



**Grant Agreement 297292**

## ***EUROPEANA INSIDE***

# **Legal, Licensing and Policy Constraints Inhibiting Organisations across Europe from Contributing their Previews and Metadata to Europeana**

<b>Deliverable number</b>	<i>Not Applicable</i>
<b>Dissemination level</b>	<i>Public</i>
<b>Delivery date</b>	<i>February 2012</i>
<b>Status</b>	<i>Final</i>
<b>Authors</b>	<i>Naomi Korn and Professor Charles Oppenheim (Naomi Korn Copyright Consultancy)</i>



This project is funded under the  
*ICT Policy Support Programme* part of the  
*Competitiveness and Innovation Framework Programme*.

**Statement of originality and licence:**

This report contains original unpublished work except where clearly indicated otherwise.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.

It is made available under a Creative Commons Attribution-ShareAlike licence (CC BY SA).

## Contents

1	CONTEXT .....	4
2	PROJECT METHODOLOGY .....	6
3	THE FINDINGS .....	9
4	RECOMMENDATIONS .....	16
4.1	For the Europeana Inside Consortium.....	16
4.2	For organisations wishing to participate in Europeana .....	17
	APPENDIX 1: NATIONAL LIBRARY OF LUXEMBOURG .....	18
	APPENDIX 2: WELLCOME TRUST LIBRARY .....	20
	APPENDIX 3: VISUAL ARTS DATA SERVICE.....	22
	APPENDIX 4: EUROPEANA INSIDE USE CASES.....	24

## 1 Context

Europeana provides a portal to Europe's rich cultural heritage, images and other content together with links to content as well as associated metadata that are provided to Europeana by cultural heritage institutions across Europe. In order to optimise

- The discoverability, access, use and enrichment of these collections and associated metadata;
- The interoperability of these collections, as well as
- The number of works and data which are available through Europeana through the release of previews (including thumbnails) as well as data (especially metadata) under Creative Commons Licences (including CC 0 for the metadata),

Europeana is dependent upon the ability of museums, libraries, archives, etc., to provide it (or portals feeding it) with previews and metadata under the terms of Europeana's Licensing Framework, including its Data Exchange Agreement

Many organisations across Europe have already submitted thumbnails of their content and metadata to Europeana. This has been supported by advocacy and tool creation by Europeana<sup>1</sup>, including the copyright calculator, a useful tool for calculating whether a work is in the public domain<sup>2</sup>, and the European Data Model (EDM) Selection Tool for helping organisations decide which rights statement might be appropriate for their previews<sup>3</sup>. Organisations such as the National Library of Luxembourg state that it is politically, technologically and culturally important to have visibility via the Europeana platform<sup>4</sup>. Other organisations, such as the Natural History Museum in London<sup>5</sup>, note that making its collections available to Europeana will help it achieve its open access agenda. A brief review of the internet has shown that there are many organisations who already subscribe to the principles of "open" by providing access to open data. These include the British Library<sup>6</sup>, Deutsche National Bibliothek, the Swedish National Library and Cambridge University Library and many others, which are listed in the JISC-funded Open Bibliographic Guide<sup>7</sup>. There are also a number of other initiatives which support the work of Europeana, such as the Discovery Project, funded by JISC and the Research Libraries UK, whose primary aim is to encourage museums, libraries and archives primarily across the UK to make their collections more discoverable by openly licensing their metadata<sup>8</sup>

However, it is apparent that there are still a number of technical and legal, licensing and policy related issues which are inhibiting many other institutions, particularly museums across Europe, from embracing open data licensing and exporting large amounts of previews and associated metadata to Europeana: those which are inhibiting Europeana from ingesting the previews and metadata, and others which restrict any enriched metadata from being fed back to the cultural heritage institutions.

*"Currently, cultural institutions are confronted with a choice between either not sharing their content or applying blanket licensing terms to whole collections which may not be appropriate to individual items (for example, higher value artefacts that are part of a wider collection).*

*The key outcome which EUROPEANA INSIDE is seeking to achieve is to put into the hands of the content providers the tools to manage IPR and licensing permissions at a granular (object, group or collection) level. This will significantly reduce the overall risk of infringement arising from participation in aggregated*

<sup>1</sup> <http://pro.europeana.eu/documents/904448/981580/Copyright+and+related+rights+framework+WP5+-+Patrick+Peiffer>

<sup>2</sup> <http://www.outofcopyright.eu>

<sup>3</sup> <http://pro.europeana.eu/web/quest/edm-rights-selection-tool>

<sup>4</sup> Via questionnaire - see below

<sup>5</sup> Via questionnaire – see below

<sup>6</sup> <http://www.bl.uk/bibliographic/datafree.html>

<sup>7</sup> <http://obd.jisc.ac.uk>

<sup>8</sup> [www.discovery.ac.uk](http://www.discovery.ac.uk)

*services such as Europeana, and will provide a more robust basis for the institution's own IP management.*

*The key metric for this outcome will be the quantity of material that is provided for aggregation into Europeana which is supplied with accurate, granular rights metadata."*

From the Europeana Inside project's *Description of Work*.

Europeana Inside, a pan-European project involving 27 partners, is developing a European Connection Kit to try and resolve these issues. An important component of this work is an inventory of the legal, licensing and policy constraints which restrict institutions from sharing their collections, links to their collections and associated metadata with Europeana, as well as an outline of the problems that they face by exporting previews including thumbnails and associated metadata under Europeana's licensing framework, including its Data Exchange Agreement<sup>9</sup>.

To address this work, Naomi Korn and Professor Charles Oppenheim, IP Consultants, were asked to provide:

- An inventory of legal, licensing and policy related constraints
- In certain areas, specify possible solutions, drawn from the Consultants' experiences, desk research and case studies.
- Draw reference to specific issues and how they have (or have not) been managed, taken from some of the 27 project partners, as well as existing clients and contacts of Ms Korn and Professor Oppenheim.

The research and case studies accompanying the report indicate that although there are numerous legal, licensing and policy related issues which can act as obstacles to the supply of content/metadata and/or links to Europeana, many of these can either be entirely addressed or their risks mitigated to acceptable levels, by the implementation of various rights management related policies and procedures at an organisational level. Whilst the report references specific use cases, the types of issues outlined are nonetheless not unusual for most cultural heritage organisations across Europe and point to a lack of suitable policies, procedures and tools/systems to understand, identify and manage rights and permissions. There are also severe problems with the legislative framework which present obstacles which are less easy to resolve.

---

<sup>9</sup> <http://pro.europeana.eu/web/guest/data-exchange-agreement>

## 2 Project methodology

### WP 1: Project Initiation (1/2 day)

This work package comprised an initial Skype meeting between Naomi Korn, Professor Charles Oppenheim, Gordon McKenna from Collections Trust and Jeanine Tielemans and Marco de Nijt from Digital Heritage Foundation in the Netherlands. The aim of this meeting was to: confirm the project's objectives and desired outcomes; discuss and confirm management and review arrangements; confirm reporting and presentation requirements (for the report), agree timeframes for WP 4 and WP 5, and agree a payment schedule.

A further priority of the kick-off meeting was to "scope the problem" in more detail in order to: understand priorities (e.g. breadth vs. depth, legal issues vs. guidance); agree the sample of project partners who might be most appropriate to be consulted and agree on the format of the guidance and how it will be delivered.

### WP 2: Information Gathering and Synthesis (4 days)

The aim of this work package was to gather and synthesise information from a wide variety of sources within a very short timescale, to form the basis of the content for the report and final recommendations (there is no deliverable for this WP). This work package focussed on the following activities:

- Desk research;
- Developing a questionnaire which was circulated to 9 organisations across Europe, including a number of project partners as well as contacts supplied by the Consultants. Among the organisations selected were those which could provide a broad heterogeneity of materials, those at which the Consultants and/or Europeana Inside team members had personal contacts to try and ensure a response within limited timescales, and those which represented organisations who had, would, could not submit metadata and previews to Europeana. These organisations and the contacts who were sent the questionnaire were:
  - Dutch National Archives, Noor Schreuder, legal advisor;
  - British Library, UK: Ben White, Head of Intellectual Property;
  - Imperial War Museums, UK, Debbie McDonnell, IP Manager;
  - National Library of Luxembourg, Patrick Peiffer, Coordinator consortial services at Bibliothèque nationale de Luxembourg;
  - Visual Arts Data Service, UK, Leigh Garrett, Director of Service;
  - Netherlands Institute for Sound and Vision, Mieke Lauwers, Policy Advisor, Project Manager;
  - British Film Institute, Richard Patterson, Head of Knowledge;
  - Natural History Museum, UK, Nancy Chillingworth, Copyright Officer;
  - Wellcome Trust Library, UK, Robert Kiley, Head of Wellcome Trust Library.
- Organisations who responded to the questionnaire either in whole or in part (as indicated), have a rich range of published (books, journals and newspapers) and unpublished text-based works, art and photographs, as well as audio visual works:
  - **Natural History Museum** who are planning to submit content and metadata (submitted in part and informally, therefore did not wish responses to be formally shared)
  - **Wellcome Trust Library** who have to date contributed modern biomedical images (over 120,000) and historical films (over 400) to Europeana via the Europeana Libraries project (full responses are provided in the Appendix)

- **British Film Institute** who are unable to supply content and metadata (due to limited time, only responded in part)
- **National Library of Luxembourg** who are supplying content and metadata to Europeana (full responses are provided in the Appendix)
- **Visual Arts Data Service** who had supplied thumbnails and metadata to Europeana, but have since withdrawn the previews and the metadata (full responses are provided in the Appendix).

The following organisations either did not respond, or cited the lack of time in order to respond, as indicated below:

- Dutch National Archives (LACK OF TIME);
- British Library, UK (LACK OF TIME);
- Imperial War Museums, UK (DID NOT RESPOND);
- Netherlands Institute for Sound and Vision (DID NOT RESPOND).

The questions that were asked were:

1. Briefly describe your organisation and the subject matter of the metadata and thumbnails which you have supplied or are intending to supply to Europeana.
2. How important to you is it that you contribute to Europeana?
3. Under what terms do you already provide access and/or use of this metadata and thumbnails on your own site and/or other sites and which ones?
4. What type of issues and/or obstacles have you had to address internally in order to provide metadata and thumbnails to Europeana?
5. How have you overcome these issues/obstacles?
6. Under what terms have you selected for access (and/or use) of your thumbnails on Europeana?
7. If you have only supplied certain metadata/thumb nail images, please describe the issues that you have encountered which have prevented you from supplying all your other metadata/thumb nail images?
8. What is your willingness to accept third party metadata into your systems?
9. What is your willingness to allow Europeana to translate your metadata and/or amend previews supplied by you to Europeana?
10. Are there any other issues that in any way affect your provision of thumbnails/metadata to Europeana, e.g., technical issues/standards, lack of resources, lack of necessary skills in your organisation?
11. Do you record copyright and licensing issues in your collection registration system? If yes, how? At collection level? At record level? If no, where else do you store this information?
12. Do you use different licensing models for contributing data to third party services like Europeana? If so, what are these licensing models?
13. Have you received any complaints or legal actions from third parties re. the publication or distribution of metadata/thumb nails/full content? If so, what follow-up actions did you take? Do you have a Notice and Take Down Policy and Procedure in place?
14. What would help you best to automate the process of providing licensing information about the objects in your collection?

**WP 3: Report and Recommendation Drafting (4 days)****WP 4: Project Review (1/2 day)**

This work package will comprise a Skype meeting with Naomi Korn, Professor Charles Oppenheim, the programme manager and any other project team members. The aim of this meeting will be to: discuss the report and recommendations, suggest enhancements as well as confirm timeframes for the final output.

**WP 5: Review and Enhancement (1 day)**

The aim of this work package is to revise the draft outputs and incorporate any changes agreed with the programme manager. The final output will be delivered according to the timeframe established in WP 1 and WP 4.

### 3 The Findings

Here the legal, licensing and policy constraints which might restrict institutions from sharing their collections (links, previews, metadata, etc.) with Europeana, as well as an outline of the problems that they face by exporting their metadata under Europeana's licensing framework, including its Data Exchange Agreement, are considered. The Report draws upon desk research, the Consultants' knowledge as well as some information collected from the case studies. Note that the National Library of Luxembourg and the Wellcome Trust Library answers are provided in Appendices 1 and 2, whilst those supplied by the British Film Institute and the Natural History Museum are not supplied following requests from these organisations not to quote their responses verbatim.

In addition, the three use cases taken from Europeana Inside Use Cases (Deliverable D2.2 from DEN, published September 2012) are briefly described and mapped against the types of issues described below<sup>10</sup>:

**Use Case 1** is a basic automated transformation and supply scenario, and describes the situation for a content supplier that can comply to a widely-used set of minimum standards and protocols for the supply of metadata and other data relevant to its collections.

**Use Case 2** describes a situation where richer data than the minimum is on offer, and the contributor has options to configure the workflow to suit its requirements.

**Use Case 3** refers to the situation where a cultural heritage institution that has delivered data to Europeana wants to re-ingest or re-use the enriched data Europeana offers back in to its own system. This enriched data will comprise both the data originally offered to Europeana and enhancements made by Europeana, possibly incorporating third party data.

The legal, licensing and policy related constraints can be broken down into three broad areas:

- A. The Restrictive Legislative Framework;
- B. Constraints associated with the objects themselves (i.e., their content and associated metadata);
- C. Constraints associated with the organisations that hold the materials.

Each of these is considered further below.

#### A. The Restrictive Legislative Framework

The legal environment in which an organisation in Europe operates can severely impede an organisation's ability to provide metadata and content to Europeana. There is no international copyright law, so whilst Europeana is pan-European, and despite the efforts of the European Commission to achieve harmonisation, copyright laws are not harmonised across Europe. Although successive copyright-related Directives, and requirements to sign up to various conventions and treaties have established a reasonable degree of harmonisation on ownership, lifetime and the exclusive rights (so-called "restricted acts") of rights owners in the field of copyright and database rights, there is virtually no harmonisation of exceptions and limitations to copyright amongst member states. The laws relating to performers rights and data protection are harmonised within the EU, but the laws relating to Moral Rights, whether contracts can over-ride exceptions to copyright, and defamation, etc., are not. Furthermore, even where there is notionally harmonisation, national Courts may choose to interpret the wording of the legislation in different ways. Finally, the various member states have different laws and traditions on topics such as freedom to enter into a contract, the powers and regulation of reproduction rights organisations, and what constitutes an unfair contractual term. As a result, organisations in different member states will be restricted by their own legislation or previous Court cases on topics as wide ranging as what might be considered as public domain, an individual's or organisation's ability to waive copyright, the types of licence

---

<sup>10</sup> See Appendix XXX for a full description of the use cases taken from the deliverable D2.2.

agreements they can enter into, and copyright limitations and exceptions. Arguably, the recently passed Orphan Works Directive creates more difficulties for organisations wishing to supply Orphan Works to Europeana<sup>11</sup>, whilst the extension of the duration of copyright in sound recordings and performers' rights recently passed in another EC Directive 2011/77/EU<sup>12</sup> will harm further the ability of organisations across Europe to share audio visual works with Europeana.

Patrick Peiffer from the National Library of Luxembourg notes that: "The instruments we have chosen and others such as the MOU on out-of-commerce works, orphan works, ECL<sup>13</sup>, etc. all require making up your tools as you go along. The basic problem is that current copyright law does not offer a satisfactory framework which is in line with our legal mission to preserve and give access to cultural heritage."

In terms of linking, the **1709 BlogSpot** has just reported on some ambiguities relating to linking which potentially could create additional difficulties in the supply of links to Europeana:

"The IPO has reported that a Swedish court, the Svea hovrätt, has made a request for a preliminary ruling from the European Court of Justice (CJEU) regarding the infringement of exclusive rights to make copyright protected work publicly available by a third party subscription search engine. The case is *Nils Svensson, Sten Sjögren, Madelaine Sahlman, Pia Gadd v Retreiver Sverige AB* (Case C-466/12)<sup>14</sup>.

The following questions have been referred:

1. If anyone other than the holder of copyright in a certain work supplies a clickable link to the work on his website, **does that constitute communication to the public** within the meaning of Article 3(1) of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society?
2. Is the assessment under question 1 affected if the work to which the link refers is on a website on the Internet which can be accessed by anyone **without restrictions or if access is restricted** in some way?
3. When making the assessment under question 1, should any distinction be drawn between a case where the work, after the user has clicked on the link, is shown on another website and one where the work, after the user has clicked on the link, is **shown in such a way as to give the impression that it is appearing on the same website**?
4. Is it possible for a Member State to give **wider protection** to authors' exclusive right by enabling "communication to the public" to cover a greater range of acts than provided for in Article 3(1) of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society?

In short, the CJEU is being asked:

1. Does linking constitute communication to the public?
2. Does it matter if the website being linked to imposes restrictions on access on its users? (It is not clear from the questions what these restrictions on access are in this case, however they are likely to be the website's T&Cs or paywalls requiring subscription.)
3. Should there be a legal distinction between linking and framing?

<sup>11</sup> <http://www.communia-association.org/policy-papers-2/>

<sup>12</sup> [http://ec.europa.eu/internal\\_market/copyright/term-protection/index\\_en.htm](http://ec.europa.eu/internal_market/copyright/term-protection/index_en.htm)

<sup>13</sup> Extended Collective Licences

<sup>14</sup>

<http://curia.europa.eu/juris/document/document.jsf;jsessionid=9ea7d0f130d591830c9cf284bbfb73a8b46beb3f469.e34KaxiLc3eQc40LaxqMbN4Oa3uRe0?text=&docid=130286&pageIndex=0&doctlang=EN&mode=lst&dir=&occ=first&part=1&cid=85859>

4. Can a Member State expand the meaning of "communication to the public" to cover more than is set out at Article 3(1) of the InfoSoc Directive?

Despite the continual raising of these issues to European-related copyright consultations by many organisations across Europe, these issues remain for the most part, unresolved, and are typical of the types of issues that are referred to the CJEU with increasing frequency. Depending upon the outcome of the case which has been referred to the European Court of Justice, if links are found to be "communication to the public", this may well have ramifications for organisations in member states across Europe that supply links to works in copyright without seeking permission from the rights owners, as well as for Europeana in hosting them.

## B. Object Level Issues

Individual objects, their digital surrogates, collections of objects and associated metadata, within the same organisation will have various legal issues associated with them. These are likely to present potential obstacles to the ease with which thumbnails<sup>15</sup>, metadata and even links, can be supplied under the same terms to Europeana by cultural heritage institutions and then accessed, re-used, enriched and potentially ingested again. Furthermore, these potential legal issues may preclude all or some of the use cases outlined above. The Natural History Museum has noted in response to the case study questions that the extent and complexity of such issues have encouraged it to carry out its own analysis of the issues.

It should be noted that all organisations contributing to Europeana are based in member states of the EU, and are therefore subject to the copyright laws of the particular member state they are in. All EU member states are signatories to the Berne Copyright Convention. Article 2 (1) of that Convention states that "artistic works" includes (and is not limited to) "works of drawing, painting, architecture, sculpture, engraving and lithography, photographic works to which are assimilated works expressed by a process analogous to photography, works of applied art, illustrations, maps, plans, sketches...."

Nothing in this definition places any requirement on the size or quality of such works, and indeed, any attempt to reduce the scope of the definition below that set out by the Berne Convention would put the country in breach of the Berne Convention.

It is clear from this, then, that thumbnails of works held by a Europeana partner enjoy copyright status as artistic works, whether or not the original item is available through the Europeana services or not.

We now consider the case of five different member states, details of the relevant clauses in their respective Copyright Acts can be found in the footnotes – UK<sup>16</sup>, France<sup>17</sup>, Germany<sup>18</sup>, Netherlands<sup>19</sup> and Luxembourg<sup>20</sup>. We ran a search for "Warhol" in Europeana and filtered the results by country of origin of the hits; the following results from the different countries were obtained:

- UK = 6
- France = 8
- Germany = 48
- Netherlands = 16
- Luxembourg = 0

<sup>15</sup> There are two potential legal issues associated with thumbnails, or previews. The first, which has been the subject of many legal cases around the world, is whether a thumbnail of an artistic work can infringe the copyright of the original work. Different Courts in different countries have provided conflicting answers to that question. The second question is whether a thumbnail is itself a copyright work. There have been no court cases anywhere on this question, because the question is clearly answered in all countries which are signatories to the Berne Convention – a thumbnail is always an artistic work that enjoys copyright as long as it has not been directly copied from another thumbnail.

<sup>16</sup> Article 1(1)(a) of the Copyright Designs and Patents Act 1988

<sup>17</sup> L112-2 of the French Copyright Act

<sup>18</sup> Article 2 (1) of Urheberrechtsgesetz, UrhG (2008)

<sup>19</sup> Article 10 of the 1912 Copyright Act, revised in 2006

<sup>20</sup> Clause 1 (1) of the Law of 2001

Other than in the case of France, thumbnails of paintings by Warhol, or thumbnails of photographs of Warhol were provided. However, there is no implication that thumbnails are not used by French or Luxembourgish organisations in Europeana, as a simple search for either “France” or “Luxembourg” showed numerous thumbnails being provided. In conclusion: in the five countries studied, thumbnails of images have been provided by institutions in those countries to Europeana; and each of the thumbnails is a copyright work in the local law, as dictated by the Berne Convention.

Data providers have the choice of providing these thumbnails or previews, defined in the Europeana Data Exchange Agreement as “A reduced size or length audio and/or visual representation of Content, in the form of one or more images, text files, audio files and/or moving image files”, and which therefore include thumbnails, as well as metadata relating to their collections. Slightly different rules apply to these; for example, the metadata must be provided under a CC0 waiver<sup>21</sup> whilst the Digital Exchange Agreement requires data providers to indicate the rights status of the digital object and this rights statement also applies to the preview (there are 12 different valid rights statements that a data provider can choose from).

The legal and licensing issues are likely to include:

*B.1. Copyright and other types of Intellectual Property Rights (IPR), such as database rights and performers' rights as well as trade marks (registered and unregistered) may apply. It is very unlikely that confidential information and registered designs will be applicable.*

For example: third party rights, such as copyright and other IPR associated with the objects (if these rights still subsist), may preclude previews being made available to Europeana. Even if any licence terms imposed at the time the item was donated, purchased, borrowed<sup>22</sup> or retrospectively, can be negotiated with right holders, such terms and conditions may apply for various time scales and with various conditions (e.g., in perpetuity, for the lifetime of the copyright in the work, or for some other time period). This issue is particularly acute for museums, who often need to supply previews of their content together with the metadata to Europeana for the metadata to be meaningful. This may, in some way explain why it is easier for libraries and archives to participate in Europeana, as opposed to museums, as bibliographic records can accurately describe a specific publication, whilst metadata about an object requires a preview of the object to take into account numerous objects of the same date and/or description etc.

If permissions are restricted in duration or indeed territoriality, as noted by the British Film Institute, this would conflict with the ability of the host organisation to provide access to the image and associated metadata under any open licence (such as the standard Creative Commons suite of licences as well as CC0) which grants permission in perpetuity. The British Film Institute also notes in its response to the questionnaire, that the resources and time required to seek permission to put much of its audio visual content online acts as a major obstacle (particularly as much of this content includes layers of rights and multiple rights holders), and even if such an obstacle could be overcome, the territorial restrictions which are likely to be imposed on it by rights holders would prevent its ability to place such content online.

For example, the collection may include works and/or metadata in which the rights holders are either unknown or cannot be traced (so called “Orphan Works”). As is well known, digitisation of Orphan Works in the absence of permission from the rights owner carries with it a certain level of risk. Although the European Union's Orphan Works Directive, which will significantly reduce the level of risk involved, was recently passed, it will probably not be implemented in member states until late 2014 and in any case, does not cover all works likely to be orphan works (such as unpublished text based works and photographs - as well as other art works not embedded within published text based works), or indeed to all

---

<sup>21</sup> Note that the CC Zero Waiver acts as licence in certain jurisdictions, such as the UK

<sup>22</sup> This may apply in Use Case 1 and Use Case 2

organisations that hold Orphan Works<sup>23</sup>. Moreover, the due diligence obligations outlined in the Directive may place severe resource pressures on organisations wishing to submit orphan works to Europeana, some of whom may as a result choose not to submit these works.

#### *B.2. Data protection and other privacy issues, and the laws relating to defamation, official secrets, terrorism, and/or incitement to racial hatred.*

There may well be other legal issues involved if the objects are offered for digitisation and dissemination via Europeana, which can relate to the nature of objects, e.g., if they include identifiable images or information about identifiable living individuals, or information or images which may be legally problematic in their own right.

Contractual and IPR issues may apply to the associated metadata, in particular where metadata has been derived from other sources, and terms and conditions are associated with its reuse<sup>24</sup>.

For example 1: the Natural History Museum notes that it has to cross-reference details of its existing funding agreements with organisations, that have or will be funding its digitisation activities in order to ensure that they allow it to make the resulting digitised images and metadata available to Europeana.

For example 2: the contractual conditions under which the objects may have been donated to or lent to the organisation may potentially preclude copies of these objects being made available to Europeana by the owners of these works.

For example 3: Government bodies who fund digitisation activities may require that any metadata is made openly available.

For example 4: metadata about objects may have been supplied to the organisation under specific terms and conditions, which could preclude sharing of this data with Europeana. Likely situations include the supply of “shelf ready” library bibliographic data by a commercial supplier under restrictive terms which prevent further dissemination.

#### *B.3. Metadata which is the digital object.*

Some organisations, particularly those which handle scientific data, have mentioned informally to Naomi Korn (at the Europeana meeting in Poland in 2012), that they struggle to provide metadata under a CC0 licence in cases where the metadata is the digital object - i.e., descriptive information about plant names. In these cases, there is a greater likelihood that this type of information may be protected by copyright.

### **C. Organisational Issues**

This is the most important area of constraint. Policy issues are likely to include:

- a. The organisation’s appetite for risk. This will affect its willingness to digitise and authorise dissemination of digital versions of its collection works, including its Orphan Works, and to take due regard (or not) of creators’ Moral Rights. In addition, the organisation’s interpretation of its national copyright laws (in particular, the applicability of exceptions to copyright to the digitisation projects it is undertaking) and other laws of relevance such as defamation, and the subsequent provision of access to Europeana will be affected by its appetite for risk, both financial and reputational. The National Library of Luxembourg notes that “We have used a “risk management” approach for newspapers which we put fully on the internet up to 1950 without article level research of author rights clearance. This was put in place in 2007 as a test after we became convinced it was the most pragmatic course of action after many presentations and meetings with relevant stakeholders. Experience so far has proven us right, as we have received zero complaints, but many compliments, including from authors. Note: We do have contracts with publishers.”

---

<sup>23</sup> This could apply in Use Case 1, Use Case 2 and Use Case 3

<sup>24</sup> This will probably apply in Use Case 1 and Use Case 2; it will particularly apply in Use Case 3

- b. The organisation's willingness to accept third party metadata into its systems. It may well decide that the quality and/or provenance of what is on offer are not sufficient to satisfy its legal, technological or quality concerns<sup>25</sup>.
- c. The organisation's willingness to allow Europeana to translate metadata and/or amend Previews supplied by it to Europeana.
- d. The organisation's willingness to waive all IPR in its metadata, as required by Clause 3.2 of the Europeana Data Exchange Agreement.
- e. The organisation's willingness and ability to ensure that all IPRs in any third party metadata it plans to submit to Europeana have been waived or licensed appropriately by that third party.
- f. The organisation's commitment and ability to precisely follow the Europeana Data Model and Europeana Rights Statements requirements when creating and submitting material to Europeana.
- g. Business policies and practices, such as the need to generate income from the digitisation programme or from individual digitised items, which may restrict the organisation's willingness and ability to provide access to previews and metadata under the Europeana Licensing Framework. This also includes the willingness of the organisation to change its policy objectives away from increasing income from the licensing of rights to materials held by it, to spill-over beneficial commercial and non-commercial effects resulting from the higher public profile associated with a large amount of the organisation's material being available under a CC licence. It also includes the importance of political and economic imperatives on organisations across the cultural heritage sector to invest resources into generating income, and therefore de-prioritising their efforts in supplying content, metadata and links to Europeana, in favour of income generation activities. The Consultants have been told about this problem from one UK art-based gallery, and we believe it is a widespread issue.
- h. Pressure the organisation may be under from its funders to participate in Europeana and/or to offer its material under Creative Commons licences.
- i. The organisation's willingness to accept the risk that there may be loss of attribution or reputation associated with opening up of information about the objects held by the organisation. It is interesting in this regard that VADS found traffic to its own web site increased significantly after it withdrew the materials it had originally offered to Europeana.
- j. The organisation's willingness and ability to commit to providing an accurate rights label with every item it submits to Europeana.
- k. The organisation's understanding of, and acceptance of, the potential benefits of the use of open metadata associated with its collections, as summarised in *The Problem of the Yellow Milkmaid* White Paper (November 2011).
- l. Licensing practices (or the lack thereof), which may mean that the organisation, whether it realises it or not, has not secured the necessary permissions for previews to then be shared with Europeana in accordance with Europeana's Licensing Framework. This may be due to a lack of experience or confidence in licence negotiations.
- m. The organisation's policy and practice on requiring attribution and/or payment in association with the use of its metadata.<sup>26</sup>
- n. The organisation's policy and practice on the use of rights management information, digital watermarking and/or other technical protection measures (TPMs) in association with its digital content and/or metadata. The use of TPMs may well conflict with a Creative Commons licence that Europeana expects to offer in association with digitised content, e.g., metadata, associated with it.<sup>27</sup>
- o. The quality of the organisation's recording and management of rights and permissions, including the use of databases to record granular level rights and permissions, the reliability of this information as well as the relevance, level of detail and the currency of what the organisation has been recording.

---

<sup>25</sup> This will apply to Use Case 3.

<sup>26</sup> This will apply to all the Use Cases.

- p. The quality and consistency of the metadata that the organisation has associated with the digitised objects. The way the metadata is recorded may or may not be in accord with agreed national and international standards, or as required by Europeana.<sup>28</sup>
- q. The quality, completeness and currency of record-keeping by the institution tracking where data and/or content have come from, where it has been sent and the timings of all actions associated with the content and metadata.<sup>29</sup>
- r. The speed and manner of response by the organisation to any request for takedown of materials or metadata because they break a law or infringe an IPR<sup>30</sup>.
- s. The willingness to use systems, which automatically embed a CC licence and credit line into a resource.<sup>31</sup>
- t. The willingness to evaluate, and the consequent speed of decision-making, as to whether data of all kinds provided to Europeana is to be marked as Public Domain (using <http://www.outofcopyright.eu> to decide), Creative Commons licensed (and if so, which CC licence is to be applied), Rights Reserved (and if so, which Rights Reserved notice – free access, paid access or restricted access), or as “status unknown” (mainly Orphan Works).
- u. The sophistication of internal rights management systems to export and then ingest metadata.
- v. The organisation’s requirement to demonstrate impact to its funders, and in particular the need to demonstrate increasing visitor numbers to its own website, which could be reduced by providing previews and metadata to Europeana.

---

<sup>28</sup> These could all apply in Use Case 1

<sup>29</sup> This will apply to all Use Cases

<sup>30</sup> In the Europeana Licensing Framework, Europeana has five working days to either reject a takedown claim, or to remove the metadata and/or previews, or to request further clarification from the person making a complaint about an alleged infringement.

<sup>31</sup> The leading service is Attribute Images, developed by Nottingham University with JISC funding (<http://www.jisc.ac.uk/whatwedo/programmes/ukoer3/rapidinnovation/attributeimages.aspx>) There appears to be no reason why this service could not be used with other types of content as well.

## 4 Recommendations

### Recommendations and Concluding Remarks

It is clear that each organisation that is, or is planning to contribute previews and/or links and/or associated metadata to Europeana faces a different set of challenges to the successful completion of its task. Some of the challenges will be major, and some minor. Whilst the issues raised apply to all three Use Cases mentioned above, Use Case 3 is likely to raise more issues than the first two, and so it is likely that organisations contributing to Europeana will be less likely to follow Use Case 3 if they do not have the confidence that all the legal, policy and licensing issues associated with individual objects and the organisations themselves as described above have been fully addressed.

Confidence to deal with some of the issues will be within the organisation's control, but others (e.g., any terms and conditions imposed on the organisation by a third party donating, selling or lending objects to the organisation, and the legal environment in which the organisation operates) are largely outside the organisation's control. These constraints can include:

- The quality of the data
- Data which is confidential or commercially sensitive
- Format of the data
- Data which is academically sensitive
- Data which is medically or ethically problematic to share

Important as these constraints are, arguably the organisation's awareness and understanding of its constraints is even more important. An organisation that recognises the constraints that it operates under can develop policies and practices to minimise the negative effects of such constraints, and can pro-actively exploit opportunities that it perceives. A potentially serious situation can, however, arise when the organisation is not aware of the restraints it should be operating under, and then makes serious errors of judgement in terms of the offerings it sends, or chooses not to send, to Europeana.

So, whilst certain data might automatically be excluded because the issues as indicated above may be insurmountable, in other instances, once issues have been identified, it is possible to manage many of the other types of arising issues and either address risks entirely or mitigate risks to acceptable levels in many circumstances, by the implementation of various rights management related policies and procedures at an organisational level.

#### 4.1 For the Europeana Inside Consortium

1. The Consortium can play an important part in making sure that individual organisations can and will participate in Europeana by developing a checklist for them stating the requirements that they have to meet concerning their legal requirements and rights management responsibilities.
2. The Consortium can help raise awareness about the value of open metadata. A recently developed open data animation<sup>32</sup> and briefing paper on licensing open data<sup>33</sup> are some of the tools which can help instigate cultural change. This includes the future-proofing of the creation of open metadata by crowd-sourcing through the provision of advice, guidance and templates.
3. The Consortium can help by referring individual organisations to existing advocacy materials and tools created by Europeana<sup>34</sup>, including the copyright calculator – a useful tool for calculating whether a work is in the public domain<sup>35</sup> and the EDM Selection Tool for helping organisations decide which rights statement might be appropriate for their previews<sup>36</sup>.

<sup>32</sup> [http://www.youtube.com/watch?v=Twp5LK\\_Wko](http://www.youtube.com/watch?v=Twp5LK_Wko)

<sup>33</sup> [http://discovery.ac.uk/files/pdf/Licensing\\_Open\\_Data\\_A\\_Practical\\_Guide.pdf](http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf)

<sup>34</sup> <http://pro.europeana.eu/documents/904448/981580/Copyright+and+rights+framework+WP5++Patrick+Peiffer>

<sup>35</sup> <http://www.outofcopyright.eu>

<sup>36</sup> <http://pro.europeana.eu/web/quest/edm-rights-selection-tool>

4. The Consortium can engage more closely with funders of digitisation projects across Europe to insist on the creation and release of open metadata about the digitised projects as part of the terms and conditions of funding
5. Moving forward, even if an organisation, like the National Library of Luxembourg, implements the necessary rights and risk management procedures, organisations across Europe will still be restricted from fully embracing Europeana in their provision of previews, metadata and even links, by several legal issues. In this way, the Consortium can help try and reframe the legislative framework across Europe by undertaking the following:
  - a. Active lobbying for clear, harmonised, media and technology neutral copyright laws across Europe which are consistent with the use and opportunities offered by digital technology.
  - b. Lobbying to ensure that exceptions to copyright are fully harmonised within the EU, and that the level of harmonisation should be such to reduce the current uncertainty to a minimum<sup>37</sup>.

## **4.2 For organisations wishing to participate in Europeana**

1. Negotiate as broad permissions as possible with rights holders in order to deal with third party rights and other legal issues which might arise. There are a number of sources of template licences, which have recently been developed in the UK as well as other European funded projects<sup>38</sup>. This will assist in dealing with any third party rights issues associated with copyright in previews.
2. Understand, identify and manage consent management issues associated with Data Protection and other privacy issues<sup>39</sup> associated with images as well as metadata again using standard terms and conditions where necessary.
3. Clarify and advance the understanding of what an organisation is entitled to do with the bibliographic records and other data supplied by third parties<sup>40</sup>
4. Conduct a compliance audit using standard risk management templates, such as that provided by Web2Rights<sup>41</sup>, which is a systematic process of identifying, analysing and responding to risk.
5. Once risks have been identified, developing an organisational position on risk, as well as appropriate risk mitigation strategies, including a robust notice and take down policy and procedures. All of this implies the development, and regular updating of, a risk register, to be signed off by senior management at least annually.
6. Deploy a rights management system, which at the very least, can record rights, permissions and other legal issues at an individual object level (e.g., title and issue level for newspapers, articles, etc) as well as at a collection level. Both the National Library of Luxembourg and Wellcome Trust Library deploy such systems. Collections Trust provides a SPECTRUM-compliant collections management systems vendor comparison tool which can assist organisations to select the most appropriate system<sup>42</sup>. Currently 36% of objects supplied by organisations to Europeana lack rights information, the provision of which will be important as Europeana develops its licensing framework to provide additional options for recording the rights status of the digital objects available via Europeana such as whether they are Orphan Works etc<sup>43</sup>.

<sup>37</sup> <http://www.ip-watch.org/2012/12/06/european-commission-embarks-on-process-to-modernise-copyright>

<sup>38</sup> [www.web2rights.com/OERIPRSsupport](http://www.web2rights.com/OERIPRSsupport)

<sup>39</sup> <http://www.jisclegal.ac.uk/Projects/ConsentManagement.aspx>

<sup>40</sup> <http://www.jisclegal.ac.uk/Projects/TransferandUseofBibliographicRecords.aspx>

<sup>41</sup> <http://www.web2rights.com/OERIPRSsupport/risk-management-calculator>

<sup>42</sup> <http://www.collectionslink.org.uk/softwaresurvey>

<sup>43</sup> Information provided by Paul Keller.

## APPENDIX 1: National Library of Luxembourg

1. Briefly describe your organisation and the subject matter of the metadata and thumbnails which you have or intending to supply to Europeana?

*National Library of Luxembourg is a small national library with ca. 1.5 million holdings ([www.bnl.lu](http://www.bnl.lu)) and hosts the biggest digital library via the scientific “Consortium Luxembourg” ([www.findit.lu](http://www.findit.lu)). We have and will continue to supply to Europeana all digitised material from [www.eluxemburgensis.lu](http://www.eluxemburgensis.lu), metadata and thumbnails. Digitisation efforts are focussed on the press (dailies, weeklies, magazines).*

2. How important to you is it that you contribute to Europeana?

*It is politically (Europeana cultural policy), technically (Europeana APIs and Europeana Network) and culturally (No national heritage is limited to national boundaries or languages) important to be visible via the Europeana platform.*

3. Under what terms do you already provide access and/or use of this metadata and thumbnails on your own site and/or other sites and which ones?

*Our Europeana Metadata is under CC0 as per Europeana DEA. Previews and Content are either marked as Public Domain using the Creative Commons Public Domain Mark, or “Free access – Rights Reserved”. The local pages on [www.eluxemburgensis.lu](http://www.eluxemburgensis.lu) are in the process of being updated with these conditions, as we have submitted a large amount of content only very recently to Europeana. The “Free access – Rights Reserved” stems from some rights being cleared with publishers and others awaiting clearance and/or diligent search. We are as transparent as possible with all involved stakeholders regarding rights clearance which remains a cumbersome process, if not otherwise possible we follow a managed risk path.*

4. What type of issues and/or obstacles have you had to address internally in order to provide metadata and thumbnails to Europeana?

*Only issue is to provide the correct XML format, as there are much improved guidelines and an online checker tool now, there are no real obstacles anymore.*

5. How have you overcome these issues/obstacles?

[n/a]

6. Under what terms have you selected for access (and/or use) of your thumbnails on Europeana?

*This question does not make much sense. The DEA (art 4.1) stipulates that if thumbnails (“Previews” in DEA terms) are provided to Europeana, these must be accessible on Europeana and will be re-usable according to same conditions as indicated in edm:rights for Content. The only option would be not to provide thumbnails. We do provide Previews for all Content.*

7. If you have only supplied certain metadata/thumb nail images, please describe the issues that have encountered which has prevented you from supplying all your other metadata/thumb nail images?

[n/a]

8. What is your willingness to accept third party metadata into your systems?

*Must be compatible with the DEA, i.e. ok to re-use under CC0.*

9. What is your willingness to allow Europeana to translate your metadata and/or amend Previews supplied by you to Europeana?

*Fine with us, we see it as a free service.*

**10. Are there any other issues that in any way affect your provision of thumbnails/metadata to Europeana, e.g., technical issues/standards, lack of resources, lack of necessary skills in your organisation**

No

**11. Do you record copyright and licensing issues in your collection registration system? If yes, how? At collection level? At record level? If no, where else do you store this information?**

*We have deployed a rights management system at title and issue level for newspapers, article level is analysed but not deployed yet. This will allow us, as rights clearance progresses for a newspaper title (up to 2007) to limit access to title/issue/article to our own building. This does include the iPad app, which only gives access to such content on our local Wifi (This may seem bizarre, but has worked fine in tests). The information is stored in the Database that is used by eluxemburgensia (Software: Digitool from Exlibris with custom viewer).*

**12. Do you use different licensing models for contributing data to third party services like Europeana? If so, what are these licensing models?**

*Different to what? To our own local rules? In that case the answer is no, we use the same licensing rules, as the scope of our clearance is always the internet and Europeana is part of the internet (unless we have to limit to local access, in which case it isn't)*

**13. Have you received any complaints or legal actions from third parties re. the publication or distribution of metadata/thumbnails/full content? If so, what follow-up actions did you take? Do you have a Notice and Take Down Policy and Procedure in place**

*We have used a "risk management" approach for newspapers which we put fully on the internet up to 1950 without article level research of author rights clearance. This was put in place in 2007 as a test after we became convinced it was the most pragmatic course of action after many presentations and meetings with relevant stakeholders. Experience so far has proven us right, as we have received zero complaints, but many compliments, including from authors. Note: We do have contracts with publishers.*

*We do have a notice and takedown policy but it has never been used so far.*

**14. What would help you best to automate the process of providing licensing information about the objects in your collection?**

*Providing licensing information would seem to imply that detailed re-use and access rules are defined, are grounded in documented processes and supported by adequate legislation. The basis for this is simply not there as European copyright law is in need of reform to re-align itself with digital technology. Thus, any streamlined process to arrive at the licensing information you wish to automate, seems an illusion for now. We proceed step by step, making sure we do not end up in cul-de-sacs through consultation with stakeholders, but must keep going in a direction for which there are few way signs. The instruments we have chosen and others such as the MOU on out-of-commerce works, orphan works, ECL, etc. all require making up your tools as you go along. The basic problem is that current copyright law does not offer a satisfactory framework which is in line with our legal mission to preserve and give access to cultural heritage.*

## APPENDIX 2: Wellcome Trust Library

1. Briefly describe your organisation and the subject matter of the metadata and thumbnails which you have or intending to supply to Europeana?

*Wellcome Library. We have contributed modern biomedical images (over 120,000) and historical films (over 400) to Europeana via the Europeana Libraries project.*

2. How important to you is it that you contribute to Europeana?

[unanswered]

3. Under what terms do you already provide access and/or use of this metadata and thumbnails on your own site and/or other sites and which ones?

*Images are from Wellcome Images which has information about e.g. copyright at this page: <http://images.wellcome.ac.uk/indexplus/page/How+do+I+%3F.html>*

*Films are from Wellcome Film. Information about terms is at this page: <http://library.wellcome.ac.uk/about-us/about-the-collections/moving-image-and-sound-collection>*

4. What type of issues and/or obstacles have you had to address internally in order to provide metadata and thumbnails to Europeana?

[unanswered]

5. How have you overcome these issues/obstacles?

*No issues or obstacles once curators understood that their metadata and thumbnails were available on the web via our web sites.*

6. Under what terms have you selected for access (and/or use) of your thumbnails on Europeana?

[unanswered]

7. If you have only supplied certain metadata/thumb nail images, please describe the issues that have encountered which has prevented you from supplying all your other metadata/thumb nail images?

[unanswered]

8. What is your willingness to accept third party metadata into your systems?

*For the Wellcome Library's digitisation project (not Europeana) we are bringing third party metadata into our LMS.*

9. What is your willingness to allow Europeana to translate your metadata and/or amend Previews supplied by you to Europeana?

[unanswered]

10. Are there any other issues that in any way affect your provision of thumbnails/metadata to Europeana, e.g., technical issues/standards, lack of resources, lack of necessary skills in your organisation?

*Resource would to be found beyond the Europeana Libraries project. Any updates/amendments to records are being handed over to our Collection Management department (who deal with our metadata).*

11. Do you record copyright and licensing issues in your collection registration system? If yes, how? At collection level? At record level? If no, where else do you store this information?

*Yes – at item level for both Wellcome Film and Wellcome Images.*

**12. Do you use different licensing models for contributing data to third party services like Europeana? If so, what are these licensing models?**

No

**13. Have you received any complaints or legal actions from third parties re. the publication or distribution of metadata/thumbnails/full content? If so, what follow-up actions did you take? Do you have a Notice and Take Down Policy and Procedure in place?**

*No, not for Europeana Libraries material.*

**14. What would help you best to automate the process of providing licensing information about the objects in your collection?**

[unanswered]

## **APPENDIX 3: Visual Arts Data Service**

### **1. Briefly describe your organisation, how you are funded and your types of collections?**

*VADS is a research centre of the Library at the University for the Creative Arts, we focus on looking at the use of digital assets to support learning, teaching, research and knowledge transfer activities in the visual arts higher education sector and beyond. Historically, we hold the national digital image repository for the sector - this consists of 120,000 items drawn from over 300 collections from the arts, culture and heritage sectors. We were funded as a public service by the AHRC and JISC until four years ago, since which we have been constituted as a research centre and funding is largely project based, both from external and internal sources, this is complemented by knowledge transfer activities and funding for deposit and preservation services.*

### **2. How important to you is it that you contribute to Europeana?**

*Until August 2012 we had 120,000 items harvest by OAI with Europeana, however with the advent of CC0 licensing requirements these records were withdrawn - we hold an educational licence and therefore do not have the appropriate licence to share core metadata required by Europeana. In addition, there are issues around non-attribution which presents higher education with particular difficulties with regards academic integrity; and the commercial and contemporary nature of the materials we handle present problems and would discourage collection holders submitting works within copyright to us on an educational basis. In addition, since the withdrawal from Europeana our website analytics have indicated a significant increase in traffic to our own site - given that our funders, host and future revenue model require evidence of impact - it is unlikely that we would submit records to Europeana again.*

### **3. Under what terms do you already provide access and/or use of your metadata and thumbnails on your own site and/or other sites and which ones?**

*This was on a non-commercial, attribution basis, but records have been removed following the requirement to provide data on a CC0 licence.*

### **4. Why are you not supplying metadata and thumbnails to Europeana?**

*Described above.*

### **5. What type of issues and/or obstacles have arisen with regards to your supply of metadata and thumbnails to Europeana?**

*Described above.*

### **6. Are there any other issues that in any way affect your provision of thumbnails/metadata to Europeana, e.g., technical issues/standards, lack of resources, lack of necessary skills in your organisation?**

*There are resourcing issues - in terms of negotiation with collection holders of current licence to enable CC0. Also I feel more work needs to be undertaken in the field to understand the impact of CC0 licensing within the education sector. I think there are also issues around commercial usage - commercial research is non-taxable and therefore subsided by the tax payer in the same way public research is, in addition I would question the right of companies to freely use public assets without attribution when also avoiding paying corporation tax.*

### **7. Do you record copyright and licensing issues in your collection registration system? If yes, how? At collection level? At record level? If no, where else do you store this information?**

*Licence agreements are stored and some variations have been permitted recently. A standard non-exclusive, educational and promotional licence is taken by VADS.*

- 8. Do you use different licensing models for contributing data to third party services like Europeana? If so, what are these licensing models?**

*No - and not without further consideration of the impact and resources to renegotiate licences.*

- 9. Have you received any complaints or legal actions from third parties re. the publication or distribution of metadata/thumbnails/full content? If so, what follow-up actions did you take?**

*We do have a take down policy and do get requests to remove items from time to time; these are referred back to the collection holder in the first instance for validation.*

- 10. What would help you best to automate the process of providing licensing information about the objects in your collection?**

*Not sure there is a solution at the moment - I think any solution would need to be future proof and proven reliable.*

## APPENDIX 4: Europeana Inside Use Cases

Below are the full descriptions of the use cases created for the Europeana Inside project as they appear in the deliverable D2.2 – Use cases.

### Use Case 1: Basic automated transformation and supply scenario

Discovery has been a major objective of content providers since the beginning of digitisation and the internet. The knowledge of how to make cultural content findable is widespread among content providers and they have applied standards for structuring, interpreting and making data available to make their collections findable also outside the original context, e.g. in portals and web search engines. Europeana has given a major impulse to the available knowledge and strengthened methods due to the large number of professionals and collections involved in it. The EDM, the aggregation infrastructure and the aim to publish all its data as Linked Open Data are important results to improve discovery and exposure of cultural data on the Web. Discovery is not the only feature in this basic use case scenario, but for various institutions it may be its main and only purpose. Users who wish to supply metadata and / or metadata to Europeana and/or other aggregators for additional purposes beyond simple discovery may be expected to opt for the advanced scenario.

The basic scenario aims at content providers who can comply to a minimum quality level, consisting of a set of minimum standards and formally accepted protocols. The ECK should be designed to enable a content provider with only basic knowledge of data standards to transform data into a Europeana compliant format.

Preconditions are that:

- a) The content provider knows his own data (what is the meaning of the value in a certain field in his database);
- b) The content provider uses the data model in a consistent way (the same form and type of information is always provided in a certain field in his database);
- b) They can make their data available to other applications (e.g. by being able to export it in a basic data format which will be defined later).

The ECK mapping tool can help with the transformation of the source format to an intermediate format or directly to the EDM format. An intermediate format might be chosen if the data is supplied first to another portal or a Europeana aggregator. Once this transformation has been achieved, the data can be supplied to Europeana by the institution itself or by the aggregator.

#### Use Case 1: Basic automated transformation and supply scenario

Describes the most basic scenario for delivering content to Europeana. Basic in this context means the minimum amount of workflow steps, taken with a minimum amount of resources that a content provider has to run through in order to provide data. It represents the lowest possible barrier for providing content and also meets the basic quality requirements for data.

#### Goal:

The content provider wants to deliver (part of) his collection data to Europeana to increase the exposure and discovery on the web. They spend a minimum of resources (time, financial investment in human capacity and technical components) on this process. Everything should be as automated as possible to save time and resources. The data transformation and supply has to be quick, smooth and with as little human intervention as possible. The content provider may compromise on the quality by focusing on minimum data requirements, but does not want to compromise on control over the process. They want to feel in charge and informed about the impact and consequence of every step taken.

<b>Actor(s):</b>
Content providers which fit in or choose to operate within the basic profile.
<b>System Components per step:</b>
CMS (manage, select, provide the data to the ECK); ECK (prepare, validate, supply)
<b>Short description:</b>
<ul style="list-style-type: none"> <li>• The content provider makes a selection of records in the CMS which shall be contributed to Europeana;</li> <li>• The selected records are exported from the CMS in an appropriate open and machine readable data exchange format;</li> <li>• The data is loaded in the ECK;</li> <li>• The ECK may recognize the user from previous interaction and suggests a mapping and necessary data enhancements (like apply license, apply unique identifier) based on input format and preferred user settings (which can be overruled by the user, if necessary);</li> <li>• The ECK detects possible problems with the data structure and flags fields and/or records as problematic;</li> <li>• The content provider sees the problematic record(s) and gets the possibility to fix them;</li> <li>• (After the fix or when no problems appear:) The content provider gets a preview of what his data will look like after the transformation;</li> <li>• The content provider approves the mapping and starts the transformation with one click;</li> <li>• (Before or after the transformation:) The content provider can choose what license to apply to the (records in the) dataset;</li> <li>• The content provider gets a preview of what his data will look like after the enhancements (e.g. with a license of choice, with a PI);</li> <li>• The content provider approves the enhancements and starts the transformation with one click;</li> <li>• The ECK transforms the data and the content provider gets to click through the result and sees a preview on record base (<i>note: as pointed out in D2.1, whether or not the preview shows the result of the mapping exactly as in Europeana might be up for discussion</i>);</li> <li>• The content provider is satisfied with what he sees and clicks the 'supply' button;</li> <li>• The ECK supplies the data to Europeana or the preferred aggregator;</li> <li>• The ECK gets an indication on how long it will take to display the data in Europeana and notifies the content provider about the expected time.</li> </ul>

<b>Assumptions:</b>
<ul style="list-style-type: none"> <li>• Prior to the supply process described in this use case, the content provider has logged on to the ECK (either to make a profile and save his preferred settings, like the data exchange format or because he has already supplied data with the ECK);</li> <li>• The content provider has already signed the Europeana Data Exchange Agreement (DEA);</li> <li>• The content provider can export a set of data in minimum standard requirements from his CMS;</li> <li>• The content provider has clear guidelines about the required data and its quality regarding the export to Europeana and thus knows what his preferred data format of exchange is;</li> <li>• The content provider agrees to overrule possible licenses already applied to his records in the CMS by the licenses needed for exchange with Europeana, as applied by the ECK. <i>Note: the licenses meant here refer to the metadata, not the objects they represent.</i></li> </ul>
<b>Technical preconditions:</b>
<ul style="list-style-type: none"> <li>• The ECK can be used either from within the CMS which has a direct connection to the ECK or the ECK can be used as standalone web tool;</li> <li>• Content providers can log on to the ECK, get recognized and only have to make their choices once;</li> <li>• There is a limited amount of data formats that can be recognised by the ECK;</li> <li>• The content provider receives feedback on errors and is able to correct the automatic mapping where necessary (this includes missing information or missing thumbnails);</li> <li>• The content provider is able to save mappings as draft and adjust them again later;</li> <li>• In case of an error the content provider is able to start the uploading process again;</li> <li>• Each record exported from the CMS contains a unique identifier (e.g. record number) needed for recognising previously uploaded records in order to transfer updated records to Europeana and to prevent double records in Europeana;</li> <li>• License field, license and unique identifier are given (or overruled) by the ECK;</li> <li>• The ECK can be used to add semantics and project related information to the data (e.g. digital asset type: Image, Sound, Video).</li> </ul>
<b>Steps:</b>
<ol style="list-style-type: none"> <li>1. Manage (in the CMS and thus not relevant for the ECK, however, the user requirements relevant to the CMS must be met by the CMS);</li> <li>2. Select (in the CMS);</li> <li>3. Prepare (mapping, transformation and data enhancement in the ECK);</li> <li>4. Validate (validation of mapping, transformation and data enhancement in the ECK, step can be repeated until all errors are restored);</li> <li>5. Supply data (one-click-operation within the ECK to supply data to Europeana);</li> <li>6. Data acceptance (the content provider receives a notification about the acceptance</li> </ol>

of his data by Europeana when it has happened and when he can expect to view his data in Europeana).

#### **Post conditions**

- The content provider can view the contributed data in Europeana (asap);
- The ECK uses unique identifiers from the source data to apply persistent identifiers (Europeana conform) and to check the data in a following supply process for later operations (e.g. deletion, updating, re-ingestion, exchange with other targets);
- When the content provider logs on next time their choices are saved and can be applied again or be manually overruled if the content provider wishes to do that;
- When the content provider logs on next time with the same dataset the ECK should recognise this and update the right records in Europeana rather than offering them as new records;
- The content provider's CMS is able to export data compliant to a minimum quality level of interoperability;
- The content provider can withdraw its records from Europeana (which might be necessary in certain cases e.g. de-selected records; collections that are inherited from other institutions which cease to exist; a new CMS is acquired which provides the institution with new record numbers; collections are merged).

#### **Remarks/Notes**

Since content providers often use different databases with different structures for different collections, the source data made accessible to the ECK can have different formats. When using the tools for the first time for a certain collection, content providers need to indicate which data model is used, in order to create a mapping scheme. Errors should be used to adjust and improve the quality of the automatic mapping.

## **Use Case 2: Advanced configuration and supply scenario**

This scenario aims at users with the advanced profile who wish to exchange richer data and have more options to configure the workflow steps and use more granular modules than the basic scenario offers. They aim at more meaningful data that can be more useful for different purposes. This use case addresses the needs and wishes of advanced profile institutions who wish to have higher granularity and more advanced settings available for the mapping, enrichment and transformation of their data.

This richer and more meaningful data needs to be converted into data that can be used and re-used by Europeana. This conversion requires decisions to be made by a data manager or other staff member who understands the provenance of the data and the working procedures of the content provider as well as the Europeana data model. This means that the conversion requires manual configuration. So in order to support this advanced scenario the ECK needs to support customisation, configuration and various modules for the most effective conversion of each content provider. On the other hand, the ECK must offer enough standardisation also within this advanced scenario in order to make the various steps of the conversion process more efficient than when they are performed without the ECK.

<b>Use Case 2: Advanced configuration of data transformation and supply scenario</b>
Describes an advanced scenario for delivering content to Europeana. Advanced in this context refers to a version of the ECK that enables content providers to configure the settings according to the content providers needs in order to exchange various richer data sets with Europeana, one or multiple aggregators or other platforms than Europeana.
<b>Goal:</b>
The content provider wants to deliver (part of) their collection data to Europeana to create new meaningful context for their collections and profit from the possibilities Europeana has to offer, such as the publishing as Linked Open Data or the creation of meaningful data visualisation. The content provider therefore invests in time, human skills and technical components for this process. His goal is to be able to customise the steps in the mapping, transformation and uploading process, including such things as the metadata scheme and the metadata fields to create the most effective and rich (or: fit) data set out of his data.
<b>Actor(s):</b>
Content providers which fit in or choose to operate within the advanced profile.
<b>System Components:</b>
CMS (manage, select, prepare, supply to the ECK); ECK (prepare, validate, supply, accept)
<b>Short description:</b>
<ul style="list-style-type: none"><li>• Content provider makes available a selection of records within his CMS;</li><li>• Content provider makes a selection of fields to be incorporated in data exchange;</li><li>• The content provider makes sure that all the licenses are correct and then logs on to the ECK (either from within the CMS: ECK has a direct connection to the CMS or can be used as separate tool);</li><li>• The CMS and/or ECK keep logs of each record processed through the ECK, so the content provider always knows which records have already been supplied to</li></ul>

Europeana and when;

- The data is loaded into the ECK;
- The ECK detects possible problems with the data structure and flags fields and/or records as problematic;
- The content provider can fix the problematic records and then return to the ECK;
- The ECK recognises when some of the records or whole data sets are already present in Europeana and checks which ones need updating;
- The content provider can agree to the updating of his records in Europeana or decline that;
- The ECK suggests necessary data enhancements on data set and/or record level (like apply license, apply unique identifier) and gives the possibility to approve or decline them;
- The content provider points out what source format the data is in and chooses a target format;
- The content provider chooses a default mapping and opens it to edit;
- The content provider checks the settings of the default mapping and configures it further to his needs. Pilot transformations and validations help him to make further adjustments until the optimal representation is achieved;
- When the content provider is satisfied, he saves the configurations so he can use them again later;
- Content provider approves to the conversion and the enhancements and validates the converted and enhanced data again;
- The content provider supplies the data to Europeana through one click in the ECK;
- The ECK gets an indication on how long it will take to display the data in Europeana and notifies the content provider about the expected time;
- The content provider reads the notification that his data has been accepted in Europeana and goes to view his data in Europeana.

#### **Assumptions:**

- The content provider has already signed the Europeana DEA;
- The CMS can make a connection to the ECK;
- The ECK consists of different modules which can be applied by advanced users;
- The content provider has clear guidelines about the required data and its quality regarding the export to Europeana;
- The content provider applied licenses to his records in the CMS. In the ECK the content provider can have this source data translated, added or replaced by target specific licenses information;
- The ECK allows the creation of multiple sets for multiple aggregators inside and outside the Europeana context.

<b>Technical preconditions:</b>
<ul style="list-style-type: none"> <li>• The ECK can be used either from within the CMS which has a direct connection to the ECK or the ECK can be used as standalone web tool;</li> <li>• The CMS or content provider can log on to the ECK, get recognized and settings can be saved;</li> <li>• The content provider receives feedback on errors and is able to correct the mapping where necessary (this includes e.g. missing information, missing thumbnails);</li> <li>• The content provider is able to save mappings as draft and adjust them again later;</li> <li>• The content provider is able to save mappings to use them again next time;</li> <li>• In case of an error: the content provider is able to start the uploading process again;</li> <li>• Each record from the CMS contains a unique identifier (e.g. record number) needed for recognizing previously uploaded records to Europeana in order to transfer updated records and to prevent double records in Europeana;</li> <li>• Fields for licensing information, unique identifier and file format are present in the CMS and can only be given by the ECK if the information has to differ from the CMS;</li> <li>• The CMS and / or ECK can be used to add semantics and project related information to the data (e.g. digital asset type: Image, Sound, and Video);</li> <li>• Europeana can handle incremental harvesting</li> </ul>
<b>Steps and responsibilities:</b>
<ol style="list-style-type: none"> <li>1. Manage (in the CMS and thus not relevant for the ECK, however, the user requirements relevant to the CMS must be met by the CMS);</li> <li>2. Select (in the CMS);</li> <li>3. Prepare (data enhancement in the CMS and / or the ECK);</li> <li>4. Connect (CMS or content provider makes connection to the ECK and loads selected data);</li> <li>5. Configure (ECK settings for mapping, transformation and further enhancements);</li> <li>6. Validate (validation of mapping, transformation and data enhancement in the ECK, step can be repeated until all errors are restored);</li> <li>7. Supply data (one-click-operation within the ECK to supply data to Europeana or aggregator);</li> <li>8. Check for updates, double records (ECK checks if records are present in Europeana, suggests updates);</li> <li>9. Data acceptance (the content provider receives a notification about the acceptance of his data by Europeana or the aggregator, when it has happened and when he can expect to view his data in Europeana).</li> </ol>
<b>Post-conditions:</b>
<ul style="list-style-type: none"> <li>• Log file with description of the changes that will have occurred in the system when the use case is fully completed (where 'system' can also refer to the content provider situation or data condition);</li> <li>• Data exchanged by the ECK must have PIDs linked to the source data in the CMS of</li> </ul>

the content provider so that they can be used for later operations (e.g. deletion, updating, re-ingestion, and exchange with other targets).

### **Use case 3: Enriched data return scenario**

Cultural heritage institutions that have delivered content to Europeana, using a basic or advanced profile, want to re-ingest or reuse enriched data Europeana offers in return, to be incorporated in their own system again and for other services.<sup>44</sup> The ECK should enable them to re-use, or get back enriched data from Europeana. Granular possibilities of integrating this data in their own system (not necessarily their CMS) should be offered. They should be able to use and/or reuse it wherever they want and for instance publish it on their own website.

As mentioned in section 2 the reason for choosing to have one return scenario for both profiles is that currently both profiles need advanced options for customisation when receiving returned enriched data from Europeana. They need to be provided with maximum granularity and given all possible mechanisms to control the enriched data return. Whether they use these or not, is up to them. An institution with an advanced profile might choose for a basic execution of this return scenario with little customisation and more basic settings because they already closely monitored the quality of the supplied data and pre-calculated the kind of data enrichment they expect to get back from Europeana. Basic profile institutions might want to control and configure the return of enriched data more closely and use more advanced settings when returning data. In doing so, basic profile institutions learn more about their own data quality from the advanced return scenario. Therefore the possibility that they might upgrade to the advanced supply scenario the next time may increase.

<b>Use case 3: Enriched data return scenario:</b>
A content provider who has contributed data to Europeana (whether it has been contributed directly by the institution itself or via an aggregator) wants to receive and reuse enriched data from Europeana either inside or outside their own CMS, for instance in another local or web based system or on their website.
<b>Goal:</b>
Using the ECK to re-ingest enriched data from Europeana for the purpose of re-use by the content provider (whether that is by reintegrating selected enriched content into their own CMS or to collect and keep it somewhere outside their CMS, like their website).
<b>Actor(s):</b>
Content providers, including aggregators, who have contributed data to Europeana, independent of the fact if they contributed data within the basic or advanced supply scenario.
<b>System components:</b>
ECK, CMS or any other database that feeds into a CMS and is aimed at re-using the returned data.

<sup>44</sup> The ECK does not have to provide extra functionality for re-using content through third parties. The reason we used the term 'parties' in this description, is that it includes heritage institutions, thus content providers, as well as aggregators and companies.

<b>Short description:</b>
<ul style="list-style-type: none"> <li>• The content provider can log on to the ECK to add settings about data re-ingestion to his profile;</li> <li>• The content provider can log on to the ECK to check if (some of) his data has been enriched within the Europeana environment. If he wants to, he can tell the ECK that he wants to receive a notification if new enriched data is available;</li> <li>• The content provider can view the enriched data the ECK found;</li> <li>• The content provider selects which data he wants to accept directly, accept after manipulation or decline. This selection can be executed on collection level, on dataset level, on record level and on field level;</li> <li>• The content provider selects the target system;</li> <li>• The content provider checks the configuration setting of the target system (e.g. where the received data will be stored: as a separate dataset or integrated with the original data, e.g. a field for user generated comments that was empty in the CMS can be filled with information from the enriched data).</li> </ul>
After the selection to accept directly:
<ul style="list-style-type: none"> <li>• Data is ingested automatically in the system of the content provider.</li> </ul>
Acceptance after manipulation:
<ul style="list-style-type: none"> <li>• Content provider edits the enriched data, if necessary;</li> <li>• Content provider maps fields of the enriched data to fields in his own data system of choice;</li> <li>• Content provider previews the transformation, mapping and possible changes;</li> <li>• Content provider is able to save the configuration settings for the next time and as a draft version;</li> <li>• Content provider allows the data to be ingested in the system or content provider declines the ingestion (either for now or definitely).</li> </ul>
<b>Assumptions:</b>
<ul style="list-style-type: none"> <li>• The content provider wants to manipulate and control the ingestion process manually and down to field level as much as possible;</li> <li>• The ECK is used as a tool for content providers to ingest enriched data from Europeana in a user-friendly way;</li> <li>• The unique identifiers (record numbers or the persistent identifiers which has been supplied by the ECK or Europeana) is used as identification of records within Europeana to check for updates on both sides;</li> <li>• The content provider has a policy that allows the ingestion of metadata from other sources besides the institution itself;</li> <li>• The content provider is able to control which data will be re-ingested;</li> <li>• The content provider will be able to edit the enriched data manually before the data is exported to its own system.</li> </ul>

<b>Technical preconditions:</b>
<ul style="list-style-type: none"> <li>• Content providers need to be able to log in to the ECK;</li> <li>• The ECK keeps log files of previously supplied and re-ingested data per content provider;</li> <li>• The content provider has an infrastructure that enables data ingestion using an appropriate protocol.;</li> <li>• The ECK is capable of supplying the data in an appropriate format and protocol.;</li> <li>• The selection process of enriched data in the ECK consists of a yes or no option and an option to edit the enriched data on a collection level, a record/field level;</li> <li>• Log functionality is in place to record when, which records or fields have been re-ingested.</li> </ul>
<b>Steps:</b>
<ol style="list-style-type: none"> <li>1. Manage (ECK checks for available enriched data and data updates in Europeana);</li> <li>2. Select (in the ECK; content provider selects enrichments on field level);</li> <li>3. Prepare (in the ECK the enriched data is mapped to a target format which the content provider chooses, also the enriched data might be edited further: e.g. provenance of the data enrichment source added);</li> <li>4. Validate (Validation of mapping, transformation and data enhancement in the ECK, step can be repeated until all errors are restored);</li> <li>5. Supply data (one-click-operation within the ECK to supply or export data for the local system of the content providers' choice);</li> <li>6. Data acceptance (The content provider ingests and reuses the data further outside the ECK).</li> </ol>
<b>Post conditions:</b>
<ul style="list-style-type: none"> <li>• The content provider is able to use the enriched data in his preferred local system of choice, in online collections and websites and in other applications.</li> </ul>
<b>Remarks:</b>
<p>How to deal locally with the enrichment of records?</p> <p>Store the enriched record separately from the original record in another system than the CMS of the content provider. This can also be another database that feeds into the online collection website or an app.</p> <p><i>Option 1: separation</i></p> <p>Update the whole record with the enriched information: provided the ECK is able to keep and check logs about the last update of a record.</p> <p>Outcome: one original record and one record with the latest enrichments.</p> <p>Problem: the target system for enriched records probably needs to be extended and equipped with extra fields to store the extra content (e.g. User Generated Content).</p>

*Option 2: integration*

Select certain fields that contain enrichments and add them to the appropriate records in the CMS of the content providers. So updating on fields level rather than on record level.

Problem: how to make sure that the right record gets enriched with the right data?

Logs must be kept in all cases.