

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



**CC creative commons**  
COMMONS DEED

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

**BY:** **Attribution.** You must attribute the work in the manner specified by the author or licensor.

**Noncommercial.** You may not use this work for commercial purposes.

**No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

## **The IPR issues facing self-archiving: key findings of the RoMEO Project**

Elizabeth Gadd, Charles Oppenheim, and Steve Proberts

### **Introduction**

The RoMEO (Rights Metadata for Open archiving) Project was funded for one year (2002-2003) by the UK JISC to investigate the IPR issues relating to self-archiving. It was specifically considering the self-archiving of research papers by academics and the subsequent disclosure and harvesting of metadata about those papers using the OAI-PMH (Open Archives Initiative, 2002) by Data and Service Providers. It aimed to develop simple rights metadata by which academics could protect their research papers in an open-access environment, and also a means by which Data and Service Providers could protect their open-access metadata. It proposed to show how such rights solutions might be disclosed and harvested under OAI-PMH. The project was divided into two phases: a data-gathering phase and a development phase. The project team have produced a series of six studies based on their work, referred to as RoMEO Studies 1-6 in the remainder of this article (Gadd, Oppenheim, and Proberts, 2003a; 2003b, 2003c, 2003d, 2003e, 2003f). However, this article aims to provide an overview of all the activities of the project and to report on some of its key findings and recommendations.

### **Understanding stakeholder requirements**

The goal of the first phase of the project was to understand the IPR issues facing the key stakeholders in the self-archiving process. Online questionnaire surveys were performed of academic authors, scholarly journal publishers, and OAI Data Providers (DPs) and Service Providers (SPs). As the response from journal publishers was poor, it was fortuitous that the project had also planned to perform an analysis of journal publishers' author copyright agreements. Such agreements provide a good overview of the contractual relationship between author and publisher, and the analysis was very enlightening.

#### *Academic authors*

The full results of the academic author survey are reported in RoMEO Studies 1, 2 and 3. In total, 542 academics responded. They were based in 57 countries and from a wide variety of subject disciplines. The main aims of the survey were to find out how academics wanted to protect their actual or potential open-access research papers, and also how they expected to use such papers. This data would inform the development of some appropriate rights metadata for this purpose. Authors were also asked for their views on the copyright ownership of research papers.

In total, 61% of respondents thought that academics owned the copyright in such papers, although 32% admitted that they did not know. When it came to assigning copyright, however, 90% of respondents did so, which must include many of those that were not sure whether they in fact owned such rights. Fifty per cent of respondents indicated that 71-100% of their papers were multi-authored. This could leave room for disagreement amongst co-authors as to if, when, and where self-archiving took place. An unexpected finding was that 25% of respondents had had to clear third-party materials before publishing a paper. Again, this would affect an author's ability to self archive as the third-party would have to agree not only to publication in a peer-reviewed journal, but also to have their work made freely available on the web.

We provided academics with a list of possible permissions (e.g. print, save, excerpt), restrictions and conditions using terms from the Open Digital Rights Language (ODRL, 2003). A restriction is a limit on the extent of a permission (i.e., you may print, but only four times), and a condition is a prerequisite that must be met before the permission is granted (i.e., you may only print if you have first paid a fee).

We asked them to specify which they should like to apply to their own open access works. We then asked which of the same list they expected to apply to their use of *others'* open access works.

The majority (60% or more) were happy for others to display, print, save, excerpt from, and give away their research papers as long as they were attributed as the author, and that all copies were exact verbatim copies of the original work. Most wanted to prohibit sales of the work and 55% wanted to limit usage of the work to certain purposes, e.g., educational or non-commercial. A comparison between these usage limits and those provided by UK copyright law and many electronic journal licences showed that the academics' conditions were far more liberal .

Interestingly, the subsequent comparison between how academics-as-authors wished to protect their open-access research papers, and how academics-as-users expected to use such research papers, showed that academics did not expect to make use of all the permissions they were prepared to grant others.

### *Journal publishers*

The academic author survey showed that many academics feared either breaking existing publisher agreements, or not being published, if they self-archived their research papers. Obviously, there would be little benefit in developing a means of protecting open-access works through rights metadata, if academics were too anxious to self-archive those papers in the first place. The analysis of 80 scholarly journal publishers' copyright agreements (CAs) aimed to see if such fears were founded.

The full results of this analysis have been written up in RoMEO Studies 4. It showed that 90% of publishers did ask for copyright transfer, with six per cent asking for exclusive licences, and 4% for non-exclusive licences. However, we found that exclusive licences could be equally as restrictive (in terms of the rights retained by the author) as copyright transfer agreements. The move towards exclusive licences may be an increasing trend as publishers respond to authors' desire to retain copyright (ALPSP, 1999). Indeed, the UK Association of Learned and Professional Society Publishers has developed a model exclusive licence which it has made available to its members.

Despite the fact that most publishers asked for copyright, authors could still self-archive their works *before* assigning copyright, if they did not fear publishers asking them to warrant that the work has not been previously published: the so-called Ingelfinger rule (Relman, 1981; Angell, 1991). Our analysis showed that 68.7% of agreements asked for copyright prior to the refereeing process, meaning that only the preprint could be self-archived in this way if the Ingelfinger rule was in place. That would not stop authors employing the Harnad-Oppenheim proposal and self-archiving a corrigenda of changes along with the original preprint, whilst not breaking the terms of the agreement (Harnad, 2001). However, 75% of agreements asked authors to warrant that the work had not been previously published. Interestingly, only two specifically stated that they considered self-archiving to be prior publication, but it could be assumed that all the CA's prohibiting self-archiving felt the same way.

Of course, the matter of copyright transfer would be of less concern to the open-access movement if publishers granted back to authors the right to self-archive. We found that although 30% of CA's did not give authors *any* rights to use their own works, just under 50% allowed authors to self-archive. However, there were no standard terms and conditions under which they could do so. Some allowed self-archiving of the preprint only, some the postprint only, some ask for preprints to be removed on formal publication, and others specified the type of site on which the self-archiving must take place.

We recommend that author copyright agreements be revisited by a committee representing the interests of all parties, perhaps with a view to developing a model agreement that all stakeholders

could sign up to. In the meantime, however, open access proponents are pleased that at least 50% of journals allow author self-archiving of some kind. The project team have compiled a directory of publishers' self-archiving policies which is available from the RoMEO web pages (Project RoMEO, 2003).

#### *OAI Data and Service Providers*

The aims of the DP and SP surveys were to ascertain their views on the necessity of metadata protection and the format it might take. In addition, DPs were asked a number of questions about their relationship with depositing authors – in particular, whether the relationship was governed by a licence agreement. The full methodology and results have been written up in RoMEO Studies 5. Thirteen SPs and 22 DPs responded.

Only one-quarter of responding DPs had licence agreements with their depositing authors. Half of all respondents either trusted their depositors only to mount documents that they had the right to, or just issued a general warning statement on this issue. In some cases this may have been because the DP had only recently been established and had yet to finalise their policies and procedures. However, it is important that DPs protect their own interests even in an open-access environment. A DP making available a copyright-infringing work will be responsible (under UK law at least) for secondary infringement.

When it came to the subject of metadata protection, an interesting picture emerged from the responses. Initially, 50% of responding DPs thought that metadata records were facts and as such there was no copyright in them, and 68% of responding DPs believed that whilst there was database right in their collection of metadata records (Council Directive, 1996), that this was “implicitly waived” in the OAI community. One-third of SPs also thought that metadata was implicitly free under the OAI and another third had never thought about it. Only a third checked a DP's policy prior to harvesting. However, when asked about how they expected their metadata to be used, 90% of DPs selected conditions they should like to apply to their metadata indicating that there were rights that they wanted protecting. Similarly, no SP stated that it was happy for its enhanced metadata to be harvested unconditionally.

The majority of DPs (68.4%) wanted the metadata to be attributed to their organisation. Fifty-eight per cent wanted the metadata to continue to be made freely available, and for non-commercial purposes. A surprising 52.6% wanted to specify that their metadata remained unaltered. Of course, were this to be implemented, it would inhibit the function of Service Providers, many of whom need to enhance the metadata (e.g., provide subject indexing or authority control) in order to provide services. Just three main conditions of use were listed by SPs: one was ‘by prior agreement’ – not a condition that could be easily automated. Another was attribution of the Provider, and the third was that subsequent harvesters disclose the metadata under the same conditions as it was harvested.

Promisingly, 77% of DPs and 77% of SPs thought a standard means of expressing the rights status of metadata would be beneficial.

#### **Rights metadata and metadata rights solution**

Having ascertained the needs of academic authors, and OAI Data and Service Providers with regards to the protection of their open access research papers and metadata, the second phase of the project set about considering the best way to express and disclose those rights. Much of this work has been documented in RoMEO Studies 6.

### *Developing rights expressions*

To develop a set of rights expressions that met the requirements of academic research papers and metadata we had three main options. Firstly, we could have developed our own expression language for the purpose. Secondly, we could have utilised an existing Digital Rights Expression Language (DREL). 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 (2003) (Open Digital Rights Language). A third option was to turn to the Creative Commons Initiative (2002) that was developing a complete rights solution for open access works. The Initiative 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.

In the interests of standardisation, the option of developing our own expression language was quickly dismissed. XrML was also dismissed on the grounds that it was a commercial patented product with unclear licensing terms, and at the time of project development it 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. By contrast, ODRL was an open source language with a form of Data Dictionary. That is, the Dictionary provides a list of terms, but no generally agreed upon meanings for those terms. ODRL was the language used in the academic author survey.

The Creative Commons (CC) solution went beyond the communication of rights through metadata to their expression through simple human-readable “Commons Deeds” with associated symbols, and detailed “Licence Codes”. However, as a result of this three-pronged approach to rights expression, the actual rights metadata records were not very descriptive of the permissions and restrictions granted by the licences. For example, each licence allows the ‘licensee’ to aggregate the work into a collection of works. However, the rights metadata instances do not specify that “aggregation” is permitted.

A comparison of the ODRL and CC solutions showed that either would meet the basic requirements of academics and Data and Service Providers as found by the RoMEO surveys, although a RoMEO application profile of ODRL would provide a higher level of granularity of expression than the simple CC metadata. As the RoMEO Project progressed, the CC 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).

The one key technical problem with adopting the CC solution was that their metadata was expressed in RDF/XML and did not have an associated XML schema – a prerequisite for any metadata disclosed under the OAI-PMH. The project therefore proposed a two-fold solution. It would work with the CC licences for expressing rights over research papers as they looked set to becoming an emerging standard. However, in addition to approaching the CC to encourage them to develop an XML schema for their RDF, they would also develop ODRL versions (XML instances) of the CC licences that would conform to the ODRL XML schema. The ODRL versions should provide a more accurate description of the content of the eleven CC licences than the CC’s own RDF.

### *Disclosing rights expressions under the OAI-PMH*

The next step was to consider how best to disclose the rights expressions under the OAI-PMH. After discussions with the OAI, it was proposed that rights expressions for both individual and collections of

metadata records and individual and collections of resources were disclosed. However, this work is to continue through the formation of an OAI/RoMEO Technical Committee, OAI-RIGHTS, which hopes to report in Spring 2004. What follows are the proposals reached by Project RoMEO's by the end of the project timeframe (Summer 2003).

The expression of rights and permissions relating to an individual resource (such as a research paper) would be expressed by the use of a separate rights metadata record. This record would consist of the XML instance (either ODRL or CC/RDF) relating to the chosen CC licence and would be accessible through an OAI-PMH GetRecord request with a specific metadataPrefix parameter, e.g. oai\_cc (see Figure 1). This instance could also be referenced by the mandatory Dublin Core metadata record relating to that document. Thus within the <dc:rights> element an OAI GetRecord verb URL could be included, which, if followed, would retrieve the rights metadata record.

```
<header>
<identifier>oai:brill:1234567</identifier>
<datestamp>2003-06-09</datestamp>
</header>
<metadata>
<oai_cc:cc
  xmlns:cc="http://creativecommons.org/metadata/schema/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://creativecommons.org/metadata/schema/
http://creativecommons.org/metadata/schema/cclicences.xsd">
  <cc:license rdf:about="http://creativecommons.org/licenses/by-nd-nc/1.0">
    <cc:permits rdf:resource="http://web.resource.org/cc/Reproduction" />
    <cc:permits rdf:resource="http://web.resource.org/cc/Distribution" />
    <cc:requires rdf:resource="http://web.resource.org/cc/Notice" />
    <cc:requires rdf:resource="http://web.resource.org/cc/Attribution" />
    <cc:prohibits rdf:resource="http://web.resource.org/cc/CommercialUse"
  />
  </cc:license>
</oai_cc:cc>
</metadata>
```

### Figure 1 CC/RDF metadata record

Rights expressions for individual metadata records would be contained within a record's optional <about> container. Again, either CC's RDF or an ODRL version of the CC licences could be used. Figure 2 shows an ODRL/XML version of the CC's Attribution Licence within the <about> container of a record. Obviously, we would recommend consistency in the choice of either CC's RDF/XML or ODRL/XML, and the use of one or the other to express rights over both resources and metadata.

```
<about>
<o-ex:rights xmlns:o-ex="http://odrl.net/1.0/ODRL-EX"
  xmlns:o-dd="http://odrl.net/1.0/ODRL-DD"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://odrl.net/1.0/ODRL-EX
http://odrl.net/1.0/ODRL-EX.xsd">
<o-ex:offer>
  < o-ex:context>
    <o-dd:uid>http://creativecommons.org/licenses/by/1.0/</o-dd:uid>
    <o-dd:uid>http://www.romeo.ac.uk/odrl-cc-licenses/by/1.0/</o-dd:uid>
    <o-dd:name>Creative Commons Attribution Licence</o-dd:name>
    <o-dd:date><fixed>2002-10-10</fixed></ o-dd:date>
  <o-ex:/context>
  <o-ex:permission id="CCplus">
```

```

    <o-dd:display/>
    <o-dd:print/>
    <o-dd:play/>
    <o-dd:excerpt/>
    <o-dd:aggregate/>
    <o-dd:give/>
    <o-dd:duplicate/>
    <o-dd:save/>
    <o-dd:modify/>
  <o-ex:requirement>
    <o-ex:attribution/>
    <o-ex:accept>
      <o-ex:context>
        <o-dd: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/1.0
        </o-dd:remark>
      </o-ex:context>
    </o-ex:accept>
  </o-ex:requirement>
</o-ex:permission>
</o-ex:constraint>
  <o-dd:transferPerm downstream="equal" idref="CCPlus"/>
</ o-ex:constraint>
</o-ex:offer>
</about>

```

**Figure 2: ODRL version of CC Attribution Licence within individual metadata record's <about> container**

Whole collections of metadata and resources would be protected by the optional <description> response to the Identify verb. The OAI's *XML schema to describe content and policies of repositories in the e-print community* (Open Archives Initiative, 2002a) recommends that eprint repository descriptions use a <metadataPolicy> and a <dataPolicy> element, each containing optional <text> and <uri> elements. Thus, <metadataPolicy> could provide a default statement describing the permissions status of metadata, and <dataPolicy> a default statement describing the permissions status of resources (see Figure 3). The default data policy would in most cases have to be a simple copyright statement, unless the repository only accepts resources meeting a minimum set of CC licence terms

```

<description>
  <eprints xmlns="http://www.openarchives.org/OAI/1.1/eprints"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/1.1/eprints
http://www.openarchives.org/OAI/1.1/eprints.xsd">
    <metadataPolicy>
      <text>The metadata records disclosed by the University of Brilliance
      Eprints Service (http://eprints.brill.ac.uk/) are made available under
      an Attribution-NonCommercial Creative Commons Licence </text>
      <URL> http://eprints.brill.ac.uk/oai-script?
      verb=GetRecord&identifier=oai:rights/by-nc</URL>
    </metadataPolicy>
    <dataPolicy>
      <text>Unless otherwise stated, the full-text documents housed by this
      repository may be used in accordance with your national copyright
      law.</text>
      <URL>http://eprints.brill.ac.uk/copyright/</URL>
    </dataPolicy>
  </eprints>

```

</description>

### **Figure 3: Repository-wide rights expressions**

#### **Conclusions and future developments**

The RoMEO Project has been a most interesting exercise. The findings, particularly the Directory of journal publishers' self-archiving policies, should encourage academics that self-archiving is a realistic approach. Nevertheless, the project has also highlighted a number of concerns about publishers' agreements, which if dealt with, could greatly improve an author's rights under the current journal publishing system. We have shown that academics do not require the level of copy protection currently provided by (UK) copyright law and/or publishers' e-journal licences. Therefore, the provision of an alternative means of protecting their works through rights metadata, such as that proposed by the project's development phase, should be a welcome one. We have also demonstrated that whilst most Data and Service Providers are happy to share metadata in the spirit of open-access, they too are interested in protecting some of their interests as rights-holders. It is hoped that the metadata protection solution proposed by the project will protect those rights. In this vein, the RoMEO Project is delighted to be working with the OAI in establishing a dedicated Technical Committee, OAI-RIGHTS, to further develop their technical proposals into generic guidelines for disclosing rights expressions under the OAI-PMH. It is hoped that the work of the committee will be available for general comment in the Spring of 2004.

#### **Acknowledgements**

The Project team gratefully acknowledge the Joint Information Systems Committee for funding this project. They are also indebted to Herbert Van de Sompel of the OAI, Renato Ianella of the ODRL, and Aaron Swartz of CC, for the invaluable advice they have freely given during the developmental phase of the project.

#### **References**

- ALPSP (1999). What authors want :the ALPSP research study on the motivations and concerns of contributors to learn. West Sussex, Association of Learned and Professional Society Publishers.
- Angell, M., J.P. Kassirer (1991). "The Ingelfinger rule revisited." New England Journal of Medicine 325(Nov.7, 1991): 1371-1373.
- 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>.
- Council Directive No. 96/9/EC of 11 March 1996 on the legal protection of databases. (1996) URL: <http://europa.eu.int/ISPO/infosoc/legreg/docs/969ec.html>
- Creative Commons. (2002) URL: <http://www.creative-commons.org>
- Gadd, E., C. Oppenheim and S. Proberts (2003a). "RoMEO Studies 1: The impact of copyright ownership on academic author self-archiving." Journal of Documentation 59(3): 243-277.
- Gadd, E., C. Oppenheim and S. Proberts (2003b). "RoMEO Studies 2: How academics want to protect their open-access research papers." Journal of Information Science 29(5): [In Press].



- Gadd, E., C. Oppenheim and S. Proberts (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., C. Oppenheim and S. Proberts (2003d). "RoMEO Studies 4: The author-publisher bargain: an analysis of journal publisher copyright transfer agreements." Learned Publishing 16(4): [In Press].
- Gadd, E., C. Oppenheim and S. Proberts (2003e). "RoMEO Studies 5: The IPR issues facing OAI Data and Service Providers." Submitted to Journal of Information Law and Technology.
- Gadd, E., C. Oppenheim and S. Proberts (2003f). "RoMEO Studies 6: Rights metadata for Open Archiving - the RoMEO Solutions." Submitted to Program.
- Harnad, S. (2001). "For Whom the Gate Tolls? How and Why to Free the Refereed Research Literature Online Through Author/Institution Self-Archiving, Now." <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>.
- Open Archives Initiative (2002a). The Open Archives Initiative Protocol for Metadata Harvesting <http://www.openarchives.org/OAI/openarchivesprotocol.html>
- Open Archives Initiative (2002b). XML Schema to describe content and policies of repositories in the e-print community, OAI Executive. <http://www.openarchives.org/OAI/2.0/guidelines-eprints.htm>.
- Open Digital Rights Language (ODRL) (2003) URL: <http://odrl.net/>
- Powell, A., M. Day and P. Cliff (2003). Using simple Dublin Core to describe eprints. Bath, UKOLN. URL: <http://www.rdn.ac.uk/projects/eprints-uk/docs/simpledc-guidelines/>.
- Project RoMEO (2003). URL: <http://www.lboro.ac.uk/departments/ls/disresearch/romeo/index.html>
- Relman, A. S. (1981). "The Ingelfinger Rule." New England Journal of Medicine 305(Oct 1, 1981): 824-826.
- Smith, Mackenzie, [DSpace Project Director] (2003) to Elizabeth Gadd. Personal correspondence. 27 May 2003.
- XrML. (2002) URL: <http://www.xrml.org/>