



europæana
food and drink

Grant Agreement 621023

Europeana Food and Drink

D2.3 Guidelines on the Europeana Content Re-use Framework

Deliverable number	<i>D2.3</i>
Dissemination level	<i>PU</i>
Delivery date	<i>January 2015</i>
Status	<i>Final</i>
Author(s)	<i>Elena Lagoudi (PS)</i>



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

Revision History

Revision	Date	Author	Organisation	Description
V1.0	21/01/2015	Elena Lagoudi	PS	First draft
V1.0	28/01/2015	Laura Gibson	CT	First review & comments
V2.0	30/01/2015	Elena Lagoudi	PS	Final document

Statement of originality:

This deliverable is significantly indebted to Europeana Creative Guidelines for creative re-use, part of the wider Europeana project. Other than adaptations from this source, it contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Contents

1 Introduction	5
1.1 Background.....	5
1.2 Role of the deliverable in the project	5
2 Approach	7
2.1 Methodology	7
2.2 Structure of the deliverable	8
3 Functional requirements	8
4 Content Re-use Guidelines	9
5 Licensing	10
7. Results and impact	12

1 Introduction

1.1 Background

The core concept of the Europeana Food and Drink Best Practice Network is to kick-start the creative and commercial re-use of digital content relating to food and drink from the culture sector to drive new commercial applications, relationships and partnerships.

Food and drink serves the dual purpose of providing a powerful thematic focus to inspire creative re-use of digital cultural content while offering sufficient breadth to support a wide range of applications and approaches.

As demonstrated by various projects within the Europeana ecosystem, the key issue with creative and commercial re-use seems to be findability and exploitability of assets.

Users of digital cultural content seek and retrieve it from aggregators and repositories where they can browse, search and find the desired content amongst a mosaic of diverse content publishers. The interfaces for content discovery often offer poor user experience and lack the functionality of intuitively presenting diverse cultural material, as metadata expressivity and overall quality varies.

There is a need for more detailed, domain-specific curation, presentation and publication and Europeana Food and Drink can demonstrate its benefits through its Content Base/Picture Library which aims to offer coherent navigation, user-friendly browsing and search functions, meaningful documentation that supports creative and commercial re-use, multi-linguality and easy-to-understand licensing choices.

1.2 Role of the deliverable in the project

WP2 as a whole will support the objective of identifying, describing, enhancing, licensing and uploading a body of high-quality digital assets and their associated metadata, to support the delivery of commercial applications and public engagement activity.

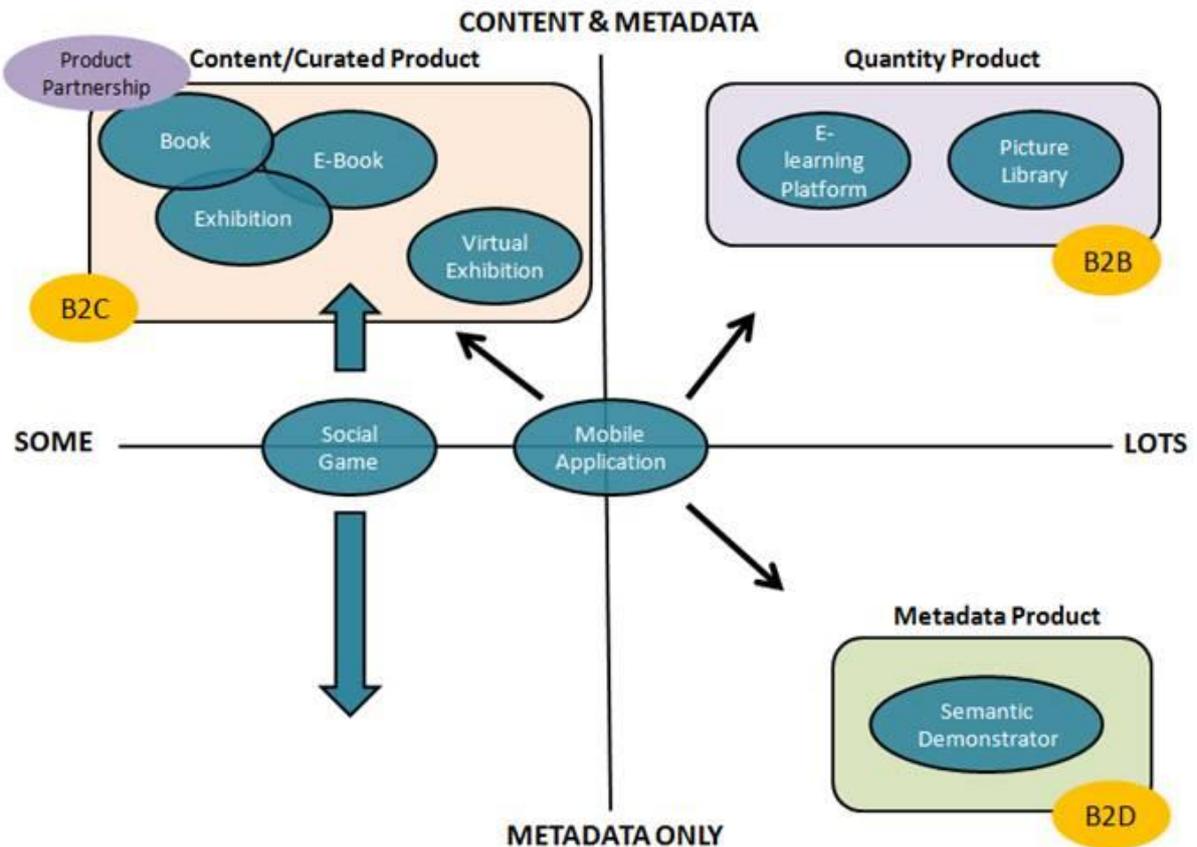
D2.3, Guidelines on the Europeana Content re-use framework will support creative and commercial re-use of Europeana Food and Drink content, offering guidelines to facilitate with the technical aspects and specifications.

The Content Base developed throughout the project will be available for cultural institutions, creative industries, professional users and third party developers in order to easily search for the cultural resources that meet their retrieval criteria so as to use and re-use them for the development of applications, products and services.

The result will be a body of high quality digital assets and semantically-enriched metadata that can support a wider range of (multi-lingual) natural language applications such as search, discovery and browse.

The Europeana Food and Drink Content Base will feed content into the applications and products the Consortium will develop. Based on the quantity of content and metadata needed, the applications were split into the following diagram:

2.1 Europeana Food and Drink D.2.3 Guidelines for the Creative Re-use



As it is crucial to build on the potential coordination and cooperation with other activities within the Europeana ecosystem (mainly Europeana Creative and Europeana Space), the Guidelines developed by Europeana Creative will be used for Europeana Food and Drink, in order to ensure the interoperability of available platforms.

This document reports on the requirements with respect to content and metadata, in particular with respect to formats, types of access and licencing.

2 Approach

The initial content that the consortium content partners had committed to was described in a table included in the DoW and analysed in detail in D2.1, Inventory of Europeana Food and Drink Content.

Here are some key findings about the Europeana Food and Drink content:

- The content comes from a variety of Cultural Heritage organizations, ranging from Ministries to academic libraries and specialist museums to picture libraries.
- The content represents a significant number of European nations and cultures
- The content comprises of objects illustrating Food and Drink heritage, recipes, artworks, photographs, some audio and video content and advertising relating to Food and Drink
- The content is heterogeneous in types and significance, but with the common thread of food and drink heritage and its cultural and social meaning
- The content is mainly documented in Dublin Core.
- The content metadata are available partly in English and native languages, with almost half of metadata only available in native languages

The content will be ingested into Europeana via MINT (NTUA) and was to be hosted in Europeana Cloud. However, since Europeana Cloud is not yet ready, this initial select body of high-quality content which has been identified by the consortium for its uniqueness and commercial potential will be ingested into a Digital Asset Management Repository in order to test the entrepreneurial models the Consortium is developing to facilitate commercial re-use.

It was deemed necessary to produce a platform where content and its metadata is discoverable and where commercial activity can be supported. Therefore, it was decided that the Content Base will effectively incorporate the Picture Library product (see schematic above), so that the commercial viability of the project can be piloted. Therefore, the body of content to be aggregated for this initial phase of the project will be forming a Picture Library for commercial re-use.

Once this initial content base/picture library has been developed and tested and the work-flow for ingestion and mapping has been established, the remaining content will be contributed from the consortium and future content providers, as the Consortium expands to include more and diverse organizations and harvested by Europeana in compliance with the Description of Work, thus enabling products such as the Semantic Demonstrator to be built on top of a mass body of thematic content.

2.1 Methodology

The project leader and other project partners participate in other projects in the Europeana ecosystem, especially in related projects such as Europeana Creative, Europeana Space, Europeana Inside and Europeana Photography. It was deemed vital to try and integrate findings that related projects reported on and good practices they recommend. Therefore, the Guidelines produced by Europeana Creative and the “Requirements for the creative use of Digital Cultural Resources” produced by Europeana Space will be drawn upon for re-use by Europeana Food and Drink, to ensure there is continuity and integration of practice and policies.

The requirements that guide the development of Europeana Space’s “Technical Space” will be used for the specifications for the Content Base/Picture Library, as “Technical Space” will essentially perform a similar role to Europeana Food and Drink’s Picture Library. Although the requirements are not guidelines for re-use, they are an important building block and were included in this deliverable, as they outline what the platform needs to be able to do to facilitate access to this content and, as it is being developed during this phase of the project, it is vital to incorporate them to inform decision making.

One key difference is that Food and Drink is not only aiming at creative re-use, but also commercial re-use. Therefore, it is important to distinguish the difference between the two business models and what they are trying to achieve. Licensing plays a key role for this, so the Guidelines for the creative re-use incorporate a business and commercially sustainable value proposition.

2.2 Structure of the deliverable

The deliverable consists of a word document outlining first the functional requirements of the content base/picture library and then the content requirements and finally the licencing developments that need to happen.

3 Functional requirements

Initially the functional requirements for the Content Base/Picture Library (from now on referred to as CB/PL) will be outlined:

Functional requirements

1. CB/PL should provide storage and access to high quality content for use by web and mobile applications.
2. Types of content include image, video, audio and text files in various formats.
3. Metadata should include rich thematic fields to facilitate browsing and searching
4. Metadata should include a rights statement to define the conditions for re-use.
5. Access for project partner and developers may follow different licensing strategies for content, which may then be available differently to a wider audience (e.g. through the project’s hackathons) or the public
6. Support and guidance for storing and accessing metadata records.
7. Access types include HTTP upload and download, OAI-PMH protocol and HTTP API
8. Any data model will be allowed when ingesting while aggregation and publication can be performed using EDM and available profiles or more domain specific standards such as LIDO.
9. Record serialization can be CSV, XML or JSON with preview interfaces for raw data and available HTML renderings.
10. The system should offer the ability to automatically convert between data models for which there are established, formalized crosswalks
11. Filtering and merging of datasets should be allowed.
12. CB/PL should offer dataset Indexing and statistics

13. CB/PL should offer 3rd party APIs integration for services such as Wikimedia Commons, Flickr, YouTube or Vimeo and plugins for crowdsourcing/user generated content addition
14. Should offer metadata cleaning and enrichment services

Non-functional requirements

1. Performance (response time, throughput, efficient resource usage for specific performance requirements)
2. Scalability (number of organizations and users, ease of resource allocation to accommodate changing load)
3. Availability and Recoverability (ability to maintain an accepted level of performance over time, recovery from errors)
4. Data Security and Integrity
5. Usability (efficiency, documentation, ease to learn, satisfying for a target user community)
6. Interoperability (use services from and provide services to other systems for Digital Cultural Heritage)

4 Content Re-use Guidelines

In alignment with Europeana Creative, we make the following recommendations for each Digital Object type:

a. Images

- The minimum quality of images is indicated at image resolution of 1,280 x 800 pixels
- Recommended file types are jpg, png and tiff
- There are no visible watermarks

These requirements are in line with the current practice and availability of content in Europeana.

b. Video

- The minimum quality of video is indicated at a resolution of 1,280 x 720 pixels (720p) However, the reality of video content provided by the heritage sector (digitised television content) is that this excludes all content from the SD era. Therefore we recommend a somewhat lower minimum quality of 704 x 576 pixels (576p).
- Recommended bitrate of 2,000 Kbs (2 Mbit). However, this quality is much higher than the previously indicated resolution of 720p. Because of the contradictory nature of this outcome, we recommend not taking this as a requirement to build upon.
- Recommended file types for video are MP4, mov, WEBM
- There are no visible watermarks, although it will be hard to enforce this requirement for digitised television content, as television broadcasters often use watermarks. We therefore do not recommend to include this as a hard

requirement, but as something that would be nice to enforce, whenever possible.

c. Audio

- The minimum quality of audio is indicated at a sample rate of 41,100 kHz
- 16 Kbits
- Recommended file type is MP3. We recommend that besides MP3 all lossless file formats like FLAC, WAV and APE should be added as requirement for the minimum technical quality of audio objects.

d. Text

- The minimum quality of texts is indicated at full text searchable
- Recommended file types are: pdf, txt, epub, xml and/or rtf

e. 3D Objects

- The minimum quality of 3D Objects is indicated at: file types OBJ and BLEND
- As too few people have answered questions in this section of the survey produced by that Europeana Creative, more research is needed to come to hard conclusions on minimal quality requirements for this type of Digital Object so we do not include this as a minimum requirement yet.
- This is subject to revision as soon as the sector agrees on desirable standards. COLLADA Digital Asset and FX Exchange Schema is an initiative to try and define standards for 3D objects. We are closely following their work to inform us of the appropriate standards for creative re-use of 3D objects.
- Likewise, there is not yet widespread standardisation of the hardware and software needed to create or display 3D digital models and so this must be monitored so that content can be shared

5 Licensing

To allow interaction between content providers and creative industries, the Europeana Licensing Framework is currently being extended with a layer that governs access and re-use conditions for the content itself in addition to metadata. The Content Reuse Framework specifies the Content Layer of the Extended European Licensing Framework¹. It is closely aligned with the work that is undertaken in the Europeana Cloud project on a cloud-based storage infrastructure.

5.1. Authorisation for Reuse

Rights edm:rights needs to be one of edm:rights statements that allow reuse (Public Domain Mark, CC0, Out of Copyright – noncommercial reuse¹ or one of the six Creative Commons licenses) or a conditional rights statement.

5.2. Modifications to the rights information in EDM

The implementation of the Content Reuse Framework requires some modifications and additions to the way rights information about Digital Objects is stored in the Europeana Data Model (EDM). The introduction of the Content Reuse Framework means that EDM needs to be able to model Cultural Heritage Objects that reference multiple Digital Objects (edm:WebResource) with heterogeneous edm:rights statements. In addition, the Content Reuse Framework supports conditional access rules that can be used by

data providers to limit access to specific Digital Objects to end users that meet predefined conditions (such as: the end user is an educational user).

The Content Reuse Framework requires that the rights status of each Digital Object is being checked and that only those Digital Objects that carry a rights statement that allows reuse are included in the Content Reuse Framework. As a result it is necessary to allow data providers to record a `edm:rights` statement for each individual Digital Object that they provide.

This requires modifying the scope of the `edm:rights` statement from the level of the metadata record (`ore:Aggregation`) to also apply at the level of the individual Digital Objects (`edm:WebResource`):²

5.3. Adding Conditional Rights Statements to the Europeana Data Model

In addition to the existing `edm:rights` statements, Digital Objects can be included in the Content Reuse Framework if they carry a conditional rights statement that limits access (and subsequent reuse) to end users that meet predefined conditions.

- Conditional rights statements (`temp:conditionalRights` or similar) apply to individual Digital Objects in the same way as `edm:rights` statements.
- An `edm:WebResource` can have both a conditional rights statement and an `edm:rights` statement. If a data provider wishes to make the same Digital Object available under both a conditional rights statement and an `edm:rights` statement, then two instances (`edm:WebResource`) of the Digital Object must be provided.
- This task includes the development of a mechanism (such as a controlled list of values for `temp:conditionalRights`) that ensures that only approved conditional rights statement can be applied to Digital Objects.
- In addition there must be a mechanism that ensures that each metadata record (`ore:Aggregation`) in Europeana contains at least one Digital Object (`edm:WebResource`) that carries a valid `edm:rights` statement. This is to prevent that Europeana contains metadata records related to Cultural Heritage Objects without a publicly available Digital Object.

5.4. Developing a Syntax for Conditional Rights Statements

A syntax to express the conditions for access and reuse of a conditional rights statement needs to be developed.

- This syntax needs to be flexible to allow the modeling of various types of access conditions.
- In addition, the rights statement needs to be able to express the conditions under which an authorised user may reuse the Digital Object.

Both elements need to be encoded in the conditional rights statement in a way that allows users to clearly understand under which conditions they can access and reuse a Digital Object. And that allows storage providers⁴ that host these Digital Objects to determine if access should be granted to a particular user. The conditional rights statement is not intended to let a storage provider or the data provider enforce the reuse conditions (they are not intended to be used as DRM/TPM).

5.5. Specifying and Implement Access Controls

Digital Objects that carry a conditional rights statement are not publicly available. As such there is a need for an access control mechanism that corresponds to the access conditions expressed in the conditional rights statement that are part of the Content Reuse Framework. The access control mechanism will reside at the storage level: A user who follows a link to a specific edm:webResource will need to provide credentials matching the conditional rights statement before access to the edm:webResource is granted by the storage providers.

5.6 Community contributed content

WP5 of the Europeana Food and Drink project is developing the community engagement strategy and will contribute to this section.

6 Results and impact

The project has already delivered the Editorial Guidelines, which is essentially a guide to curating Food and Drink content for the Content Base and commercial products and partnerships. The Editorial Guidelines deal with curating aspects and selecting content. The Guidelines for Creative Re-use deal with the technical aspects of content selection for ingestion in the Content Base/Picture Library and the Functional Requirements will inform the decision making on selecting the appropriate platform for this content and the deriving applications and products.

This way, Europeana Food and Drink has outlined the main specifications and guidance for creative and commercial re-use and can now move on to delivering the Content Base/Picture Library.