→ **Europeana Publishing Guide v1.3**

A guide to the metadata requirements for data partners publishing their collections via Europeana

Table of Contents

1.   Background and Purpose

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

The *Europeana Publishing Framework* 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 organization, or particular collection, whether you want to provide minimum level 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 over 45 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][1] and the [Business Plan 2015][2].

So 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][3] to the [Public Domain Charter][4], but there has been no single document that clarifies all our policies for publication. This document brings all the existing information into one place. It is intended to help aggregators and data partners share their data, become fully standardized and interoperable if desired and to improve Europeana in a uniform and consistent way.

The *Europeana Publishing Guide* outlines the criteria for submission of metadata to Europeana. We will work with you to make sure that your datasets meet our 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 seen to be authentic, trustworthy and robust by our audiences. The acceptance criteria are based on real life examples as well as existing documentation and policies (e.g. [EDM mapping guidelines][5], [Europeana Licensing Framework][6]).

As of 2016, the acceptance criteria will also be applied to legacy material (material already published on Europeana). This means data quality will be improved and metadata not

---

[1] http://strategy2020.europeana.eu/

[2] http://pro.europeana.eu/publication/make-the-beautiful-thing-business-plan-2015

[3] http://pro.europeana.eu/get-involved/europeana-ipr/the-data-exchange-agreement

[4] http://pro.europeana.eu/publications/the-europeana-public-domain-charter

[5] http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation

[6] http://pro.europeana.eu/get-involved/europeana-ipr/europeana-licensing-framework

compliant with the criteria can 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*.

The acceptance criteria are divided into three sections:
- Process: how to provide data to Europeana.
- Technical: the minimal technical criteria.
- Legal: how digital objects should be labelled.

2.   Process: how to provide data to Europeana

This section specifies the acceptance criteria with regards to the ingestion workflow. It gives exact deadlines for each part of the process and guidelines for prioritizing your metadata for publication. Figure 1 provides a clear visualization of the process.
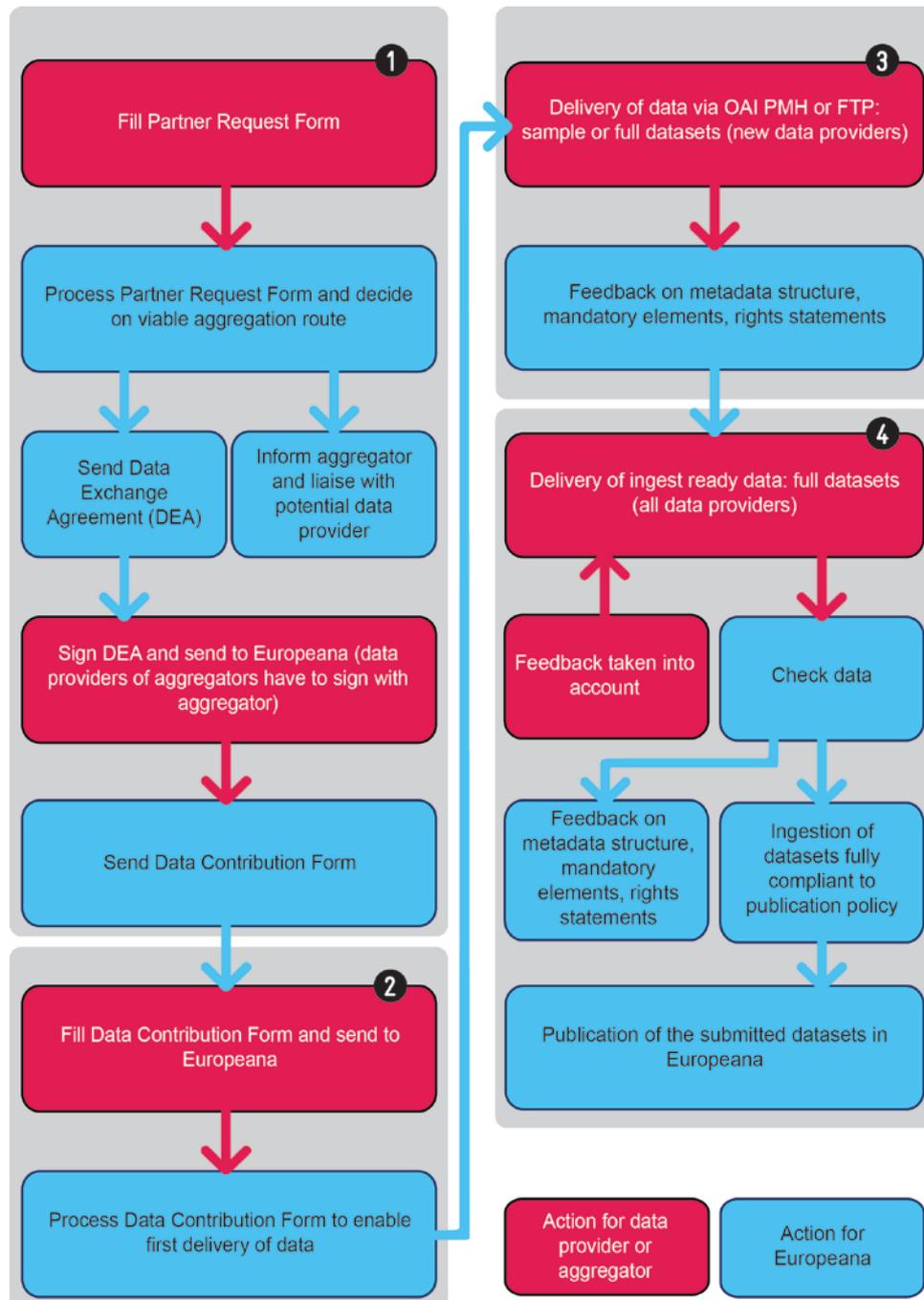

Fig. 1. *The process for becoming a data partner and publishing datasets on Europeana. The four steps labelled in the chart are explained in more detail in the text (chapters 2.1-2.3).*

2.1.   Potential data partners

Every cultural institution in Europe with digitized collections can provide them to Europeana. Currently, more than 3,500 data partners, working largely through 110+ aggregators, publish collections on Europeana. To keep this number of partners and datasets manageable, all institutions (interested in) providing data to Europeana need to follow the process we've developed. (See Fig. 1 and http://pro.europeana.eu/share-your-data/how-to-contribute-data for more detail).

A potential data partner must first fill in the Partner Request Form (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 can be established between them.

2.2.   Data Exchange Agreement

The Europeana Data Exchange Agreement (DEA)[7] establishes the terms under which Europeana can make use of the previews and descriptive metadata provided by cultural institutions. Before their 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.

Potential aggregators (including European Commission funded projects) must ensure that all their data partners have either signed the DEA or an agreement with the aggregators that reflects the terms of the DEA (see more information and the template clauses for aggregators on Europeana Professional[8]). New data partners should always include an accompanying letter with the signed DEA, stating through which aggregator they will be delivering their metadata.

2.3.   Data contribution workflow

Before the first submission of data to Europeana, a data partner must fill in the Data Contribution Form (see step 2 in Fig. 1). Europeana will evaluate the form and invite the data partner to deliver data to Europeana.

The first delivery of data should be a full sample dataset (not just 2-3 records) 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 *Europeana Publishing Guide*. If it is not, the dataset will have to be corrected and be re-submitted. We will provide feedback within two weeks after the date of submission. From your second delivery onwards, full datasets smaller than 200,000 records can be submitted once a month. Larger datasets can only be submitted quarterly.

---

[7] http://pro.europeana.eu/get-involved/europeana-ipr/the-data-exchange-agreement

[8] http://pro.europeana.eu/get-involved/europeana-ipr/the-data-exchange-agreement

All data will be checked within two weeks after the date of submission. If everything is fine, it 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.

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. Although, some priority may be given to specific collections that have a strict project deadline.

Data will be processed and published continuously throughout the year. We will inform you if unexpected technical problems prevent us from properly processing all the data submitted.

Acceptance criteria for data and how we prioritize data publication - summary

- 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 Data Exchange Agreement 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 Data Exchange Agreement and must ensure their data partners have signed an agreement that equally reflects the terms of the DEA. Europeana has to operate a clean hands policy here, so trusts that the aggregator is maintaining full and proper records.

- Metadata is accepted for publication after the 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 quality accepted datasets. We will prioritize data that is required for specific projects.

3. Technical: what are the minimal technical criteria?

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, to make the data more discoverable. This section specifies 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 format specified in the Europeana Data Model (EDM)[9], metadata submitted using the Europeana Semantic Elements (ESEv3.4)[10] format is still accepted.

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 further processed for publication. We will provide you with feedback to help you 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. For example, 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 how to apply EDM or ESE, please see the technical documentation for EDM and ESE on the Europeana Pro website.

3.2. Mandatory elements

There are 10 mandatory elements in EDM that must be present for the metadata to be further processed for publication (see below and the technical documentation for EDM[11] for more details). Existence of metadata elements are validated automatically, while the quality of the metadata provided in these fields is manually checked and validated.

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. 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) Every metadata record describing a text object (e.g. book, manuscript, letter) must have the language of the document in the metadata (dc:language). *This allows us to present more material against the language requirements of the user.*

---

[9] http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation

[10] http://pro.europeana.eu/share-your-data/data-guidelines/ese-documentation

[11] http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation

3) Every metadata record must have the type of 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. *Many users want to search according to media, accurate completion of this field allows them to do so.*

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). *The more data you can give here the more opportunity the user has to find it against their search terms.*

5) Every metadata record describing digital objects contributed by users (e.g. during a public digitization 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) Every metadata record must have 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) 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 partner can give a link to the digital object in the context of the organization's website or a link to a (book or image) viewer that shows the digital object (edm:isShownAt). However, it is 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). This generates a lightbox in Europeana Collections (edm:isShownBy), improving the user experience within Europeana. *We highly recommend that you provide both edm:isShownAt and edm:isShownBy along with the data, so that the user is better satisfied and more likely to use your 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 Pro](http://pro.europeana.eu/share-your-data/rights-statement-guidelines/available-rights-statements)[12]. 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, encouraging them to do the right thing. It also makes machine readable use of the data possible, improving its interoperability enormously.*

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 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, again improving the user experience and likelihood of use of your material.*

---

[12] http://pro.europeana.eu/share-your-data/rights-statement-guidelines/available-rights-statements

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

Acceptance criteria for technical validation - summary

- 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 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, dcterms:spatial, dc:coverage).

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

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

- Every metadata record must give the name 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

Every digital object must be supplied with a valid rights statement in the edm:rights field and a corresponding URI to the rights statement. The list of valid rights statements[13] is published on Europeana Pro. Use the Rights Statement Selection tool[14] to help you choose the most appropriate statement.

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 legal copyright status of the work.

4.1.   Validating edm:rights

Providing a valid rights statement isn't just about passing our technical validation. It's 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. So that every person who comes into contact with Europeana's collections knows exactly what they can (and can't) do with every item.

We consider a valid rights statement to be one that is based on the existence or absence of copyright. Cultural heritage institutions are largely public and wish to uphold the law of their country – it is therefore important that the process for establishing the existence or absence of copyright should be undertaken by each organization during their rights clearance process, prior to submitting metadata for publication. Tools such as the Public Domain Calculator[15] should help you to identify the existence or absence of copyright.

Europeana has a clean hands policy here and will assume that the data partner has undertaken the correct level of due diligence and labelled the digital objects correctly. However, because we also wish to help the user and help the data partner to improve and conform to standards, the use of certain rights statements will prompt a manual review during the ingestion process (prior to publication) and we may at this point question some rights statements. Please see the examples in the following sections.

4.2.   Public Domain Mark (PDM)

Europeana is committed to the principle that the digitization of public domain content does not automatically create new rights over it. According to the Public Domain Charter,[16] works that are in the public domain in analogue form should continue to be in the public domain once they have been digitized. A work is in the public domain when its copyright does not exist or has expired.

---

[13] http://pro.europeana.eu/share-your-data/rights-statement-guidelines/available-rights-statements

[14] http://pro.europeana.eu/get-involved/europeana-ipr/tools-resources/tools/the-europeana-rights-statement-selection-tool

[15] http://outofcopyright.eu/

[16] http://pro.europeana.eu/publications/the-europeana-public-domain-charter

For instance a work created in 1820 will no longer be protected by copyright and should be labelled with the rights statement: 'Public Domain Mark'.

It may be that your national laws sometimes allow you to create a new copyright or similar temporary restriction (which is not copyright-based) as a result of digitization. (We have undertaken research into these national scenarios and they are published on the [Out Of Copyright](#)[17] website). For such works we prefer the use of the 'Rights Reserved - Free Access' statement but are looking at how to best represent these. Another exception is where a commercial contract has been entered into for the mass digitization of your collection, here the 'Out of Copyright – Non Commercial Re-Use' statement may be applicable. This is described in the following section.

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 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.   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 digitized 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.

If a data partner wishes to use OOC-NC, there should be a contract that specifies these restrictions. A data partner should also try to 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.

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 partner 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 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.

---

[17] http://outofcopyright.eu/rights-after-digitisation/

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.](url)[18] 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 partner. The data partner should have undertaken and be able to point to due diligence to the standard defined in the national orphan works legislation. We are happy to work with you to verify the correct application of the rights statement.

4.6.   Unknown rights statement

Data partners submitting datasets with digital objects labelled 'Unknown' should be able to explain why no conclusive information pertaining to the rights status of the digital object is available. The application of this rights statement should be based on the explanation for the ['Unknown' rights statement](url)[19]. As the 'Unknown' rights statement does not provide any legal clarity for users of Europeana or your work, it should only be used as a last resort.

4.7.   Consistency of rights statements

The dc:rights field is often used to describe additional information about the rights in the digital object. 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'.

---

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

[19] http://pro.europeana.eu/share-your-data/rights-statement-guidelines/available-rights-statements

Acceptance criteria for rights statements - summary

- 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 'Out of Copyright – Non Commercial Re-use' (OOC-NC) for edm:rights need to show a valid contract that imposes a restriction on the commercial re-use of the digitized public domain works. An end date to the restrictions on commercial re-use 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 showing '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 partner and if the data partner has undertaken due diligence to the standard defined in their national orphan works legislation.

- Digital objects with 'Unknown' for edm:rights should be justified by the data partner as to 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	• 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 public domain works; links updated (new Europeana Pro).
v1.2	Henning Scholz	11/05/2015	• Copy editing and revision by Beth Daley and Jill Cousins. This resulted in language / tone of voice changes to make it more clear and readable. It also resulted in 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.

6.   Acknowledgements