**BYZART - Byzantine Art and Archaeology on Europeana Art Collection**

**(Implementation of the Action n°2016-EU-IA-0094)**

**Activity 4: Metadata Modelling, metadata guidelines and metadata quality check (Task 4.1; 4.2; 4.3).**

**Milestone 2 (Contact Europeana Core Platform, with regard to both technical and legal aspects about metadata modelling, by 30/04/2018). Accomplished.**

**Means of verification: an e-mail exchange with Europeana contact point.**

**Accomplished.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Revision** | **Date** | **Author** | **Format** | **Description** |
|  |  | **(institution)** |  |  |
| 0.1 | 2018- | UNIBO | Excel doc | First Draft |
|  | 01-25 |  | (BYZART\_metadata\_structure) |  |
| 0.2 | 2018- | NHRF | Excel doc | Updated Draft |
|  | 02-12 |  | (BYZART\_metadata\_structure) |  |
| 0.3 | 2018- | IU-AVARTS | Excel doc | Updated Draft |
|  | 02-15 |  | (BYZART\_metadata\_structure) |  |
| 0.4 | 2018- | UNIBO, | Excel doc | Updated Draft |
|  | 02-16 | NHRF | (BYZART\_metadata\_structure) |  |
| 1.1 | 2018- | IU-AVARTS | Word doc | Updated Draft |
|  | 02-21 |  | (BYZART\_Metadata\_documentation); |  |
|  |  |  | Excel doc |  |
|  |  |  | (BYZART\_metadata\_structure) |  |
| 1.2 | 2018- | UNIBO | Word doc | Updated Draft |
|  | 02-27 |  | (BYZART\_Metadata\_documentation); |  |
|  |  |  | Excel doc |  |
|  |  |  | (BYZART\_metadata\_structure) |  |
| 1.3 | 2018- | IU-AVARTS | Word doc | Updated Draft |
|  | 02-28 |  | (BYZART\_Metadata\_documentation); |  |
|  |  |  | Excel doc |  |
|  |  |  | (BYZART\_metadata\_structure) |  |
| 1.4 | 2018- | UNIBO | Word doc | Updated Draft |
|  | 03-01 |  | (BYZART\_Metadata\_documentation); |  |
|  |  |  | Excel doc |  |
|  |  |  | (BYZART\_metadata\_structure) |  |
| 1.5 | 2018-  03-14 | IU-AVARTS | Word doc  (BYZART\_Metadata\_documentation);  Excel doc  (BYZART\_metadata\_structure) | Updated Draft |
| 1.6 | 2018-  03-15 | UNIBO | Excel doc  (BYZART\_working\_template)  Word doc  (BYZART\_Metadata\_documentation) | Updated Draft |
| 1.7 | 2018-  03-19 | IU-AVARTS | Excel doc  (BYZART\_working\_template)  Excel doc  (BYZART\_metadata\_structure)  Word doc  (BYZART\_Metadata\_documentation) | Updated Draft |
| 1.8 | 2018-07-03 | IU-AVARTS | Excel doc  (BYZART\_working\_template)  Excel doc  (BYZART\_metadata\_structure)  Word doc  (BYZART\_Metadata\_documentation) | Updated Draft |



**BYZART\_Metadata\_documentation (v1.8.2)**

**The BYZART metadata format**

This document contains details concerning the BYZART metadata format. The main goal of the BYZART format is to contain all the information needed to accurately describe an object of cultural heritage from the Byzantine period, while strictly adhering to the rules of the Europeana Data Model (EDM).

**Short introduction to EDM**

The EDM is build around three core classes:

**edm:providedCHO:** This core class is used to describe an object of cultural heritage(CHO). All the information contained here refers to the actual object (painting, mosaic, hymn etc).

**edm:webResource:** This core class is used to describe a digital representation of anobject of cultural heritage. Such representations are digital images, digital recordings and so on.

**ore:Aggregation:** This core class is used to describe the connection between an objectof cultural heritage and a digital representation.

Please note that there is no provisioning for entailing data concerning the physical medium which might be depicting or otherwise representing the object of cultural heritage (what is usually referred to as “support”). That is so because the digital representation takes the place of the analog-physical medium. The object of cultural heritage is no longer captured in a physical format (tape, photograph etc) but in a digital one which is the result of the digitization part of the project.

**Structure of the BYZART metadata format**

In the excel file that accompanies this document, you will find a table-based representation of the structure of the BYZART metadata format. The core classes equivalent for the BYZART metadata format are the “**object of cultural heritage**” (CHO), the “**digital representation**” (WebResource) and the “**aggregation**” (OreAggregation). It should be noted that, despite the fact that the structure is represented as a two

dimensional table, an object of cultural heritage can be connected to multiple digital representations through multiple aggregations. In addition to that, many of the fields of each core class can hold more than one value and can be duplicated (for example there can be two creators for a single object, or an object can represent more than one subject).

On the “labels” column of the excel file we see the name of the metadata fields. What each field represents will be explained further down in detail. On the “EDM” column we see what EDM field each field belongs to. This will help with the transfer of the data to Europeana and only concerns the technical aspect of the format. The “repeatable” column indicates whether a column can have multiple values. Finally, the “value type” column indicates what sort of value each field requires. The value types are as following:

**Automatic implementation**: This value will be automatically generated by the BYZARTsystem when the metadata or files are imported into the system.

**Closed list**: This value can be picked from a predetermined set of values which areexplained on each individual field’s description below.

**Controlled terms - vocabularies**: This value must refer to a controlled vocabulary term,either BYZART vocabulary (under construction by our colleagues from MAR and AUTH) or other existing online vocabulary, in this case in the form of a link (i.e. Getty AAT, GeoNames, VIAF etc).

**Free text**: This value can be simple text of any kind.

The BYZART metadata format has the following requirements that are dictated by the

EDM:

The fields marked with **red** (**Europeana type**, **Europeana rights**, **Provider**, **DataProvider** and **Language for text objects**)*must always*be provided for every CHO orAggregation.

The fields marked with **yellow** (**Title** and **Description**) are important for defining the cultural heritage object. At least one of them *must* be provided, preferably both. If an object is part of an unidentified whole or otherwise lacks a distinct title, a detailed description will suffice and the title field can be omitted.

The fields marked with **green** (**Type**, **Subject**, **Chronology of the item**) are used to define what the subject of the object of cultural heritage is. Since not all objects have an easily defined subject just declaring the type will be sufficient for most. At least one of the four fields *must* be provided.

The fields marked with **blue** are used to provide access to the digital representation on the web. When the digitalized media have been properly uploaded on the BYZART repository, those fields will contain the appropriate links.

**Fields of the BYZART metadata format**

**1. Object of Cultural Heritage fields**

**1.1 Contributor**

A person involved in or contributing to the creation and the life of a CHO. For example, this can be used for the instrument player of a song or the video maker or the patron of a work of art. Please, be accurate in reporting the personal name of the contributor (for example, not Agnello but Agnellus Ravennatis). This is a repeatable field, that can be used to provide multiple contributors.

**1.2 Creator**

The creator of the CHO, for example an artist (if known) or a composer of music. Please, be accurate in reporting the personal name of the creator (for example, not Rublev but Andrei Rublev). This is a repeatable field that can be used to provide multiple creators of the CHO.

**1.3 Description**

A Description of the CHO. If there is no description available the Title field must be provided instead. This field and the Title field must always be distinct. This is a repeatable field that can be used to provide a description of the aspects of a CHO.

**1.4 ID code**

An identification code used to refer to this individual CHO. This code is automatically generated when a new CHO’s metadata are inserted into the BYZART database

**1.5 Language**

The language of text CHOs (for example, a book) and also for other types of CHO if there is a language aspect. For example this can be used for the language of the lyrics of a song or an inscription on a mosaic or on an icon. Mandatory for TEXT objects, strongly recommended for other object types with a language element. Best practice is to use ISO 639 two or three letter primary language tags. Repeat for multiple languages. We also

recommend the use of the ISO 639-2 code for no linguistic content (ZXX). You can find the ISO 639 codes here: [http://www.loc.gov/standards/iso639-2/php/code\_list.php.](http://www.loc.gov/standards/iso639-2/php/code_list.php)

**1.6 Current Ownership**

Used to provide the name of the rights holder of the CHO if possible or for more general rights information.

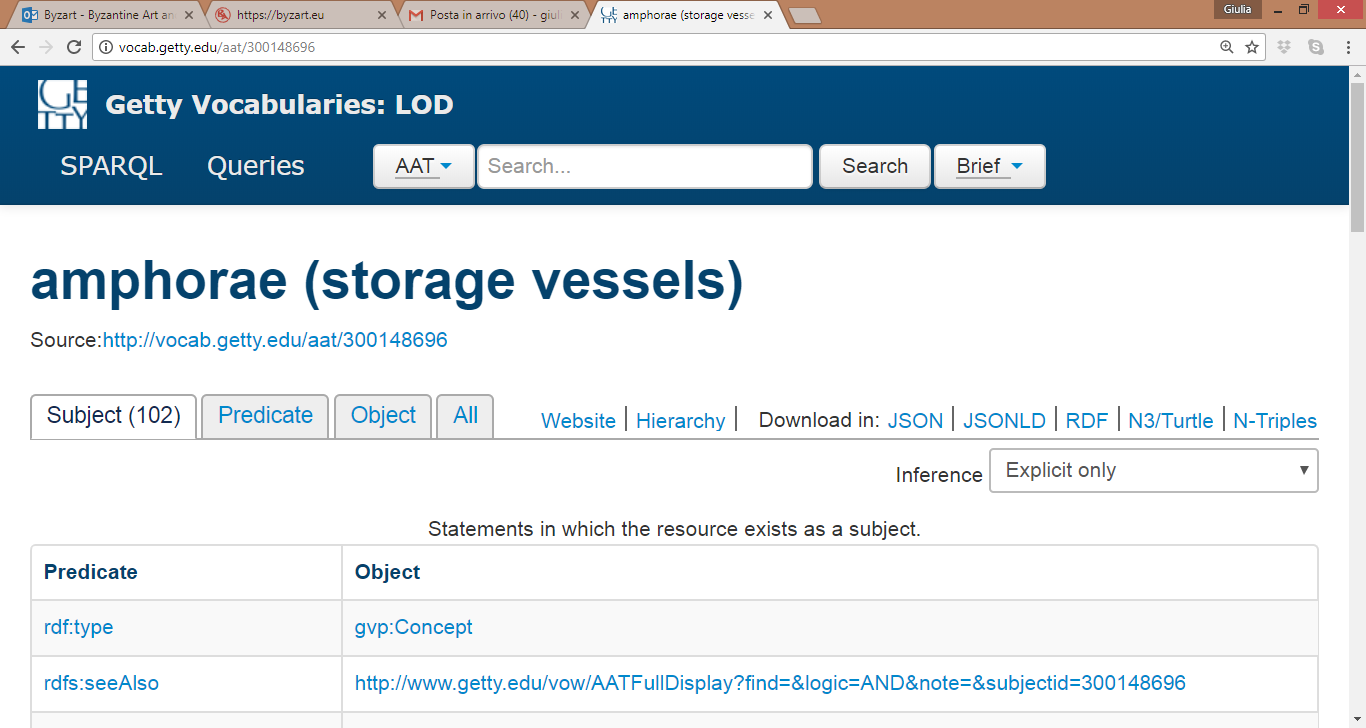
**1.7 Subject**

The subject of the CHO. This field should be provided if the CHO refers to a specific subject. For example this can be used for what a painting depicts or what a hymn is about. This field should be populated with a reference to <http://iconclass.org/help/outline>. The links from iconclass must follow the pattern http://iconclass.org/NUMBER where “NUMBER” is the identifier of the iconclass entry.

**1.8 Title**

The name of the CHO. This field can be populated with the original archival caption of the record (possibly, in original language). This field should always be provided but if no name is available for the CHO then the description field must be provided instead. This field and the description field must always be distinct.

**1.9 Type**

The nature or genre of the object in vocabulary terms. This category can be populated with both general and more specific terms. This is a repeatable field that can be used to provide multiple types the object belongs to. Each term has to be retrieved on the Getty Art and Architecture Thesaurus (<http://vocab.getty.edu/aat/>) and supplied with the corresponding URI. If the proper term to describe the “type” of a CHO is not included in the vocabulary list, you can add it. In any case, the reference to the AAT is mandatory. For example: Type (1): Pottery, and Type (2): Amphora/bowl; Type (1): Architecture, and Type (2): architectural sculpture/ecclesiastical architecture.

**1.10 Chronology (exact date or time range)**

Used to provide a precise date or time range for the creation of the CHO. For example, a chronology “530 -547” for the church of St. Vitale in Ravenna. Europeana recommends date conforming ISO 8601 starting with the year and with hyphens (YYYY-MMDD), in case of exact date (i.e., 2016 -03-01).

**1.11 Extent**

It is used to describe the size of the CHO, if known (e.g., the dimension of an icon). Use the metric system for all measurments. Sizes in m, weights in Kg and time in s.

**1.12 Material**

Used to provide the material or physical carrier of the CHO in vocabulary terms.This is a repeatable field that can be used to provide multiple materials or physical means that the CHO consists of. Each term has to be retrieved on the Getty Art and Architecture Thesaurus (<http://vocab.getty.edu/aat/>) and supplied with the corresponding URI. If the proper term to describe the “type” of a CHO is not included in the vocabulary list, you can add it. In any case, the reference to the AAT is mandatory.

**1.13 Chronology (as periodization)**

Used to provide a time period or era significant to the creation of the CHO.

In this label, we can include the periodization in the form CENTURY BC/AD (+ period), according to the following values: early/middle/late/post Byzantine etc). We can consider as main periods:

* 2nd-3rd c. AD (Late Roman)
* 4th-7th c. AD (Early Byzantine Period)
* 8th-12th c. AD (Middle Byzantine Period)
* 13th-15th c. AD (Late Byzantine Period)
* 7th-15thAD (Western Middle Ages)
* 16th-20th c.AD (Post-Byzantine)

The same field can be used in case of indication of the exact century of creation of the CHO (for example: 5th c. AD).

**1.14 Current location**

Used to indicate the place/building/collection/museum in which the object is currently preserved. This is the geographic location whose boundaries presently include the CHO. This location must have a position within an established positioning system: a location with coordinates or address or inside another location that has a position, such as a room within a (museum) building. Ideally this position should be provided with the value of the property, either by using a reference (to a Place entity) that has coordinates or an address attribute, or as a simple literal. For these reason, it is necessary to use controlled vocabularies to locate the object (please, refer to [www.geonames.org](http://www.geonames.org/)).

Links to geonames must follow the pattern: http://sws.geonames.org/NUMBER/ where “NUMBER” must be the identification number of the geolocation the CHO is located on. If such geolocation does not already exist, the partners should add it to the geonames map themselves.

**1.15 Technique**

Used to specify what technique was used during the creation of the CHO in vocabulary terms. This is a repeatable field that can be used to provide multiple techniques that contributed to the creation of the CHO.Each term has to be retrieved on the Getty Art and Architecture Thesaurus (<http://vocab.getty.edu/aat/>) and supplied with the corresponding URI. If the proper term to describe the “type” of a CHO is not included in the vocabulary list, you can ask for it to be added. In any case, the reference to the AAT is mandatory.

**1.16 Europeana Type**

The value must be one of the types accepted by Europeana as it will support portal functionality: TEXT, VIDEO, SOUND, IMAGE, 3D. This field is mandatory and not repeatable.

**2. Digital representation fields**

**2.1 Creator - Institution**

The institution that provided the digital representation, which can be chosen from a closed list including all BYZART data providers.

**2.2 Format**

Used to provide the format of the digital representation (TIFF, JPEG etc). It is automatically implemented by the system after the upload of the digital item.

**2.3 Collection**

A related resource from which the Web resource is derived in whole or in part. This field has to be populate with the name of the collection that the CHO is part of, in reference to the archival collection in which the CHO is included. The name of the archive collection has to be as precise as possible. We can create a closed list for this field: at this purpose, all the data providers have to furnish to IU-AVARTS a list of the names of the archival collections they refer to.

**2.4 Extent**

Used to describe the size or duration of the digital representation. It will be automatically implemented by the system after the upload of the digital item.

**3. Aggregation Fields**

**3.1 Object the aggregation applies to**

The ID code of the object that this aggregation is referring to. This will be automatically implemented after the metadata for a specific object have been uploaded to the BYZART repository and a new digital representation for that CHO is being created.

**3.2 Data provider**

The name or identifier of the data (content) provider of the digital representation) which can be chosen from a closed list including all BYZART data providers.

.

**3.3 Link to additional web resource files**

This field will be automatically implemented when a new digital representation is uploaded to the BYZART servers and linked to a CHO that already has a primary web resource file.

**3.4 Link to the web resource information**

This field will be automatically implemented when a new digital representation is uploaded to the BYZART servers and linked to a CHO.

**3.5 Link to the web resource file**

This field will be automatically implemented when a new digital representation is uploaded to the BYZART servers and linked to a CHO.

**3.6 Link for preview generation**

This field will be automatically implemented when a new digital representation is uploaded to the BYZART servers and linked to a CHO.

**3.7 Provider**

The name or identifier of the provider of the object (i.e. the organization providing data directly to Europeana). This will be automatically implemented and will contain the Alma Digital Library-University of Bologna identifier.

**3.8 Europeana rights**

The value in this element will indicate the copyright, usage and access rights that apply to the digital representation that this aggregation is referring to. The value in this element is a URI taken from the set of those defined for use in Europeana. A list of these can be found at [https://pro.europeana.eu/available‐rights‐statements.](https://pro.europeana.eu/availableârightsâstatements)

**3.9 Intermediate provider**

The name or identifier of the intermediate organization that selects, collates, or curates data from a Data Provider that is then aggregated by a Provider from which Europeana harvests. This value will be automatically implemented and will contain the BYZART identifier.

**Using the BYZART metadata format**

Collaborators who are expected to provide digital representations and metadata do not necessarily have to provide information for every field separately in the Web-resource

and Aggregation core classes. For example the data provider and Europeana rights fields in the aggregation core class can be automatically filled with the values every collaborator desires for all his contributions. Similarly, in case needed, the type and format of the files can be inferred during the file upload to the BYZART system.

Over 29 required fields, only 10 have to be filled in with free text: we have reduced as much as possible the free text fields in order to speed up the filling of the form for each item.

The mandatory EDM fields will be highlighted and the system won’t validate the record unless they are completed.

**BYZART\_metadata\_structure (v1.8.2)**

(see the attached excel format)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Label** | **EDM** | **Repeatable** | **Value type** |
| **1. PROVIDED CHO** | | | | |
| 1.1 | Contributor | dc:contributor | Yes | Free text |
| 1.2 | Creator | dc:creator | Yes | Free text |
| 1.3 | Description | dc:description | Yes | Free text |
| 1.4 | ID code | dc:identifier | No | Automatic implementation |
| 1.5 | Language | dc:language | Yes | Closed list |
| 1.6 | Current Ownership | dc:rights | No | Free text |
| 1.7 | Subject | dc:subject | Yes | Free text |
| 1.8 | Title | dc:title | No | Free text |
| 1.9 | Type | dc:type | Yes | Controlled terms-vocabularies |
| 1.10 | Chronology (exact date or time range) | dcterms:created | No | Free text |
| 1.11 | Extent | dcterms:extent | No | Free text |
| 1.12 | Material | dcterms:medium | Yes | Controlled terms-vocabularies |
| 1.13 | Chronology (as periodization) | dcterms:temporal | Yes | Free text |
| 1.14 | Current location | edm:currentLocation | No | Controlled terms-vocabularies |
| 1.15 | Technique | edm:hasType | Yes | Controlled terms-vocabularies |
| 1.16 | Europeana type | edm:type | No | Closed list |
| **2. EDM:WEBRESOURCE** | | | | |
| 2.1 | Creator - Institution | edm:webresource dc:creator | No | Closed list |
| 2.2 | Format | edm:webresource dc:format | No | Automatic implementation |
| 2.3 | Collection or Archival Fund | edm:webresource dc:source | No | Closed list |
| 2.4 | Extent of the digital resource | edm:webresource dc:extent | No | Automatic implementation |
| **3. ORE:AGGREGATION** | | | | |
| 3.1 | Object the aggregation applies to | edm:aggregatedCHO | No | Automatic implementation |
| 3.2 | Data (content) Provider | edm:dataProvider | No | Closed list |
| 3.3 | Link to additional webresource files | edm:hasView | Yes | Automatic implementation |
| 3.4 | Link to the webresource information | edm:isShownAt | No | Automatic implementation |
| 3.5 | Link to the webresource file | edm:isShownBy | No | Automatic implementation |
| 3.6 | Link for preview generation | edm:object | No | Automatic implementation |
| 3.7 | Provider | edm:Provider | No | Automatic implementation |
| 3.8 | Europeana rights | edm:rights | No | Closed list |
| 3.9 | Intermediate Provider | edm:intermediateProvider | No | Automatic implementation |