



→ **Europeanana Publication Policy v1.0**

A guide for aggregators and data providers on metadata and content requirements



## Table of Contents

<b>1.</b>	<b>Background and Purpose</b>	<b>3</b>
<b>2.</b>	<b>Process: how to provide data to Europeana</b>	<b>4</b>
<b>2.1.</b>	<b>Potential data provider</b>	<b>5</b>
<b>2.2.</b>	<b>Data Exchange Agreement</b>	<b>5</b>
<b>2.3.</b>	<b>Data contribution workflow</b>	<b>5</b>
<b>2.4.</b>	<b>Prioritisation of metadata for publication</b>	<b>6</b>
<b>2.5.</b>	<b>Acceptance criteria</b>	<b>6</b>
<b>3.</b>	<b>Technical: what are the minimal technical criteria?</b>	<b>7</b>
<b>3.1.</b>	<b>Metadata schema and structure</b>	<b>7</b>
<b>3.2.</b>	<b>Mandatory elements</b>	<b>7</b>
<b>3.3.</b>	<b>Acceptance criteria</b>	<b>9</b>
<b>4.</b>	<b>Legal: how digital objects should be labelled</b>	<b>10</b>
<b>4.1.</b>	<b>Validating edm:rights</b>	<b>10</b>
<b>4.2.</b>	<b>Public Domain Mark (PDM)</b>	<b>10</b>
<b>4.3.</b>	<b>Out Of Copyright - Non Commercial Re-use (OOC-NC)</b>	<b>11</b>
<b>4.4.</b>	<b>The Creative Commons CC0 1.0 Universal Public Domain Dedication (CC0)</b>	<b>11</b>
<b>4.5.</b>	<b>Orphan Works (OW)</b>	<b>11</b>
<b>4.6.</b>	<b>Unknown rights statement</b>	<b>12</b>
<b>4.7.</b>	<b>Consistency of rights statements</b>	<b>12</b>
<b>4.8.</b>	<b>Acceptance criteria</b>	<b>12</b>
<b>5.</b>	<b>Document history</b>	<b>14</b>
<b>6.</b>	<b>Acknowledgements</b>	<b>14</b>



# 1. Background and Purpose

Europeanana has brought together over 36 million digital objects from the online collections of more than 3,000 galleries, libraries, museums and archives from across Europe. While we will continue to be as inclusive as possible and collect the diversity of digital cultural heritage that is available, we will focus on the improvement of the quality of data; accessibility, accuracy or consistency of metadata and content are key components of the service we want to develop together with our data providers. This focus on improving quality has been highlighted as a key priority in [Europeanana's Strategy 2020](#)<sup>1</sup>.

So how can we achieve this? One of the things that has become clear in our consultation with aggregators and data providers is that there is a lack of clarity about what the minimum standards or acceptance criteria are for publication. Over the years, these acceptance criteria have been discussed and laid down in various documents and policies ranging from the [Data Exchange Agreement](#)<sup>2</sup> to the [Public Domain Charter](#)<sup>3</sup>, but there is not one single document that clarifies all our policies for publication. This document therefore brings all of the existing information into one concise document. It is intended to guide and help aggregators and data providers to share their data and improve Europeanana in a uniform and consistent way.

We developed a set of clear criteria as a reference for aggregators and data providers to understand what we need in order to ensure that our digital data is always authentic, trustworthy and robust. As the acceptance criteria were developed based on real life scenarios, it also facilitates their implementation in the day-to-day work of the Europeanana Aggregation Team. In addition, the acceptance criteria were developed based on existing documentation and policies (e.g. [EDM mapping guidelines](#)<sup>4</sup>, [Europeanana Licensing Framework](#)<sup>5</sup>, etc.) to ensure the publication policy is a concise reference document.

This publication policy outlines the acceptance criteria for the submission of metadata to Europeanana. Where metadata is submitted to Europeanana and does not meet the acceptance criteria, Europeanana reserves the right to withhold publication until the acceptance criteria are met. Starting 2016, the acceptance criteria will also be applied to legacy material to improve the data quality and remove metadata not compliant with the criteria. A plan will be developed in consultation and collaboration with the affected data providers to manage and implement the application to legacy data over time to ensure a consistent improvement of data quality. In order to consider and incorporate new developments on data quality, all acceptance criteria will be reviewed quarterly and the addition of new acceptance criteria may be necessary too.

The acceptance criteria currently are composed of 3 sections:

- 1) Process: how to provide data to Europeanana.
- 2) Technical: the minimal technical criteria.
- 3) Legal: how data should be labelled.

---

<sup>1</sup> <http://strategy2020.europeana.eu/>

<sup>2</sup> <http://pro.europeana.eu/documents/900548/380f8794-6db3-45de-acf4-3d5721138d26>

<sup>3</sup> [http://pro.europeana.eu/c/document\\_library/get\\_file?uuid=d542819d-d169-4240-9247-f96749113eaa&groupId=10602](http://pro.europeana.eu/c/document_library/get_file?uuid=d542819d-d169-4240-9247-f96749113eaa&groupId=10602)

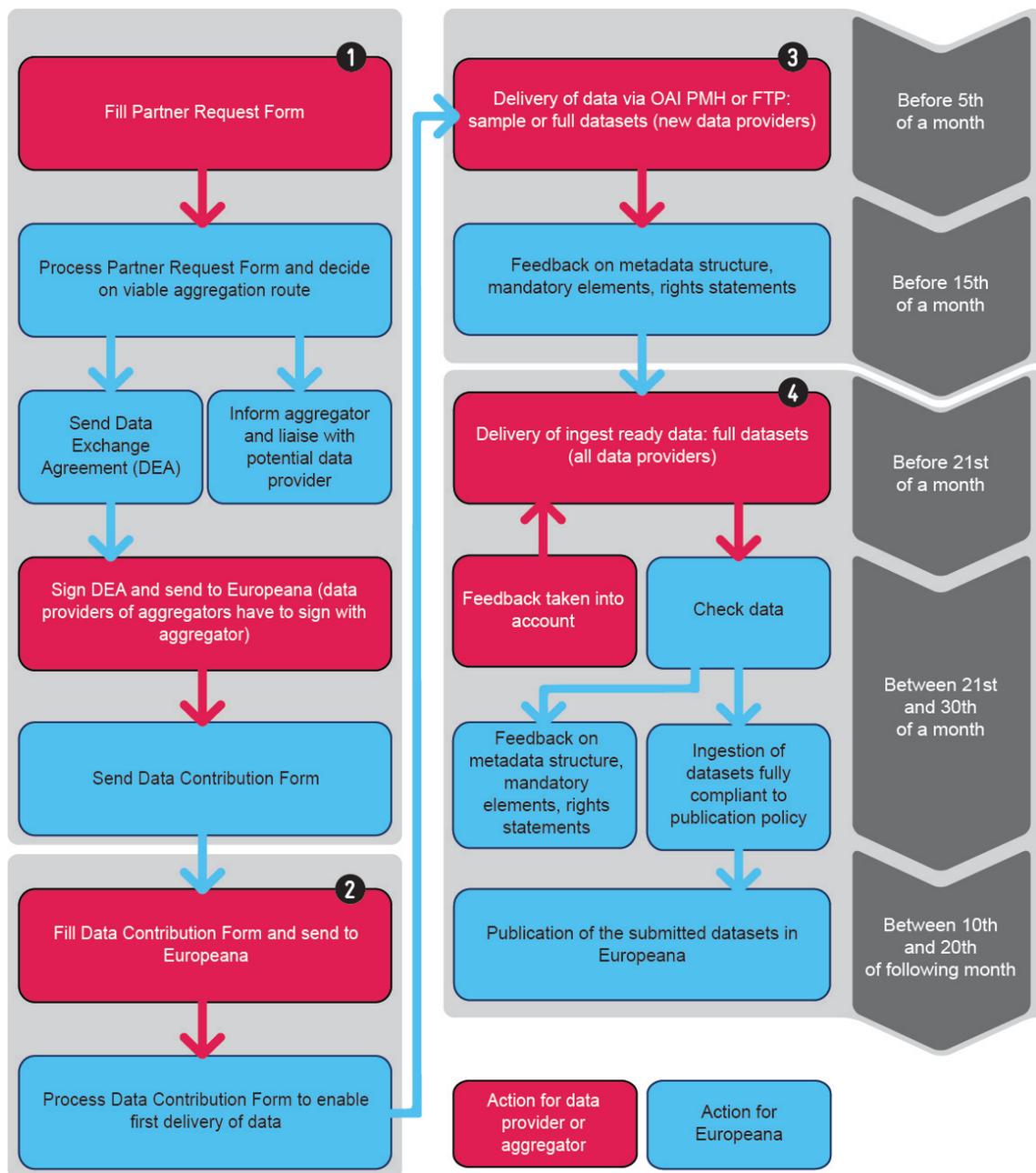
<sup>4</sup> <http://pro.europeana.eu/edm-documentation>

<sup>5</sup> [http://pro.europeana.eu/c/document\\_library/get\\_file?uuid=b16bdaf6-4e53-4f58-968a-9d4943a5d297&groupId=858566](http://pro.europeana.eu/c/document_library/get_file?uuid=b16bdaf6-4e53-4f58-968a-9d4943a5d297&groupId=858566)



## 2. Process: how to provide data to Europeana

This section specifies the acceptance criteria with regards to the ingestion workflow, e.g. meeting deadlines for metadata submission and prioritisation of metadata for publication. See also Fig. 1 for a visualisation of the process.



**Fig. 1.** How to become a data provider of Europeana.



## 2.1. Potential data provider

Every cultural institution in Europe that digitised items of their collections can provide these to Europeana. As Europeana is managing more than 3,000 data providers through about 130 direct data providers, all institutions (interested in) providing data to Europeana have to follow a process to keep it manageable (see Fig. 1 and <http://pro.europeana.eu/procedure> for more detail).

A potential data provider must fill the partner request form (see step 1 in Fig. 1). With this form, we will also collect some information about the data provider and its digital collections that are available. Based on this information, Europeana advises, whether the data provider should join Europeana directly or an aggregator is better positioned to process the data and make it available to Europeana. If we route the potential data provider to an aggregator, we make the connection between both institutions and make sure that a working relationship can be established between them.

## 2.2. Data Exchange Agreement

The [Europeana Data Exchange Agreement \(DEA\)](#)<sup>6</sup> establishes the terms under which Europeana can make use of the previews and descriptive metadata provided by cultural institutions. All data providers must either sign the DEA or an agreement with their aggregators that reflects the terms of the DEA before their metadata will be accepted for publication.

Potential aggregators (including European Commission-funded projects) must ensure that all their data providers have either signed the DEA or an agreement with the aggregators that reflects the terms of the DEA (more information and the [template clauses for aggregators on Europeana Professional](#)<sup>7</sup>).

## 2.3. Data contribution workflow

Before the first submission of data to Europeana, a data provider must fill the Data Contribution Form (see step 2 in Fig. 1). Europeana will evaluate the form and invite the data provider to deliver data to Europeana. The first delivery of data from a new data provider must happen before the 5th of every month (see step 3 in Fig. 1) so that the data can be processed for publication within the same month (in the event that all the data is compliant with all aspects of the publication policy - otherwise the publication will be delayed for at least a month). This first delivery should be a real sample dataset (not just 2-3 records) using the technical infrastructure envisaged for data delivery (preferably OAI PMH). We provide feedback by the 15th of every month and then proceed with accepting full datasets from all providers. The submission deadline for data is the 21st of every month (see step 4 in Fig. 1). Submitting after that deadline means that the data won't be considered for publication that month. From the 21st to the end of the month, all data will be checked, and if everything is fine, will be ingested and prepared for publication. If the data is not correct, another round of feedback and re-submission will begin to receive the correct data. If this feedback loop can be

---

<sup>6</sup> <http://pro.europeana.eu/documents/900548/380f8794-6db3-45de-acf4-3d5721138d26>

<sup>7</sup> <http://pro.europeana.eu/ensuring-permissions-for-aggregators>



finished in time, the data will still be considered for publication the same month. More likely, the publication of the data will be delayed at least by another month. The following month, the publication is finalised and will be live between the 10th and 20th. All these steps are currently recurring on a monthly basis. The entire process is subject to change if unexpected technical problems prevent us from properly processing the data of all our data providers. We inform them if submission or publication deadlines are significantly different from what is described above.

## **2.4. Prioritisation of metadata for publication**

Increasing the amount of openly licensed material is a key task for Europeana to service end-users and creatives better, resulting in corresponding higher returns for the contributing partners. Because of the high quantity of data that is submitted for publication to Europeana every month, it is necessary to set priorities to process it efficiently and ensure that the publication is of high quality. Therefore, data describing digital objects that are openly licensed (edm:rights) and have direct links to digital objects (edm:isShownBy) will be processed first by the Europeana operations officers. Data describing content that comes with heavier copyright restrictions will be processed later, but may also be postponed to the following publication (i.e. one month's delay) if it is not possible to process it in time for a publication to go live. In this case, data providers will be informed when they submit the metadata, in order to manage expectations.

## **2.5. Acceptance criteria**

- Data providers should complete and submit the Partner Request Form and the Content Contribution Form.
- Data providers must sign the Data Exchange Agreement or an accordant agreement with their aggregators before metadata is published.
- Aggregators must sign the Data Exchange Agreement and must ensure their data providers have signed an agreement that equally reflects the terms of the DEA.
- Metadata is accepted for publication after the feedback from Europeana is taken into account and all data is compliant with the publication policy.
- Datasets are prioritised for publication if the edm:rights in the majority of the metadata of the dataset is PDM, CC0, OOC-NC, CC BY, CC BY-SA, CC BY-ND, CC BY-NC, CC BY-NC-SA, CC BY-NC-ND.
- Datasets are prioritised for publication if the edm:isShownBy in the majority of the metadata of the dataset is a direct link to the digital object.



### 3. Technical: what are the minimal technical criteria?

Europeana checks and validates the metadata during the ingestion process, in order to ensure and verify a minimum level of metadata quality. This section will specify the acceptance criteria for metadata applied during the technical validation.

#### 3.1. Metadata schema and structure

While Europeana prefers metadata to be submitted using the [Europeana Data Model \(EDM\)<sup>8</sup>](#), metadata submitted in [Europeana Semantic Elements \(ESEv3.4\)<sup>9</sup>](#) is still accepted. Aggregators and data providers have to make sure to follow the specifications and guidelines for EDM or ESE as the metadata will automatically be validated according to the schema during the ingestion process. All metadata must pass this validation step before being further processed for publication, but Europeana provides feedback to aggregators and data providers to help them improve the data.

In addition to the automatic validation, Europeana manually validates the type of values provided in the metadata and the semantic structure of the data. That means for example that in cases where a URL is expected (e.g. for `edm:isShownAt` and `edm:isShownBy`), a URL has to be in the metadata. For more information on this aspect, you can consult the technical documentation for [EDM<sup>10</sup>](#) and [ESE<sup>11</sup>](#) schemas on the Europeana Professional website.

#### 3.2. Mandatory elements

The mapping guidelines for EDM define 10 mandatory elements in the metadata that must be populated before qualifying the metadata to be further processed for publication (see below and the [technical documentation for EDM<sup>12</sup>](#) for more details). The population of the metadata elements is validated automatically, while the quality of the metadata provided in these fields is checked and validated manually.

- 1) Every metadata record must have either a title (`dc:title`) or a description (`dc:description`). The values in these fields need to be unique and meaningful across the submitted dataset. It is not acceptable that all metadata records in the dataset have an identical title or description that does not specify the cultural heritage object sufficiently.
- 2) Every metadata record describing a text object (e.g. book, manuscript, letter) must have the language of the document in the metadata (`dc:language`).

---

<sup>8</sup> <http://pro.europeana.eu/edm-documentation>

<sup>9</sup> <http://pro.europeana.eu/ese-documentation>

<sup>10</sup> <http://pro.europeana.eu/edm-documentation>

<sup>11</sup> <http://pro.europeana.eu/ese-documentation>

<sup>12</sup> <http://pro.europeana.eu/edm-documentation>



- 3) Every metadata record must have the type of the digital object specified in `edm:type`. This metadata field can only be populated with one of the following five fixed values: TEXT, IMAGE, SOUND, VIDEO, 3D.
- 4) Every metadata record must give some context to and details about the objects described by the metadata. This additional information can be either the subject of the cultural heritage object (`dc:subject`), its nature or genre (`dc:type`), the location or place the object depicts (`dcterms:spatial`) or the spatial or temporal topic of the object (`dc:coverage`).
- 5) Every metadata record describing digital objects contributed by users (e.g. during a public digitisation campaign as part of a collection day organised within the project Europeana 1914-18) must be provided with `edm:ugc = true`.
- 6) Every metadata record must have the information of the actual cultural heritage institution that provides the data to an aggregator (`edm:dataProvider`).
- 7) Every metadata record must have information about the direct data provider to Europeana (`edm:provider`). The value will be identical to the data provider information if the cultural heritage institution that owns the object also provides the digital representation of the object to Europeana. If the cultural heritage institution collaborates with an aggregator in order to deliver the data to Europeana, the direct data provider to Europeana is the aggregator.
- 8) Every metadata record must have at least one link (URL) to the digital object on a website of a cultural heritage institution or an aggregator. The data provider can give a link to the digital object in the context of the organisation website or a link to a (image or book) viewer that shows the digital object (`edm:isShownAt`). It is however strongly recommended to give a link to a web view of the digital object, which is a direct link to the actual file that is ready for download (e.g. jpg, mp3, pdf) and generates a lightbox in the Europeana portal (`edm:isShownBy`). Instead of giving just one of these links, we highly recommend that you provide both links along with the data.
- 9) Every metadata record must be supplied with a valid rights statement using `edm:rights` and the corresponding URI to the rights statement. The list of valid rights statements is published on [Europeana Professional](http://pro.europeana.eu/available-rights-statements)<sup>13</sup>. For more details on this, see the following section (rights statements for digital objects).
- 10) Every metadata record must have a unique and persistent identifier (`rdf:about` of the `ProvidedCHO` class in EDM, and `edm:aggregatedCHO`) that will be used to generate the permalink to the record on the Europeana portal as well as the Europeana Identifier for both Europeana portal and API. The persistence of this identifier will guarantee that the links to each object on the portal remain when the metadata record is updated.

Every metadata record must be submitted in UTF-8 character encoding to ensure a readable display of the data in both the Europeana portal and API. Otherwise, no encodings like html brackets are allowed in the metadata. Whitespaces like a horizontal space (e.g. tab) or vertical space (line break) must be avoided in metadata values.

---

<sup>13</sup> <http://pro.europeana.eu/available-rights-statements>



### 3.3. Acceptance criteria

- Aggregators and data providers must follow the specifications and guidelines for EDM or ESE as the metadata is automatically validated according to the schema during the ingestion process.
- Ten mandatory metadata elements (specified by the EDM mapping guidelines) must be populated with meaningful and correct values before submitting data to Europeana.
- Every metadata record must have either a title (dc:title) or a description (dc:description).
- Every metadata record describing a text object (e.g. book, manuscript, letter) must have the language of the document in the metadata (dc:language).
- Every metadata record must have the type of the digital object specified in edm:type.
- Every metadata record must provide some context and details about the objects described by the metadata (dc:subject, dc:type, dc:coverage).
- Every metadata record describing digital objects contributed by users must be provided with edm:ugc = true.
- Every metadata record must have the information of the actual cultural heritage institution that provides it to an aggregator (edm:dataProvider).
- Every metadata record must have the information of the direct data provider to Europeana (edm:provider).
- Every metadata record must have at least one link (URL) to the digital object on a website of a cultural heritage institution or an aggregator (edm:isShownAt or edm:isShownBy).
- Every metadata record must be supplied with a valid rights statement for the digital object using edm:rights and the corresponding URI to the rights statement.
- Every metadata record must have a unique and persistent identifier (rdf:about of the ProvidedCHO class in EDM, and edm:aggregatedCHO).



## 4. Legal: how digital objects should be labelled

It is mandatory that each digital object is supplied with a valid rights statement supplied using `edm:rights` and the corresponding URI to the rights statement. The list of valid rights statements is published on [Europeana Professional](#)<sup>14</sup>. To assist you with selecting the most appropriate rights statement, you can use the [Rights Statement Selection tool](#)<sup>15</sup>.

### 4.1. Validating `edm:rights`

The process of validating the `edm:rights` field goes beyond the technical validation process, and underpins the mission of Europeana: that all digital objects are published with a rights statement that is valid.

We consider a valid rights statement to be one that is based on the existence, or absence, of copyright. The process for establishing the existence, or absence of copyright should be undertaken by each organisation during their rights clearance process. This should occur, before submitting metadata for publication. You can use an online tool, such as the [Public Domain Calculator](#)<sup>16</sup> to identify the existence, or absence of copyright.

Europeana will continue to follow our clean hands approach, assuming the data provider undertakes the correct level of due diligence and labels the digital objects correctly. However, the use of certain rights statements will be manually reviewed during the ingestion process (prior to publication). This process is undertaken where our experience and research indicates that rights statements may be misapplied, or where explicit criteria must be met before a rights statement can be used. These are described in further detail below.

### 4.2. Public Domain Mark (PDM)

Europeana is committed to the principle that digitisation of Public Domain content does not create new rights over it: works that are in the Public Domain in analogue form should continue to be in the Public Domain once they have been digitised. A work is in the public domain when its copyright does not exist or has expired.

A good example of this is a work created in 1820 that is not protected by copyright and so should be labelled with the rights statement: 'Public Domain Mark'. There is currently one exception to this rule, the 'Out of Copyright - Non Commercial Re-Use' statement and is described in the following section.

We will query the submission of objects that appear to be in the Public Domain and are incorrectly labelled. We pay particular attention to metadata fields like `dc:date`, `dcterms:created` and `dc:creator` to verify the existence of copyright. Europeana will withhold

---

<sup>14</sup> <http://pro.europeana.eu/available-rights-statements>

<sup>15</sup> [http://www.pro.europeana.eu/web/guest/edm-rights-selection-tool/v2/still\\_protected\\_by\\_copyright.html](http://www.pro.europeana.eu/web/guest/edm-rights-selection-tool/v2/still_protected_by_copyright.html)

<sup>16</sup> <http://outofcopyright.eu/>



the publication of the dataset containing objects that appear to be in the Public Domain, but are not labelled as such, whilst we work with you to determine the correct rights statement.

We are aware of certain national legislative provisions that may in some case give a right, based on copyright or other law, to works that are ordinarily in the Public Domain. We are undertaking research into these national scenarios and will publish the research in the first half of 2015.

### **4.3. Out Of Copyright - Non Commercial Re-use (OOC-NC)**

The use of this rights statement is limited to the digital representations of public domain works that have been digitised in a public-private partnership (PPP). And wherein the partners have agreed to contractual limitations to take reasonable steps to limit or discourage commercial re-uses.

If a data provider wishes to use OOC-NC, the contract that specifies these restrictions must be shown to Europeana. A data provider must also specify a year of expiration in the metadata of the digital objects, to indicate the first calendar year in which the digital object can be used by third parties without restrictions on commercial use. Otherwise, Europeana will withhold the publication of the dataset containing objects labelled 'OOC-NC' until the data provider proves that these objects have been digitised in a PPP.

### **4.4. The Creative Commons CC0 1.0 Universal Public Domain Dedication (CC0)**

By applying the CC0 public domain dedication to one or more digital object, a data provider waives all rights in the objects in question, allowing them to be used by anyone without any restrictions. Note that by applying CC0 to a digital object, a data provider can only waive rights held by the data provider (or third party rights for which the data provider has explicitly received permission to apply CC0 from the rightsholder(s)).

Europeana will verify prior to publication the correct application of this rights statement, e.g. is it likely that the data provider is the rights holder and is able to waive all rights. In case of doubt, Europeana will withhold the publication of the dataset containing objects labelled as 'CC0', whilst we work with you to see if you have the necessary rights to apply a CC0 public domain dedication to the digital objects in question. If not, we will work together to determine the correct rights statement.

### **4.5. Orphan Works (OW)**

The use of this rights statement is limited to digital objects that fall within the scope of the [Orphan Works Directive](#)<sup>17</sup>. In order for digital objects to be eligible for this rights statement, orphan works legislation implementing the Orphan Works Directive must exist in the country of origin of the data provider and the data provider must confirm having undertaken due diligence

---

<sup>17</sup> <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32012L0028>



to the standard defined in the national orphan works legislation. Europeana will withhold the publication of the dataset containing objects labelled as 'Orphan Works' whilst we work with you to verify the correct application of the rights statement.

#### **4.6. Unknown rights statement**

Data providers submitting datasets with digital objects labelled 'Unknown' must explain why no conclusive information pertaining to the rights status of the digital object is available. Europeana will verify the application of this rights statement based on the explanation against the conditions specified on Europeana Professional for the 'Unknown' rights statement. Meanwhile, Europeana will withhold the publication of the dataset containing objects labelled as 'Unknown'. As the unknown rights statements does not provide any legal clarity for users of Europeana, it should only be used as a last resort.

#### **4.7. Consistency of rights statements**

Populating dc:rights is often used to describe additional information about the rights in the digital object. Data providers should ensure that the values in both rights related elements (dc:rights and edm:rights) do not contradict each other. A good example of a contradictory scenario is where edm:right is 'Public Domain', and dc:rights contains a statement such as '(c) Europeana 2014'.

#### **4.8. Acceptance criteria**

- Digital objects must be submitted with valid edm:rights and the corresponding URI to the rights statement
- Digital objects that are in the public domain must be submitted with 'Public Domain Mark' for edm:rights unless a verified exception applies
- Digital objects submitted with 'Out of Copyright - Non Commercial Reuse' (OOC-NC) for edm:rights will only be accepted for publication upon proof of a valid contract that imposes a restriction on the commercial re-use of the digitised Public Domain works. An end date to the restrictions on commercial re-use must be recorded in the metadata of the digital objects.
- Digital objects submitted with CC0 for edm:rights will only be accepted for publication once verified by Europeana.
- Digital objects of 'Orphan Works' (OW) for edm:rights will only be accepted for publication if orphan works legislation exists in the country of origin of the data provider and if the data provider confirms having undertaken due diligence to the standard defined in their national orphan works legislation.



- Digital objects with 'Unknown' for edm:rights will only be accepted for publication if data providers submit an explanation that justifies why no conclusive information pertaining to the rights status of the digital object is available.
- The information in dc:rights must not contradict the rights statement in edm:rights.



## 5. Document history

Version	Editor	Date	Comments
v1.0	Henning Scholz	20/11/2014	<ul style="list-style-type: none"><li>• Final version of v1.0 incorporating previous drafts and contributions as well as comments and suggestions from colleagues and reviewers.</li></ul>

## 6. Acknowledgements

This document was reviewed and amended by Marie-Claire Dangerfield, Cecile Devarenne, Julia Fallon, David Haskiya, Lisette Kalshoven, Paul Keller, Chiara Latronico, Gina van der Linden, Susan Muthalaly, Joris Pekel and Harry Verwayen who either wrote individual paragraphs or suggested useful improvements.