

DOI in a DRM environment

A white paper for Copyright Agency Limited (CAL)
DOI EPICS project

by James Dalziel
Director, Macquarie E-Learning Centre of Excellence (MELCOE)
Macquarie University
james@melcoe.mq.edu.au

<http://dx.doi.org/10.1275/1002>

Contents

Contents.....	2
Application of DOI in a DRM environment.....	3
What is DOI?	3
Benefits of DOI	5
DOI and DRM.....	7
DOI and DRM for the EPICS Project	10
Other challenges for DOI	11
Specific DRM issues in DOI Coursepacks business model.....	13
1. DOI for DRM across the value chain	13
2. Standards and interoperability	14
3. How does DOI relate to existing custom publishing practices?.....	14
4. Additional possible stakeholders.....	15
5. End user customisation	15
6. Importance of dissemination of findings.....	16
7. The role of digital delivery	16
8. The relationship between student information system and bookshop	16
Specific use cases for DOI Coursepacks model	17
Lecturer	17
CAL.....	17
Publisher	17
University.....	18
Typesetter/renderer.....	18
Printer	18
Bookseller	18
Student.....	19
Summary.....	19
Implications for CAL.....	20
COLIS, MAMS and DOI	21
Implications for CAL EPICS project revisited	21
References.....	27
Appendix A: Learning object identifier requirements.....	30
Appendix B: Kinko's Coursepacks press release	32
Appendix C: Recent international conferences/workshops of relevance to DOI	35

Application of DOI in a DRM environment

What is DOI?

The digital object identifier (DOI) is not simply a number, it incorporates:

- an open standard describing a syntax for unique identifiers (ANSI/NISO, 2000)
- a technical implementation infrastructure, based on persistent identification and resolution, using the Handle system (Paskin, 2003a, 2003c)
- a 'social' implementation infrastructure covering governance, policy, membership, etc (Paskin, 2003b)
- the ability to associate metadata and services with identifiers using the extensible indecs ontology

The core features of DOI are its persistence and uniqueness over time, and its ability to reliably resolve ('point') to resources. Each of these features is a significant advance over the problems associated with URLs and the 'dead links' which plague the Internet. The additional components of the overall DOI model outlined above have been provided by the International DOI Foundation (IDF) to ensure a complete and self-sustaining implementation infrastructure for DOIs. The breadth of the endeavour is neatly summarised by Paskin (2003a) as follows: *DOI is the UPC (Bar Code) for objects of intellectual property on the Internet.*

While the full DOI infrastructure provided by IDF has many advantages (discussed further below), the fundamental concept of persistence in and of itself can be valuable for content owners:

Persistent identification alone is a good enough reason to adopt identifiers such as DOI which provide a means by which potential customers can find your digital offering even if a 'broken link' URL of a retailer or other intermediary intervenes. (Paskin, 2003c)

The benefits of a broadly adopted, unified scheme for digital object identifiers are illustrated by the following example:

Consider what would have happened to the EAN/UCC bar code system if the grocery trade, furniture trade, shoe trade etc. had each devised their own solutions in isolation. Yet that is almost what is happening now, and IDF has made strenuous efforts to reach out to many different groups. Fortunately, there are signs that this message is getting across. (Paskin, 2003b)

Continuing the bar code analogy, Paskin illustrates the importance of the broader activities of IDF beyond the creation of the identifier itself as follows:

Equally importantly, the recognition that standards alone are not a working implementation; a sustainable infrastructure both of technology and policies was

needed to harness them for practical deployment. A useful analogy is with the physical bar code system, where the standard bar code symbology and code readers are only a small part of a business infrastructure of managed allocation, policies and commercial application tools, allowing a wide range of uses (both commercial and non-commercial). (Paskin, 2003b)

IDF is a membership-supported not-for-profit organisation. This does not mean everything is 'free', as noted by Paskin (2003b):

The DOI Foundation is a not-for-profit organisation, not a commercial operation; however, the system has costs that need to be met. Persistence is a function of organisations, not technology: to support a persistent identifier system, a persistent organisation needs to exist.

As Paskin (2003b) notes in a quote from Rosenblatt, even when a standard is free and open, there are significant hidden costs which need to be acknowledged to ensure a realistic approach to sustainability is adopted.

Free open standards are a nice idea, but just as there is no such thing as a free lunch, there is no such thing as a free standard. Even many of the programmers who contributed to Linux did so through the beneficence of their employers or educational institutions, who furnished their equipment, Internet connections and so on ... it's not just enough to throw a spec out there; it has to be resourced with market development, training, reference applications ... and so on — all of which costs somebody money. The only question is who will pay for it. Plenty of technologically worthwhile would-be standards have withered and died because nobody would.

However, this acknowledgement of realistic costs does not mean that DOIs can only be used in commercial settings:

DOIs can be used to identify free materials and transactions as well as entities of commercial value. (Paskin, 2003b)

and

Commerce does not necessarily mean the exchange of money: any environment where creations are made or used employing electronic means is encompassed by commerce in this sense. (Paskin, 2003c)

This is an important requirement for the education sector, where both commercial and non-commercial content models typically exist side-by-side due to the presence of both Government funding of public education and commercial publishing. Unlike many standards and specification bodies (including e-learning organisations such as IMS), IDF has recognised that implementing a standard in a consistent and reliable way requires significant policy and governance infrastructure over and above the technical infrastructure required to make the standard 'work'.

What this means for practical implementation: irrespective of Internet specifications, to make use of persistent identification schemes in useful ways will usually require more than a simple technical implementation. Especially, policy and governance issues (such as scope, authority to issue), and control of assigned metadata (quality control, interoperability considerations, etc) will be important components in adding value in practical implementations. (Paskin, 2003c)

There are important lessons to be learned by other standards initiatives from the work of DOI in developing this comprehensive technical and 'social' infrastructure to enable the use of the standard.

All of the above is not to suggest that DOI has addressed all issues related to its use, or that there are no misunderstandings of, or ongoing challenges, for DOI. For a review of various 'myths' about DOI, together with a list of 'open issues and difficulties', see Paskin (2003b). This section will conclude with a discussion of other challenges for DOI.

Benefits of DOI

The fundamental benefits of DOI have been noted above – that is, the persistent and unique identification of digital objects, and a persistent resolution mechanism to help 'find' objects. However, the broader DOI/IDF infrastructure has a number of additional benefits which may not be immediately obvious such as these:

- The use of the Handle system makes DOIs 'actionable', that is, clicking the identifier can link to the resource or to related metadata/services.
- The use of metadata Application Profiles associated with DOIs allows not only for the easy updating of information for individual resources, but where relevant, the updating of an Application Profile itself will cascade to all DOI-enabled resources using that Profile – as Paskin (2003b) points out:
If the Acme Registration Agency registers one million DOIs all of Application Profile 9 and a year later adds one more service to the three services that were available from the start, only the single 10.AP/9 record needs updating, and not the one million DOIs already tagged with AP9.
- Another significant efficiency arising from the use of DOIs occurs when the organisation affiliated with a DOI changes – this requires a simple updating of the relevant DOI field, not a 'by hand' change to every resource – (Paskin, 2003b):

Another advantage is flexibility over time as the document origins reflected in a hierarchy lose meaning, such as a change in ownership. (If acme.com sells some assets to newco.com, all URL filenames beginning acme.com/ that pertain to the sale need to be changed.) This benefit has already been seen in the case of CrossRef, where millions of DOIs identified through the Academic Press IDEAL system were merged into Elsevier's Science Direct system when the companies merged.

- A recent integration of DOI into PDF documents (Paskin, 2003b) allows the reader to be automatically alerted to newer versions of the document that may be available.
- Using DOIs can result in significant cost savings for publishers and related organisations – in Rosenblatt (2002), four scenarios are outlined, each indicating significant savings from the adoption of DOI for management of resources.
- The mapping of DOI metadata to a central ontology (the indecs framework) assists with developing cross-walks to other metadata schemas, as the number of required mappings is greatly reduced by the presence of a central framework. A similar case exists for mapping of cross-referencing between journal vendors. For example, Sidman (2001) points out that if 50 publishers use a central cross-referencing infrastructure (such as DOI), then each publisher only needs to create one cross-linking implementation (from their own system to the central scheme). Without DOI, each publisher would need to establish bilateral cross-referencing agreements, which would lead to a total of 1225 (not 50) different cross-linking implementations. This feature has been of significant value to the highly successful CrossRef implementation of DOI for journals.
- Unlike some other persistent identifier systems (such as PURL), DOIs are not dependent on the http protocol. In the future, DOIs can easily be used with other protocols, such as in the area of mobile devices.

This is not meant to be an exhaustive list of DOI benefits, but rather an overview of some additional benefits that may not be immediately obvious from the basic description of DOI. There are specific benefits for the Copyright Agency Limited (CAL) Coursepacks project that come from its adoption of DOI – these include:

- The CAL business model is based on a printer accessing the relevant coursepack from the CAL website, and at the time of access, specifying the approximate number of copies to be printed. This approach could provide CAL with detailed, near real-time record keeping of copyright usage. This, in turn, provides a basis for more efficient, regular royalty distribution and auditing.
- Due to the fact that the Coursepack model provides licensing of content from publishers using agreed licence formats and royalty distribution, there is no need for any special restriction on the amount of content from a particular volume that can be included (e.g. no more than 10%).
- Efficient support for coursepacks that combine content from a variety of publishers within the same coursepack, rather than custom publishing of content from only a single publisher.
- A growing library of 'pre-approved' DOI-enabled content which keeps a record of copyright clearance and royalty requirements, and hence allows for rapid incorporation into coursepacks using automated reporting procedures.
- In the case of digitally delivered student coursepacks, there is the potential for rights-enabled management of content to address publisher concerns about piracy, etc. (provided that suitable enforcement technologies are adopted for the coursepack).

In summary, there are many benefits associated with the adoption of DOI. The following section explicitly addresses DOI and DRM.

DOI and DRM

The following three quotes illustrate the nature of the digital object identifier (DOI) and its potential for use in digital rights management (DRM):

The Digital Object Identifier (DOI®) is a system for identifying and exchanging intellectual property in the digital environment. It provides a framework for managing intellectual content, for linking customers with content suppliers, for facilitating electronic commerce, and enabling automated copyright management for all types of media. Using DOIs makes managing intellectual property in a networked environment much easier and more convenient, and allows the construction of automated services and transactions for e-commerce. (DOI web site, 2003a)

DOIs provide essential tools for driving the growth of Australia's digital content industry. Their adoption within Australian content development, publishing and printing industries will significantly increase the competitiveness of the industry as a whole as the system specifically addresses one of the greatest constraints to growth – namely rights and permissions infrastructure requirements. (p.4, EPICS Grant Application)

Both pilots are designed to facilitate DRM (digital rights management) transactions across the value chain and include options for delivering content online or in print using print-on-demand (POD) facilities through strategic alliances (p.4, CAL Coursepacks business plan)

It is clear from these examples that the type of digital rights management envisaged here is broader than the management of copyright clearances and royalties – it incorporates 'automated services and transactions for e-commerce', 'rights and permissions infrastructure' and 'delivering content online or in print'. For most digital rights management experts outside the DOI community, these requirements indicate a need for digital delivery of rights-enabled content. This, in turn, generally requires access and identity management, and/or encryption/watermarking technology.

Historically, the focus of DOI was initially on managing creator/owner rights (incorporating tasks such as copyright clearance). For example, Paskin (2003b) notes: *The original conception of DOI was simply as a tool to aid in managing content for rights owners.*

The difficulty with DOI for end-to-end DRM is that while DOI is a necessary (required) condition for DRM, it is not a sufficient (complete) solution for DRM in and of itself. The recognition of DOI as necessary but not sufficient is acknowledged in the DOI handbook:

Assigning an identifier does not of itself change any aspect of a digital object's use; it is a prerequisite for copyright management — necessary but not sufficient — hence copyright infringement could occur. However, removing such an identifier from material constitutes 'modification or removal of copyright management information' and is prohibited under legislation implementing the WIPO Copyright Treaty, such as the US Digital Millennium Copyright Act. (DOI, 2003b)

In a related example, Paskin (2003c) begins his detailed DRM article as follows: *Identifiers (unique labels for entities) and metadata (structured relationships between identified entities) are prerequisites for DRM.*

While this is an entirely accurate statement, it does not explain that DOI on its own is insufficient to address 'delivery-end' DRM in contexts where protection of digital resources is required. While delivery-end issues are addressed briefly towards the end of Paskin's article, it is my sense from DOI literature that most of the focus on DRM within the DOI community is on managing rights using digital means (e.g. electronic management of rights workflow tasks), rather than on managing digital rights. This has the potential to mislead naïve readers into believing that DOI provides a sufficient solution for end-to-end DRM.

Contrast the following quote from a recent article by Paskin (2003c) including a quote from Godfrey Rust:

Mapping complex concepts is possible, but concepts like 'digital rights management' are not currently consensually precisely defined; there is a majority view that it is digital management of rights, rather than management of digital rights, but beyond that 'DRM is something to do with managing, something to do with rights and something to do with the digital environment. But not necessarily.'

with the view of DRM articulated by Rightscom (2003) on behalf of the MPEG community:

These two strands of work, the Management of Digital Rights (identification and metadata, including the expression of rights) and the Digital Management of Rights (encryption, watermarking, authentication etc), have been gradually pulled together in a single standards forum, the Moving Picture Experts Group, MPEG.

Despite these issues, persistent and unique identifiers are an essential part of a successful DRM end-to-end solution. As Sidman (2000) notes:

Why is DOI an enabler of DRM?

- *Globally-unique, machine-readable ID enables interoperability among all the 'moving parts' of DRM process.*
- *So does DOI's linking capability: persistent links which don't break, perform one-to-many linking instead of single linking, are centrally maintained/updatable.*
- *Use of DOI to represent (and link to) digital certificates, digital signatures, authentication records, etc.*

The first two points are important reasons for considering a persistent, unique identifier (such as DOI) as part of any DRM infrastructure. The final point relates directly to delivery-end issues, but there is little further discussion of these issues in subsequent DOI literature to my knowledge. The likely reason for this is that these are some of the most difficult issues of DRM, as they require links to other enterprise systems (such as access and identity management) which are difficult to implement. It was these problems that were among the most challenging elements of the COLIS project's use of delivery-end DRM that were noted in my earlier paper. See the 'Implications for CAL' section (below) for further discussion of potential collaboration between CAL and COLIS in this area.

Part of the reason for difficulties in delivery-end DRM has been the plethora of competing proprietary systems which have made it a nightmare for publishers of materials such as eBooks to manage competing formats. Rightscom clearly identifies this problem in the quote below, including the helpful analogy with CD players:

From a management perspective, one of the major perceived barriers to the development of extensive secure online markets for content is that the security applications available from vendors appear to be non-interoperable. Whereas in the analogue space, content owners are used to being able to choose technology and be relatively sure that the way they package the content will not prove an impediment to its consumption (a CD produced on one pressing system is interoperable with all CD players). This does not seem to be the case in the digital environment. Worse still, different encryption and player systems often dictate that content owners have to develop multiple work flows for content packaging, dramatically increasing production and distribution costs. (Rightscom, 2003)

In terms of a solution to the need for end-to-end DRM, there are two main components related to delivery – the (secure) expression of rights, and their enforcement. While enforcement technology is currently fragmented, proprietary and immature, expression of rights is making progress, particularly through the development of rights expression language standards. Rightscom (2003) outlines the basic components of rights expression as:

- *The principal to whom the grant is issued*
- *The right that the grant specifies*
- *The resource to which the right in the grant applies*
- *The condition that must be met before the right can be exercised*

N.B. There are two types of identity information assumed here – the first point assumes people/party identifiers, whereas the third point assumes object identifiers.

Lest there be any doubt that the work on rights expression is relevant to DOI, the following points from Sidman (2000) unknowingly describe the exact purpose of a rights expression language in relation to DOI for DRM:

DRM is more than Anti-Piracy.

It allows the Publisher to specify all the things that the reader can do with content:

- *Read*
- *Sample/preview*
- *Print*
- *Copy/paste*
- *Re-use*
- *Modify*
- *Forward*
- *Re-sell*
- *Syndicate*

They can also specify:

- *How many times*
- *For how long a period*
- *For what price*
- *To whom (forwarding)*

Not just negative (locking content up), but also affirmative (new ways to sell, great marketing potential)

Notably, the MPEG-21 Rights Expression Language (REL) group has adopted XrML from ContentGuard as the basis for its REL, but has also adopted the indecs ontology as the basis for a Rights Data Dictionary (RDD) as noted below:

In order to support the REL, MPEG-21 is also defining a specification for a Rights Data Dictionary, with a methodology for defining and cataloguing terms necessary for the description of rights. The MPEG-21 RDD specification, though independent, is closely linked to the REL specification in order to ensure the interoperability of rights semantics across different content types. This will greatly enhance the value of the REL to content owners.

The link between indecs and rights expression may provide the foundation for the DOI community to incorporate end-to-end DRM in the future, provided that the technical challenges of enforcement can be resolved. This is an area that requires further investigation and close monitoring.

DOI and DRM for the EPICS Project

The Coursepacks business model has been designed to avoid many of the 'delivery-end' DRM challenges noted above by relying on a 'Business to Business' (B2B) model rather than a 'Business to Consumer' (B2C) model. In practice, this means that

the use of coursepacks arises from an agreement between CAL (on behalf of publishers) and a given university to require mandatory purchasing of coursepacks via a fee collected at enrolment from all relevant students. As a result of this business model, there is little incentive for students to misuse coursepack digital content, as all students receive this content as a result of the mandatory fee. On this basis, both physical and digital versions of the content could be distributed without necessarily requiring delivery-end DRM such as web-based authorisation or encryption.

The challenge for this model would be if the mandatory fee approach were not possible. Without this feature, the B2B model may collapse, resulting in a return to a more typical B2C course content model equivalent to purchasing a textbook from the bookshop. If digital content were available only to those students who had paid an appropriate fee via the bookshop (or online equivalent), then delivery-end DRM may be required to reduce the likelihood of students who have purchased the digital content providing it to students who have not purchased it. In this case, the concerns raised above about DOI and delivery-end DRM would be relevant.

Other challenges for DOI

Apart from DRM issues noted above, there are some other challenges ahead for DOI which are worth briefly noting here, as they may impact on the future success or failure of the proposed Coursepacks business model.

- The main focus of DOI is on object identifiers. While it is occasionally noted that DOI may be applicable to people/parties, this is not its main function. IDF acknowledges its collaboration with the European Interparty project on people/party identifiers as part of its engagement with these issues. However, there is a large community focused on enterprise identity management as represented by the Digital Identity World Conference (see Appendix C for details). This community has focused mainly on people identifiers, but also on identifiers for objects, processes/services and applications. As a community, it is increasingly coming to the view that all kinds of identifiers should be managed through locally-managed enterprise directories. This is an important challenge for DOI as it represents a different model for managing identifiers of digital objects, and their relationships to other kinds of identifiers, such as people and processes. It is important that IDF investigate this approach to better understand its points of similarity and difference with DOI, and the potential for interoperability (or otherwise) in the future.
- While IDF is a not-for-profit organisation working towards a cost recovery model, some identifier experts within the education community have raised concerns about the cost of DOIs – e.g. Friesen (2002) *Note that the use of [DOI] involves significant initial and ongoing costs, and is consequently not often being considered by many educational infrastructure initiatives.* While this concern may arise from insufficient consideration of the complete costs of well-managed identifiers, it at least acknowledges that further discussion and justification of costs within the education community may be required. Perhaps the greatest concern in this community is that DOIs are currently provided on a variable

pricing model (i.e. 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 (Dan Rehak estimates over one million per student for ePortfolios within the next few years); 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.

- Given the mandatory fee proposal combined with a digital delivery option (in addition to print), could students be given a choice of receiving both print and digital versions for a fee, or just the digital version for a discounted fee (with the discount equivalent to the cost savings in paper and ink)? What would the student demand for a digital-only version be if provided with a discount, and how would the size of the discount impact on student demand? These questions may be appropriate research questions during the Coursepack initial trials.
- There is a need for more innovative use cases about the combination of print and online learning materials according to student needs. At the recent AICTEC industry forum, I proposed the following use case to try to capture what I see as the best of all worlds for authors, teachers, printers and students:

A student uses a secure e-book reader to view a large teacher-created 'coursepack' of readings sourced from many different publishers. The student can choose to print out individual sections themselves, or send off a request to their university printing service for overnight custom printing of large selections of the coursepack. Publishers receive an automated, fair return for their relevant contributions to the coursepack.' (Dalziel, 2003)

N.B. 'Automated' in the context of the Coursepacks business model could mean the figure supplied by the printer when the coursepack is downloaded from the CAL website (see above).

Specific DRM issues in DOI Coursepacks business model

This section is based on the 'general comments' section of the 'Reflections on the Coursepacks EPICS project' presentation from the second project steering committee meeting. It explores general issues with DOI Coursepacks business model from a DRM perspective. Following it is a section on issues related to the particular actors in the business model and the use cases associated with these, based on the 'Specific business model issues' section of the presentation.

The presentation listed eight general comments related to the project and DRM – each of these is developed below:

1. Use DOI for DRM across the value chain

This phrase (which comes from p.4 of the CAL business model document) implies that an end-to-end solution for digital rights management is possible using DOI. In one sense this is true – DOI acts as a persistent identifier throughout the value chain, from (potentially) creation, through management by publishers, selection of resources by lecturers, packaging of resources into coursepacks for typesetters/renderers, printing of resources by printers, and the distribution of coursepacks to students through bookshops. The presence of this identifier, together with easily discovered information (metadata) and services related to this resource can encourage efficient production of coursepacks.

For example, DOI can assist rapid incorporation of resources into coursepacks where there is existing copyright clearance and royalty distribution requirements are associated with DOI (i.e. stored as metadata with the identifier). DOI can also provide a location for sourcing technical details for typesetters/renderers in the conversion from digital to printed materials. It may even be possible for universities and bookshops to use DOI for tracking distribution of printed coursepacks in a way similar to a product barcode. These features of DOIs are likely to result in considerable efficiencies and savings once DOIs are widely adopted and used at scale.

However, there is a second sense in which the phrase may be misleading. In a digital environment, management of rights across the value chain implies that this extends to DRM at the 'delivery end' for digital resources. In practice, this often requires technologies such as encryption and access management, or at least web-based authorisation. Today's students are clear that they want the option of both printed and digital forms of learning resources. Hence, if the proposed B2B business model (including the mandatory fee) were not possible, the CoursepackC project would need to explore options for digital delivery of resources which incorporate DOI and include 'delivery-end' DRM.

2. Standards and interoperability

DOI itself is an open standard (ANSI/NISO, 2000), and IDF has been heavily involved in a number of standards efforts, particularly indecs, MPEG-21 RDD and ONIX. As Paskin (2003b) notes:

IDF has placed a strong emphasis on communicating with, and participating in, a wide range of related standards and consortia activities.

This commitment to open standards in process and technical implementation is to be highly commended. It provides a solid foundation for engaging with vendors and other standards initiatives.

However, the standards world is large and complex, and in the current age, many areas of digital content are converging, and hence there is an even greater need for liaison between standardisation efforts and harmonisation where relevant. In the context of the Coursepacks project, the use of coursepacks content for digital delivery may require liaison with the relevant e-learning standards initiatives such as IMS and IEEE LTSC. In particular, the work of the IEEE LTSC on Digital Rights Expression Languages (DREs) is directly relevant to DRM and DOI. In this context, it is recommended that IDF engage further with the IEEE LTSC about the use of DOI for DRM, and that the EPICS DOI project contribute one or more use cases to the DREL study related to digital delivery of DOI-enabled learning content. Given the near to complete nature of the current IEEE work, this recommendation should be considered urgent. See the comments on question 10 in the section 'Implications for CAL' for further details.

Agreed standards are a foundation for interoperability, but there are always significant implementation issues to be agreed before an area becomes truly 'plug and play'. For discussion of interoperability issues for IDF with DOI and Handle, see the comments on question 3 in the section 'Implications for CAL'.

3. How does DOI relate to existing custom publishing practices?

Publishers such as McGraw-Hill have invested significant resources in custom publishing systems which provide similar overall outputs to the proposed Coursepacks model. In a related context, Kinko's has recently announced that it will be offering coursepacks as part of a revised business strategy (see Appendix B for details). It is important that the DOI approach demonstrates its advantages over existing custom publishing/coursepack approaches – these may include:

- Ability to easily aggregate content across many publishers for a single coursepack
- Centralised management of copyright and royalty distribution across multiple publishers
- Greater depth and variety of content available to lecturers
- Standardisation of typesetter/renderer information across different publishers via DOI technical metadata

The success of the CrossRef model in the field of journal publishing may be a valuable parallel example in making the case for DOI over custom publishing.

4. Additional possible stakeholders

The Coursepacks project has engaged a wide range of stakeholders across the value chain to ensure representation of different views in the development of the business model, and the value of this approach has been demonstrated at steering committee meetings. Given the current business model, it may be necessary to include some additional stakeholders due to the impact of the proposed model on their areas. These additional potential stakeholders are:

- Student information/records system vendors (e.g. PeopleSoft, Callista and Technology One) – the systems of these vendors manage student records, and hence may be required to record information (e.g. payment for coursepack). Given the high cost of these systems and of making any change to them, this stakeholder group should not be overlooked.
- Bookshops – the tasks on the bookshop/coursepack distributor are quite onerous in terms of verifying the identity of students requesting a coursepack, keeping records of who has received a coursepack, managing physical stock that is not 'sold' using normal sales methods and tracking (e.g. electronic cash register linked to inventory), etc. Given the potential complexity of managing the workflow associated with coursepacks, they are an important stakeholder.
- Librarians – if the coursepacks are to be made available in digital form, then issues relating to e-reserve access will require engagement with librarians. Even without a digital format, the issue of libraries keeping copies of coursepacks in closed reserve will need to be addressed.
- Enrolment administrators – apart from the technical requirements on student information system vendors, the processes for administering the Coursepack model during enrolment will require further investigation, for example, will vouchers be used, and if so, who will manage their creation and distribution.

5. End user customisation

The Coursepacks business model refers to 'end user customisation', but it is unclear whether the end user envisaged is only the initial lecturer (who can choose which resources go into the coursepack) or other future users, such as corporate trainers or educators from other sectors. It is important to set the right expectations within the relevant communities about who will be able to benefit from end user customisation, and how this will operate.

6. Importance of dissemination of findings

One of the most useful aspects of the COLIS project has been the dissemination of project findings to assist others in addressing similar issues. This is just as important for 'positive' (successful) aspects of the project as it is for 'negative' (difficult or failed) aspects. Given that the Coursepacks project is government funded, it is important that the findings are widely disseminated regardless of success or failure. In particular, this applies to the use of DOI itself – any problems that arise with its implementation should be addressed in the dissemination of project findings. Ideally, this should also include an independent review of the project outcomes and its implications for future use of identifiers in coursepacks.

7. The role of digital delivery

As noted, digital delivery poses important challenges for DRM, and yet it is highly valued by students, who desire both printed and digital versions of learning content (printed so that they can avoid the problems of extended screen reading and make notes on the printed page; digital so that they can take advantage of digital functionality such as search, copy and paste). This is explicit in the use case in question 10 (at page 29).

The use of DOI for DRM indicates that most of the focus of DOI to date has been on the creator/owner end of the DRM spectrum, and less so on the digital delivery end of DRM. This suggests that a successful coursepacks model either needs to avoid this issue (via a B2B business model) or solve the challenges of delivery DRM for a B2C model (if the mandatory fee approach was not possible). While aspects of DOI infrastructure may be useful in assisting this (e.g. Handle), additional technology and processes will be required (e.g. access and identity management – that is, mapping people identifiers to object identifiers to make decisions about access). If encryption of content is required, additional technical challenges related to this requirement will need to be addressed.

8. The relationship between student information system and bookshop

It appears more investigation of the relationship between the student information system and the bookshop is needed so as to ensure that the right students get the right coursepack reliably within the right time frame. The specific problems with this element are discussed in the next section. The inclusion of a student information system vendor and a bookshop manager on the steering committee (as noted above) would provide valuable feedback on the potential challenges related to this element of the business model.

Specific use cases for DOI Coursepacks model

This section complements the general comments made in the previous section with specific questions/concerns related to the actors/stakeholders in the CAL Coursepacks business model. This analysis takes into account the new use cases presented at the second steering committee meeting.

Lecturer

The process of searching and selecting content items for coursepacks will require further detail (how is this achieved, what search protocols are used, etc.), although this issue is mainly applicable to the software development stage. The current lecturer use case is appropriate as a high level abstraction, but will require more explicit workflow description to ensure the correct functioning of the software for selecting coursepack items.

CAL

CAL has a key facilitation role in ensuring that issues related to other actors in the value chain are addressed to ensure the overall success of the coursepacks model. The inclusion of representatives from across the value chain is an important step in this process, although as noted above, additional stakeholders (particularly student information system vendors and bookshop managers) may need to be included to ensure there are no gaps in the proposed model.

It would be interesting to compare the proposed CAL licensing agreements with those already in use for the COLIS Demonstrator project and the VET-sector AShareNet project. Harmonisation of learning content licensing frameworks would be of significant value to the education sector if approaches like AShareNet, COLIS and CAL Coursepacks are widely adopted.

Publisher

The publisher use case provides a useful high-level abstraction of publisher issues. The metadata tagging process associated with submission should be described in more detail for publishers as they engage with the Coursepack model, and the advantages of this approach (e.g. items only need to be described once, and then can be re-used many times across many coursepacks) should be emphasised. The potential for automation of metadata creation should also be explored (e.g. TSO in the UK made very significant time savings through automation of DOI metadata creation for its 450,000 records).

The benefits of the DOI approach over existing methods (such as custom publishing, Kinko's Coursepacks) should be clearly outlined for publishers to assist in the uptake of the Coursepack model.

University

This actor may require inclusion due to issues associated with tracking data about students and coursepacks. Issues to consider further include how coursepack data is recorded in the student information system (e.g. where, by whom, at what stage of the enrolment process); how payment data is transferred to the bookshop (e.g. vouchers, electronic lists over email); and any privacy issues associated with payment information.

It remains unclear to me whether the mandatory payment approach is permitted in Australian higher education institutions for on-campus students. This issue requires a clear response from DEST as it is a fundamental component of the Coursepack model.

Typesetter/renderer

The use case for this step is appropriate. The only issue to observe here is the related question of digital coursepacks for students and delivery-end DRM.

Printer

The use case for this step is appropriate. Further detail on how the printer provides reports to CAL, and any related privacy issues, will need to be addressed. The use of the phrase 'print-on-demand' requires further clarification. The model proposed for the trial may be interpreted by some as being 'traditional' printing rather than 'print-on-demand'. Further discussion of the cost impact of alternative printing techniques is required. This could include discussion of the practicalities and costs of the scenario described in the AICTEC Industry Forum use case.

Bookseller

This currently appears to be the most problematic area of the Coursepack model. The use case requires further details of how the coursepacks are made available to students (addressing the concerns raised below), and how this affects the business processes of the bookshop. Given the potentially onerous nature of coursepack distribution once the model is operating at scale, the incentives for bookshops to be involved in this process may need to be revisited.

The quote from the CAL Coursepacks business model:

The student logs on to the bookshop's web site, selects the relevant coursepacks and gets them printed on a print-on-demand facility located within the bookshop. (p.5)

will require significant further investigation and development to ensure this is a workable scenario. If it is not, then alternatives to this scenario that still achieve the objectives of the grant will need to be described.

Some of the problems with this element of the model include: student identity information required for bookshop web site login; matching of enrolment information to coursepack distribution information; the evidence required by a student to acquire a coursepack (e.g. voucher, student card, other credential); the recording of distribution of coursepacks by the bookshop (to avoid giving out multiple copies); reporting of distribution back to the university and/or CAL; and management of a 'non-sale' item of inventory by the bookshop.

If the problems associated with the current model are not easily resolved (or if mandatory payment for coursepacks is not permitted by DEST), an alternative would be to treat coursepacks just like any other printed learning content and sell them through the bookshop using traditional sales and reporting processes.

Student

A student use case should be developed to describe the acquisition of the printed coursepack (to complement the expanded bookshop use case described above). If digital delivery of coursepacks is included, the process of acquiring and using the digital coursepack should be described.

It would be interesting to explore the workflow and cost issues associated with the AICTEC use case example of individual overnight print-on-demand requests.

Summary

The main point of concern with the current model is the relationship between university enrolment processes and the distribution of coursepacks through bookshops. While these concerns can be addressed relatively easily 'by hand' during the initial trial, this approach will not be sustainable if the Coursepacks model is successful. Further investigation of these issues is required to ensure that the Coursepack model is not limited to a small number of 'by hand' implementations.

The issue of whether mandatory payment for coursepacks is permitted by DEST also needs to be addressed.

Implications for CAL

In this section, I provide further comments on the earlier 'Reflections on digital rights management for education' paper based on the review leading up to this white paper, together with commentary on the concluding questions from the earlier paper.

In terms of the original paper, the review of DRM for DOI and recent discussion about identifiers with international experts has confirmed the general points made earlier. The breadth and complexity of DRM is an important challenge for any initiative which takes DRM as a focal point of its work, and hence this applies to DOI. In particular, discussion of DOI within the DOI community tends to focus predominantly on the creator/owner 'end' of the spectrum (copyright management), rather than the 'delivery' end. While this may be an appropriate focus given the historical origins of DOI and its current areas of successful application (e.g. CrossRef), it is important that the whole DRM spectrum be addressed for the future if DOI is to provide an end-to-end solution. Alternatively, the DOI community needs to become more explicit about which parts of the end-to-end solution it addresses, and where solutions to the remaining components can be found.

The CAL Coursepacks project has expanded the traditional DOI focus to include delivery of physical resources (printed coursepacks) following management of digital content via DOI – this is a valuable contribution to the wider implementation of DOI. However, there remains the important problem of digital delivery of rights-enabled content in contexts that require a B2C (not B2B) model. It is clear that students desire both printed and digital versions of their learning content (printed so that they can avoid the problems of extended screen reading and make notes on the printed page; digital so that they can take advantage of digital functionality such as search, copy and paste) so there is a need to conduct further work to investigate options for providing both formats as part of a single, seamless process.

As noted in the earlier paper, digital delivery which incorporates DRM will involve addressing directly the difficult issues of access management (and the people/party identity information that is required for access management solutions) and, potentially, encryption/watermarking enforcement technologies. While CAL cannot be expected to solve these problems itself, it is important that they be identified and made part of a wider DOI agenda to ensure the potential for broad take-up of the DOI approach.

In this broader context, the role of Digital Rights Expression Languages (DREs) and, in particular, the systems that implement them will play an increasing role.

The MPEG Rights Expression Language (REL), the Open Mobile Alliance use of Open Digital Rights Language (ODRL), and the IEEE LTSC DREL working group for e-learning content are key forums that the DOI community should investigate. In particular, the current discussion of use cases within the IEEE group provides an opportunity for the DOI community to ensure its own community's needs are represented.

COLIS, MAMS and DOI

The earlier paper discussed the work of the COLIS project on DRM including delivery of rights-enabled resources. Since this paper, MELCOE has been successful in gaining significant funding for the 'MAMS' access management project (which grew out of the COLIS project) under ARIIC. This three-year project is to research and implement next generation access and identity management infrastructure for Australia's research community, and by extension, the e-learning community. It includes the areas of digital rights management, metadata management and integration with repositories. Given the scale and importance of this project, it presents a valuable opportunity for the DOI community to engage with other communities which have an interest in solving the problems of identity, access and delivery of rights-enabled content.

From my review of DOI literature, there is a fascinating opportunity for the combination of Handle and DOI with web-based authorisation using sophisticated access management (potentially also including a DREL) to provide a demonstration of digital delivery of DRM content using DOI. There are fleeting references to this kind of approach in early DOI literature, for example:

This ability of DOI to 'phone home' for access (e.g. to obtain the server half of a required key pair) is a very powerful tool for DRM. (Sidman, 2001a)

This does not appear to have been explored further in recent work. Given that the COLIS project provides a significant existing testbed to investigate this approach in conjunction with the opportunity represented by MAMS for next generation access and identity infrastructure, a combination of COLIS, MAMS and Handle/DOI would provide a fascinating and potentially crucial proof-of-concept for digital delivery. Further exploration of this possibility between IDF, CAL and MELCOE is recommended.

Implications for CAL EPICS project revisited

The concluding questions from the earlier paper are presented in italics, with comments below in normal type.

1. What are the actual requirements or use cases for DRM within the project? Have these been well documented? Are they comprehensive for the project's needs? Do they identify all relevant user and software requirements?

The recent work by Eric Moore on use cases for the project is an important step in this direction. Once the initial use cases are well established, additional work should be conducted to provide more detail for each case (particularly related to implied workflow) to assist with software development and/or business process re-engineering. Analysis of any 'missing' use cases or actors should also be conducted (e.g. the role of student information system vendors in managing university records of student access to coursepacks).

2. What components of DRM 'big picture' are relevant? Is the focus mainly on rights expression and rights enforcement? What are the interdependencies between systems that arise from the desired outcomes?

As noted, the concept of DRM that exists within the DOI community has a tendency to focus more on the creator/owner end, and less on the delivery end. It is recommended that clear expectations be set about what CAL will solve now in the Coursepacks project (especially in relation to the claims made in the grant proposal), what it will solve in the future, and what elements it will not attempt to solve – this relates particularly to digital delivery of coursepacks. More generally, the DOI community should also be explicit about the same issues in relation to digital delivery, i.e., what will it solve now, what will it solve in the future, and what elements will it not attempt to solve?

If DOI wishes to provide an end-to-end DRM solution, it will need to engage with additional standards communities and vendors who are working in this field, such as the IEEE DREL working group (for standards), and various vendors involved in DRM, such as Microsoft and ContentGuard. Further work will be required exploring the various interdependencies between object identifiers and people/party identifiers (beyond the existing work with InterParty, and towards the communities represented at the Digital Identity World conference), and on the relationship of Handle and DOI to technology for rights management at the 'delivery' end.

3. What system integration requirements will arise in the project? Will these be proprietary or open standards-based integration points? Can other vendors 'plug and play' if they conform to the relevant standards?

System integration requirements for digital delivery will become clearer as investigation of the points made above is conducted. IDF's general commitment to open standards is to be highly commended, and this will prove an important part of any strategy to solve end-to-end DRM issues.

It remains to be seen whether other components of the DOI approach are 'plug and play'. For example, the use of the Handle system seems an appropriate choice given the desired functionality of DOI, however it is unclear whether non-CNRI organisations are free to set up alternative Handle domains using Handle technology – the CNRI patent may be relevant to this issue. Similarly, it is unclear whether organisations are free to recreate 'Handle-like' systems which are not based on Handle, but achieve similar functionality. The TSO report (2003) notes in Footnote 1 to Table 2 that alternatives to Handle are theoretically possible, but have not been explored to date.

A similar question exists for DOI itself – are organisations free to set up alternative identifiers frameworks that implement DOI-like functionality (either with or without Handle)? Given that the syntax for DOI is an open standard, this should be possible theoretically. If the answer is yes, how would an alternative DOI-like identifier domain interoperate with DOI? Would this be encouraged or discouraged? For example, will DOI seek to interoperate with PURL or next generation URI (e.g. DDDS), or will DOI seek to remain a 'closed' network within itself? If interoperability is possible, what would the implications be for the DOI business model? If identifiers could be created outside of the DOI framework under a 'free' framework provided by, let's say, a Government Department, and these identifiers interoperate freely with DOI, what incentives would remain for DOI members to continue to fund the DOI community?

4. What standards are relevant to the project? Are there multiple relevant standards, and if so, what is the relationship between these standards? How open are the standards, and are they encumbered by patents?

The DOI community has been involved in significant standards work, both for DOI itself, and for connections of DOI to other frameworks (e.g. indecs). However, there are many standards and standards communities relevant to identifiers and DRM, so there are additional points of engagement required for DOI, such as IEEE LTSC, IMS, DCMI and CETIS. It is worth noting that the indecs framework appears to have had limited impact outside the communities already linked to DOI (e.g. W3C, OASIS, etc.), so it should not be assumed that the standards current within DOI-related communities will be adopted more broadly by significant vendors such as Microsoft, IBM, Sun and Oracle (who work mainly in W3C and OASIS).

5. What is the role of authentication, authorisation and people identifiers and object identifiers in the project? How will this impact on both demonstrator and enterprise examples of the proposed solution?

As noted in the discussion of the Digital Identity World conference (see Appendix C), there is significant convergence occurring between people and other identifier approaches in 2003. This will gain further momentum in 2004. Given the limited connection between the DOI community and the enterprise identity community, there is an urgent need to further investigate this area to determine appropriate points

of engagement, and areas which do not require close liaison. Key standards in enterprise identity include SAML, XACML, Liberty and WS-*

As noted above, there are some fascinating points of potential connection between Handle/DOI and COLIS/MAMS, so this may be worthy of further exploration in the future. This could help address not just object identifier issues, but also the issues of people identifiers (through the MAMS work on EDUCAUSE EduPerson and WALAP auEduPerson), and authentication and authorisation in MAMS.

6. Is there a need for desktop enforcement/encryption, or is web-based authorisation sufficient? If desktop enforcement/encryption is required, how will this be implemented?

It is interesting to note that premium library journals have relied primarily on web-based authorisation (initially based on machine identity such as IP address, but more recently on either machine or people identity such as university enterprise LDAP directories) for rights management without requiring desktop-level encryption – even despite the potential for copying of journal resources. Given this, there are grounds for exploring web-based authorisation over the top of Handle as an initial solution to ‘delivery-end’ rights management of digital resources.

However, it can be expected that in at least some cases desktop-level encryption will be required, in which case further investigation of incorporating DOI and rights management with technologies is recommended (e.g. secure Adobe eBooks and the new Windows Rights Management server).

7. What are the different DRM requirements of any searchable repository components of the DRM value chain as opposed to delivery/enforcement components of the value chain?

An important distinction in rights management is rights over metadata versus rights over resources. There is some recognition of this issue already with DOI via the option for ‘restricted’ metadata for DOIs. The issue of metadata rights is gaining momentum, and was a topic of discussion at the CETIS Identifiers and Digital Repositories meetings (see Appendix C). Within the UK, the JISC RoMEO project has begun to examine these issues, and they will be of relevance to MAMS and the other ARIIC projects. Addressing this issue will help to solve the questions of DRM for searchable repositories. For delivery-end DRM issues, see the discussion above.

8. How does the definition of a learning object or learning resource for this project compare with other definitions within the education sector? In particular, what are the implications (if any) for ‘run-time’ aspects of learning object delivery (e.g. tracking)?

The limited information about the SCORM use of DOIs relates only to the learning object metadata (descriptive metadata) component of SCORM, not the run-time API metadata (technical and structural). Nonetheless, these issues have a conceptual

similarity to the technical metadata information provided to typesetter/renderers in a print context, except that run-time metadata for learning objects is provided to assist learning management systems in delivery and tracking of learning objects (rather than in rendering of printed material). The issues associated with run-time metadata are even more explicit in my own work on learning design and LAMS (Dalziel, 2003b), in which most of the metadata relates to run-time issues, but content metadata is also required.

More generally, the nature of software development is currently in a period of disruptive change as systems move towards service-oriented architectures such as web services. Web services constitute the core of e-government interoperability approaches, so they can be expected to figure prominently in the future of education. This means that management of services (rather than simply content) will become a key feature of future software infrastructure. As a result, the content-centric DOI model may require extensions or revision to adjust to the evolving nature of software and its identity needs. It is recommended that IDF investigate the evolving web services framework and its implications for the future evolution of DOI and Handle.

9. To what extent (if any) is copyright management or intellectual property (IP) workflow management relevant to this project? If relevant, how will these elements be combined with other elements of the DRM value chain?

Based on my reading of DOI literature, it appears that the various facets of DRM related to managing copyright for creators/owners, and its related workflow, are the main focus of DOI to date. While this is appropriate given its origins and immediate areas of application, there will need to be further investigation of 'delivery-end' DRM issues for digital content for B2C-style business models. As noted in the first section, contrast the following quote from a recent article by Paskin (2003c) including a quote from Godfrey Rust:

Mapping complex concepts is possible, but concepts like 'digital rights management' are not currently consensually precisely defined; there is a majority view that it is digital management of rights, rather than management of digital rights, but beyond that 'DRM is something to do with managing, something to do with rights and something to do with the digital environment. But not necessarily'.

with the view of DRM articulated by Rightscom (2003) on behalf of the MPEG community:

These two strands of work, the Management of Digital Rights (identification and metadata, including the expression of rights) and the Digital Management of Rights (encryption, watermarking, authentication etc.), have been gradually pulled together in a single standards forum, the Moving Picture Experts Group, MPEG.

N.B. Rightscom recently acquired Godfrey Rust's company Ontologics, and Rust has taken up a position with Rightscom.

10. Does the project break new ground in DRM for education? If so, does it have valuable new use cases to contribute to the wider international work on DRM and DREL for education (e.g. to IEEE)?

If the CAL Coursepacks project is unable to use a B2B model, and hence needs to address B2C digital content delivery in addition to physical print delivery, then it may have contributions to make to DRM for education through groups such as the IEEE DREL working group. At the AICTEC Industry Forum, I presented a use case based on my experience of what students and teachers want in relation to DOI/CAL approach (see Dalziel, 2003a):

A student uses a secure e-book reader to view a large teacher-created 'coursepack' of readings sourced from many different publishers. The student can choose to print out individual sections themselves, or send off a request to their university printing service for overnight custom printing of large selections of the coursepack. Publishers receive an automated, fair return for their relevant contributions to the coursepack.

N.B. 'Automated' in the context of the Coursepacks business model could mean the figure supplied by the printer when the coursepack is downloaded from the CAL web site. This use case, along with some others, was recently raised by Dr Evan Arthur at the VET*Networking 2003 Conference. It was derived, in part, from the Coursepack business plan developed by CAL:

Both pilots are designed to facilitate DRM (Digital Rights Management) transactions across the value chain and include options for delivering content online or in print using print-on-demand (PoD) facilities through strategic alliances. (p.4)

If DOI and CAL are unable to use a B2B model (based on mandatory fees), then a B2C model will require a solution to DRM for learning content. The issues outlined in this use case need to be addressed either to make explicit those issues that will not be solved (and why), or how they will be solved and in what timeframe. This will help to set appropriate expectations among the relevant communities about the potential benefits of the CAL Coursepacks model.

References

ANSI/NISO (2000) *Syntax for the digital object identifier*
http://www.doi.org/handbook_2000/appendix_1.html

Brand, A; Daly, F; and Meyers, B (2003) *Metadata demystified: a guide for publishers*
The Sheridan Press
http://www.niso.org/standards/resources/Metadata_Demystified.pdf

CoverPages (2003) *Patents and open standards*
<http://xml.coverpages.org/patents.html>

Dalziel, JR (2003a) *Building common services infrastructure* (presentation to the AICTEC Industry Forum, Sydney, 3-4 September 2003)
http://www.melcoe.mq.edu.au/documents/DESTIndustryForum_CommonServices.pdf

Dalziel, JR (2003b) *Implementing learning design: the learning activity management System (LAMS)*
<http://www.melcoe.mq.edu.au/documents/ASCILITE2003%20Dalziel%20Final.pdf>

Dalziel, JR (2003c) *Discussion paper for learning activities and meta-data*
http://mdlet.jtc1sc36.org/doc/SC36_WG4_N0043.pdf

DOI (2003a) *DOI web site*
<http://www.doi.org/>

DOI (2003b) *DOI handbook* (Edition 3.3.0)
<http://www.doi.org/hb.html>

DOI (2003c) *Project announced to develop DOIs for scientific data: German National Library of Science and Technology joins IDF*
<http://www.doi.org/news/TIBNews.html>

Duncan, C (2003) *Identifiers for learning objects* (presentation for CETIS Identifiers Workshop, October 2003, London)

Dyson, E (2003) *Online registries: the DNS and beyond ...*
<http://www.edventure.com/release1/abstracts.cfm?Counter=6805767>

Friesen, N (2002) *Recommendations for globally unique, location-independent, persistent identifiers*
<http://www.cancore.ca/Resourceids.doc>

Guth, S; Neumann, G; and Strembeck, M (2003) *Experiences with the enforcement of access rights extracted from ODRL-based digital contracts* (DRM'03, 27 October 2003, Washington DC)
<http://nm.wu-wien.ac.at/research/publications/b328.pdf>

IFLA (1998) *Functional requirements for bibliographic records*
<http://www.ifla.org/VII/s13/frbr/frbr.htm>

IMS (2001) *IMS persistent, location-independent, resource identifier implementation handbook*
http://www.imsglobal.org/implementationhandbook/imsrid_handv1p0.html

Manifest Solutions (2003) *Unofficial DOI FAQ*
<http://www.manifestsolutions.com/doi/doifaq.html>

Masie, E (2003) *Making sense of learning specifications and standards: a decision maker's guide to their adoption* (2nd edition)
http://www.masie.com/standards/s3_2nd_edition.pdf

OASIS (2003) *Extensible resource identifier (XRI) syntax and resolution specification* (working draft 08, 30 September 2003)
<http://www.oasis-open.org/committees/xri>

Paskin, N (2003a) *Digital object identifier*
http://www.doi.org/doi_presentations/030203shortDOI.ppt

Paskin, N (2003b) *DOI: A 2003 progress report* (*D-Lib Magazine*, 9, 6)
<http://www.dlib.org/dlib/june03/paskin/06paskin.html>

Paskin, N (2003c) *Components for DRM systems: identification and metadata* (in *Digital rights management: technical, economical, juridical and political aspects in the European Union*)
http://www.doi.org/topics/drm_paskin_20030113_b1.pdf

Powell, A (2003) *Identifiers for learning objects - a discussion paper*
<http://www.ukoln.ac.uk/distributed-systems/lo-identifiers/>

Rosenblatt, B (2002) *Enterprise content integration with the digital object identifier* (DOI: 10.1220/whitepaper5)

Sidman, D (2000) *What is the digital object identifier (DOI)?*
<http://www.contentdirections.com/materials/WhatistheDOI.ppt>

Sidman, D (2001a) *The digital object identifier (DOI): the keystone for digital rights management (DRM)* (Content Directions Inc. White Paper)
<http://www.contentdirections.com/materials/SIIA-DOIandDRM-DavidSidman.doc>

Sidman, D (2001b) *The digital object identifier (DOI): The keystone for digital rights management (DRM)* (presentation for SIIA Annual Spring Conference, 12 March 2001, San Diego)

<http://www.contentdirections.com/TopLeftLogo.htm>

TSO (2003) *Digital objects identifiers for publishing and the e-learning community*. September 2003 (a report for the JISC from TSO)

<http://www.tso.co.uk/download/requestDocument.aspx?documentID=44>

Appendix A: Learning object identifier requirements

(Powell, 2003)

The following list forms a proposed statement of requirements for the identifiers of learning objects within the UK HE and FE community. Learning object identifiers should be:

1. Persistent: Learning object identifiers should be expected to work reliably for 10-15 years after they have been assigned.
2. Unique: Learning object identifiers should be unique, i.e. the same identifier should not be assigned to more than one learning object. (Note that a learning object may have more than one identifier assigned to it.)
3. Resolvable: All identifiers assigned to learning objects should be resolvable, i.e. there should be a 'resolution service' that accepts the identifier and returns a URL for a current location of the learning object.
4. Usable in web browsers: All identifiers assigned to learning objects should be actionable in current web browsers, i.e. they should take the form of 'clickable' links without the need for additional plug-ins. (Ideally, the 'resolution service' should not make use of HTTP redirects which leave the URL of the current location of the learning object in the browser location bar. The reason for this is that if the end-user then bookmarks the learning object, they bookmark the URL, not the identifier. However, this may not be possible to achieve technically.)
5. Transportable: Learning objects should be able to move between repositories (locations) without the identifier needing to be changed.
6. Simple to assign: The process of assigning identifiers to learning objects should be as simple as possible. Assignment should be independent of the workflow associated with creating and managing a learning object. In particular, assignment should be independent of the process of depositing the learning object in a repository. For example, if a person creates a learning object on their PC, and packages it using a desktop tool, creating metadata about it in the process, then they should be able to assign an identifier to both the learning object and the metadata record and have those two identifiers honoured (i.e. used) by a repository (or several repositories) when the learning object is deposited. The creator doesn't want the repository to assign a different identifier to the object and metadata when they deposit them, nor do they want to have to wait until the point of deposit before they can assign an identifier.
7. Assignable in devolved environments: Assigners of identifiers should be able to work independently of each other, without reference to a central service, in such a way that guarantees uniqueness of each identifier.
8. Usable in non-digital environments: Learning object identifiers should be usable in non-web contexts, i.e. it should be possible to print identifiers in paper articles or dictate them over the phone. For this reason, learning object identifiers should be reasonably short and reasonably intelligible to people.

9. URI compliant: Learning object identifiers must conform to the URI specification. (Ideally, identifiers should be based on existing standards and technologies, however, they should also be independent of any particular protocol.)
10. Free at the point of use: There should be no cost (at least to the end-user) associated with the assignment or use of learning object identifiers.

Metadata record identifier requirements

The requirements for identifiers of metadata records about learning objects are the same as those listed above. However, it could be argued that requirements 4 and 8 are of much lower priority in the case of identifiers for metadata records.

Appendix B: Kinko's Coursepacks press release

Kinko's goes back to school with Coursepacks

Tuesday 16 September 2003, 5:02 pm ET

Kinko's and USC University Custom Publishing Team to provide Coursepack solutions

DALLAS and LOS ANGELES, 16 Sept /PRNewswire/ - This fall, Kinko's® returns to its campus roots with two new and improved versions of Coursepacks, a product the company hasn't focused on in more than a decade.

Coursepacks are custom-made compilations used by educators to supplement traditional textbooks in the classroom. They typically contain up-to-date information from newspapers, magazines or book excerpts that require permission from the copyright owners in order to be legally produced. Kinko's re-enters the market with Kinko's Coursepack Solutions, including two services designed to meet the needs of educators and students. For Kinko's Coursepacks Complete, Kinko's has teamed with University of Southern California's University Custom Publishing (UCP) to provide a full-service option that includes copyright clearance and formatting by UCP, as well as custom covers, printing and delivery from Kinko's. The company also introduced Kinko's Coursepacks Direct, an option for schools that obtain their own copyright permissions.

'Kinko's Coursepack Solutions represent a new, customized outsourcing option for Coursepacks that speeds production, is cost-effective for educators and students, and protects the rights of content creators,' said Chris Gibson, Kinko's vice president of marketing and a former teacher. 'The introduction of Kinko's Coursepack Solutions also reflects our commitment to Kinko's Commercial Solutions, which delivers variable-cost and print-on-demand outsourcing for educational institutions and other businesses.'

Kinko's Re-enters the Coursepack market

Kinko's was founded near the University of California Santa Barbara in 1970 to meet the needs of faculty and students, and Coursepacks quickly became a key driver in Kinko's early success. However, Kinko's exited the Coursepacks business in the early '90s after its involvement in a landmark suit regarding the scope of 'fair use' for copyrighted material as defined in the Copyright Act of 1976. At that time, Kinko's opted to discontinue Coursepacks until it could create a better process for obtaining copyright clearances.

A better solution

Kinko's Coursepack Solutions capitalize on some of the company's core strengths service, speed and convenience while creating an enhanced process that strictly complies with copyright laws. To ensure proper copyright permissions and royalty payments, Kinko's has formed a strategic alliance with UCP, which is a division of the University of Southern California. Kinko's Coursepack Solutions launches with pilot programs at seven universities this fall.

'University Custom Publishing has more than a decade of experience in managing centralized copyright clearance for clients throughout California and the Northwest,' said Daniel Archer, Director of UCP. 'We are especially pleased that Kinko's has placed such a high value on campus bookstores as a retailing partner. Our expertise, combined with Kinko's digitally connected network of locations conveniently located near campuses everywhere, creates an unparalleled nationwide Coursepacks capability.'

Flexible options for educators, students

To meet the varying needs of Coursepacks customers, Kinko's offers two flexible options Kinko's Coursepacks Complete and Kinko's Coursepacks Direct.

Kinko's Coursepacks Complete is a full-service option including copyright clearance and formatting by UCP, custom covers, printing and delivery.

Benefits of Kinko's Coursepacks Complete include:

- Initial Coursepacks copyright clearance and production in five weeks or less (with an average of 12 working days), compared to six to eight weeks among some competitors.
- Re-orders are ready in 24 hours in most cases.
- Coursepacks are produced locally with free pick-up and delivery available. No shipping charges or delays.
- Production is managed entirely by Kinko's with copyright clearance and formatting by UCP. Mastering process that includes 'clean-up' of low- quality originals.
- Competitive pricing along with the ability to return any unsold Coursepacks for full credit in most cases.

For colleges and universities that obtain copyright permissions themselves, Kinko's Coursepacks Direct offers Coursepack production, but only after copyright clearance certification.

Campus bookstore executives, educators and administrators responsible for Coursepacks are invited to call 1-888-KINKOS-1 for referral to a local Kinko's Account Manager.

About Kinko's

Kinko's is the world's leading provider of document solutions and business services. Its global network of more than 20,000 team members and 1200 digitally connected locations offers access to technology for color printing, finishing and presentation services, Internet access, videoconferencing, outsourcing, facilities management, web-based printing, and document management solutions. Dallas-based Kinko's is a privately held corporation with locations in nine countries. For more information please visit www.kinkos.com. Products, services and hours vary by location.

About USC University Custom Publishing

Located in Los Angeles, Calif., University Custom Publishing is a division of the University of Southern California specializing in reprinting lecture notes, journal reprints, newspaper articles, case studies, lab manuals, syllabi, workbooks and out-of-print books. UCP is one of the largest campus-based custom publishing operations in the country with more than 300,000 cleared permissions and a wide range of site licenses.

Appendix C: Recent international conferences/workshops of relevance to DOI

This section describes recent conferences I have attended which are relevant to the Coursepack project and the general field of learning content and identifiers.

Digital Identity World (14-17 October 2003, Denver)

Digital Identity World (www.digitalidworld.com) is the world's leading conference for enterprise identity management and related issues. Approximately 1000 attendees and over 20 vendors (including Microsoft, Sun and Novell) participated in this three-day event. The conference is the yearly culmination of the 'Digital Identity World Newsletter', a weekly newsletter dedicated to leading-edge implementations and news about enterprise digital identity. The first conference was held in 2002, and the 2003 event was upbeat and optimistic, with a shift from theory to practical results in the presentations (when compared to 2002).

Digital Identity World emphasises that all forms of identity need to be managed as part of a coherent enterprise approach. Types of identity discussed included people/parties, digital/physical objects, processes/services and applications – and while the main focus of the conference is on people/party identifiers, there is a movement towards managing all types of identity under a unified framework, such as an enterprise directory. Identity management topics included use of directories and meta-directories, implementation of single sign-on, federation of identity and trust, password management, provisioning and de-provisioning, and security.

Digital Identity World included discussion of enterprise identity management standards such as SAML, XACML, EPC-RFID, the Liberty Alliance, and the WS-* family of standards. Despite differences between Liberty and WS-*, there was a sense that convergence of these standards was more likely than divergence due to client needs for interoperability. Hence the current split between Liberty/Sun/Novell and Microsoft/IBM/BEA may not lead to market duality for identity over the medium term.

There did not seem to be any sessions about DOI (or indecs, RDD, etc.), and there was no discussion of DOI in any of the sessions I attended. The closest topic to DOI was EPC-RFID, although there remain many differences between these two areas. Digital rights management was discussed in some sessions, and was highly contentious when compared to all other identity topics. In terms of the education sector, the presentation by Scott Cantor about the Internet2 Shibboleth project indicated that higher education is one of the leading sectors in federated identity.

Given the importance of this conference and newsletter, the breadth of attendees and vendors and the standards under discussion, IDF should review this area to determine appropriate points of engagement with this community concerning DOI. The lack of connection between Digital Identity World and DOI is a point of concern.

CETIS workshops related to learning design (20, 22, 28-29 October 2003, UK)

There were various workshops held by CETIS in October which addressed the new IMS learning design standard – there is a strong interest in the UK (and Australia) in e-learning systems which move beyond a simple content model to a model that combines both content and activities (such as discussion groups, chat and assessment).

I have been heavily involved in this area, and have led the development of one of the first systems to implement this approach (LAMS). Sequences of activities created in LAMS can be captured as a single file which contains all the properties required to run the sequence, so this file can be stored in a repository, and hence could be assigned an identifier (this was demonstrated in practice using the intrallect learning object repository at the 28-29 October meeting). It would be interesting to explore the potential relevance of DOI for learning design sequences, and this would have some conceptual similarity to the challenges of the run-time (not learning object metadata) elements of SCORM. For further information about LAMS, see Dalziel (2003b). For a discussion of new types of descriptive metadata that may be required for learning designs, see a recent submission to ISO/IEC JTC1 SC36 WG4 (Dalziel, 2003c).

CETIS identifiers meeting (21 October 2003, London)

This meeting of approximately 30 international experts was focused on the requirements for identifiers for learning objects. It was facilitated by Lorna Campbell (Assistant Director, CETIS), with assistance from Ed Walker (CEO, IMS Global Learning Consortium) during the afternoon. As part of the meeting, many participants contributed use cases, issues documents and/or papers about identifier issues. IDF (Norman Paskin) and TSO were represented at the meeting.

Initial discussions were based on Charles Duncan's paper (2003) about needs for identifiers – of particular note is Duncan's exploration of the needs for one identifier for the 'item' (e.g. learning object), and a separate identifier for its metadata. This approach allows for changes of one component but not the other to be tracked via identifier changes. The need for separate identifiers of objects and metadata is also acknowledged in Powell (2003), and the TSO report on DOI:

Identifiers can be associated with almost anything, but the two main types of items that are considered are:

- *digital content objects for e-learning communities*
- *the metadata records associated with learning objects (p.4, TSO, 2003)*

Following Duncan, a discussion of fundamental requirements for identifiers was led by Wayne Hodgkins, who promoted a minimal set of requirements for an identifier.

Following extensive difficult discussion, the meeting adopted the URI specification for identifiers as a foundation.

Later discussion covered topics such as additional services over URI that may be desirable. Dan Rehak noted that given the need for identification of contributions to a student ePortfolio, the number of identifiers required would exceed one million per student within a few years. This point is relevant to a comment from the TSO report (p.3):

The informal sharing of information resources may have different digital identifier requirements to that of the more formal traditional publishing and dissemination process.

DOI advocates discussed the benefits of DRM, resolution, etc. but there was uncertainty within the rest of the group as to whether these should be closely associated with identifiers, or should be dealt with as other parts of the overall software infrastructure for learning. Business and cost models for management of identifiers also proved contentious. Powell (2003) outlined a set of requirements for identifiers which was proposed as a starting point for further development (see Appendix A).

It was clear from this discussion that gaining agreement on identifiers is a difficult issue, with considerable differences of opinion among experts. There are many complex yet subtle problems in defining an identifier specification, and addressing the possible implementations that may arise from the specification. Apart from the adoption of URI, and the proposal to continue work based on Powell's list of requirements (reproduced in Appendix A), there was little further agreement within the group. Notes on the meeting are expected to be made available by Lorna Campbell of CETIS from the CETIS website (www.cetis.ac.uk).

Learning object economy conference – sponsored by Intrallect (23 October 2003, Edinburgh)

This conference, hosted by Charles Duncan and colleagues from Intrallect, explored the challenges of creating a viable learning object economy. It included presentations from higher education, K-12 schools and corporate training contexts. Despite the fact that consistent identification of learning objects is a requirement for a successful learning object economy, there was little discussion of identifiers during the conference, and no discussion of DOI. This is typical of similar learning object discussions in the UK and elsewhere.

CETIS digital repositories workshop (27 October 2003, Glasgow)

Workshop devoted to discussion of digital repositories and digital libraries in UK e-learning. There was little discussion of identifiers, and no discussion of DOI.

Joint JISC/OKI/CETIS/IMS/IMS Australia coordination meeting (31 October 2003, London)

This meeting explored collaboration and co-ordinated development across US, UK and Australian projects in the e-learning standards arena. Despite comprehensive frameworks for service-oriented architectures, there was little discussion of identifiers, and no discussion of DOI.

EDUCAUSE (3-7 November 2003, Anaheim, California)

EDUCAUSE is the world's leading higher education IT, library and e-learning conference, attracting over 6000 attendees and over 100 vendors. While many topics were canvassed across the various sessions, there was little if any discussion of identifiers, and none of the paper titles or abstracts mentioned DOI.

While at EDUCAUSE, I had the opportunity for informal discussions with two key players in the DRM space – Harry Picarello from ContentGuard, and Robby Robson, Chair of the IEEE LTSC and co-lead of the DREL working group.

Discussion with ContentGuard indicated that they view their patent portfolio as applying to any rights expression language, and that they would potentially take legal action against entities which breach their patents. It is possible that ContentGuard would regard the application of indecs to rights management as being a potential infringement of their patents.

Discussion with Robby Robson indicated that the IEEE DREL working group was making good progress, and had received many responses to its use cases. There did not appear to have been any responses related to DOI or persistent identifier and resolution technologies as part of the working group. Given the likely impact of the IEEE DREL outcomes on e-learning, there is an urgent need for IDF to investigate and if appropriate contribute use cases to the IEEE DREL process.

IMS Global Learning Consortium (10-14 November 2003, San Jose, California)

This was a regular quarterly meeting of the IMS e-learning specification development organisation. Apart from some work on identifiers in 2001, the organisation has formally addressed identifiers or DRM in the past few years. There was no special discussion of identifiers at this meetings (apart from some informal discussion about the outcomes of the CETIS Identifiers meeting), and no discussion of the use of DOIs.