

# DELIVERABLE

**Project Acronym:** Europeana Libraries  
**Grant Agreement number:** 270933  
**Project Title:** Europeana Libraries: Aggregating digital content from Europe's libraries

## D2.2. IPR requirements for full-text delivery

**Revision:** v.1

**Authors:**

**Charles Oppenheim (Loughborough University Enterprises Ltd)**  
**Martin Moyle (UCL)**

<b>Project co-funded by the European Commission within the ICT Policy Support Programme</b>		
<b>Dissemination Level</b>		
<b>P</b>	<b>Public</b>	<b>P</b>
<b>C</b>	<b>Confidential, only for members of the consortium and the Commission Services</b>	

## Revision History

Revision	Date	Author	Organisation	Description
0.1	26/07/12	Charles Oppenheim, Martin Moyle	LUEL UCL	First draft
0.2	06/08/12	Charles Oppenheim	LUEL	Comment from case study partner
0.3	06/08/12	Charles Oppenheim	LUEL	Comment from case study partner
0.4	06/08/12	Charles Oppenheim	LUEL	Comment from case study partner
0.5	06/08/12	Charles Oppenheim	LUEL	Comment from case study partner
0.6	08/08/12	Charles Oppenheim	LUEL	Comment from case study partner and external reviewer
1	10/08/12	Charles Oppenheim, Martin Moyle	LUEL UCL	Comments from second external reviewer and final editing

### Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

## D2.2. IPR requirements for full-text delivery

### Table of Contents

1. Introduction	4
1.1. Background and overall approach	4
1.2. The EU and copyright principles	4
1.3. Scope of copyright law	4
1.4. Full text and copyright	5
1.5. Main recommendations to The European Library	5
2. Ownership, and owners' rights	6
3. Users' rights	7
4. Orphan Works	7
5. The position of legal deposit libraries	9
6. The position of full text	9
7. Protection of databases	10
8. Licences	10
9. Europeana Data Exchange Agreement	11
10. Ensuring that digitisation projects are conducted with a minimum of risk	12
11. Case studies	12
11.1. An aggregator: Directory of Open Access Journals (DOAJ)	12
11.2. A national library: Bibliothèque Nationale De France (BNF)	13
11.3. A 'Google library': University of Ghent	15
11.4. Open data: Bayerische Staatsbibliothek	16
11.5. Conclusions from the case studies	17

## 1. Introduction

### 1.1. Background and overall approach

This report originated in Task 2.4.1 of the Europeana Libraries project's Work Package on landscape modelling and sustainability, in which the partnership committed to

*review current work on legal, copyright and licensing issues...paying particular attention to full-text delivery.*

The report documents the copyright, licensing and authors' rights issues that are relevant to The European Library and its interaction with Europeana. The report begins with an overview of the main copyright and related issues that pertain to the work of the project. These sections are followed in the report by four case studies, each intended, through description and commentary, to draw out and highlight elements of the main body of the report and any implications for The European Library. Four project partners were selected for case study, each representing a different type of content provider: the studies highlight, in turn, an aggregator, a national library, a 'Google library', and a library with aspirations for open data publication. Case study partners introduced their collections, and responded to follow-up questions about, for example, the copyright status of the material, the process (if any) by which permissions were sought, the relevance of national laws (if any) to the collections, and so on.

Professor Charles Oppenheim of Loughborough University Enterprises Ltd, UK, carried out the review and case studies under the overall direction of the Europeana Libraries Business Planning Group, with valuable input from the Europeana Libraries Aggregation Steering Group. The authors are especially grateful to the four partners who agreed to participate in the case studies.

### 1.2. The EU and copyright principles

The same basic principles of copyright apply throughout the EU because a number of EU Directives have imposed common basics of the legal regime. Europeana partners based in non-EU Member States will find that, with minor differences, their laws follow the same ground rules as well. This is because all participating countries, including all EU Member States, are signatories to the two major international copyright conventions (The Berne Convention and the Universal Copyright Convention) and all have also signed up to the TRIPS Agreement under the WTO. These three agreements between them impose considerable constraints on the flexibility of individual countries to have divergent copyright laws in respect of ownership and owners' rights, though they allow considerable leeway on exceptions to copyright, as noted below. This report covers the broad principles that will apply to all participants, but readers should satisfy themselves about details unique to their particular countries' laws.

### 1.3. Scope of copyright law

The basics of copyright law apply to metadata, previews, full text, sound, images, moving images, broadcasts, etc. EU Member States all provide for database rights in addition to copyright for certain collections of data, and subject to provisos regarding the effort that was required to collate and present the data. Database rights are explored in section 7 of this report. Non-EU member States do not generally offer such rights, though in practice databases are often protected under generic copyright legislation. Metadata can cause problems due to inadequate protection in some

countries, which means that third parties based in such countries may be able to copy and exploit such metadata without having to ask for permission.

#### **1.4. Full text and copyright**

As shown at 1.3, in terms of the application of copyright law, full text differs little from any other content. If there is a difference, it is one of practice, in so far as the owners of rights in full text might be expected to be more protective of their assets, increasing the risk that perceived infringements would be pursued by rights holders. Throughout the report, IPR issues relating to full text are drawn out wherever possible. Section 6, in particular, discusses the issues of full text management in more detail.

#### **1.5. Main recommendations to The European Library**

Recommendations are made to The European Library concerning the necessary rights frameworks for the indexing and dissemination of full text content, and for the transformation of metadata. The recommendations are discussed in full in sections 6 and 9 respectively. The main substance of each recommendation is given here, as follows:

##### ***Recommendation: full text***

The European Library will need to assure itself that what is offered to it does not infringe any third party copyright. If the contributing library using human skill and judgement has created the indexes or other metadata, then arguably the copyright and database rights in those indexes belong to the library submitting them. If the indexes have been created using computer algorithms, then the copyright ownership of them is less certain. Either way, services such as The European Library will need to obtain from the contributing library a legal disclaimer along the lines of the following:

*The Contributor hereby warrants to Europeana and its assignees, licensees and distributors that EITHER it is the owner of the rights in the Contribution OR that the Contribution is in no way an infringement of any existing copyright, database right or licence, or duty of confidentiality, or any other right of any person whatsoever.*

This, or similar wording, should be present in the agreement that The European Library should enter into with each contributing organisation.

##### ***Recommendation: metadata***

The European Library, as an aggregator, intends to merge the metadata it receives, and to transform it for consumption by Europeana. The European Library should ensure it gets a royalty-free licence from each contributor for any metadata or indexes the contributor provides, and that it makes explicit that it will own the rights to any merged metadata or indexes. The European Library also needs to decide how it will make available the data it receives in the form of metadata or indexes, and under what terms and conditions.

For full discussion of these issues, see section 9.

## 2. Ownership, and owners' rights

The fundamental rules on these aspects of copyright law are nearly uniform throughout the EU. The first owner of copyright in a work is generally the original creator, though rules regarding creations made by employees do vary slightly. In some cases, ownership of copyright in materials created by an employee in the course of their employee duties automatically belongs to the employer, but in other cases it belongs to the employee unless there is a contract specifying that the ownership be transferred to the employer. Freelancers and students own the copyright in the work they create unless they have signed an agreement to the contrary.

The degree of creativity required in order to enjoy copyright also varies slightly. In some countries, a minimum of creativity is required, just as long as what is created is new and involved some minimal degree of effort by the creator. In other countries, the creator has to have shown some level of skill and judgement before the work enjoys copyright. In practice, the differences are slight, but it can mean that sometimes something that enjoys copyright if made in the one country might not have enjoyed copyright if it had been made in another.

The lifetime of copyright for most works is uniform throughout the EU and in most non-EU countries. The default period is life of the creator + 70 years, though there are exceptions to this basic rule, e.g., for films, broadcasts, unpublished works and sound recordings, where the lifetime is not 70 years from the death of the creator, or indeed can differ between Member States. Some non-Member States might have lifetime rules that are shorter than life + 70. The Europeana Public Domain Calculator (<http://outofcopyright.eu/calculator.html>) is an excellent resource for assessing if a given object is in copyright or not, but readers are advised that it does not necessarily cover every subtlety of each country's laws on lifetimes.

The rights enjoyed by the rights owner (who of course is not always the creator, as the creator may have assigned rights to a third party) are uniform throughout the EU and in non-EU Member States. These include the right to authorise (or refuse to authorise) certain restricted acts, such as copying the work, issuing physical copies of the work to the public (N.B. "Public" includes any sub-set of the public throughout this text), broadcasting the work, performing the work in public, and communicating the work to the public. The latter right covers all forms of electronic dissemination, including placing the work on an Intranet, on the Web, attaching it to an e-mail, etc. Some non-EU Member States do not explicitly offer communication to the public as an explicit owner right, but in practice their laws will also protect against unauthorised placing on the Web, etc.

Joint ownership of a given work is common (see Section 4, Orphan Works, for more background). This can arise in two different ways. The first is where the work is the creation of more than one person, and it is impossible to tease out who contributed what. Typical examples are jointly authored texts, and jointly created artistic works. The second is where a whole work in practice consists of a series of separate works. An example might be a book, written by one person, but including photographs created by another person, an index created by a third person, a foreword written by a fourth person, and so. In such cases, multiple permission requests for permission to digitise will be necessary if the relevant parts of the work are still in copyright.

### 3. Users' rights

In sharp contrast to ownership and owners' rights, exceptions to copyright are hardly harmonised at all in the EU. This is caused by a combination of different traditions in the various Member States and the permissive nature of the EU Directive, which provides a long list of POSSIBLE exceptions to copyright, which Member States can choose to adopt, or not adopt, as they see fit. There is only one compulsory exception, i.e., for short-term caching of a work on a computer when the caching is a necessary part of some technical function and the caching has no economic significance. Non-EU Member States also offer considerable variation in exceptions to copyright.

Although this section is entitled "Users' rights", in practice these exceptions to copyright are not rights granted to users, but are rather defences against an infringement action that a user is entitled to employ if sued. Many users are concerned that these exceptions are not spelt out as rights, as well as the fact that the exceptions are not consistent across the EU and elsewhere. However, it is worth noting that the World Intellectual Property Organisation (WIPO), the UN Agency responsible for the running of the Berne Convention, as well as some other copyright-related treaties, has recognised this as an issue. As a result, WIPO is proposing to introduce a Treaty in 2012 setting minimum standards for exceptions to copyright as they apply to libraries and archives. This is important to the Europeana projects, as for the first time minimum international standards will be set for the exceptions to copyright that Europeana projects will be exploiting. The current state of play is that a WIPO Committee is examining the wording of a possible draft Treaty, but nothing is likely to happen for quite a few years to come.

In general, exceptions to copyright permit users to make copies of copyright materials for non-commercial research or private study. In some countries, copying for educational purposes is also available as an exception. In other countries, such as the USA, the purposes of copying covered by exceptions are not specified, but each use is considered on a case-by-case basis. In addition to copying, some countries' exceptions permit redissemination of the copied work, whilst others permit copying and amendment for the purposes of parody. Some countries' exceptions explicitly permit the provision of links to third party materials, but others do not. The key point to note is that the law relating to exceptions does vary from country to country. Thus, TEL may well find (say) users from the USA being free to copy materials obtained via TEL in quite different ways to (say) a user from the UK. It is up to each user to obey the law of their home country. The law of the contributing library, or of TEL's home country itself, is irrelevant to what a user can or cannot do.

TEL is of course a user in its own right, providing links to contributing libraries, but since its relationship with contributing libraries is governed by a series of formal agreements, TEL does not need to worry about whether what it is doing is permitted under an exception to copyright.

### 4. Orphan Works

Orphan works (OW) are works that are in copyright, but where the rights owner cannot be identified, or if they can be identified, cannot be traced to contact for permission to carry out a restricted act of their work. The problem of OW is growing steadily as more and more material is placed on the Web by people and organisations who do not bother to identify themselves, e.g., in the metadata associated with the creative work. The legal position is that any copying of an OW, e.g., by digitisation, and any communicating of that OW to the public, e.g., by placing the digitised version

on the Web, is potentially infringement. The European Commission is well aware of the difficulties that this legal position raises for projects such as Europeana and, driven by this and by the opportunities raised by the failure of the Google Books settlement (which gives the EU an opportunity to be the front runner in the digitising of OW), has introduced a Draft Directive on Orphan Works. Unfortunately, this fails to address the issues raised by a large-scale digitisation programme such as Europeana.

The strictures of Berne mean that copyright is obtained without having to go through any formalities such as registration (Article 5(2) of the Convention). So one obvious approach for a draft Directive – that all rights owners are obliged to register to obtain (or renew) copyright (and by that registration can then be traced for approaches for permission to copy to be made), is not possible. The Berne Convention also is worded in a way that *seems to prevent* compulsory licences, though this is not explicitly stated in the Convention. However, compulsory licences are very unpopular – understandably – with rightsholders, and so the Draft Directive was carefully constrained to minimise opposition. Obvious constraints, some of which have been adopted in the Draft Directive, include the following:

1. Restricting the media types that could be subject to the new rules
2. Restricting the time periods to materials older than a certain age
3. Only allowing a limited number of types of organisation to take advantage of the licence, e.g., specified publicly-funded libraries, archives or museums
4. Requiring minimum standards of effort by those applying for a licence in attempting to track down the rights owners in question

Restrictions such as 1, 2 and/or 4 would have the effect of increasing the workload on any library, archive or museum wanting to digitise orphan works. However, they would be popular with rights owners, especially the photographers' lobby, which is especially vocal in its opposition to compulsory licences for orphan photographs.

In a quite separate development, a Memorandum of Understanding (MoU) was signed in July 2011 by representatives of European libraries, publishers and Reproduction Rights Organisations agreeing key principles on the digitisation and making available out of commerce works (see <http://www.libereurope.eu/news/liber-signs-mou-on-out-of-commerce-works-on-behalf-of-european-research-libraries>). Although “out of commerce” and “orphan” are not synonymous, there is significant overlap between the two. The MoU, which strongly encourages bureaucracy-free digitisation of out of commerce works by libraries and cultural institutions, was developed with encouragement from the European Commission, and indeed recommends that the Commission drafts legislation to help the MoU achieve its objectives. It provides a good definition of what is meant by “out of commerce”, despite minor issues with the Key Principles document in its definition of “rightsholder” and its failure to take into account problems that arise if more than one Reproduction Rights organisation exists in a country for a particular media type. This is an important step forward, but time will tell how willing individual publishers and rightsholders will be to abide by the Principles.



## 5. The position of legal deposit libraries

Legal deposit libraries have an obligation to collect materials published in their countries. Typically, the publisher of the material is required to provide one or more copies free of charge to the deposit library in question. Whilst the laws vary from country to country regarding the media types that must be deposited and the number of items to be deposited, and the types of preservation copies of deposited items that can be made, the various countries' laws are agreed that once obtained, the legal deposit library is not permitted to offer widespread access to its deposited materials by electronic or other means.

## 6. The position of full text

The law does not distinguish between full text materials and shorter items based upon such full text, such as abstracts or metadata, including indexes. Thus, all the comments made in this report apply equally to all original works. However, because often the owners of rights in full text are more concerned about protecting those rights than they might be in regard to (say) abstracts or metadata associated with that full text, they will be more assertive when they believe their full text has been copied without permission. Thus, the key difference in practice between full text and surrogates for that full text is the likelihood of being sued for infringement if such full text is copied and/or re-disseminated without permission.

What if the full text has been indexed? There may well be database rights (see section 7) in the index itself. Any agreements entered into regarding licensing of the full text would have also to include reference to the licensing of the metadata associated with the full text, including the index terms used.

In a typical aggregation scenario, a central agency such as The European Library will wish to retrieve a selection of full text materials, index the contents and then offer the indexes to discovery services, including Europeana. The European Library will need to assure itself that what is offered to it does not infringe any third party copyright. If the contributing library using human skill and judgement has created the indexes or other metadata, then arguably the copyright and database rights in those indexes belong to the library submitting them. If the indexes have been created using computer algorithms, then the copyright ownership of them is less certain. Either way, services such as The European Library will need to obtain from the contributing library a legal disclaimer along the lines of the following:

*The Contributor hereby warrants to Europeana and its assignees, licensees and distributors that EITHER it is the owner of the rights in the Contribution OR that the Contribution is in no way an infringement of any existing copyright, database right or licence, or duty of confidentiality, or any other right of any person whatsoever.*

This, or similar wording, should be present in the agreement that The European Library should enter into with each contributing organisation.

Further questions arise regarding the ownership of copyright and database rights in any merged indexes that The European Library might create from its various constituent contributions. This is a

matter of business development and is outside the remit of this report, though there is an argument that whilst individual contributors should retain copyright in their particular indexes, they should grant The European Library a royalty free licence to both exploit such indexes and to merge them with others' indexes to create a master index of all contributions. Copyright and database right (see section 7) in such a master merged index would belong to The European Library.

## 7. Protection of databases

All EU member states, but few other countries, provide special protection for databases. A database is any collection of data or other materials that are organised in a systematic manner of some sort and where each item is individually accessible in some way. Each individual item may or may not be copyright (if it is a single piece of data, that is too small to justify copyright protection; anything written by someone who died more than 70 years ago is out of copyright), but as long as the collection was created relatively recently, the creator of the collection may enjoy intellectual property rights. Within the EU, two rights are possible – copyright for the database, and/or database right for the database. Copyright is associated with the collection if the selection and arrangement of the materials involves creativity. Database rights are associated with the collection if there was significant resource expended in obtaining, verifying or presenting the collection. Important European Court of Justice cases have confirmed that “obtaining” refers solely to getting hold of the original data, and not to the effort of creating the data in the first place. The distinction is subtle, but means that if the data in question just fell out of other activities being undertaken in the normal course of its activities by the organisation that owns the data, then there may not be any protection under database rights. Thus, for example, if metadata were also generated as a necessary part of the creation of a digital object, the collection of such metadata might not enjoy database right.

Consider a project for the digitisation of a series of manuscripts. It is likely that the digitised collection of manuscripts will enjoy copyright because someone used their intellect and creativity to decide which items were to go in the collection and which were not. The collection may well also enjoy database rights, though because database rights are weaker than copyright in terms of length of protection and the list of restricted acts, most will ignore that factor. What if metadata is routinely added to each item? The individual metadata records may or may not enjoy copyright, depending on how much information is within the metadata record; each metadata record might enjoy database right if it comprises a series of fields and populating those fields required effort in obtaining the data; and the metadata collection as a whole may not enjoy either database right or copyright. It is difficult to generalise, and each case would have to be examined on its own merits.

## 8. Licences

Where a project intends to digitise materials that are owned by an identifiable third party, it is necessary to approach the owner with a request for a licence. The material may already be available under a licence, such as a Creative Commons licence (<http://creativecommons.org>), or some similar licence; in such circumstances, there is no need to make a separate approach to the licensor unless the uses envisaged fall outside the scope of the licence on offer. But frequently, there is no licence obviously on offer, so the owner has to be approached for one. Model wording

for an approach can be found at

<http://www.jisc.ac.uk/publications/programmerelated/2009/scaiprtoolkit/2emailtemplate.aspx>.

It is important to note that if the rights owner fails to reply, that must be interpreted as a rejection. One cannot approach a rights owner on the basis of “I intend to digitise this and unless I hear from the contrary, I will assume you have agreed” (the approach of Google that led to the collapse of its Google Books project). It is important that the request for the licence is drafted as broadly as possible, so that all possible future use of the digitised materials are anticipated; this ensures that there will be no need to get back in touch with the owner later on with further requests for permissions. In many cases, the rights owner will not require a fee for the licence, but in some cases they will do. If they do demand a fee, then it is up to the project to decide whether it considers the fee value for money.

Countries vary greatly in their approach to whether a licence can over-ride an exception to copyright. Many countries adopt the approach that if a user has willingly entered into a licence with a rightsholder, and that licence in effect removes some user rights enshrined in legislation, then that is the user’s problem. Others have enshrined in legislation that no licence can reduce the rights of a user to enjoy exceptions to copyright.

## 9. Europeana Data Exchange Agreement

The final draft was issued on 15 July 2011 – see [www.version1.europeana.eu/web/europeana-project/newagreement/](http://www.version1.europeana.eu/web/europeana-project/newagreement/). A summary of how it will be applied can be found at <http://pro.europeana.eu/documents/900548/978614/Europeana+LOD+pilot>. The parties have agreed to submit Content, Metadata and Previews of material submitted to the project. The Agreement notes that the Metadata must be accurate in respect of Intellectual Property Rights associated with the Content. All Metadata submitted to Europeana has to be made available under a CC0 Universal Public Domain licence (note the caveats about the use of this licence in the Case Studies – see section 11.4), though this does not stop a Data Provider entering into commercial arrangements with third parties separately from Europeana. All Data providers also agree that wherever possible they shall not make any claims to own IPR in the metadata they provide. In contrast, the Data providers can impose whatever terms they like on the use of Previews. The Agreement also includes provisions regarding what happens if a third party complains that material contributed by a Data Provider infringes their Intellectual Property Rights or is otherwise illegal, and that in particular, Europeana shall maintain a notice and take down procedure. It is not made clear who is responsible for paying damages or fines if it turns out that a claim of illegality or infringement is proven.

TEL, as an aggregator, intends to transform and merge metadata it receives for consumption by Europeana. The rights issues involved (other than those mentioned above) are as follows: TEL must get warranties from each organisation contributing metadata that either the copyright is theirs or that nothing in what is supplied infringes any third party rights; it must obtain this material under a royalty-free licence agreement (CC0, CC-BY, CC-BY-NC, etc. are appropriate; as noted above, CC0 is the preferred option); assuming the transformation and merging of data involves skill and judgement, then TEL will own the copyright and any database rights (see section 7) in any transformed metadata, but will presumably offer the transformed metadata under a CC licence; and

if the transformed data is used in any way contrary to the CC licence it is offered under, TEL will have to put into place methods for checking for abuse, and for initiating procedures should abuse occur.

## 10. Ensuring that digitisation projects are conducted with a minimum of risk

There are a number of useful resources to help ensure that material that is digitised is done so at minimum risk to individuals involved in the project or their employers. One useful resource is a set of guidance notes from the Web2Rights project, <http://www.web2rights.org.uk/documents.html>, in particular documents 2.1, 2.2, 2.3, 2.4 and 2.11. Although aimed at Web 2.0 projects and based upon UK law, these documents (which are available under a Creative Commons licence) can easily be adapted for Europeana projects. Similar material, still with a slight UK law bias but not aimed at Web 2.0 applications, can be found at the Strategic Content Alliance web site (<http://www.jisc.ac.uk/publications/programmerelated/2009/scaiprtoolkitarchivist.aspx>). The advice given by the Strategic Content Alliance (SCA) on Orphan Works is also extremely helpful – see <http://www.jisc.ac.uk/publications/programmerelated/2011/orphanworksbp.aspx>. The SCA site contains a wealth of other helpful advice, e.g., on the use of Creative Commons licences in digitised materials made available to the public: (<http://www.jisc.ac.uk/publications/programmerelated/2011/scaembeddingcclicencesbp.aspx>).

The key issue for all Europeana projects is risk assessment and management. Indeed, it can be argued that copyright is less to do with formal legal statements than it is to do with the management of risk. The Strategic Content Alliance provides much helpful guidance on this, too – see <http://www.jisc.ac.uk/publications/programmerelated/2009/scaiprtoolkit/2riskassessments.aspx>. All the SCA materials are available under Creative Commons, and so can be adapted for use for particular circumstances for specific projects.

## 11. Case Studies

### 11.1. An Aggregator: Directory of Open Access Journals (DOAJ)

This well-known service – see <http://www.doaj.org/> - is a collection of nearly 8,000 Open Access journals. DOAJ, run by Lund University, checks the journals according to its definitions and selection criteria, indexes the journals and provides links to the publisher's site. Around 45% of the journals are also searchable on article level in the directory; the publishers themselves upload the metadata. One can harvest content, on journal title level or on article level, into any library's own catalogue. All editors/journal owners are recommended to choose a CC-licence, which is shown in the record. If they choose the most permissive license, CC-BY, and upload all metadata back to the start year listed in DOAJ, they receive the SPARC Europe Seal for Open Access Journals. DOAJ receives funding from a wide variety of sources, including Lund University, but is always on the lookout for sponsorship from individuals and organisations.

About 1,400 of the journals listed emanate from the USA, followed by 700 from Brazil and 500 from the UK. However, these figures should be taken with a pinch of salt. For example, some of the US journals are published by Hindawi Publishing, which is in fact based in Egypt, but happens to have a

branch office in the USA. DOAJ's "country of origin" stats can be a bit misleading, because all its journals have an ISSN, which is connected to a certain country. Many OA journals from, e.g., India and Arabian countries have a US ISSN. It seems that many editors think a journal is more interesting to readers if its country of origin appears to be USA.

The DOAJ does not cover so-called hybrid journals, where some of the articles in any given issue are OA, but others are only available to subscribers.

### **Commentary**

There is little to comment on this service in terms of IPR. DOAJ staff check to ensure that the journals (and/or the articles within them) really do abide by the criteria set down for "Open Access". There will no doubt be a minority of cases where although the journal, or a particular article, is claimed to be "Open Access", for some reason the rights to the item have been reserved and the material should not be considered OA. Under such circumstances, someone copying and re-disseminating the material might find themselves threatened with a copyright infringement action. However, the chances of any serious problem arising are very low, and DOAJ has withdrawn an article, or an entire journal from the DOAJ if they have received notice that it should not be in its listing; either a reader/user informs them, or the editor her/himself. The reason very often is that the journal has turned into subscription mode. DOAJ staff write to the editor after having checked the journal and tell her/him DOAJ will remove the journal from list. For example, 22 journals were removed for the reason: "no longer OA" in the period Jan.- April 2012.

There appear to have been no cases where someone has taken information from DOAJ in good faith, copied the material, re-disseminated it, etc., and then received a complaint from a rights owner that they were infringing copyright.

It is understood that DOAJ does not enter into a formal agreement of any kind with publishers. If so, it would be better that Lund University enter into a formal contractual arrangement with each contributing publisher whereby each publisher grants Lund University permission to reproduce what information it supplies, and warrants that what it supplies to Lund does not infringe any third party rights. Copyright and database rights (if any) in the entire DOAJ listing belong to Lund University. It should grant The European Library a royalty-free licence to use its materials by means of a written agreement, including a warranty as described earlier.

### **11.2. A National Library: Bibliothèque Nationale De France (BNF)**

The BNF is one of the largest public and research libraries in the world. Its collections encompass all areas of culture and knowledge in a great variety of languages. It offers access to its digital library *Gallica* based on the digitisation of selected items of its collections – see <http://gallica.bnf.fr>. *Gallica* now contains over 1.9 million digitised documents: manuscripts, sound materials, music scores, books, images and over 800,000 newspaper issues, in French and other languages. They cover all subject domains, but with an emphasis on literature and history. In addition to public domain materials, *Gallica* gives access to digitised documents belonging to French partner libraries as well as a set of copyright documents in collaboration with the French Publishers Association, some publishers and retailers. In 2010, *Gallica* received 9.5 million visits (up 25% compared to 2010) and 150 million pages were viewed. BNF plans to test a new digital mode of cooperation with other French libraries: digitisation of partners' materials using the production lines of the BNF, made

available on *Gallica* with reference to partners and with customised desktop environments, digital preservation and sustainable delivery of a digital copy for local use.

"Conditions for the use of Gallica's contents" are available at:

<http://gallica.bnf.fr/html/editorial/conditions-use-gallica-s-contents>

The BNF is also involved in several European projects: *IMPACT* which aimed at high quality OCR, full-text indexing and subject multilingual issues, *BHL -Europe* which intended to make available online Europe's biodiversity information as well as *Europeana Collections 1914-1918* which aims at creating a European corpus of digitized materials about the First World War. It is also a member of the *Europeana v2.0* thematic network. Besides these projects, the BNF was in charge of the co-ordination of *KEEP* as well as *Europeana Regia* until June 2012. It is also member of two new Europeana Projects: *Europeana Newspapers* and *Europeana Awareness*. *Europeana Newspapers* aims at the aggregation and refinement of newspapers for The European Library and Europeana.

### **Commentary**

Most of the material available through *Gallica* on the Internet is out of copyright. For copyright documents, three approaches are in use, as follows:

- (i) The BNF provides access to the documents held by some publishers, but they are delivered to users from the website of the relevant publisher.
- (ii) Born-digital documents received under legal deposit are available only on site within the premises of the BNF.
- (iii) Finally, the BNF is the rights holder (exclusive or not) for some of the documents on line on *Gallica* (for example some photography collections).

Where material is in copyright, contracts of assignment of rights are signed to allow BNF to digitise these works and to put them online. So far, no rights holder has refused to sign a contract of assignment of rights. In case of refusal, the work would not be digitised, nor put online.

In France, an "Out of Commerce law" was adopted into the Intellectual Property code on 1<sup>st</sup> March 2012. (Out of commerce was defined as pre-2001 books published in France, but which are not currently being published or distributed). For more information, see:

<http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=?cidTexte=JORFTEXT000025422700&dateTexte=&oldAction=rechJO&categorieLien=id>

BNF provided some commentary on this law and its implications for *Gallica*:

The law on the exploitation of out-of-commerce books of the twentieth century, adopted on March 1st, 2012, aims to make accessible digital forms of out-of-commerce books of the twentieth century still protected under copyright. To be considered as out-of-commerce within the meaning of this law, a book had to be published in France before January 1st, 2001, is no longer subject to a commercial distribution by a publisher and is not currently the subject of a publication in printed or digital form. This law provides that the National Library of France identifies in a public database all these works, whose digital exploitation will be managed by a Collective management organisation, which will pay publishers and authors equally.

The rights holders may oppose the registration of their work on the database within six months. In case of opposition from the author, the work cannot be exploited in a digital form. If the opposition comes from the publisher of the printed book, he/she will have to exploit the book in digital form within two years. In the absence of opposition from the publisher or the author, it is proposed that the publisher may exploit the work in digital form. The latter then has a period of three years to act. In the absence of digital exploitation by the publisher of the printed work, the Collective management organisation grants to a third party a non-exclusive licence to use the book in a digital form, in return for payment, for a period of five years renewable. If no right holder of the book in printed form is found, after 10 years within the 1st authorisation of exploitation, the CMO shall authorise libraries that are open to the public to reproduce and distribute in digital form any such books free of charge to its readers. 500,000 books could be digitised under this law. This law could lead to some out-of-commerce books appearing on *Gallica*.

Overall, BNF has taken a model approach to respecting and clearing copyright. Nonetheless, The European Library will require a warranty from BNF along the lines of that described earlier in this report to ensure that it takes on no risk when reproducing BNF contributions.

### **11.3. A 'Google' Library: Ghent University**

Ghent University is contributing a large number of digitised texts to Europeana. One of the most important collections is the Google books project. Google started digitising many of its out-of-copyright books at the end of 2007, including some of the greatest works of Dutch and French literature. As one of Belgium's largest libraries, the Ghent University Library will greatly increase the number of Dutch and French-language books available through the Google Book Search Library Project.

Google scans the books for free. The University Library invests time and effort in describing the books, in picking them from the shelves, in preparing shipments, etc., and receives the digitised materials. All printed books out of copyright, and selected manuscripts dating from before 1872 and works printed before 1942 of which the author has died more than 70 years ago, are included in the digitisation project - in total, some 300,000 volumes. Part of the University library's pamphlet collection, known as Tiele-Meulemann collection, is also being digitised. The collection focuses on the history of the Low Countries from 1490 until 1800. The collection comprises 25,000 items (approx. 1,000,000 pages). The collection provides a rich source for researchers studying political, historical, ecclesiastical, religious or socio-economic subjects.

As all these collections comprise out of copyright materials only, no legal problems arise. Nonetheless, the University will have entered into a formal agreement with Google which, while confidential, no doubt mentions copyright and warranties.

In addition, the University is contributing the following collections to Europeana:

Giovanni Battista Piranesi (1720-1778) was one of the greatest artists of the 18<sup>th</sup> century. As an architect, archeologist and etcher, he set the scene of Roman culture and the city of Rome. Ghent University Library possesses and has digitised almost the complete collection of Piranesi's prints.

The Ghent World Fair, 1913, was an exciting event that represented the state of the art of science, architecture and society at that point in time. The digitised collection contains pictures, posters, maps, medallions, post cards, brochures and many more ephemera in connection with the World Fair in Ghent. The large majority of the items are out of copyright.

A collection of pictures, newspaper cuttings, pamphlets and many other documents from World War I. Again, the large majority is already out of copyright. If items are clearly still protected by copyright (e.g., World War I photos by brothers Anthony), the digital objects are not open to the public, but access is limited to the institution. If someone detects items that are still protected by copyright, they are removed instantly.

24 masterpieces of the Ghent University Library collection have been recognised as cultural heritage by the Flemish government. They were created between the 12<sup>th</sup> and 19<sup>th</sup> centuries. 19 of them have been digitised completely. All these items are out of copyright. Details can be found at:

<http://search.ugent.be/meercat/x/all?start=0&q=%22Topstuk+Vlaamse+Gemeenschap%22>.

### **Commentary**

As can be seen from the above, the vast majority of Ghent University's digitised materials were out of copyright. No doubt the University checked the copyright status of some of its 20<sup>th</sup> century materials, but, like all Universities pursuing similar digitisation initiatives, it would be appropriate to have mitigation strategies in place should a claim of infringement occur. Such mitigation strategies could include any of the following:

- a clearly identified email address to send any complaints of alleged infringement to
- a suitable form for making such a claim so that the problem item can be quickly identified
- a rapid response procedure so that a potentially offending item is taken down or access is blocked pending investigation of the complaint
- a clear appeal procedure for the complainant should his complaint be rejected
- sums of money held in reserve for payment of fees, legal advice, etc.

The Jorum notice and takedown policy is recommended as a model for this – see <http://www.jorum.ac.uk/policies/jorum-notice-and-takedown-policy>. Indeed, ALL Europeana partners should use this policy, or something closely based on it.

It is difficult to comment in detail about the Google Books arrangement without sight of the agreement between Ghent University and Google.

Again, it is necessary for Ghent University to provide a warranty to The European Library regarding copyright before The European Library should accept Ghent University materials.

### **11.4. Open data: Bayerische Staatsbibliothek**

The digital collections of the Bayerische Staatsbibliothek (BSB) in Europeana can be found at: [http://www.europeana.eu/portal/search.html?query=\\*&qf=PROVIDER:Bayerische+Staatsbibliothek](http://www.europeana.eu/portal/search.html?query=*&qf=PROVIDER:Bayerische+Staatsbibliothek)

The BSB collection in Europeana consists of more than 67,000 digital items from the Munich Digitisation Centre that are not subject to intellectual property rights (those subject to IPR have not been made available), and more than 398,000 digital items deriving from the Public Private



Partnership with Google. That includes works from 100 AD up to the 21st century: monographs, periodicals, music manuscripts, printed music, maps, atlases, manuscripts and incunabula.

The BSB has made its complete catalogue available as linked open data since November 2011 (<http://lod.b3kat.de>). Linked open data involves publishing structured data so that it can be interlinked and become more useful. It builds upon standard Web technologies, but rather than using them to serve web pages for human readers, it shares information in a way that can be read automatically by computers. This enables data from different sources to be connected and queried.

### **Commentary**

To comply fully with the concept and prerequisites of linked open data, it is necessary to provide these data without any restrictions under a completely free licence. The bibliographic records of Bayerische Staatsbibliothek (including the link to a thumbnail) are therefore published under a CC0 licence. CC0 waives all copyrights and related rights over the work, such as moral rights (to the extent waivable), publicity or privacy rights, rights protecting against unfair competition, and database rights and rights protecting the extraction, dissemination and reuse of data. It is by far the most generous of all the Creative Commons suite of licences and is only appropriate for materials where the owner is well aware of the implications of using it. Whilst it may well be appropriate for digitised versions of out of copyright works, it may not be appropriate for any indexing or metadata that is included. Europeana needs to be aware that if indexing and metadata is also made available by the Staatsbibliothek under CC0, then it can do nothing to stop third parties using that material for its own purposes, including making charges for access to it.

The vast majority of the items that have been digitised were out of copyright already. The library has taken a good approach to respecting and clearing copyright. Nonetheless, The European Library will require a warranty from it along the lines of that described earlier in this report to ensure that it takes on no risk when reproducing any of the Staatsbibliothek's contributions. It is worth stressing that for anything to be offered under a CC0 licence, the organisation making the offer must be 100% confident that either the material is out of copyright, or that it owns the copyright and all related rights in the material.

### **11.5. Conclusions from the case studies**

It is clear from these case studies that the participating libraries have understood and respected copyright issues and have avoided or greatly reduced any risk of possible infringement. Their efforts represent a role model for other Europeana partners. It remains important, nonetheless, for The European Library to obtain a warranty of some kind from each contributing library to ensure that The European Library itself does not have to carry any legal risk for reproducing full text materials. The European Library should also ensure it gets a royalty-free licence from each contributor for any metadata or indexes the contributor provides, and that it makes explicit that it will own the rights to any merged metadata or indexes.

The European Library will have to think carefully about how far it wishes to go down the route of using CC0 for linked open data, as promoted by the Staatsbibliothek. Data intended to be used for such purposes is best made available using a generous Creative Commons licence, such as CC0, CC BY, CC BY NC, or CC BY SA. The European Library needs to decide how it will make available the data it receives in the form of metadata or indexes, and under what terms and conditions.