RoMEO Studies 6: Rights metadata for open archiving

Elizabeth Gadd, Charles Oppenheim, and Steve Probets

Biographical note

Charles Oppenheim is a Professor of Information Science, Dr Steve Probets a Lecturer, and Elizabeth Gadd a Research Associate, in the Department of Information Science at Loughborough University, Loughborough, LE11 3TU, UK. Prof. Oppenheim is the Director of the RoMEO project, and Dr Probets the Technical Director.

Email: C.Oppenheim@lboro.ac.uk

This article has been accepted for publication in Program 38(1) 2004.

Abstract

This is the final study in a series of six emanating from the UK JISC-funded RoMEO Project (Rights Metadata for Open-archiving) which investigated the Intellectual Property Rights (IPR) issues relating to academic author self-archiving of research papers. It reports the results of a survey of 542 academic authors showing the level of protection required for their open-access research papers. It then describes the selection of an appropriate means of expressing those rights through metadata and the resulting choice of Creative Commons licences. Finally it outlines proposals for communicating rights metadata via the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH).

Why rights metadata?

The open access movement promotes the free availability of research papers via one of two models: 1) freely available electronic journals, and 2) author self-archiving of research papers on institutional or subject-based repositories (such as ArXiv (2003)). Although they are seen as complementary, the institutional repository model has been promoted as the fastest way to make open access a reality (Crow, 2002). If the current scholarly journal literature became open-access overnight (and there are estimates that this consists of 2,000,000 articles per annum (Harnad, 2001)) there would undoubtedly be enormous benefits. However, it would leave the scholarly journal literature facing the same problems that other open-access web pages have faced since the birth of the Internet, namely, the misunderstanding that 'publicly available' means 'in the public domain'. Of course, while the protection of copyright law applies to web pages, there is a common misconception that posting to the internet gives end-users an 'implied licence' to use - or abuse. Concerns about abuse have led many web-based information providers to invest in digital rights management systems to provide technical protection for their copyright material (Cope and Freeman, 2000). There has also been a proliferation of 'click-through' licences which force end-users to agree to conditions of use (Johnson, 2003). However, it is questionable whether such measures are appropriate for "give-away" research literature.

The development of free-ware and share-ware as software distribution mechanisms brought with them a range of 'alternative' copyright licensing regimes by which such works could be protected. The most prominent is the GNU Public Licence (Anon., 2002) (GPL) which states that whilst software may be copied and re-used, any resulting derivatives must be made available under the same generous terms as the source software. A number of other alternative regimes are now being developed for other types of work such as research papers.

The OpenContent (2002) movement was inspired by the GNU Public Licence and has developed licences along similar principles to protect any open-access content. A "Counter Copyright" campaign was launched by the Berkman Centre for Internet & Society in 2002. They suggested that "If you place the [cc] icon at the end of your work, you signal to others that you are allowing them to use, modify, edit, adapt and redistribute the work that you created." This campaign has now been superceded by the Creative Commons (2002) Initiative, which provides creators with a series of 11 licences under which they may make their open-access work available. The licences have three incarnations: a simple "human-readable" version, a "lawyer-readable" licence document, and machine-readable rights metadata. Such

initiatives suggest that whilst restrictive copyright protection may not be appropriate to open access works, a certain level of protection is still required: creators are not prepared simply to bequeath their works to the public domain.

The major difference between commercial digital rights management systems, and open-access alternatives is that the former are focussed on preventing unauthorized use whilst the latter are focussed on communicating authorized use. The one element common to both approaches is an understanding of the importance of making such rights information machine-readable. The commercial content industries have developed a number of standards for this purpose, including the comprehensive <indecs> metadata framework (2003), which specifies the metadata necessary for e-commerce in Intellectual Property (IP) in the network environment. <indecs> itself has been utilised by ONIX (Online Information Exchange) (2003) standard and the DOI (Digital Object Identifier) Foundation (2003). The open access communities have also made steps in this direction with the development of the Open Digital Rights Language (ODRL, 2002) – which is a Digital Rights Expression Language (see below) and by Creative Commons' rights metadata approach.

With such initiatives already underway, the UK JISC (Joint Information Systems Committee) funded the RoMEO (2003) (Rights Metadata for Open-archiving) Project to consider how the rights status of one particular niche of giveaway content, namely, open-access academic research papers, might be communicated digitally through rights metadata (Gadd, Oppenheim, and Probets, 2003a, 2003b, 2003c, 2003d, 2003e). We were concerned with how such rights metadata might be disclosed and harvested under the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH, 2002). This open-source protocol has been one of the principal enablers of the open-access movement. It allows metadata about freely-available resources (and closed-access resources) to be shared simply and easily by so-called Data Providers, and then harvested by Service Providers, who may then build services upon that metadata by which end-users can access it.

Methods adopted

Before considering how best to communicate the rights and permissions status of research papers, it was first necessary to gather information on exactly how academics wished to protect those papers. Thus, an online questionnaire was designed and divided into three sections: A – About You, B – Your research papers, and C – How you use other peoples' research papers. Section A collected demographic information. Section B collected information on how academics wished to protect their own freely available research papers and Section C collected information on how academics usually used other peoples' research papers.

To provide a framework for collecting information for sections B and C, we looked to an existing Digital Rights Expression Language (DREL). DRELs provide models for expressing rights and permissions information over content. The chosen DREL was the Open Digital Rights Language (ODRL, 2002). (More information is provided on DRELs below). This provides a model of permissions, restrictions and conditions over works. A restriction is a constraint or limit on the extent of the permission being offered (e.g., you may print, but only four times), whereas a condition is a prerequisite requirement that must be met before the permission may be performed (e.g., you may print four times, if you pay a fee).

Academic authors were asked to select from a list of permissions, restrictions and conditions (P, R and C's) those they would like to apply to the use of their own freely available works. For the list of permissions academics could specify whether they would allow that activity to be performed freely, with limits or conditions, or not at all. Where over 60% of academics agreed on a P, R or C it was considered important enough to be made a mandatory element in the RoMEO rights metadata solution. Those that received between 50-59% agreement would be considered optional elements.

Survey results

The full methodology and results of the survey are reported in two earlier studies (Gadd et al, 2003b, 2003c). However, a summary of the P, R and C's selected by over 50% of academics is given in Table 1 below. Respondents were given examples of the meanings of the terms and these are supplied in brackets below.

Permissions	Restrictions	Conditions
Display (e.g. may be viewed on screen)	**Exact replicas** (e.g. the text must not be altered in any way)	**Attribution** (e.g. your name should always be clearly displayed on the article)
Give (e.g. copies may be forwarded to colleagues)	**For non-commercial purposes (optional)**	
Print (e.g. copies may be printed out)		
Excerpt (e.g. a short passage may be quoted)		
Save (e.g. may be saved to disk)		
Aggregate (optional) (e.g. may be compiled into an anthology)		
Sell (prohibit) (e.g. either on a cost-recovery basis or as a commercial enterprise)		

Table 1 The permissions, restrictions and conditions required over open-access works

Generating rights metadata

Having discovered what permissions, restrictions and conditions were required over open-access research papers, the next step was to develop a means of expressing those rights in metadata. Expressing rights is a complicated activity. Rights are bestowed on the rightsholder by law, and national copyright laws grants rights holders different rights depending on the type of work they produce. In the UK for example, the rights conferred on an author of a literary work are to authorise or prevent copying the work; issuing copies of the work to the public; performing the work in public; broadcasting the work or including it in a cable programme service; and to make an adaptation of the work or do any of the above in relation to an adaptation, in most cases for the duration of their lifetime plus seventy years (Great Britain, 1988). The duration of copyright and interpretations of the term 'copy' (for example) vary for different types of work.

However, the rightsholder is also able to waive her rights, to give them away, or to license them to third parties. Thus, the rights conferred by law can be 'sliced' up and licensed in a number of ways: geographically (e.g., the right to publish in the UK only), by media (broadcast rights, print rights, etc.), by time period (number of days, years, etc.), and so on (McCracken, 1995). Expressing such rights even in a written contract for a human being to understand can be a difficult enough task. See, for example, Owen (2002) for examples of the many types of licence agreements that are possible. Expressing such rights in a way that a machine can understand is even more complicated.

Digital Rights Expression Languages

For rights specifications to be machine readable they must be formally expressed so that permissions, restrictons and conditions can be unambiguously determined. To this end a number of attempts have been made to model the seemingly endless number of permutations of rights expressions. These models manifest themselves in Digital Rights Expression Languages (DRELs). (As Iannella (2001) has pointed out, such languages are concerned with the "'digital management of rights' and not the 'management of digital rights'". There are currently two main DREL players: XrML (2002) (eXtensible Rights Mark-up Language), and ODRL (2002) (Open Digital Rights Language).

The first ever DREL was developed by Mark Stefik of Xerox. It was called the Digital Rights Property Language (DRPL) and it formed a foundation for both XrML and ODRL. Xerox subsequently spun out

their Digital Rights Management (DRM) work to a separate company called ContentGuard and developed DRPL into XrML (Extensible Rights Markup Language). XrML has been cited as the most promising of all DRELs and it is the only DREL to have been used in a shipped product to date (McAllister, 2002). However, people have been cautious about taking it up, as ContentGuard hold a number of patents over XrML and companies are concerned about building systems around a competitor's product (Rosenblatt et al, 2002, Ch. 6).

The RoMEO Project decided against XrML on three main grounds. Firstly, XrML was developed purely as a component of DRM systems (i.e., technical protection measures). Such measures are concerned with controlling and limiting access, as opposed to enabling open-access. As such, XrML does not attempt to specify usage restrictions that cannot be implemented by software. For example, it does not allow for "non-commercial use" or "educational use", because the contextual information necessary to enforce such a restriction is unlikely to be available. Secondly, at the time of project development, XrML did not have a Data Dictionary component. Thus, although the grammar of the language was available (how rights expressions would fit together) it had no generally agreed upon words or terms to give those expressions meaning. A standard Rights Data Dictionary (RDD) is currently being developed by the <indecs> consortium for this purpose, but it is not available yet (Indecs Consortium, 2002). The third concern of the project team was that the licensing conditions of XrML were unclear. As we were developing a solution for the open access community, we did not want to be locked up, either now or in the future, by an unknown licensing agreement.

ODRL is favoured by open source and educational communities because it "is freely available and has *no licensing requirements*" (ODRL, 2002). It incorporates what it describes as a Data Dictionary, but this is currently just a listing of terms; it does not provide standardised meanings for those terms such as are promised by the <indecs> RDD work. The ODRL Data Dictionary did however, provide the Project with something to work with. Indeed, ODRL terms were used in the design of RoMEO's author survey by which academics could choose what permissions restrictions and conditions they would like to assert over their 'giveaway' literature. The main complaint about ODRL is that it is not as well developed as XrML, and some predict that it does not have much of a future (Rosenblatt et al, 2002, Ch.6).

Other options

Instead of choosing an existing DREL there were other options open to us. Firstly, we could have developed our own method of expressing rights. However, this would be time-consuming and possibly unnecessary considering the number of existing DRELs available. It would also be another bespoke solution, not in-keeping with the JISC's desire to work with existing standards. Secondly we could work with the Creative Commons Initiative's means of expressing rights.

Creative Commons

Whilst DRELs are designed to express any rights agreement, independent of the type of intellectual property, rights holder, business model, etc., the Creative Commons (2002) have developed a model which deals purely with 'giveaway' or open content. They subscribe to the 'Public Domain Plus' school of rights protection which aims essentially to protect authors' moral rights but not the economic rights that copyright law gives them. Creative Commons (CC) have developed a series of 'licences' that can be selected by creators to describe the terms under which their work can be used. The licences allow display, public performance, reproduction, and distribution of a work whilst providing creators with four optional restrictions (see Figure 1): attribution, non-commercial use, no derivative works, or permitting derivative works under a "sharealike" condition (meaning that subsequent works have to be made available under the same terms as the original). Creators select the restrictions they wish to apply. In total there are a possible eleven alternative licences.

Attribution. You let others copy, distribute, display, and perform your copyrighted work — and derivative works based upon it — but only if they give you credit.


Noncommercial. You let others copy, distribute, display, and perform your work — and derivative works based upon it — but for noncommercial purposes only.

 **No Derivative Works**. You let others copy, distribute, display, and perform only verbatim copies of your work, not derivative works based upon it.

 **Sharealike**. You allow others to distribute derivative works only under a license identical to the license that governs your work.

Figure 1. Creative Commons licence restriction options
(Developed by Creative Commons and reproduced under licence)

Each licence consists of a brief "human-readable" statement called the Commons Deed to communicate the terms quickly to end-users; a full licence document describing the conditions in legal code; and some machine-readable rights metadata specified in RDF/XML. The Resource Description Framework (RDF) provides a means of structuring metadata to provide unambiguous, meaningful and reusable expressions. It does not specify the metadata that is to be maintained, rather it ensures that the metadata descriptions are both interchangeable and scalable (Miller, 1998).

ODRL or Creative Commons?
Having discounted XrML, we were left with two main options for a rights metadata solution. We could either develop a 'perfect fit' solution by creating an application profile of the flexible ODRL for the academic research community, or we could adopt the Creative Commons solution which was rapidly gaining momentum. The advantages and disadvantages of each option were carefully weighed.

Benefits of ODRL
As mentioned above, ODRL terms were used in the academic author survey, so a rights metadata solution based on ODRL could be sure to match author's exact requirements. ODRL has also been used in a small number of other educational projects, proving it's relevance to the sector. A solution based on ODRL would be flexible and extensible due to the comprehensive nature of the language. ODRL also has it's own XML schema, something essential to all metadata exposed under the OAI-PMH.

Benefits and disadvantages of Creative Commons
Creative Commons on the other hand, whilst designed with all open access creators in mind (visual artists, film makers, musicians, etc), did prove to be quite a good match with the RoMEO author requirements (see Figure 2). There were only two main concerns for academic authors. Firstly, all CC licences allow works to be incorporated into one or more collective works and for that collective work to be reproduced. Sixty-seven per cent of RoMEO author questionnaire respondents wanted to limit or prohibit this activity (what ODRL calls "aggregation"). Secondly, CC licences do not explicitly forbid third parties from selling the work. Although one of their optional restrictions is "for non-commercial purposes", this would still not forbid the non-commercial sale of a work.

Permissions		Restrictions/conditions	
RoMEO	**Creative Commons** (all mandatory)	**RoMEO**	**Creative Commons** (all optional)
Display	Publicly display	Attribution	Attribution
Give	Distribute	Exact replicas	No derivative works
Print	Reproduce	For non-commercial purposes (optional)	Non-commercial purposes
Excerpt	Reproduce		
Save	Reproduce		
Aggregate (optional)	Incorporate the work into one or more collective works		
Sell (prohibit)	-		

However, as the course of the RoMEO Project went on, the Creative Commons initiative increased in momentum, as did the level of support from open access proponents. The Open Archives Initiative developed a keen interest in adopting the CC solution, as did the Dublin Core Metadata Initiative (Powell et al, 2003). DSpace (Bass, 2002), the open-source institutional repository software developed at MIT also expressed its intention to adopt the CC licences (Smith, 2003). One of the key benefits of CC is that it provides not just rights metadata, but a whole rights system with human readable statements for end users, legal licences for lawyers, and rights metadata for machines! If Project RoMEO simply developed some rights metadata based on ODRL, somewhere along the line these other elements (i.e. how to communicate the terms simply to end users, and the legal interpretation and validity of the metadata) would still need to be addressed. Perhaps the only drawback to the Creative Commons solution was that their metadata is expressed in RDF/XML which does not currently have either an official generic XML schema or a specific CC one. All metadata disclosed under the OAI-PMH has to conform to an XML schema **.**

Solution

As the benefits of working with what looked set to become an emerging standard outweighed the minor concerns it raised, it was decided to align RoMEO's work with the Creative Commons Initiative. However, as the issue of the XML schema for the CC RDF/XML still had to be resolved, the project took two approaches. Firstly, the CC were asked whether they would consider writing such a schema. Secondly, the project decided to write ODRL versions of the Creative Commons licences which would conform to the ODRL XML schema. Interestingly, the ODRL versions provide a slightly better 'fit' with the CC licences than the CC's own RDF/XML. This is because the CC RDF is intended to provide a basic approximation of the main permissions and restrictions of each licence, which is designed to work in conjunction with the Commons Deeds and licence documents . However, the ODRL versions are designed to stand alone and thus are more descriptive of the full licence content. Examples of the CC RDF version and the RoMEO ODRL version of the Attribution-NoDerivs-NonCommercial licence are given below. You can see that the CC RDF does not state that all CC licences allow aggregation, whereas the RoMEO ODRL version does.

```
<License rdf:about="http://creativecommons.org/licenses/by-nd-nc/1.0/">
  <permits rdf:resource="http://web.resource.org/cc/Reproduction" />
  <permits rdf:resource="http://web.resource.org/cc/Distribution" />
  <requires rdf:resource="http://web.resource.org/cc/Notice" />
  <requires rdf:resource="http://web.resource.org/cc/Attribution" />
  <prohibits rdf:resource="http://web.resource.org/cc/CommercialUse" />
</License>
```

Figure 3 RDF XML instance for CC Attribution-NoDerivs-NonCommercial licence

```
<offer>
        <context>
                <uid>http://creativecommons.org/licenses/by-nd-nc/1.0/</uid>
                <uid>http://www.romeo.ac.uk/odrl-cc-licenses/by-nd-nc/1.0</uid>
                <name>Creative Commons Attribution-NoDerivs-NonCommercial
                  Licence</name>
                <date><fixed>2002-10-10</fixed></date>
        </context>
        <permission id="CCCore">
                <display/>
                <print/>
                <play/>
                <excerpt/>
                <aggregate/>
                <give/>
                <duplicate/>
                <save/>
        <constraint>
                <quality>
                        <context>
                                <uid>urn:romeo.ac.uk:vocab:quality:exactreplicas</uid>
                        </context>
                </quality>
                <purpose>
                        <context>
                                <uid>urn:romeo.ac.uk:vocab:quality:noncommercial</uid>
                        </context>
                </purpose>
        </constraint>
        <requirement>
                <attribution/>
                <accept>
                        <context>
                                <remark> I agree to use this eprint under the terms and
                                conditions stipulated in the Creative Commons licence found
                at http://creativecommons.org/licenses/by-nd-nc/1.0
                                </remark>
                        </context>
                </accept>
        </requirement>
        </permission>
        <constraint>
                <transferPerm downstream="equal" idref="CCCore"/>
        </constraint>
</offer>
```

Figure 4 RoMEO ODRL XML instance for CC Attribution-NoDerivs-NonCommercial licence

Expressing rights via the OAI-PMH

Having decided on how to express the rights and permissions relating to open-access research papers through metadata, the final step was to consider how such metadata might be disclosed and harvested under the OAI-PMH (2002).

The OAI-PMH specifies a set of requests which are sent to data providers and which contain a list of arguments in the form of key-value pairs. Upon receipt of the http-based requests metadata records are returned by the data providers. This is best explained using one of the possible requests (GetRecord) as an example: A request of:

`http://oai-repository.com/OAI-script?verb=GetRecord&identifier=eprint-1&metadataPrefix=oai_dc`

will return a metadata record containing oai-dc (Dublin-Core) metadata describing the resource eprint-1. Metadata records describing different resources could be returned passing different resource identifiers in the '`identifier`' argument or different metadata formats could be returned by specifying a different `metadataPrefix`.

Service Providers might wish to harvest rights metadata in two ways. They would certainly expect each resource to provide rights and permissions information either embedded within its Dublin Core metadata or as a separate record available in its own right. However, they may also want to check if a minimum set of terms and conditions apply to all the resources in a repository (say if a Data Provider only allows authors to deposit papers if the author is prepared to allow them to be used by third-parties in a particular way).

Expressing rights over an individual resource via it's metadata may be done in two ways. All items in an OAI-compliant repository must, as a minimum, provide an unqualified Dublin Core metadata record for harvesting. One of the 15 Dublin Core elements is <dc:rights>. In a document entitled *Using simple Dublin Core to describe eprints* by Powell, Day and Cliff (2003), they recommend that <dc:rights> contains, "A human-readable statement about the rights held in and over the eprint, the URI of a Creative Commons licence or the URI of a machine-readable statement." However, as well as using the URI of the Creative Commons licence (which actually links to the Commons Deed which in turn links to the licence), Project RoMEO proposes that the rights metadata made available under the OAI-PMH is *machine readable*. There are various ways this can be achieved. One way is to embed the rights metadata within the dc:rights element, another is to link to the rights metadata from within the <dc:rights> element. One way to implement the latter approach is to utilise a OAI GetRecord request within the dc:rights element. By passing a parameter such as 'oai-cc' in the metadataPrefix argument of the GetRecord request, actioning the request would result in a rights record for the resource being returned. The returned rights record should describe one of the 11 CC licence either in RDF/XML form or the Romeo/ODRL form outlined earlier. This approach means that in addition to a DC record being available for each resource (eprint) an additional rights metadata record would also be available for each resource and this fact should be communicated in the data provider's response to the OAI-PMH ListMetadataFormats request.

A default rights expression covering an entire repository could work in one of three ways. Firstly, if a repository only accepts items conforming to a particular rights expression - no more, no less - then the default expression would actually describe the rights status of all works in the repository. This would be very useful for Service Providers, as they could harvest entire collections of records knowing exactly how end-users should be allowed to use them. However, this scenario may deter two types of depositor: those requiring more protection than the default rights expression, and those requiring less. To cater for those requiring more, a second type of default expression could be used to provide a minimum level of protection. Those requiring more protection could be allowed to add further terms and conditions. Again, this doesn't cater for those requiring less protection. It would be a shame if a depositor was happy for their work to be used without terms or conditions, to have the unwanted minimum terms and conditions imposed upon them. Thus a third type of default expression is recommended, where the default applies unless otherwise stated. This allows for those requiring more or less protection than the default.

In some cases, repositories may wish the default level to be a particular CC licence. (The CC 'Choose a Licence' engine makes public domain the default (Creative Commons, 2003)). However, many repositories may find it simpler to use a simple copyright statement, such as "Unless otherwise stated, the items in this repository may be used in accordance with the national copyright law in the country of use."

Default rights expressions may be harvested as part of the optional <description> element that can be contained in the Identify response. For example, the OAI Executive have written an *XML Schema to describe content and policies of repositories in the e-print community* (2002). It allows the <description> element to contain both a <metadataPolicy> and a <dataPolicy> element. It is proposed that the rights

expression forms part of the <dataPolicy> element. <dataPolicy> elements can contain <text> and/or <URI> elements. Thus either the <text> element would contain the default copyright statement and the <URI> element a link to a copy of national copyright law, or the <text> element would state the name of the default CC licence and the <URI> element either a link to the appropriate Commons Deed, and/or an http GetRecord request linking to the rights metadata record.

Future work

In addition to considering how to express rights over resources (in this case, research papers) via the OAI-PMH, Project RoMEO has also developed a means of expressing rights over metadata. A short piece on this work has been written up in Ariadne magazine (Gadd et al, 2003f). However, the project is currently exploring a collaboration with the OAI aimed at developing a specification and guidelines for disclosing rights information (about both metadata and resources) under the OAI-PMH. The exact nature of the collaboration and its scope remain to be decided upon, but our expectations are that results will become available in the first half of 2004

Acknowledgements

The RoMEO Project should like to thank the UK Joint Information Systems Committee for funding this research. We should also like to thank Herbert van de Sompel of the OAI, Renato Iannella of the ODRL, and Aaron Swartz of the Creative Commons for correspondence which made a significant contribution to our work.

References

Anon. (2002) What is copyleft?, Free Software Foundation Inc., URL: http://www.gnu.org/copyleft/copyleft.html.

ArXiv.org eprint archive. (2003) URL: http://arxiv.org/

Bass, M., et al (2002). DSpace: a sustainable solution for institution digital asset services: spanning the information asset value chain: ingest, manage, preserve, disseminate: internal reference specification: functionality. Cambridge MA, Hewlett Packard Company: 10 http://dspace.org/technology/functionality.pdf.

Cope, B., and Robin Freeman, Ed. (2000). Digital rights management and content development. Altona, Vic., Common Ground.

Counter copyright, (2002) Berkman Center for Internet & Society. URL: http://cyber.law.narvard.edu/cc/cc.html.

Creative Commons. (2002) URL: http://www.creative-commons.org

Creative Commons. (2003). Choose a Licence. URL: http://creativecommons.org/license/

Crow, R. (2002). The case for institutional repositories: a SPARC position paper. Washington DC, The Scholarly Publishing and Academic Resources Coalition: URL: http://www.arl.org/sparc/IR/ir.html.

Digital Object Identifier. (2003) URL: http://www.doi.org/

Gadd, E., Charles Oppenheim and Steve Probets (2003a). "RoMEO Studies 1: The impact of copyright ownership on academic author self-archiving." Journal of Documentation **59**(3): 243-277.

Gadd, E., Charles Oppenheim and Steve Probets (2003b). "RoMEO Studies 2: How academics want to protect their open-access research papers." Journal of Information Science 29(5): [In Press].

Gadd, E., Charles Oppenheim and Steve Probets (2003c). "RoMEO Studies 3: How academics expect to use open-access research papers." Journal of Librarianship and Information Science 35(3): [In Press].

Gadd, E., Charles Oppenheim and Steve Probets (2003d). "RoMEO Studies 4: The author-publisher bargain: an analysis of journal publisher copyright transfer agreements." Submitted to Learned Publishing.

Gadd, E., Charles Oppenheim and Steve Probets (2003e). "RoMEO Studies 5: The IPR issues facing OAI Data and Service Provicers." Submitted to Journal of Information Law and Technology.

Gadd, E., Charles Oppenheim and Steve Probets (2003f). "The RoMEO Project: protecting metadata in an open-access environment." Ariadne 36, [In Press].

Great Britain. (1988). Copyright designs and patents act, 1988, The Stationery Office.

Harnad, S. (2001). For Whom the Gate Tolls? How and Why to Free the Refereed Research Literature Online Through Author/Institution Self-Archiving, Now. URL: http://cogprints.soton.ac.uk/documents/disk0/00/00/16/39/index.html.

Iannella, R. (2001). "Digital rights management (DRM) architectures." D-Lib Magazine 7(6): URL: http://www.dlib.org/dlib/june01/iannella/06iannella.html.

Indecs Consortium (2002). <indecs>rdd White Paper: a standard Rights Data Dictionary, London: Rightscom. URL: http://www.rightscom.com/Indecs2RDD_White_Paper_May2002.pdf

<indecs> framework. (2003) URL: http://www.indecs.org/

Johnson, P. (2003). "All Wrapped Up? A Review of the Enforceability of "Shrink-wrap" and "Click-wrap" Licences in the United Kingdom and the United States." European Intellectual Property Review 25(2): 98-102.

McAllister, N. (2002). "Freedom of expression: emerging standards in rights management." New Architect Magazine March. URL: http://www.newarchitectmag.com/documents/s=2453/new1011651985727/index.html.

McCracken, R., and Madeleine Gilbart (1995). Buying and clearing rights. London, Chapman and Hall.

Miller, E. (1998). "An introduction to the Resource Description Framework." D-Lib Magazine May URL: http://www.dlib.org/dlib/may98/miller/05miller.html.

Online Information Exchange. (2003) URL: http://www.editeur.org/onix.html

Open Archives Initiative Protocol for Metadata Harvesting (2002). OAI Executive; OAI Technical Committee. URL: **http://www.openarchives.org/OAI/openarchivesprotocol.html**.

Opencontent, (2002) URL:http://www.opencontent.org/index.shtml.

Open Digital Rights Language. (2002) URL: http://odrl.net/

Owen, Lynette. (2002) Clark's publishing agreements. Butterworths : London.

Powell, A., Michael Day and Peter Cliff (2003). Using simple Dublin Core to describe eprints. Bath, UKOLN. URL: http://www.rdn.ac.uk/projects/eprints-uk/docs/simpledc-guidelines/.

RoMEO Project. (2003). URL: http://www.lboro.ac.uk/departments/ls/disresearch/romeo/index.html.

Rosenblatt, W., William Trippe and Stephen Mooney (2002). Digital rights management: business and technology. NY, M&T Books.

Smith, Mackenzie (2003) to Elizabeth Gadd. Personal correspondence. 27 May 2003.

XML Schema to describe content and policies of repositories in the e-print community, (2002) OAI Executive. URL: http://www.openarchives.org/OAI/2.0/guidelines-eprints.htm.

XrML. (2002) URL: http://www.xrml.org/