is permitted to do.

Confidentiality: Ensuring that a resource may only be used by its intended recipient.

Content Repository: A secure storage facility that is capable of handling a wide variety of content types, and enabling access to authorized users.

Copyright: Rights granted to an original work under applicable copyright law (in the U.S. by the 1976 Copyright Act and other relevant law).

Derivative Work: A new work that is created by altering, transforming or building upon another work.

Digital Asset Security: Technologies for ensuring that digital assets are used only as authorized, particularly when they are delivered outside the company firewall.

DMCA: Digital Millennium Copyright Act.

Digital Rights Management (DRM): The process of defining, tracking and enforcing permissions and conditions for digital content through digital means.

Digital Signatures: A digital signature provides information about the source of a digital object, and allows the receiver to determine if the object has been altered in transit.

Directory Services: Services that provide identity, demographic and authorization information on a user.

DREL: Digital Rights Expression Language

DRM: Digital Rights Management

eBook: A book in digital format that you can download to your computer and read using a software program. Depending on the specific format, the eBook can be read on a computer, PDA, or dedicated reader device with the proper software.

Encryption: The process of encoding information so it cannot be accessed without first being decrypted through the use of an encryption key.

Fair Dealing: The term used in place of Fair Use in many parts of the world.

Fair Use: Under US law, Section 107 of the Copyright act allows reproduction and use of a work for the purposes of research, teaching and reporting, subject to restrictions.

Fingerprinting: The process of extracting information about content to create a fingerprint for later matching and metadata retrieval. These work for audio and visual content, not so well for textual.

IEEE LTSC: IEEE Learning Technology Standards Committee

Integrity: Establishing that an object has not been altered in any way.

Intellectual Property Rights (IPR): The rights given to people over the creations of their minds (WTO, 2003). In practice IPR most often refers to ownership rights of individuals, including copyrights, patents and trademarks.

License: A license is a legal vehicle for granting an individual or organization an explicit collection of rights and conditions for the use or distribution of a copyrighted work.

LCMS: Learning Content Management System

LMS: Learning Management System

MPEG: Motion Pictures Experts Group

Multi-Channel Publishing: Enables content to easily be reused or repurposed, and delivered to different channels. These could include mobile devices, PDAs, print, and interactive television.

Non-repudiation: Proof that an action was taken (e.g. an email sent) at a particular time and by a particular person or agent.

ODRL: Open Digital Rights Language

OeBF: Open eBook Forum

Offer: An offer to grant specific rights, or a request to be granted specific rights.

Open Source Software: Software that is freely available for use and distribution by anyone. This term usually also implies the existence of an open community of developers who contribute to the code base, accept specific open source licensing terms and share information about the initiative. Some open source licenses require that any derivative works based on open source software also be made available to others under the same open source license terms.

Persistent Rights Management: Associating rights with a work in a way that persists as the work moves through a network and is used by different applications, platforms and people.

Persistent Unique Identifier (PUI): A centrally registered identifier system that provides both a globally unique identifier for every separate piece of content, and a registry for accessing and maintaining accurate location information for that asset.

Public Domain: A rights holder can place their content completely into the public domain. However, in the academic community accepted practice still dictates that attribution should occur when public domain content is used.

Public / Private key encryption: An encryption approach whereby encryption is done using a user's public key, but decryption can only be done through the user's private key, which is never shared outside of the user's environment. Knowing the public key does not make it possible to derive the private key and decrypt the content.

Protection: Protecting a work from unauthorized use.

REL: Rights Expression Language

Right: A right or permission is "the most that one can do with a resource." It specifies how one may access or utilize a resource.

Rights Data Dictionary (RDD): A standardized vocabulary for expressing rights and conditions in a rights expression language

Rights Enforcement: Using technology to ensure that rights are not violated. For example, creating an Adobe PDF with protection that prevents it from being printed is not an act of enforcement. The enforcement occurs when you try to print the file and can't.

Rights Expression: The expression of IPR (including copyright, distribution rights, licenses and license requirements, and attribution and attribution requirements) associated to a resource. The expression of rights is separate from their enforcement.

Rights Expression Language (REL): A human and machine interpretable language for expressing rights, licenses, offers and other rights concepts.

Role-based Access Control: Access control that is granted based on your role within an organization (e.g. teacher, student, system administrator). Roles may be established through authentication and directory services or by other means, such as logging on through a campus IP address.

SCORM: Sharable Content Object Reference Model

Security: The prevention of unauthorized access and use via a combination of some or all of the other functions in this section.

Session Key: An encryption key that only works for a particular session.

Single key encryption: An encryption approach where the same key is used to encrypt and decrypt content.

Trusted Application: An application that interprets and enforces DRM rules.

Watermarking: Data embedded in content in a way that is imperceptible to the user, but that can be detected by special watermarking software. Allows content to be identified if it is copied.

Workflow. Supports the routing of documents and content between individuals and processes. Enables features such as document approval.

References

1. ADL (2004). *Advanced Distributed Learning initiative, Sharable Content Object Reference Model (SCORM)*. Available from the Web site <http://www.adlnet.org>
2. Adobe (2004). "Document Control", Adobe web site, accessed Nov. 2004. <http://www.adobe.com/security/doccontrol.html>
3. AEShareNet (2004). Web site, Licensing Overview. <http://www.aesharenet.com.au/coreBusiness/>
4. Australian Copyright Council (2004). Online Information Centre Web site. <http://www.copyright.org.au/>
5. C. Barlas, J. Cunard and K. Hill (2003). "Current Developments in the Field of Digital Rights Management", World Intellectual Property Organization, August 1, 2003. http://www.wipo.int/documents/en/meetings/2003/sccr/pdf/sccr_10_2.pdf
6. Berne Convention (1979). "Berne Convention for the Protection of Literary and Artistic Works", WIPO Web site. <http://www.wipo.int/clea/docs/en/wo/wo001en.htm>
7. J. Bormans and K. Hill (editors) (2002). "MPEG-21 Overview", International Organisation for Standardisation, ISO/IEC JTC1/SC29/WG11, version 5, October 2002. <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>
8. T. Byrne (2002). "The CMS Report", CMS Works Inc., 3rd Edition, Autumn 2002.
9. Center for the Public Domain (2004). web site. <http://www.centerpd.org/>
10. CEN/ISSS (2003). "Digital Rights Management – Final Report", CEN, 30 Sept 2004. <http://europa.eu.int/comm/enterprise/ict/policy/doc/drm.pdf>
11. C. Chadwick and A. Otenko (2002). "The PERMIS X.509 Role Based Privilege Management Infrastructure", Pre-print version of Future Generation Computer Systems. 936 (2002) 1–13, December 2002, Elsevier Science BV. <http://sec.isi.salford.ac.uk/download/FutureGenCompSyst.pdf>
12. CNRI (2004). "Handle System Introduction", Corporation for National Research Initiatives Web site. <http://www.handle.net/introduction.html>
13. COLIS (2004). Collaborative Online Learning and Information Services Web site. <http://www.colis.mq.edu.au>.
14. G. Collier (2002). "eLearning Application Infrastructure", Sun Microsystems, March 2002. http://www.sun.com/products-n-solutions/edu/elearning/eLearning_Application_Infrastructure_wp.pdf
15. G. Collier, H. Piccariello, and R. Robson (2004a). "Digital Rights Management: An Ecosystem Model and Scenarios for Higher Education" EDUCAUSE Center for Applied Research, Research Bulletin, Issue 21, 2004, <http://www.educause.edu/LibraryDetailPage/666?ID=ERB0421>.

16. G. Collier, H. Piccariello, and R. Robson (2004b). "Digital Rights Management: Tools and Applications for Implementing DMR in a Digital Ecosystem" EDUCAUSE Center for Applied Research, Research Bulletin, Issue 22, 2004. Scheduled for publication October 26, 2004.
17. ContentGuard (2004a). *ContentGuard*. Web site. <http://www.contentguard.com>
18. ContentGuard (2004b). list of ContentGuard's current patents, on their web site. <http://www.contentguard.com/patents.asp>
19. B. Cope and R. Freeman, editors (2001). *Digital Rights, Management and Content Development*, Common Ground Publishing, 2001.
20. Cover Pages (2004a). "*XML and Digital Rights Management*", Cover Pages web site, Updated April 28, 2004. <http://xml.coverpages.org/drm.html>
21. Cover Pages (2004b). "*Extensible Access Control Markup Language (XACML)*", Cover Pages web site, updated July 19, 2004. <http://xml.coverpages.org/xacml.html>
22. Cover Pages (2004c). "*MPEG Rights Expression Language*", Cover Pages web site, updated June 19, 2004. <http://xml.coverpages.org/xacml.html>
23. Cover Pages (2002). "*OASIS Rights Language*", Cover Pages web site, updated September 21, 2002. <http://xml.coverpages.org/oasisRightsLanguage.html>
24. Cover Pages (2004d). "*Patents and Open Standards*", Cover Pages web site, updated Sept 24, 2004. <http://xml.coverpages.org/mpegRights.html>
25. K. Coyle (2004). "*Rights Expression Languages – A Report for the Library of Congress*", Library of Congress web site. http://www.loc.gov/standards/Coylereport_final1single.pdf
26. Creative Commons (2004). Web site. <http://creativecommons.org>.
27. J. Dalziel (2002). "*Reflections on the COLIS (Collaborative Online Learning and Information Systems) Demonstrator project and the Learning Object Lifecycle*", In A. Williamson, C. Gunn, A. Young & T. Clear (Eds). *Winds of Change in the Sea of Learning: Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education*. Auckland, New Zealand: UNITEC Institute of Technology.
28. J. Dalziel (2003a). "*The Learning Object Lifecycle*", Macquarie E-learning Centre of Excellence, May 2003. www.melcoe.mq.edu.au
29. J. Dalziel (2003b). "*DOI in a DRM Environment*", CAL Coursepacks project white paper, 25 Nov 2003.
30. Digital ID World (2004). "*Assuring Networked Data and Application Reliability*", Jan / Feb 2004. Available from https://www.trustedcomputinggroup.org/press/1-3412425E_SC.pdf
31. DOI (2004). The Digital Object Identifier System, Web site. <http://www.doi.org>

32. S. Downes, et al (2004), "*Distributed Digital Rights Management: The Edusource Approach to DRM*", Retrieved from Stephen Downe's web site, 15 Apr 2004. http://www.downes.ca/files/DDRM_19April2004.pdf
33. DRM Watch (2004). web site <http://www.drmwatch.com>.
34. Duncan, Barker, Douglas, Morrey, Waelde (2004). "*Digital Rights Management*", report prepared by Intrallect for JISC, 25 August 2004. <http://dewey.intrallect.com/drm-study/>
35. ESP@cenet (2004). Europe's network of patent databases, web site. <http://gb.espacenet.com/>
36. EU (2001). "*European Union 2001 Copyright Directive*", European Communities, 22 May 2001. http://europa.eu.int/comm/internal_market/copyright/documents/documents_en.htm
37. FIPR (2001) Foundation for Information Policy Research, "*Implementing the EU Copyright Directive*", 2001. www.fipr.org/copyright/guide/eucd-guide.pdf
38. E. Gadd, C. Oppenheim, and S. Proberts (2003a). "*RoMEO Studies 6: Rights Metadata for Open Archiving*", 2003. <http://www.lboro.ac.uk/departments/ls/disresearch/romeo/index.html>
39. E. Gadd, C. Oppenheim, and S. Proberts (2003b). "*The Intellectual Property Rights Issues Facing Self-archiving: Key Findings of the RoMEO Project*", D-Lib Magazine, 2003, Volume 9 Number 9, ISSN 1082-9873. <http://www.dlib.org/dlib/september03/gadd/09gadd.html>
40. Giant Steps (2004). *References in the Giant Steps Bibliography*. <http://www.giantstepsmts.com/drmbiblio.htm>
41. S. Guth, G. Neumann, and M. Strembeck (2003). "*Experiences with the Enforcement of Asset Rights Extracted from ODRL-based Digital Contracts*", DRM'03, October 27, 2003, Washington, DC, USA. Copyright 2003 ACM. <http://wi.wu-wien.ac.at/people/Guth/p21-guth.pdf>
42. L. Halliday (2004). "*JORUM Scoping and Technical Appraisal Study: Volume VII – Digital Rights Management*", The JISC Online Repository for [learning and teaching] Materials, January 2004. http://www.jorum.ac.uk/docs/Vol7_Fin.pdf
43. G. Hulme (2003). "*Who Needs to Know?*" Information Week, June 9, 2003 www.informationweek.com/story/showArticle.jhtml?articleID=10300293
44. R. Iannella (2001). "*Digital Rights Management (DRM) Architectures*" D-Lib Magazine, June 2001, Volume 7 Number 6. <http://www.dlib.org/dlib/june01/iannella/06iannella.html>
45. IFLA (1998). "*Functional Requirements for Bibliographic Records*", International Federation of Library Associations and Institutions, UBCIM publications. N.S., Vol. 19, 1998. <http://www.ifla.org/VII/s13/frbr/frbr.pdf>
46. ISBN (2004). *International Standard Book Number*. Web site. <http://www.isbn.org>

47. ISSN (2004). *International Standard Serial Number*. Web site. <http://www.issn.org>
48. IEEE LTSC (2004). IEEE Learning Technology Standards Committee. Web Site. <http://ltsc.ieee.org>.
49. Internet2 (2004). *The Shibboleth Project*. Web site. <http://shibboleth.internet2.edu>
50. Intertrust (2004). Web Site. <http://www.intertrust.com>
51. ISO/IEC (2004). *Information technology -- Multimedia framework (MPEG-21) - Part 5: Rights Expression Language*. ISO/IEC 21000-5:2004. <http://www.iso.ch/iso/en/CombinedQueryResult.CombinedQueryResult?queryString=21000-5>
52. T. Jennings (2002). "Defining the Document and Content Management Ecosystem," Butler Direct Limited, September 2002
53. R. Koenen, J. Lacy, M. Mackay, & S. Mitchell (2003). "The Long March to Interoperable Digital Rights Management", white paper on InterTrust web site. http://www.intertrust.com/main/research/whitepapers/Interoperable_DRM.pdf
54. W. Kraan (2004). "ADL to Make a "Repository SCORM." Centre for Educational Technology Interoperability Standards. <http://www.cetis.ac.uk/content2/20040219153041>
55. J. Kunze & R. Rogers (2004). "The ARK Persistent Identifier Scheme", California Digital Library Web site. <http://www.cdlib.org/inside/diglib/ark/arkspec.pdf>
56. B. LaMacchia (2002). "Key Challenges in DRM: An Industry Perspective", ACM Workshop on Digital Rights Management, Washington. <http://www.farcaster.com/papers/drm2002/drm2002.pdf>
57. L. Lessig & J. Valenti (2000). "The Future of Intellectual Property on the Internet: A Debate" October 1, 2000, Ames Courtroom, Harvard Law School, Cambridge, MA. <http://cyber.law.harvard.edu/futureofip/>
58. LexMercatoria (2004). "Intellectual Property (protection of)", Law Faculty of the University of Oslo. http://lexmercatoria.org/intellectual_property/toc
59. Lorch, Proctor, Lepro, Kafura, Shah (2003). "First Experiences using XACML for Access Control in Distributed Systems", ACM XML Security Workshop, 31. October 2003. <http://zuni.cs.vt.edu/publications/xml-security-xacml-experiences.pdf>
60. LSAL, (2004). "CORDRA: Technical Introduction and Overview", Carnegie Mellon Learning Systems Architecture Lab, Web site. <http://www.lsal.cmu.edu/lsal/expertise/projects/cordra/intro/intro-v1p00.html>
61. C. Lynch (2003). *Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age*. Association of Research Libraries. Monthly report 226. February 2003. <http://www.arl.org/newsltr/226/ir.html>.

62. MELCOE (2004) Macquarie E-Learning Centre of Excellence, "Meta-Access Management System (MAMS) Project Overview",
<http://www.melcoe.mq.edu.au/projects/MAMS/index.htm>
63. Martin, Agnew, Kuhlman, McNair, Rhodes, Tipton (2002a). "*Federated Digital Rights Management: A Proposed DRM Solution for Research and Education*", D-Lib Magazine, July-August 2002. http://www.dlib.org/dlib/july02/martin/07martin.html#*
64. Martin, Agnew, Boyle, McNair, Page, Rhodes (2002b). "*DRM Requirements for Research and Education*", The NSF Middleware Initiative and Digital Rights Management Workshop, 4 Sept 2002. <http://lists.oasis-open.org/archives/rights/200209/pdf00000.pdf>
65. D. Maxwell (2004). "*IBM Takes Open-Source Commitment to the Next Level*", Linux Insider, Aug 5, 2004. <http://www.linuxinsider.com/story/35598.html>
66. MELCOE (2004). "*Meta Access Management System (MAMS) Project Overview*", Macquarie E-Learning Centre of Excellence, Web site.
<http://www.melcoe.mq.edu.au/projects/MAMS/index.htm>
67. Microsoft (2003). "*Microsoft Windows Rights Management Services for Windows Server 2003*", Microsoft Web site, October 2003.
<http://www.microsoft.com/windowsserver2003/techinfo/overview/rm.mspx>
68. Microsoft (2004). "*Next-Generation Secure Computing Base*", Microsoft Web site, 2004. <http://www.microsoft.com/resources/ngscb/default.mspx>
69. MPEG LA (2004a). "*MPEG LA Announces Plan for Joint Patent License for DRM Technology*", News release on the MPEG LA site. http://www.mpegla.com/news/n_03-10-02_drm.html.
70. MPEG LA (2004b). "*DRM Reference Model*", MPEG LA, LLC, Version 2, July 16, 2004.
<http://www.mpegla.com/pid/drm/>.
71. NCSU (2004). "*Copyright Tutorial*", Web Site. North Carolina State University Libraries, Web site, Office of Legal Affairs. <http://www.lib.ncsu.edu/scs/tutorial/basicsintro.html>
72. Norris, Mason, Robson, Lefrere & Collier. (2003). "*A Revolution in Knowledge Sharing*". Educause Review, September/October 2003, 38 (5). 14 – 26.
<http://www.educause.edu/ir/library/pdf/erm0350.pdf>
73. NSDL (2004). National Science Digital Library, Web site. <http://www.nsdl.org>
74. OAI (2004). Open Archives Initiative, Web site. <http://www.openarchives.org>
75. OCLC (2004). *Persistent Uniform Resource Locator*, Online Computer Library Center. Web site. <http://purl.oclc.org>
76. ODRL (2004). *Open Digital Rights Language Initiative*, Web site. <http://www.odrl.net>
77. OMA (2003) Open Mobile Alliance, "*Rights Expression Language*", Candidate Version 1.0 – 31 Oct 2003.
http://www.openmobilealliance.org/release_program/enabler_releases2.html#DRM

78. Oracle (2004). "*Deploying Fine-Grained Access Control*", Guide to Oracle 8i Features, from Oracle web site, http://www.unix.org.ua/oreilly/oracle/guide8i/ch08_01.htm
79. PADI (2002). "*Persistent Identifiers*", Preserving Access to Digital Information Initiative, National Library of Australia, Web site, updated August 2002. <http://www.nla.gov.au/padi/topics/36.html>
80. Penn State (2003). "*LionShare: Connecting and Extending Peer-to-Peer Networks*", Penn State proposal to the Andrew W. Mellon Foundation, Sept 2003.
81. PERMIS (2004). PrivilEge and Role Management Infrastructure Standards Validation, Web site. <http://www.permis.org/>
82. N. Paskin (2003). "*DOI – A 2003 Progress Report*", D-Lib Magazine, June 2003. <http://www.dlib.org/dlib/june03/paskin/06paskin.html>
83. D. Rehak (2003). "*Do We Need a Rights Expression Language?*", Carnegie Mellon Learning Systems Architecture Lab, web site. <http://www.lsal.cmu.edu/lsal/expertise/papers/notes/drel20062002/index.html>
84. Rights Express (2004). *Rights Express*. Web site. <http://www.rightsexpress.com>
85. R. Robson (2002). "*The 'Rights' Stuff*", Techlearn presentation, November 2002. <http://content.masie.com/techlearn/2002/followupsite/layout/default.cfm?page=sessiondetail&session=1237>
86. R. Robson, (2003). "*Digital Rights Management and Educational Content*", Eduworks, workshop given at EdMedia June 23, 2003 - workshops.eduworks.com/EdMedia2003/
87. Robson, Norris, Lefrere, Collier, & Mason. (2003) *Share and Share Alike: The E-Knowledge Transformation Comes to Campus*. Educause Review, September/October 2003. Online only. <http://www.educause.edu/ir/library/pdf/erm0351.pdf>
88. Robson, Muramatsu & Collier (2004). The Reusable Learning Project Web site. <http://www.reusablelearning.org>.
89. B. Rosenblatt (1997). "*The Digital Object Identifier: Solving the Dilemma of Copyright Protection Online*," in the Journal of Electronic Publishing (University of Michigan Press). volume 3, issue 2, December 1997. <http://dois.mimas.ac.uk/DoIS/data/Articles/doijappapv:03:i:02:p:285.html>
90. B. Rosenblatt (2004). "*Coral Consortium Aims to Make DRM Interoperable*", DRM Watch, 7 Oct 2004. <http://www.drmmwatch.com/standards/article.php/3418741>
91. B. Rosenblatt and G. Dykstra (2003). "*Integrating Content Management with Digital Rights Management*," Giantsteps Media Technology Strategies and Dykstra Research, May 2003. <http://www.xrml.org/reference/CM-DRMwhitepaper.pdf>
92. Rosenblatt, Trippe and Mooney (2002). *Digital Rights Management – Business and Technology*, M & T Books 2002.

93. D. Scharf (2002). "*The DOI is coming: tracking digital information*", Information Outlook Magazine, Sept 2002.
94. C. Shapiro (2001). "*Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting*", University of California at Berkeley, March 2001.
<http://faculty.haas.berkeley.edu/shapiro/thicket.pdf>
95. J. Simon & J. Colin (2003). "*Celebrate: A Federated Model for Exchange of Learning Objects*", Sun Microsystems and OxyS, 2003.
96. H. Van de Sompel & C. Lagoze (2002). "*A Progress Report on the Open Archives Initiative*", 6th European Conference on Research and Advanced Technology for Digital Libraries, September 2002, Rome. Draft at
<http://www.openarchives.org/documents/ecdl2002-oai.pdf>
97. R. Stallman (2002). "*The Right to Read*", Communications of the ACM, Feb 1997, updated in 2002 and available on the GNU site.
<http://www.gnu.org/philosophy/right-to-read.html>
98. Stanford University, (2004). Copyright and Fair Use Center Web Site, Stanford University Libraries. <http://fairuse.stanford.edu/>
99. Sun Microsystems (2003a). "*XACML: A New Standard Protects Content in Enterprise Data Exchange*", from the Sun web site, June 24, 2003.
<http://java.sun.com/developer/technicalArticles/Security/xacml/xacml.html>
100. Sun Microsystems (2003b). "*Federated Schools Architecture for National Collaborative Environments*", Sun Microsystems Web site, 11 June 2003.
http://www.sun.com/products-n-solutions/edu/whitepapers/pdf/federated_schools.pdf
101. Thomson Corporation. (2004). South-Western Book Division of Thomson Corporation: eBook Glossary. <http://www.swcollege.com/ebooks/glossary.html>.
102. Trusted Computing Group (2004). web site. <https://www.trustedcomputinggroup.org>
103. U.S. Copyright Office (1998). "*Digital Millennium Copyright Act (Summary)*".
<http://www.copyright.gov/legislation/dmca.pdf>
104. U.S. Copyright Office (2001). "*Copyright Law of the United States*", Text Revised as of July, 2001. <http://www.copyright.gov/title17/>
105. U.S. Copyright Office (2002). "*The TEACH Act Text*".
<http://www.copyright.gov/legislation/pl107-273.html#13301>
106. USPTO (2004) U.S Patent and Trade Office web site, patent search function.
<http://www.uspto.gov/patft/index.html>
107. WIPO (2004) World Intellectual Property Organization, web site. <http://www.wipo.int>
108. WTO (2004) World Trade Organization, "*Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement)*", available at
http://www.wto.org/english/tratop_e/trips_e/trips_e.htm.

109. W3C (2002). "*Team Comment on the ODRL Submission*", 19 Sept 2002.
<http://www.w3.org/Submission/2002/06/Comment>
110. W3C (2004). Web site. <http://www.w3c.org>
111. XRML (2004). *Extensible Rights Markup Language*. Web site. <http://www.xrml.org>.