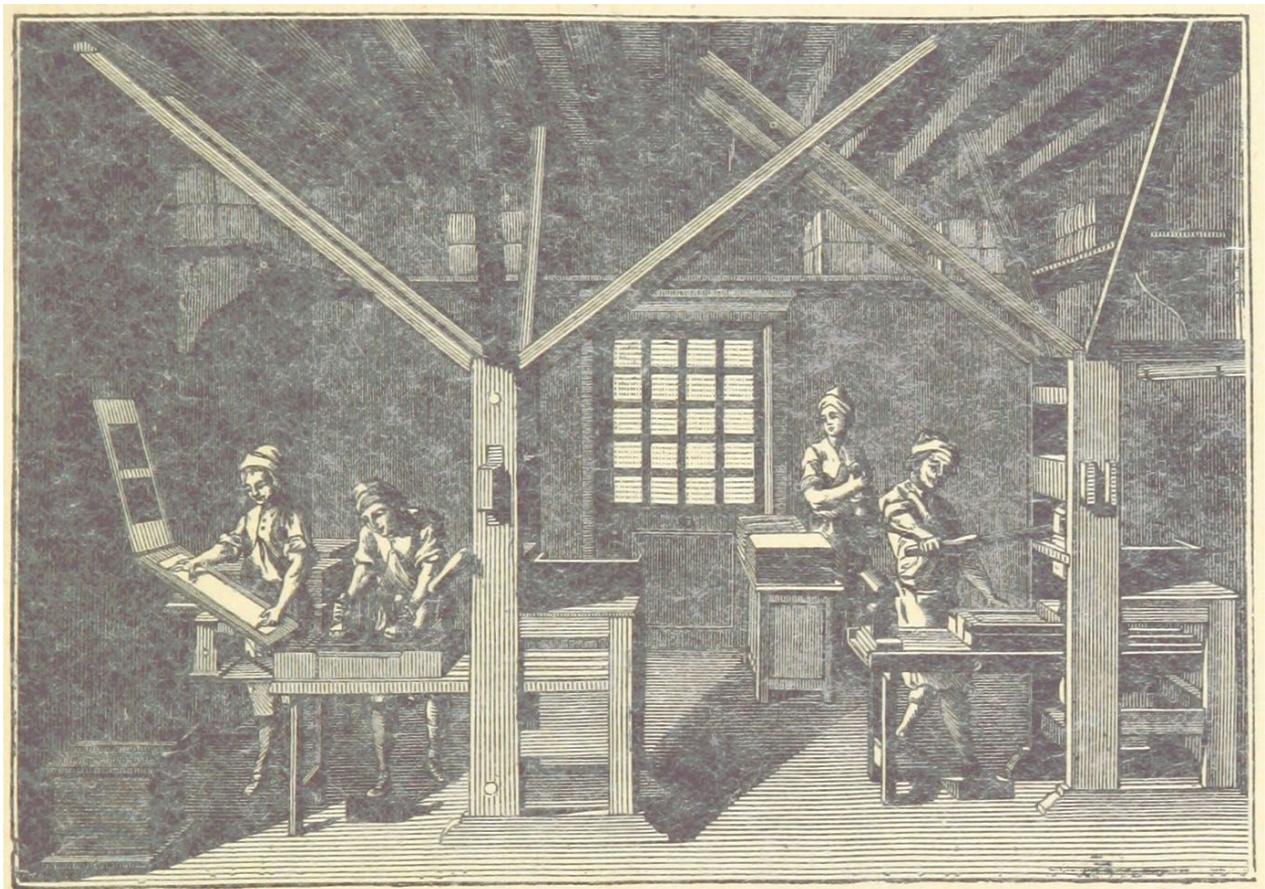# Europeana Publishing Guide v1.5

A guide to the metadata and content requirements for data partners publishing their collections in Europeana

Printing house, "La Veille de la Révolution", 1886, British Library, PDM

Europeana Foundation
March 2017

Author: Henning Scholz

Table of Content

1. Background and purpose

The *Europeana Publishing Guide* is a resource for data partners who share collections in Europeana. It outlines the minimum metadata required to incorporate your data in Europeana.

The *[Europeana Publishing Framework (EPF)](.)*[1] goes beyond the minimum metadata requirements and also addresses content quality. It outlines the four tiers of participation in Europeana and will help you to decide which tier suits your organisation, or particular collection, whether you want to provide a minimum level of metadata to make your collections available in Europeana, or higher quality data with direct links to content and open licences that people can build on.

Europeana has brought together more than 53 million digital objects from the online collections of more than 3,500 galleries, libraries, museums, audiovisual collections, and archives from across Europe. We will continue to be as inclusive as possible and collect a full range of diverse digital cultural heritage, but we must also focus on improving the quality of data. Accessibility, accuracy, and consistency of metadata and content are hugely important for the service we want to develop with you, our data partners. This focus on improving quality is a key priority in [Europeana's Strategy 2020](.)[2] and the [Europeana Content Strategy](.)[3].

How can we achieve this? Our consultation with data partners highlighted a lack of clarity about the minimum standards or acceptance criteria for publishing collections on Europeana. Over the years, these acceptance criteria have been discussed and laid down in various documents ranging from the [Data Exchange Agreement](.)[4] (DEA) to the [Public Domain Charter](.)[5], but there has been no single document that clarifies all our policies for publication.

The Europeana Publishing Guide brings all the existing information in one place and outlines the criteria for submission of metadata to Europeana. It is intended to help aggregators and data partners share their data, become fully standardised and interoperable if desired, and to improve Europeana in a uniform and consistent way. We will work with you to make sure that your datasets meet the publication criteria. Because good data gives audiences a better experience and a greater connection with *your* collections.

We hope that the clear criteria will help you as data partners understand what we need to ensure that your digital data is always considered to be authentic, trustworthy and robust by our audiences. The acceptance criteria are based on

---

[1] [http://pro.europeana.eu/publication/publishing-framework](.)
[2] [http://strategy2020.europeana.eu/](.)
[3] [http://pro.europeana.eu/publication/content-strategy](.)
[4] [http://pro.europeana.eu/get-involved/europeana-ipr/the-data-exchange-agreement](.)
[5] [http://pro.europeana.eu/publications/the-europeana-public-domain-charter](.)

real-life examples as well as existing documentation and policies (e.g. EDM mapping guidelines[6], Europeana Licensing Framework[7]).

The acceptance criteria will also be applied to legacy material (material already published on Europeana). This means data quality will be improved and metadata which is not compliant with the criteria might be removed. We will consult any affected data partners to manage and implement this over time, with the aim of ensuring a consistent improvement in data quality.

To accommodate new developments, all acceptance criteria will be reviewed quarterly and new criteria may be added where necessary. Partners will be notified of any new versions of the *Europeana Publishing Guide* in due time.

The acceptance criteria are divided into three sections to answer the following questions:

- Process: how should data be provided to Europeana?
- Technical: what are the minimum technical criteria?
- Legal: how should digital objects be labelled?

---

[6] http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation
[7] http://pro.europeana.eu/get-involved/europeana-ipr/europeana-licensing-framework

2. Process: how to provide data to Europeana

This section specifies the acceptance criteria with regard to the ingestion workflow. It gives exact deadlines for each part of the process and guidelines to help prioritise your metadata for publication.

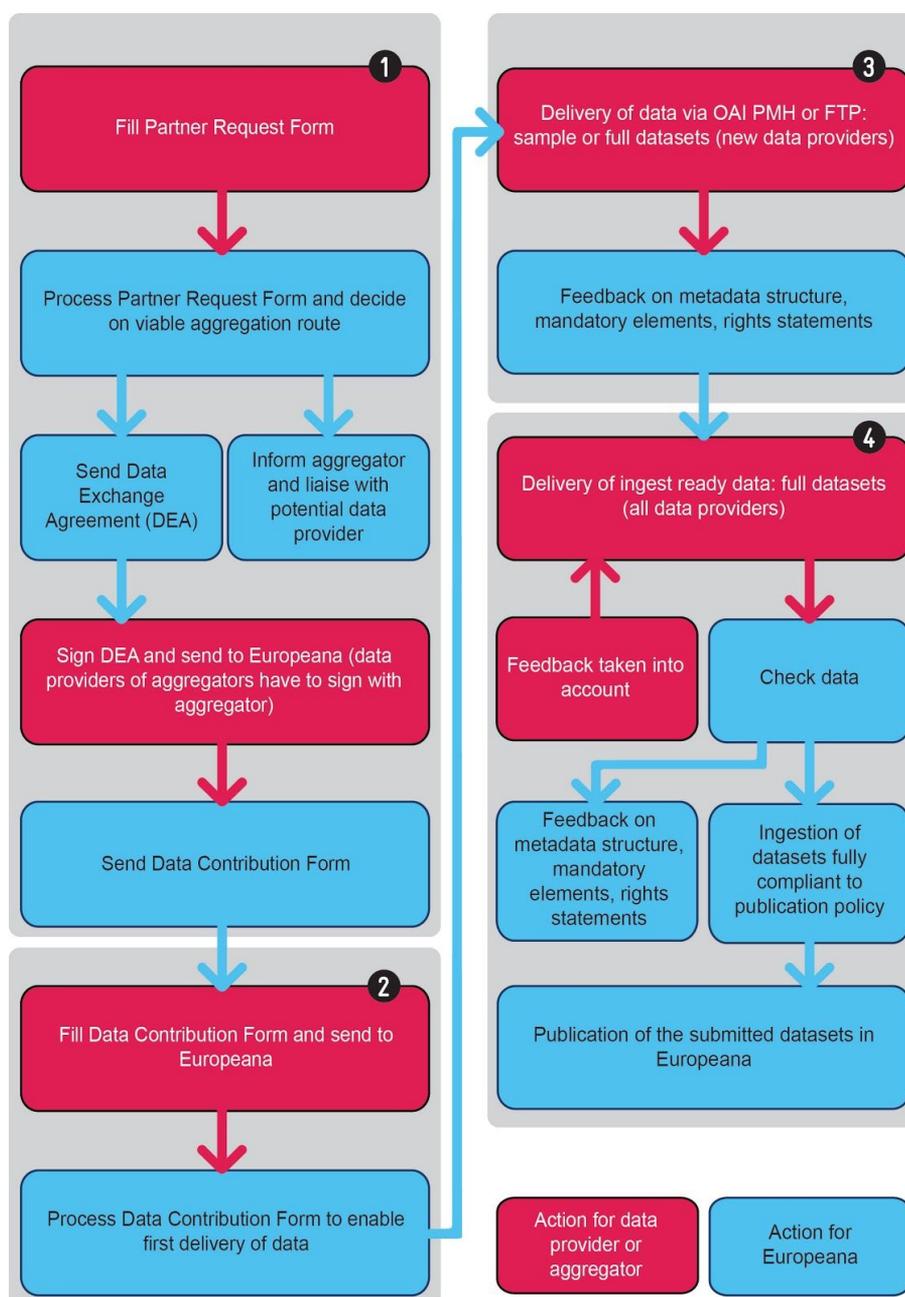Figure 1 provides a clear visualisation of the process.

Fig. 1. The process for becoming a data partner and publishing datasets in Europeana. The four steps labelled in this chart are explained in more detail in the text (see chapters 2.1-2.3).

2.1. Potential data partners

Every cultural institution in Europe with digitised collections can provide their collections to Europeana. Currently, more than 3,500 data partners, working largely through more than 110 aggregators, publish collections in Europeana. To keep this number of partners and datasets manageable, all institutions (interested in) providing data to Europeana need to follow the process we have developed. (See Fig. 1 and [Europeana Pro](#)[8] for more details.)

A potential data partner first must fill in the [Partner Request Form](#)[9] (see step 1 in Fig. 1), which gives us information about the data partner and its digital collections. Europeana can then advise whether the data partner should join Europeana directly or whether an aggregator is better positioned to process the data and make it available to Europeana. If we route the potential data partner through an aggregator, we will connect both institutions and make sure that a working relationship is established between them.

2.2. Data Exchange Agreement

The [Europeana DEA](#)[10] establishes the terms under which Europeana can make use of the previews and descriptive metadata provided by cultural institutions (see also chapter 4 for the rules relating to the intellectual property rights (IPR) in the metadata and content). Before metadata will be accepted for publication, all data partners must sign either the DEA or an agreement with their aggregator that reflects the terms of the DEA. New data partners should always include an accompanying letter with the signed DEA, stating through which aggregator they will be delivering their metadata.

Potential aggregators (including European Commission-funded projects) must ensure that all their data partners have signed either the DEA or an agreement with the aggregators that reflects the terms of the DEA. (Find more information and the template clauses for aggregators on [Europeana Pro](#)[11].)

2.3. Data contribution workflow

Before the first submission of data to Europeana, we work with the data partner to get a broader understanding of the data (e.g. amount of metadata, type of content, technical aspects) using the Data Contribution Form (see step 2 in Fig. 1).

The first delivery of data should be a full sample dataset sent to us using the data delivery method you will continue using (preferably OAI-PMH). The data should be compliant with all aspects of the most recent version of the *Europeana Publishing*

---

[8] http://pro.europeana.eu/share-your-data/how-to-contribute-data
[9] http://pro.europeana.eu/share-your-data/how-to-contribute-data
[10] http://pro.europeana.eu/get-involved/europeana-ipr/the-data-exchange-agreement
[11] http://pro.europeana.eu/get-involved/europeana-ipr/the-data-exchange-agreement

Guide. We will provide feedback on the sample set within two weeks after the date of submission. If the set is not compliant with the *Europeana Publishing Guide*, it will have to be corrected and re-submitted.

Once you submit your complete sets, all data will be checked within two weeks after the date of submission. If everything is in order, the dataset will be ingested and prepared for publication. If the data does not meet the criteria for publication, another round of feedback and re-submission will begin. It is important to note that full datasets smaller than 200,000 records can be submitted once a month. Larger datasets can only be submitted quarterly.

Before the data is published, the data partner can check it in the preview version of Europeana Collections and decide to accept the publication or withdraw the data for improvement. Depending on the outcome, the dataset will either go live or be re-submitted after the desired changes have been made.

In general, incoming data is treated on a first-come first-served basis. If we receive an extraordinarily large number of datasets at a given time, Europeana reserves the right to delay the publication of data-quality accepted datasets. Metadata and content quality, but also user demands and business priorities, will define which collections receive more attention during the publication process. We will, for example, invest more in collections that fit one of the themes to support the development of the Europeana Thematic Collections and prioritise them over collections that do not fit the themes. We will also invest more in collections that fit the higher tiers of the EPF: Tier 4 content will be prioritised over Tier 3 content, which will be prioritised over Tier 2 content, which, in turn, will be prioritised over Tier 1 content. We will still publish Tier 1 collections, but the investment in feedback and quality improvements will be reduced to a minimum to allow a focus on open data and on higher quality collections and make them fit for the intended audience.

Data will be processed and published continuously throughout the year, with the exception of the following months: January, April and July and August. In these months we identify quality issues (e.g. broken links), work with data partners to fix them, clean up the database and apply larger scale improvements without publishing new data. We will otherwise inform you if unexpected technical problems prevent us from properly processing all the data submitted.

Summary: acceptance criteria for data and how we prioritise data publication

- Data partners should complete and submit the Partner Request Form and the Data Contribution Form. These provide essential contact and content information for Europeana.

- Data partners must sign the DEA or an accordant agreement with their aggregators before metadata is published. The provision of metadata under CC0 is essential for the functioning of Europeana.

- Aggregators must sign the DEA and must ensure their data partners have signed an agreement that equally reflects the terms of the DEA. Europeana has to adhere to a "clean hands" policy here and trusts the aggregator maintains full and proper records.

- Metadata is accepted for publication after feedback from Europeana is taken into account and all data is compliant with the Europeana Publishing Guide.

- If we receive an extraordinarily large number of datasets at a given time, Europeana reserves the right to delay publication of data-quality accepted datasets. Metadata and content quality, but also user demands and business priorities, will define which collections receive more attention during the publication process.

3. Technical: what are the technical criteria to publish with Europeana?

Europeana checks and validates the metadata during the ingestion process, to ensure and verify a minimum level of metadata quality and to improve standards across our cultural heritage, and to make the data more discoverable. Europeana also checks the digital objects linked from the metadata, to ensure and verify a minimum level of content quality. This section specifies the acceptance criteria for metadata applied during the technical validation and the technical criteria to be compliant with the individual tiers of the EPF (per media type). Following the principle of the EPF, we always invite and encourage data partners to not only submit a minimum of metadata and content quality, but to aim for rich metadata and the highest possible data quality: the more you give, the more you get.

3.1. Metadata schema and structure

While Europeana prefers metadata to be submitted using the format specified in the [Europeana Data Model (EDM)](#)[12], metadata submitted using the [Europeana Semantic Elements (ESEv3.4)](#)[13] format is still acceptable.

Aggregators and data partners need to follow the specifications and guidelines for either EDM or ESE as the metadata will automatically be validated according to these schemas during the ingestion process. All metadata has to pass this validation step before being processed further for publication. We will provide you with feedback to help you improve your data.

In addition to the automatic validation (presence of metadata), Europeana manually validates the type of values provided in the metadata and the semantic structure of the data. For example, where a URL is expected (e.g. for edm:isShownAt and edm:isShownBy), a URL has to be included in the metadata. For more information on how to apply EDM or ESE, please see the technical documentation for [EDM](#) and [ESE](#) on the Europeana Pro website.

3.2. Mandatory elements in EDM

There are 10 mandatory elements in EDM that must be present for the metadata to be processed further for publication. The list below is an up-to-date summary of these mandatory elements. For more details and the full guidelines see the [EDM reference documents](#)[14]. This documentation also goes into all the other metadata

---

[12] http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation
[13] http://pro.europeana.eu/share-your-data/data-guidelines/ese-documentation
[14] http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation

elements that are recommended to be used to create a rich and full description of your objects.

1) Each metadata record must contain 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. All metadata records in the dataset cannot be given identical titles or a description that does not specify the cultural heritage object sufficiently *because your object would consequently not be discoverable*.

2) Each metadata record describing a text object (e.g. book, manuscript, letter) must provide the language of the document in the metadata (dc:language). In other cases where languages are represented (some audio or video objects or photographs of inscriptions for example) we recommend that you include the language in the metadata. *This allows us to present more material against the language requirements of the user.*

3) Each metadata record must contain the type of digital object (edm:WebResource) specified in edm:type. This metadata field can only be populated with one of the following five fixed values: TEXT, IMAGE, SOUND, VIDEO, and 3D. *Many users want to search according to media; accurate completion of this field allows them to do so.*

4) Each metadata record must provide 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 temporal topic of the object (dcterms:temporal). *The more data you can provide here the more opportunity the users have to find it against their search terms.*

5) Each metadata record describing digital objects contributed by users (e.g. during a public digitisation campaign within the project Europeana 1914-1918) must show edm:ugc = true. *This allows us to maintain Europeana's reputation as an authoritative resource, conferred by the work of the cultural heritage institutions, separating user knowledge and curated or professional knowledge.*

6) Each metadata record must contain the information of the actual cultural heritage institution that provides the data to an aggregator (edm:dataProvider). *Users want to know where the data comes from, and we want to attribute properly.*

7) Each metadata record must contain 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) Each metadata record must provide at least one link (URL) to the digital object either in the context of of a cultural heritage institution or an aggregator's website or as a direct link. It is strongly recommended to provide 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) (edm:isShownBy). The data partner can also assign a link to the digital object in the context of the organisation's website or a link to a (book or image) viewer that shows the digital object (edm:isShownAt). *We highly recommend that you provide both edm:isShownAt and edm:isShownBy along with the data, so that the user is more likely to use your data.*

9) Each metadata record must contain a valid rights statement for the object using edm:rights and the corresponding URI to the rights statement. The list of valid rights statements is published on [Europeana Pro](http://pro.europeana.eu/available-rights-statements)[15]. For more details on this, see the following section (rights statements for digital objects). *A valid rights statement tells the user what they may or may not do with your digital objects. Rights statements are also machine readable, which makes them fully searchable via the API.*

10) Each resource (edm:ProvidedCHO, edm:WebResource, ore:Aggregation and each contextual entity) described in a metadata record must contain a unique and persistent identifier (e.g. rdf:about of the ProvidedCHO class in EDM and edm:aggregatedCHO) that will be used to generate the permalink to the record on Europeana Collections as well as the Europeana identifier for both Europeana Collections and API. *The persistence of this identifier will guarantee that the links to each object remain when the metadata record is updated, which, again, will improve user experience and likelihood of your material being used.*

Each metadata record must be submitted in UTF-8 character encoding to ensure a legible display of the data in both Europeana Collections and API. Encodings like HTML brackets are not allowed in the metadata. White spaces, such as a horizontal space (e.g. tab) or vertical space (line break), must be avoided in metadata values.

3.3. Content featured in Thematic Collections

The purpose of this chapter is to provide partners publishing data on Europeana with the information they need to ensure their collection records are featured in a thematic collection on Europeana.

3.3.1. What are thematic collections?

Thematic collections are "sub-portals" of Europeana Collections. They feature, and via curated content like exhibitions contextualise, records that match a topic or type-based theme which users can view, read and play immediately and directly ("direct media access") in their browser in one click.

---

[15] http://pro.europeana.eu/available-rights-statements

Note that both requirements need to be fulfilled for an item to qualify for inclusion in a thematic collection.

Note also that the check is made on the item level and *not* on the dataset level. The item level check is done by checking for the presence of certain keywords in the item's metadata (and the presence of a direct media access link), typically in a subject or type field but sometimes also in item titles or other metadata fields — this varies from theme to theme. The keywords per theme are regularly reviewed and updated to include more relevant content, and, where necessary, exclude content that should not be included.

In some cases, Fashion and the planned Newspapers collections, the selection of which items to feature differs from the above-mentioned selection procedure (see below).

3.3.2. Direct media access (Tier 2+)

To allow users to directly view/read/play ("direct media access") in their browser in one click you need to do the following:
- If the item is represented by an image provide at least one direct link to the image file in edm:isShownBy (and in edm:hasview if there are multiple images for the item). The image should have a minimum pixel width of 800 pixels.
- If the item is represented in the form of text, sound recordings, video or in 3D provide a direct link to a file in edm:isShownBy (and in edm:hasview if there are multiple representations for the item). The file should be in a format that allows the user to read/view/play the file directly in their browser.
- If you cannot or do not want to provide a direct link to a file then you have two options[16]:
    - For images (incl. multiple sequential images like the pages of a book) provide links in an IIIF-compliant manner.
    - For any media, provide a link to an embeddable player in an oEmbed- compliant manner.

3.3.3. Keywords per theme

Art
The intention of the Art Collections is to showcase a broad selection of art-related content. For that reason not only Fine Art (e.g. paintings and sculpture) is featured but also Decorative Art or Industrial Design (e.g. jewellery, furniture, ceramics, and glassware). Also featured in the theme are books, TV-recordings, and radio broadcasts for example about art history. In anticipation of future thematic

---

[16] When using these options please notify your Europeana data officer beforehand.

collections, works of architecture and fine arts photography are not explicitly featured in the Art Collections. This high-level definition of what to feature in the thematic collection is decided by the [Collection Manager](#) in consultation with the Editorial Board of the thematic collection and the Product Owner of Europeana Collections development.

In practical terms, the selection of items to be featured in a theme is made by checking for certain keywords in the type, subject and concept class label fields.

These keywords include the following terms[17]: *Fine art, Decorative art, art history, painting, sculpture, poster, tapestry, jewellery, miniature, old master print, woodcut, lithography, engraving, illuminated manuscript, porcelain, mannerism, rococo, impressionism, expressionism, romanticism, Neo-Classicism, Pre-Raphaelite, Symbolism, Surrealism, Cubism, Art Deco, Dadaism, De Stijl, Pop Art, Art Nouveau*.

Any item that also meets the direct media access requirement and which has any of these keywords in dc:type, dc:subject or the preferred or alternate labels for controlled vocabulary terms (e.g. Getty AAT) will be included in the Art Collections.

These keywords are regularly reviewed and modified[18] by Europeana[19].

Music
The intention with the Music Collections is to showcase a broad selection of music related content. For that reason not only are music recordings included but also e.g. sheet music, music manuscripts, libretti, images of music instruments, images of musicians and images of music recording and playing devices. Also featured in the theme are works on music or the history of music e.g. books, TV-recordings, and radio broadcasts about music.

In practical terms, theme alignment means checking for certain keywords in the type, subject and concept class label fields.

These keywords include the following terms[20]: *music, sheet music, notated music, score, libretto, musical instrument, concert, symphony, orchestra, ballet, gradual, gramophone, phonograph, wax cylinder, jukebox, cassette player, composer.*

---

[17] Here in English only, but the terms are typically checked in German, French, Italian, Spanish, Polish too and sometimes in Romanian and Swedish.
[18] An example is that we are currently reviewing the keywords in an attempt to include mosaics and more pottery too.
[19] The Collection Manager of the Art Collections and the Product Owner of Europeana Collections development review and modify the keywords.
[20] Here in English only, but the terms are typically checked in German, French, Italian, Spanish, and Polish too and sometimes in Romanian and Swedish.

Also included as keywords are finer grained music genres and forms: *classical music, popular music, folk music, opera, jazz, waltz, sonata, fugue, motet, saltarello, organum, ballade, chanson, laude, madrigal, pavane, toccata, cantata, minuet, partita, sarabande, sinfonia, hymnes, lied, music hall, quartet, quintet, requiem, rhapsody, scherzo, sinfonia concertante.*

Any item that also meets the direct media access requirement and which has any of these keywords in dc:type, dc:subject or the preferred or alternate labels for controlled vocabulary terms will be included in the Music Collections.

These keywords are regularly reviewed and modified by Europeana[21].

Fashion
The Fashion Collections offer very specific functionality, allowing users to filter and facet on designers, fashion item types, date, and more. The basis for that functionality is that all the partners in Europeana Fashion have agreed to and actually use the same vocabularies. Using these vocabularies is a requirement.

3.3.4. Advice for planned thematic collections

First World War
The planned First World War Collection will be based on integrating the standalone Europeana 1914-1918 site into Europeana Collections.

Keyword requirements for inclusion will be similar or identical to those that are currently in place for Europeana 1914-1918. This means that any item that contains the keywords 'First World War' or 'Great War' in either dc:subject or in the label of a SKOS-compliant vocabulary based term will be included.

Photography
To be determined.

Newspapers
To be determined.

3.4. Digital objects and previews in Europeana

Europeana defines a digital object as a digital representation of an object that is part of Europe's cultural and/or scientific heritage. The digital object can also be the original object when born digital.

What type of digital representations are acceptable for Europeana and how objects should be represented depends on the type of the objects. Audio/video snippets or a subset of pages for a multipage document (e.g. book, report) are not an adequate

---

[21] The Collection Manager of the Music Collections and the Product Owner of Europeana Collections development review and modify the keywords.

representation of a digital object in order to fulfil the user demands. In most cases it is also not meaningful for users to submit the individual pages of a multipage document with separate metadata records as individual digital objects[22]. A 300-page book is expected to be represented in Europeana as one single object and not as 300 separate objects. Digital representations of paintings or artefacts are photographs or digitisation of the objects. If the cultural object is a photograph of a painting or an artefact (for instance, a gelatin dry plate negative representing a 19th century painting), it has to be clear from the metadata that the described object is not the painting or artefact itself.

Europeana has established requirements for media files which are linked to from the provided metadata in order to allow Europeana to generate thumbnail images based on media files and extract technical metadata from said media files. They are specified in the Europeana Media Policy[23].

3.5. Non-digital objects

In the context of the Europeana Content Strategy, non-digital objects are any objects for which a digital representation is not available. The value of these objects reside within the informative potential of their metadata and descriptions. For example, we may consider the case of an object that has not been digitised or whose digital representation is deemed not suitable, but for which a metadata record or finding aid is available (e.g. a finding aid about a non-digitised collection of photographs at Archives Portal Europe).

We acknowledge that there are cases where non-digital objects play an important role in Europeana (e.g. in usability) due to the informative potential that non-digital objects have over other digital objects. In hierarchical metadata it is possible that not every level within a hierarchy has a digital representation attached to it. For this metadata to be present in Europeana, the objects described at these levels of the hierarchy must be ingested as non-digital objects. Without the possibility to deliver the metadata of these non-digital objects to Europeana, the quality of the data's hierarchy is compromised[24], which will negatively affect the end-user browsing experience. Another example of hierarchical metadata are bibliographic records of newspaper titles, which have no digital surrogate, but are essential for the interpretation of the metadata of the individual newspaper issues, which do have digital surrogates.

In both cases described above non-digital objects are included in a collection that contains hierarchical metadata, and the metadata of the non-digital objects contains explicit hasPart/isPartOf relations with other digital objects. It is important to keep in mind that these cases are exceptional and that non-digital object aggregation will otherwise not be implemented in Europeana.

---

[22] The exception are manuscripts and illuminated manuscripts where it can be meaningful to users to have individual pages as separate metadata records.
[23] http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/FAQs/Europeana_Media_Policy_052016.pdf
[24] More information about this issue, please see the Recommendations for the representation of hierarchical objects in Europeana

3.6. Minimum requirements for digital objects

Tier 1 of the EPF is the lowest of the four publishing tiers in Europeana, which allows you to make your organisation and your content more visible in a European context via Europeana Collections and the Europeana APIs. The criteria for tier 1 define the minimum requirements for digital objects to be published with Europeana. These criteria are different depending on the media type that is foreseen to be made available and are therefore specified accordingly.

3.6.1. Minimum requirements for image files (Tier 1)

Images are visual resources for users to look at. It is best practice for image files to be published in a quality for a user to tell what the image is about. Image quality is also expected to reflect the changes in technologies over the last years: displays' screen resolutions have increased significantly since then. Correspondingly, user expectations have changed too, which requires image quality to be significantly higher than in 2008 when Europeana first started.

If you are contributing image material because you want it to be discoverable in Europeana Collections, then you need to provide at least a link to the image file or a website on which the image file can be accessed. You should also provide a direct link to an image file of at least 0.1 megapixel in size to have the means to create ~400 pixel-wide preview images to illustrate search results in Europeana Collections.

3.6.2. Minimum requirements for text files (Tier 1)

Text materials are meant to be read and not looked at; therefore, best practice for text-based materials is to publish documents in full, scanned in an adequate resolution to support their legibility, and, if possible, with an added plain-text layer to allow for searching within the content of document. For easy reuse of your text materials it is recommended to allow users to download the documents in file formats that allow users to extract text.

If you are contributing text documents to Europeana because you want them to be discoverable in Europeana Collections, then you need to provide at least a link to the file(s) of the document, a document viewer or a website on which the document can be accessed. We recommend that a link to a still image file is at least 0.1 megapixel in size to have the means to create ~400 pixels preview images to represent the text material on Europeana Collections.

3.6.3. Minimum requirements for audio files (Tier 1)

It is best practice for audio files to be published in a quality that preserves the quality of the underlying recording. On a web-based platform like Europeana Collections, an important criterion is direct access to an audio file that can be played in a web browser without the need for additional software (plugins).

If you are contributing audio material to Europeana because you want it to be discoverable in Europeana Collections then you need to provide a minimum of a link to the audio file or a website on which the audio file can be accessed. We recommend that a link to a still image file is at least 0.1 megapixel in size to have the means to create ~400 pixels preview images to represent the audio file on Europeana Collections (such as an album cover or similar).

3.6.4. Minimum requirements for video files (Tier 1)

The quality of streaming video depends on a number of factors, such as the quality of the source file, the type of encoding chosen for compression, the framerate and bitrate of the file, and the transfer rates that the network supports. Video streaming relies on a number of standards whose support varies across different browsers. The quality criteria of a video file are therefore as much dependent on the efficiency of the file compression as on the setup of the end-user. In equal parts the robustness and connectedness of the video playout service come into play.

Moving image cultural heritage materials should be offered in their original aspect ratio. The codec and file format used are preferably in an open format that can be played in a web browser without the need for specific proprietary software or plugins.

If you are contributing video material because you want it to be discoverable via Europeana Collections then you need to provide at least a link to the video file or a website on which the video file can be accessed. You should also provide a link to a still image file of at least 0.1 megapixel in size to have the means to create ~400 pixels preview images to illustrate search results on Europeana Collections. Usually, this is a representative still from the video or, for example, a from a film poster.

3.7. Recommendations for digital objects

Digital objects compliant with Tier 1 are sufficient to use in Europeana as a search engine and allow people to find your collections. The EPF specifies three additional publication scenarios (or tiers) for you to choose from: depending on what you want and what you give in terms of quality, Europeana can deliver different services to you and you will see different actions as a result of it. If you want to use Europeana as a showcase and to be considered for inclusion in one of our Thematic Collections[25], your digital objects need to be compliant with Tier 2 of the EPF. If you want to use Europeana as a distribution platform for non-commercial reuse, your digital objects need to be compliant with Tier 3 of the EPF. If you want to use Europeana as a free reuse platform, your digital objects need to be compliant with Tier 4 of the EPF.

With regards to the different publication scenarios mentioned above and specified in more detail in the EPF, the following recommendations per media type should be taken into account.

---

[25] Product Specifications and Policy for Europeana Thematic Collections.

3.7.1. Technical criteria for image files (Tiers 2-4)

Tier 2 (Europeana as a showcase): If you want to enable Europeana to use your image material as part of thematic collections and make it more accessible on the Europeana Collections site, then you need to provide at least a direct link to an image file of at least 0.5 megapixel in size (e.g. ~800 pixels wide). The metadata should be as descriptive and rich as possible.

Tier 3 (Europeana as a distribution platform for non-commercial reuse): If you want to make use of Europeana as a distribution platform that enables the use of your image files by private individuals, educators and researchers, then you need to provide at least a direct link to an image file of at least 1.5 megapixel in size (e.g. ~1,200 pixels wide). In addition, you also need to make sure that the image file comes with one of the five rights statements that allow reuse (four Creative Commons licences: CC BY-NC, CC BY-NC-ND, CC BY-NC-SA, CC BY-NC-ND, or NoC-NC). The metadata should be as descriptive and rich as possible.

Tier 4 (Europeana as a free reuse platform): If you want to make use of Europeana as a platform that enables the free reuse of your image files elsewhere then, in addition to the criteria described for Tier 3 above, you also need to make sure that the image file comes with a rights statement that allows free reuse (CC BY, CC BY-SA CC0, or PDM).

3.7.2. Technical criteria for text files (Tiers 2-4)

Tier 2 (Europeana as a showcase): If you want to enable Europeana to present your text materials as part of thematic collections and make them more accessible on the Europeana Collections site then you need to provide at least a direct link to the file(s) of the document, preferably in PDF file format), which will be made directly available on thematic collections pages via a PDF viewer. Although Europeana currently does not support a full-text search across documents, we recommend adding an embedded text layer in PDF files to allow for searching inside the document. Additionally, the International Image Interoperability Framework (IIIF)[26] is also supported by Europeana and can be used for displaying text materials on Europeana.

Tier 3 (Europeana as a distribution platform for non-commercial reuse): If you want to make use of Europeana as a distribution platform that enables the use of your text materials by private individuals, educators and researchers then, in addition to the criteria described for Tier 2 above, you also need to make sure that the text documents come with one of the five rights statements that allow reuse (four Creative Commons licences: CC BY-NC, CC BY-NC-ND, CC BY-NC-SA, CC BY-NC-ND, or NoC-NC).

---

[26] http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_profiles/IIIFtoEDM_profile_042016.pdf

Tier 4 (Europeana as a free reuse platform): If you want to make use of Europeana as a platform that enables the free reuse of your text materials then, in addition to the criteria described for Tier 2 above, you also need to make sure that the text documents come with a rights statement that allows free reuse (CC BY, CC BY-SA CC0 or PDM).

3.7.3. Technical criteria for audio files (Tiers 2-4)

Tier 2 (Europeana as a showcase): If you want to enable Europeana to use your audio material as part of thematic collections and make it more accessible on the Europeana Collections site then the audio needs to be embeddable[27], or you need to provide at least a direct link to an audio file in a format that can be played directly by modern browsers and supported in HTML5.[28] The quality of the audio file should be as high as possible; in other words, as close as possible to the digital source file encoded in a lossless file format. Audio files that meet these requirements will be accessible directly on the thematic collection pages as embedded audio files.

Tier 3 (Europeana as a distribution platform for non-commercial reuse): If you want to make use of Europeana as a distribution platform that enables the use of your audio files by private individuals, educators and researchers then, in addition to the criteria described for Tier 2 above, you also need to make sure that the audio file comes with one of the five rights statements that allow reuse (four Creative Commons licences: CC BY-NC, CC BY-NC-ND, CC BY-NC-SA, CC BY-NC-ND, or NoC-NC).

Tier 4 (Europeana as a free reuse platform): If you want to make use of Europeana as a platform that enables the free reuse of your audio files elsewhere then, in addition to the criteria described for Tiers 2 and 3 above, you also need to make sure that the audio file comes with a rights statement that allows free reuse (CC BY, CC BY-SA CC0 or PDM).

3.7.4. Technical criteria for video files (Tiers 2-4)

Tier 2 (Europeana as a showcase): If you want to enable Europeana to use your video material as part of thematic collections and make it more accessible on the Europeana Collections site then the video needs to be embeddable[29] or you need to provide at least a direct link to a video file in a format that can be played directly by modern browsers and supported in HTML5.[30] The video file should have a minimum vertical resolution (height) of 480 pixels. Video files that meet these requirements will be accessible directly on the thematic collections pages as embedded videos.

---

[27] Any oEmbed compliant player is supported by Europeana.
[28] https://developer.mozilla.org/en-US/docs/Web/HTML/Supported_media_formats
[29] Any oEmbed compliant player is supported by Europeana.
[30] https://developer.mozilla.org/en-US/docs/Web/HTML/Supported_media_formats

Tier 3 (Europeana as a distribution platform for non-commercial reuse): If you want to make use of Europeana as a distribution platform that enables the use of your video files by private individuals, educators and researchers then, in addition to the criteria described for Tier 2 above, you also need to make sure that the video file comes with one of the five rights statements that allow reuse (four Creative Commons licences: CC BY-NC, CC BY-NC-ND, CC BY-NC-SA, CC BY-NC-ND, or NoC-NC).

Tier 4 (Europeana as a free reuse platform): If you want to make use of Europeana as a platform that enables the free reuse of your video files then, in addition to the criteria described for Tiers 2 and 3 above you also need to make sure that the video file comes with a rights statement that allows free reuse (CC BY, CC BY-SA, CC0 or PDM).

3.8. Metadata and content from Wikimedia Foundation projects

Integration of metadata and content from Wikimedia Foundation projects[31] into Europeana products can help improve user experience of these products. Depending on the source of data, Europeana will select the most appropriate integration approach.

Europeana can use Wikidata as a source from which to annotate existing Europeana objects, for example, to add the name or title of a work (painting, musical composition, etc.) in multiple languages. If Europeana uses Wikidata as a source in annotations, it will be clear in API responses and in the Europeana Collections display that these are annotations coming from Wikidata, not metadata provided by a Europeana data partner.

Europeana will not publish and display but may link to digital files from Wikimedia Commons or Wikisource that have been uploaded from cultural heritage institutions' websites (or scanned from publications) by individual Wikimedians without consent from same institutions.

Europeana can publish digital files from Wikimedia Commons media that have been crowdsourced in initiatives like Wiki Loves Monuments[32]. As the metadata of Wikimedia Commons is officially CC-BY-SA (not CC0), this needs to be published with support of the Wikimedia Commons community and with longer descriptive text fields shortened.

3.9. How we work with data partners to correct data post publication

Europeana not only checks and validates the metadata and digital objects linked from the metadata during the ingestion process, we also analyse the data post publication to ensure and verify a minimum and consistent level of data quality across the database.

---

[31] Primarily Wikipedia, Wikimedia Commons, Wikidata, and Wikisource.
[32] And actually has done so already.

Post publication there may be situations where it becomes clear that the metadata's quality of mandatory fields does not meet the acceptance criteria (e.g. links in edm:isShownAt and/or edm:isShownBy do not resolve, respectively, a web page or a digital object). We will discuss these issues with you when they occur. If we cannot resolve these issues within three months we may remove the data from Europeana API responses and thus also from Europeana Collections.

Summary: acceptance criteria for technical validation

- Aggregators and data partners must follow the specifications and guidelines for EDM or ESE as the metadata is automatically validated according to the schema during the ingestion process.

- The 10 mandatory metadata elements (specified by the EDM mapping guidelines) must be populated with meaningful and correct values before submitting data to Europeana.

- Each metadata record must have either a title (dc:title) or a description (dc:description).

- Each metadata record describing a text object (e.g. book, manuscript, letter) must have the language of the document in the metadata (dc:language).

- Each metadata record must have the type of the digital object specified in edm:type.

- Each metadata record must provide some context and details about the objects described by the metadata (dc:subject, dc:type, dcterms:spatial, dcterms:temporal).

- Each metadata record describing digital objects contributed by users must show edm:ugc = true.

- Each metadata record must give the name of the actual cultural heritage institution that provides it to an aggregator (edm:dataProvider).

- Each metadata record must give the name of the direct data provider to Europeana (edm:provider).

- Each metadata record must have at least one link (URL) to the digital object on a website of a cultural heritage institution or to an aggregator (edm:isShownAt or edm:isShownBy).

- Each metadata record must be supplied with a valid rights statement applicable to the digital object, using edm:rights and the corresponding URI to the rights statement.

- Each resource described in a metadata record must have a unique and persistent identifier (e.g. rdf:about of the ProvidedCHO class in EDM, and edm:aggregatedCHO).

- Each digital object must be an adequate digital representation of an object that is part of Europe's cultural and/or scientific heritage. Audio/video snippets or a subset of pages for a multipage document (e.g. book, report) are not an adequate digital representation of a cultural heritage object.

- Digital objects that are image or video files must be provided with a direct link to an image file of at least 0.1 megapixel in size to have the means to create ~400 pixels preview images.

4. Legal: how digital objects should be labelled

Every digital object must be labelled with a valid rights statement in the edm:rights field and a corresponding URI to the rights statement. This will make it clear to the user as to how your digital object may be used. The rights statement you select may depend on a number of factors, such as the date and type of work and knowing if you have permission to share someone else's work.

From November 2016, data partners will be able to choose from one of 14 rights statements[33]. This transition[34], which replaces the existing Europeana specific rights statements with statements provided by rightsstatements.org, is the outcome of a collaboration between Europeana, DPLA, Kennisland, and Creative Commons to create international interoperable rights statements. This update will make it easier for data partners to correctly communicate the copyright status of their works online, and also help users understand what they are able to do with the digital objects they discover in Europeana and other digital platforms.

Under the DEA, all metadata must be submitted under CC0. This is different to the rights for the digital object, which should always be in accordance with the copyright status of the work. Use the Rights Statement Selection tool[35] to help you choose the most appropriate statement.

4.1. How do I know my edm:rights statement is correct?

Providing a valid rights statement is not just about passing our technical validation. It is fundamental to the philosophy and mission of Europeana, which is to ensure that all digital objects are published with a rights statement that is valid. This means that every person who discovers Europeana's collections knows exactly what they can (and cannot) do with every item.

A rights statement is valid if it accurately describes the copyright status (in or out of copyright) of the digital object. Cultural heritage institutions have a responsibility to respect the copyright holders of the works they represent in order to enable public access. It is therefore the responsibility of each organisation to comply with their national law to establish the copyright status of each work during their rights clearance process, prior to submitting metadata for publication. Tools such as the Public Domain Calculator[36] should help you identify the existence or absence of copyright.

Europeana takes a "clean hands" approach and assumes that the data partner has

---

[33] http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_DSI/Milestones/europeana-dsi-ms21-review-of-rights-statements-under-the-europeana-licensing-framework.pdf
[34] http://pro.europeana.eu/blogpost/finding-the-right-fit-to-share-your-cultural-content
[35] http://pro.europeana.eu/get-involved/europeana-ipr/tools-resources/tools/the-europeana-rights-statement-selection-tool
[36] http://outofcopyright.eu/

undertaken the correct level of due diligence and labelled the digital objects correctly. However, we also wish to support data partners to make sure rights statements are applied correctly and in line with their respective national laws and sector standards. In order to do this the use of the rights statements described in the following chapters will prompt a manual review during the ingestion process (prior to publication), and we may at this point question the use of some rights statements. This approach enables a consistent standard to be reached when applying rights statements, which ultimately helps users to be clear on how the objects and collections can be used. Please see the examples in the following sections.

4.2. Public Domain Mark (PDM)

Europeana is committed to the principle that the digitisation of public domain content does not automatically create new rights over it.

The Europeana Public Domain Charter[37] establishes that 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.

For instance, in the EU copyright usually lasts for 70 years after the death of the last surviving author. This means, for example, that all the works of Piet Mondriaan (who died in 1944) are now in the Public Domain.

Some national laws allow you to claim a new copyright or copyright like rights as a result of digitisation. (We have undertaken research into these national scenarios; they are published on the Out Of Copyright[38] website). In these cases we encourage you to waive such rights via CC0 (see section 4.4). If you do not want to waive these rights you can apply one of the Creative Commons licenses or one of the In copyright statements (as described in section 4.6 and 4.7) to the digital object.

Another exception is where a commercial contract has been entered into for the mass digitisation of your collection, here the 'No Copyright - Non-Commercial Use Only' statement may be applicable. This is described in section 4.3.

Finally it may be possible that laws other than copyright restrict how an out of copyright digital object can be re-used. Such laws include public sector information laws (in France) or national heritage laws (for example in Italy). In such cases the 'No Copyright - Other Known Legal Restrictions' statement (see section 4.5) may be applicable.

It is likely that, before we publish your data, we will query the submission of any objects that appear to be in the public domain but which are not labelled as such. We will work with you to determine the correct rights statement and subsequently

---

[37] http://pro.europeana.eu/publications/the-europeana-public-domain-charter
[38] http://outofcopyright.eu/rights-after-digitisation/

publish the data. We pay particular attention to metadata fields like dc:date, dcterms:created and dc:creator to verify the existence of copyright.

4.3. No Copyright - Non-Commercial Use Only (NoC-NC)

Update November 2016: Previously titled 'Out of copyright - non-commercial re-use', definition unchanged.

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). For this statement to be applicable, the partners in the PPP will have agreed to contractual limitations to take reasonable steps to limit or discourage commercial re-uses.

In order to use NoC-NC, the data provider should be able to supply a copy of the contract that specifies these restrictions. If such information is publicly available, the data partner should also specify a year of expiration in the digital objects' metadata to indicate the first calendar year in which the digital object(s) can be used by third parties without restrictions on commercial use.

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 objects, a data partner waives all possible rights to 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 partner can only waive rights held by the data partner (or third party rights for which the data partner has explicitly received permission to apply CC0 from the rights-holder(s)).

Europeana will verify the correct application of this rights statement prior to publication. If there is doubt that the data partner is the rights-holder and is able to waive all rights, we will work with you to see if you have the necessary rights to apply a CC0 public domain dedication or to determine the correct rights statement, prior to any publication on Europeana.

4.5. No Copyright - Other Known Legal Restrictions (NoC-OKLR)

Update November 2016 - New rights statement added.

The NoC-OKLR statement should be used for objects that are in the public domain, but that cannot be freely reused due to known legal restrictions other than copyright that prevent the data partner from allowing free reuse. An example of where a law restricts free reuse is Italy: the Code of the Cultural and Landscape Heritage, Article 107 where reuse of national heritage requires permission even when copyright has expired.

In order to support the correct implementation of this statement, a data partner

must communicate the legal restriction that applies to the Europeana Operations Officer working with you.

4.6. In Copyright (IC)

Update November 2016 - New rights statement added. Subsumes Rights Reserved-Free Access and Rights Reserved-Paid Access.

The IC statement is for use with objects that are subject to copyright.

Any conditions or restrictions relating to the reuse of the object should be communicated with the user via dc:rights.

4.7. In Copyright - Educational Use Permitted (InC-EDU)

Update November 2016 - New rights statement added.

This rights statement should be used where the rightsholder has authorised the use of the object for educational purposes.

We ask data providers to confirm that they have correct permission from the rights holder. When they do, we ask them if permission is granted for anything in addition to educational use. In such instances it may be that a different rights statement such as a CC licence is more suitable. Additional information on the educational uses permitted by the rightsholder is encouraged to be added to the dc:rights field.

4.8. In Copyright - EU Orphan Work (InC-OW-EU)

Update November 2016: Previously titled 'Orphan Work' — definition unchanged.

The use of this rights statement is limited to digital objects that fall within the scope of the Orphan Works Directive.[39] For digital objects to be eligible for this rights statement, orphan works legislation implementing the Orphan Works Directive must exist in the data partner's country of origin.

In order to apply the InC-OW-EU statement, data providers should be able to confirm that the required due diligence has been undertaken in accordance with the appropriate national legislation, and that they have taken steps required to register the Orphan Work with their national representative. Where available, a link to the published record of the Orphan Work in the National or European registry should be provided in dc:rights.

4.9. Copyright not evaluated (CNE)

Update November 2016: Previously titled 'Unknown', definition updated.

---

[39] http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32012L0028

The CNE statement is for use with objects where the copyright has not been evaluated.

A data provider should always evaluate the copyright status of an object and obtain permission from the copyright holder where necessary prior to making it available online. For this reason, Europeana expects the use of this rights statement to be limited, and any use by a data provider requires prior discussion and agreement of the Europeana Operations Officer.

4.10. Consistency of rights statements

The dc:rights field is often used to describe additional information about the rights in the digital object (than provided in the standardised rights statements).

We intend to enable edm fields in 2017 that allow you to submit additional information for the NoC-NC, NoC-OKLR and InC-EU-OW. Until then you can use this field for additional information, such as the year of expiration for objects labelled NoC-NC and other information necessary for some of the rights statements presented above.

Data partners 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:rights is 'Public Domain', and dc:rights contains a statement such as '© Cultural Heritage Institution 2014'.

In addition, it is best practice to ensure that the information presented in both edm:rights and dc:rights is consistent with the information presented on the data provider's website.

4.11. How we work with data partners to correct data post publication

Europeana not only checks and validates the accuracy of the rights statements prior to publication as set out above, we analyse the data also post publication to ensure and verify a consistent level of accuracy of rights statements across the database.

There may be situations post publication where it becomes clear that the chosen rights statements do not meet the acceptance criteria. We will discuss these issues with you when they occur. If we cannot resolve these issues within three months we may remove the data from Europeana API responses and thus also from Europeana Collections.

Summary: Acceptance criteria for rights statements

- Digital objects must be submitted with valid edm:rights and a URI corresponding to the rights statement.

- Digital objects that are in the public domain must show 'Public Domain Mark' in the edm:rights field unless a verified exception applies.

- Digital objects submitted with 'No Copyright - Other Known Legal Restrictions (NoC-OKLR)' for edm:rights will only be accepted for publication if the data partner provides information in their metadata about the legal restriction and when it expires.

- Digital objects submitted with 'No Copyright - NonCommercial Use Only' for edm:rights need to show a valid contract that imposes a restriction on the commercial reuse of the digitized public domain works. An end date to the restrictions on commercial reuse should be recorded in the metadata of the digital objects.

- Digital objects submitted with CC0 for edm:rights will only be accepted for publication once double checked by Europeana.

- Digital objects submitted with 'In Copyright - Educational Use Permitted (InC-EDU)' for edm:rights will only be accepted for publication if the data partner is able to explain the decision in being able to make works available using this rights statement.

- Digital objects showing 'In Copyright - EU Orphan Work' for edm:rights will only be accepted for publication if orphan works legislation exists in the data partner's country of origin and if the data partner confirms that they have undertaken due diligence to the standard defined in their respective national orphan works legislation.

- Digital objects with 'Copyright not evaluated' for edm:rights will only be accepted with prior discussion and agreement with Europeana.

- 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	● Final version of v1.0 incorporating previous drafts and contributions as well as comments and suggestions from colleagues and reviewers.
v1.1	Henning Scholz	04/03/2015	● Corrected wording in two places; reference to the research on national legislative provisions for PD works; links updated (new Europeana Pro).
v1.2	Henning Scholz	11/05/2015	● Copy editing and revision by Beth Daley and Jill Cousins: language / tone of voice changes to make it more clear and readable; removing the section on prioritizing of metadata for publication when digital objects are openly licensed and directly linked. With this edition we widened the background to refer to the *Europeana Publishing Framework*. We also added some more context to the mandatory elements. We boxed the acceptance criteria to make them more easily accessible as a quick guide. The name changed from 'Europeana Publication Policy' to 'Europeana Publishing Guide', as part of the work on the *Europeana Publishing Framework*.
v1.3	Henning Scholz	09/11/2015	● Amendment by Jeroen Cichy to reflect the change of the publication process from monthly cycles to continuous publication. Minor correction of spelling.
v1.4	Henning Scholz	26/10/2016	● Amendment by Julia Fallon, Lisette Kalshoven, Roxanne Peters to adapt the legal section with rightsstatements.org rights statements. Review by Annick van der Schoot and Marjolein de Vos to improve the document.
v1.5	Henning Scholz		● Amendment by Julia Fallon, Paul Keller to adapt the legal section; amendment by Henning Scholz, Joris Pekel, Erwin Verbruggen, Marco Rendina, Uldis Zarins to the technical criteria per EPF tier; addition of a guide to depublish from Europeana by Henning Scholz, Roxanne Peters, Julia Fallon. Copy editing and proofreading by Annick van der Schoot.

6. Acknowledgements