The NSF Middleware Initiative and Digital Rights Management Workshop

DRM Requirements for Research and Education

Discussion Paper

September 4, 2002

Mairéad Martin (The University of Tennessee)
Grace Agnew (Rutgers, The State University of New Jersey)
Jeanne Boyle (Rutgers, The State University of New Jersey)
John McNair (The University of Tennessee)
Mary Page (Rutgers, The State University of New Jersey)
William Rhodes (The University of Tennessee)

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 1
INTRODUCTION

The goal of this document is to present some key issues to support discussion and articulation of Digital Rights Management (DRM) requirements for Research and Education (R&E). Articulating R&E needs in the DRM space will serve several purposes: it will support analysis of existing DRM systems and Rights Expression Languages to determine how well they can meet R&E needs; it will communicate to industry the unique DRM requirements of R&E, and the incentive that exists to develop solutions to support those requirements; and it will hopefully lead to the development of DRM solutions to accommodate needs that are not being met by existing solutions.

This document consists of definitions of key concepts and terms, a discussion of scope, representative scenarios, general DRM requirements, a discussion of middleware in DRM, and concludes with requirements for a Rights Expression Language.

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 2
DEFINITIONS

One of the issues concerning DRM is the lack of a standardized definition for the process and for the key concepts involved. Different interpretations of the term abound, including:

> "Digital Rights Management refers to *controlling* and *managing* rights to digital intellectual property."[1]
>
> Digital Rights Management is "the description, identification, trading, protection, monitoring and tracking of all forms of rights usages over both tangible and intangible assets including management of rights holders relationships…it is the "digital management of rights" not the "management of digital rights."[2]

---

[1] Rosenblatt, B., Trippe, B. and Mooney, S. *Digital Rights Management: Business and Technology*.

[2] Iannella, Renato. "Digital Rights Management (DRM) Architectures." *D-Lib Magazine* v.7, no. 6 (June 2001). Available at http://www.dlib.org/dlib/june01/iannella/06iannella.html.

Key DRM concepts and components include:

Rights Data Dictionary: A collection of standardized data elements required to identify entities and relationships in a rights transaction, including the rights or permissions extended to users for consumption of a resource; the constraints on the exercise of those rights; the agents involved in a transaction, such as the rights holder, the user, or the distributor; the application processing the resource; and the storage device housing the resource.

Rights Expression Language (REL): A REL communicates rights, obligations, and pertinent information, including identification of entities participating in a rights transaction. In addition, a REL facilitates and documents rights transactions among entities, is standardized according to documented rules, and employs a rights data dictionary and a standard syntax, such as XML.

Right: A right or permission is "the **most** that one can do with a resource."[3] It specifies how one may access or utilize a resource.

Constraint: A constraint or obligation is "the **least** one needs to do in order to be granted the right to access or use a resource."[4] A constraint limits or imposes a requirement that must be met to exercise a right.

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 3
SCOPE

A key issue in preparing a requirements document is determination of the scope of that document. Issues for consideration in this regard include:

3.1 Who is the audience for the requirements document?

3.2 How are the requirements of the various communities represented at the NSF Middleware Initiative and DRM Workshop to be included?

3.3 Should requirements be articulated "generally", i.e., higher-level requirements, at the application level, ontologically, or based on some other construct? What organizing principle will convey the most meaning in a general model of R&E requirements? If multiple constructs are employed, can

[3] Parrott, David, responding on behalf of Reuters. *Requirements for a Rights Data Dictionary and Rights Expression Language*. In response to ISO/IEC JTC1/SC29/WG11 N4044*: Reissue of the Call for Requirements for a Rights Data Dictionary and a Rights Expression Language*. MPEG-21, March 2001. (1 June 2001). Version 1.

[4] Ibid.

all of these exist in one document or would a series of documents be more suitable?

3.4 Should general DRM and REL requirements be presented in the same or separate documents?

3.5 A substantial amount of requirements documentation already exists for RELs. However, only some of these include or infer requirements for DRM architectures or solutions. In addition, it seems that in some cases the REL mirrors the functionalities and capabilities of DRM technologies, rather than vice versa. The documentation on REL requirements has also been designed to meet DRM requirements for the commercial sector. How can we leverage existing requirements documentation to insure that the unique needs of R&E are also met?

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 4
SCENARIOS

The following are a selection of scenarios that illustrate the complexity of rights management in R&E.

4.1 Collaborative Database Development

A university library is creating a state- and grant-funded digital archive of environmental indicator data. Published reports and original research results are contributed from government, academic, and industry researchers. The database design and functionality are developed collaboratively but implemented and managed by technical staff of the university. The university also provides content storage. The project is intended to be a model for other states to replicate. Source acknowledgements and terms of use for published reports, integrity and sharing of original research, and availability of the database design and functionality to other developers are some of the rights management issues that must be resolved.

4.2 Ambiguous Permissions

4.4 Fair Use

For an undergraduate course on Developmental Psychology, Professor Atwood has prepared a series of PowerPoint™ presentations, and she usually begins each class with a presentation that illustrates the concepts to be discussed that day. The presentations include photographs of children at different stages of development engaged in various activities, which were downloaded from a commercial photo archive that the university library makes available to its students, faculty, and staff. Professor Atwood's classroom use of photographs from this database fall under the fair use guidelines, and they it is expressly permitted under the terms of the site license between the university library and the database producer.

Professor Atwood is invited to be a visiting scholar at another university, and she is asked to give a lecture on her teaching methods to its faculty. The nature of this invited lecture falls under the fair use guidelines. However, the site license for the photo archive at her home university specifically prohibits using materials from its database outside of the home university's authorized user community of students, faculty, and staff.

4.5 Research Involving Human Subjects

Professor Weiner of the Human Ecology department is planning a research project to study seat belt use among teenagers. He is collaborating with faculty in psychology at his own and another university. The study will randomly select students at three high schools to complete a questionnaire. Students who return the completed questionnaire will be contacted for further study. The project will involve videotaped segments of interviews with students and weekly journal entries by the students that will be recorded online. The researchers intend to share and analyze the video segments via a website for the project. The web site will also record the student's journal entries.

4.6 Commercial Database

A university library subscribes to collection of e-journals. The license between the publisher and the library states that excessive or systematic downloading of articles can result in suspension of the subscription. On one occasion, it appeared that systematic process was employed to download a large number of articles. The traditional response has been for the publisher to contact the library's e-journals license manager, who works with the university systems office to identify the end user. The e-journal manager contacts the end user, who did not fully understand the parameters of fair use. The user is apologetic and agrees to be more careful. Now, however, the library has been notified that the publisher is implementing a "state-of-the-art" DRM system that will allow the publisher to identify directly any end user who downloads articles systematically. Misuse will require an additional fee to be paid by the end user, or the end user will be removed from the subscription database permanently.

4.7 Educational Repository

A professor contributes a course lesson, including a pre- and post-test, a digital video and a flash tutorial illustrating a scientific concept, to a multi-institution education repository. A professor at another institution downloads the course, replaces the pre- and post-tests with his own tests and adds textual discussion slides

related to the concept. The following year, the first professor replaces the course lesson with new pre- and post-tests and a new flash tutorial referencing recent discoveries that impact understanding of the concept.

4.8 Medical Scenario [5]

Tom, a Professor in Cell Biology, wants to call Jane, Chief of the Division of Infectious Diseases in the Department of Medicine, to discuss sensitive matters dealing with their work on the Faculty Promotions and Tenure Committee. Jane has her clinics and office in the hospital, which is a separate corporation linked to the university by an affiliation agreement. Her directory information resides on the firewall-protected hospital network. Tom has his lab and offices in a university-owned building and is connected to the "open" university network with all of his contact information posted on the publicly-accessible university directory.

Physician-related online directory information at the hospital is restricted for internal use but an agreement between the hospital corporation and the university gives authenticated medical school faculty, students and staff rights to access this information. Tom accesses the online university directory and enters Jane's name. He is authenticated as a medical school faculty member and presented directory information from the hospital directory that include Jane's phone number, office location and email address. The public and all non-authorized users at the university are only presented with Jane's hospital, department and division affiliations as well as her academic rank in the School of Medicine.

[5] This scenario is presented here with the permission of the authors of the draft *Internet2 Medical Middleware (MedMid) Working Group: Workplan Scenarios*, available at http://middleware.internet2.edu/medmid/draft-internet2-medmid-workplan-scenarios-latest.htm.

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 5
GENERAL DRM REQUIREMENTS

5.1 It is widely considered that several key legal and policy issues fundamental to the use of information resources in R&E are undermined by some current commercial DRM systems. These include fair use, the "first sale" principle, and privacy of the end user. The case for supporting these policies and principles in a DRM system is succinctly presented in the paper *Supporting Limits on Copyright Exclusivity in a Rights Expression Language Standard*.[6] In addition to the issues identified in this paper, what other policy and legal policies need to be accommodated in a DRM system?

5.2 Policies that support sharing of information with respect for intellectual property, academic honesty, and the integrity of the resource are an integral part of the academic culture and value system. How can DRM systems be designed to both leverage and support these policies?

5.3 What attributes does a DRM system need to have to support the fact that the academic community is a producer, an intermediary/distributor and a consumer of information?

5.4 Is there an ontological model that might be applied to describe DRM in R&E? Do existing models handle the potential for derivation, complex digital objects, and the distributed collaborations that are typical in R&E?

5.5 How can a DRM system support the heterogeneous nature of collaborations in R&E? A collaboration between an institution of Higher Education and a Department of Defense entity, for example, may require sharing of classified information. Can hybrid policies be developed to support these heterogeneous collaborations?

5.6 How can DRM systems accommodate the international aspect of education and research and support differences in laws, cultures, languages, and mores about what constitutes piracy, for example?

5.7 How can DRM systems be designed to be flexible and extensible enough to accommodate developments in e-learning, digital libraries, and online collaboration as they emerge? How can local needs for data elements and values within data elements co-exist with the requirement for interoperability in DRM?

---

[6] Samuelson Law, Technology & Public Policy Clinic, the Electronic Privacy Information Center, and John Erickson, HP Labs. *Supporting Limits on Copyright Exclusivity in a Rights Expression Language Standard.* Submitted to the OASIS Rights Language Technical Committee Call for Requirements for a Rights Expression Language. August 13, 2002. Available at http://xml.coverpages.org/OASIS-SLTPPC-EPIC-8-13-02.pdf

5.8 How can we ensure a DRM system interoperates with digital repository architectures already in place, particularly the administrative, technical and descriptive metadata that already populate these repositories?

5.9 What DRM requirements do new scholarly publishing models have?

5.10 How are collaborative digital objects managed in a DRM system, particularly complex objects such as educational resources concatenated as a "sharable content object" (SCO), (e.g., a sequenced digital lesson in a course consisting of many different unique separately-authored digital objects)?

5.11 How can the concept of duration be accommodated in DRM systems, i.e., how long will rights be documented and maintained? This is a critical concept in the digital library arena.

5.12 How is the concept of public domain respected in DRM systems? How are objects that under time fall into the public domain managed?

DRM systems will need, however, to facilitate communication between resource creators, between rights holders and downstream distributors, between users and the DRM administrators/rights holders and between users involved in the downstream transfer of resources. In *Supporting Limits on Copyright Exclusivity in a Rights Expression Language Standard,* the authors state that a digital rights communication protocol is also required in addition to a rights expression language and data dictionary.[7] Can a common protocol support both complex and simple transactions without unnecessary overhead? How would such a protocol interoperate with policy expression (XACML), security assertion languages (SAML), or a messaging protocol such as SOAP?

5.19 Can existing protocols such as Open Archives Initiative, SOAP, XACML, and OpenURL be extended and leveraged in an application suite to support DRM, rather than creating a new protocol which repositories will be required to support? What role can resolver protocols such as OpenURL play for resolving conflicts between multiple DRM systems in use?

5.20 Who is going to be responsible for DRM services on our campuses? Since a DRM instance entails an interaction of attributes about a user, content and usage, collaboration between the departments responsible for maintaining those attributes is necessary, at the very least. Most entities producing resources in universities often do so as participants in consortia, such as IEEE or the Research Libraries Group. How can DRM services be designed to support seamless interoperability between services offered at the consortial level and those offered at the local level, perhaps by the university's IT department or library? Without such integration, end users will be exposed to unnecessary complexity and even conflicts between rights and obligations offered by competing systems for the same resource.

[7] Ibid.

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 6
MIDDLEWARE AND DRM

In Higher Education (HE) today, the working definition of middleware is "a broad array of tools and data that help applications use networked resources and services." [8] Middleware services create an OS of sorts for the network, enabling users to transparently and securely access network resources; these services include authentication and authorization, identity management, policy management, resource discovery, and security. Since directories form the core of these services, a lot of work in this area is being conducted in directory development. In addition, the distributed collaborative nature of R&E activities makes inter-institutional, cross-realm access of critical import in middleware development.

In a discussion of the relationship between middleware and DRM, and a consideration of how that relationship impacts DRM requirements for R&E, some of the issues to be considered are:

6.1 Where do middleware and DRM architectures intersect with regard to technical and policy issues?

6.2 Are we proposing the development of middleware to support existing commercial DRM systems or vice versa – the development of DRM systems that work with existing and emerging middleware infrastructure? What is the merit of either approach?

6.3 What attributes do DRM technologies share with other applications that are currently enabled by the HE middleware infrastructure?

6.4 If DRM management is functionally separate from other middleware applications, how will changes be tracked and resolved across systems?

6.5 In what other ways will the HE middleware infrastructure shape DRM requirements for R&E?

6.6 How can the middleware fabric being developed by the NSF Middleware Initiative be leveraged towards implementing scalable and interoperable DRM solutions?

6.7 HE has made a significant investment in middleware infrastructure. Scalability, cross-realm access and extensibility are key criteria that have driven the development of this infrastructure. DRM systems should interoperate with this infrastructure. How easily will commercial DRM scale for inter-institutional sharing of resources?

---

[8] I2 Middleware Initiative: *Middleware FAQ*. Available at http://middleware.internet2.edu/overview/middleware-faq.shtml.

6.8 Infrastructural middleware is frequently managed by IT departments in academic environments. DRM is an intellectual property and content management activity that has not traditionally been managed by IT. How will integration of DRM with other middleware impact the management of DRM? Will this management be collaborative across multiple departments having legal and digital repository expertise?

6.9 Proprietary DRM systems typically focus on media packaging, encryption and enforcement at the client. Middleware in HE focuses on scalable, distributed, cross-realm content delivery and collaboration. If a DRM system were to be developed with these attributes, would it make sense to use the HE middleware infrastructure as a framework?

6.10 Are there other components evolving in the IT and networking infrastructure at our institutions, such as wireless computing and portals, that need to be included in this discussion?

6.11 What are the future middleware developments, such as Global PKI, HEPKI, and Role Based Access Control in directories, that are likely to affect implementation of DRM systems?

6.12 Is there development work in DRM that will impact enabling middleware? Is there future development work in middleware that might impact implementation of DRM systems?

6.13 Middleware capabilities are greatly extended with the establishment of communities of trust. Do DRM communities of trust have particular requirements?

DRM REQUIREMENTS FOR RESEARCH AND EDUCATION

SECTION 7
REQUIREMENTS FOR A RIGHTS EXPRESSION LANGUAGE

In addition to the requirements identified here, what would make a REL flexible and extensible enough to support R&E requirements? Is there incentive for the R&E community to develop its own REL or REL core schema?

A REL must actively support a number of DRM functions:

7.1 Identify and define all entities (agents) involved or potentially involved in rights transactions involving a resource. These entities include the resource itself, the creator, publisher, distributor, downstream rights holders, and the resource users. Entities can also include the applications and machines used to access, use or store the resource.

7.2 Identify, define and reference the resource for which rights and obligations are being managed.

7.3 Define rights and obligations, as well as attributes (temporal, locational, etc.) belonging to rights and obligations.

7.4 Bind rights and obligations about a resource to the resource.

7.5 Identify and reference obligation protocols and systems that apply conditions on the exercise of rights, such as authorization and authentication mechanisms, watermarks and third-party use tracking systems, encryption schemes, and third-party payment mechanisms or referrals.

7.6 Support description and management of entities and transactions across entities through transactional subschema and through data elements that identify and reference transaction types.

7.7 Provide documentation, such as copyright attribution, for resources without applying constraints on rights.

7.8 Document and explain licenses and other transactions between rights holders and users for the exercising of rights bound to objects.

7.9 Document trusted systems and applications for managing the digital rights and obligations.

7.10 Support downstream transactions, including rights transfer, use tracking, etc.

7.11 Support privacy for all entities involved in a rights transaction.

7.12 Log statistics on the exercise of rights and constraints.

In addition to robust support for DRM functionality, a REL must support modular constructs to handle complex situations and promote integration and interoperability with other resource management systems. These include:

7.13 Administrative metadata that documents the creation, modification, and lifecycle of the rights metadata governing a resource. This includes version control for the metadata schema in use.

7.14 Inheritance and/or other rules of precedence and application for hierarchical or modular metadata statements so that rights and obligations do not need explicit repetition.

7.15 Extensibility for the addition of local data elements, qualifiers, attributes and value lists (controlled vocabularies).

7.16 A core element set that interoperates with other RELs to enable minimal semantic interoperability across repositories.

7.17 Explicit mapping across data elements to other RELs to enable higher level semantic interoperability across repositories.

7.18 The ability to reference or incorporate in a modular manner the descriptive and administrative information, not explicitly bound to rights, that also documents the resource (e.g., the descriptive metadata in Dublin Core) or documents other entities (e.g., a Vcard record for a rights holder). This information includes administrative metadata managing the lifecycle of the resource that may impact constraints on rights, such as security classification level, retention schedule, etc. Support for integration of other associated metadata must include the schema name, version and associated namespace.

7.19 Integration of descriptive, administrative and rights metadata governing the resource. This integration should facilitate automatic resolution of conflicts across metadata records that result from a modification in any record (e.g., an institutional document that exceeds its retention schedule and is discarded must be removed from discovery and access and rights management systems). A document that is unclassified and thus available to general users must have its authorization constraints modified or removed. Protocols that bind metadata about an object such as METS (Metadata Encoding and Transmission Protocol) could be utilized by the rights management system to facilitate resolution of conflicts.

7.20 Support for default rights and obligations and methods for automatically adjudicating among conflicts.

7.21 Support for hierarchies of alternative rights tied to descending obligations to enable the automated resolution of obligations leading to the exercise of a right to a resource (e.g., authenticating as a student provides access to a high resolution video, or alternatively, inability to authenticate as a student provides the option to pay for a one time viewing of a lower-resolution video.)

7.22 Two-way communication, allowing the end user to request a right that cannot be explicitly or automatically provided by meeting an online obligation. This

request might trigger an automatic decision tree or transfer the request to the human administrator or rights holder. This might entail interaction with a communication protocol.