

EUROPEAN COLLECTED LIBRARY OF ARTISTIC PERFORMANCE

www.ECLAP.eu
Grant Agreement No 250481

DE4.2.2

Content and Metadata, Selection, Aggregation and Augmentation

Version: 0.3

Date: 05/07/2013

Project Title: ECLAP Project Number: ICT-PSP-250481 Deliverable Number: DE4.2.2 Accessibility: public Work-Package contributing to the Deliverable: WP4 Nature of the Deliverable: report Status: Final

Contractual Date of Delivery: M36 (30/06/2013) Approve for quality control by: 05/06/2013 Finally approved by coordinator: 05/06/2013 Actual Date of Delivery: 05/06/2013
--

Document responsible: P. Bellini Email address: pbellini@dsi.unifi.it Affiliation acronym: DSI, UNIFI, DINFO

Authors:

- Pierfrancesco Bellini, Paolo Nesi (DSI)
- Theofilos Malis, Natasa Sofou (NTUA)
- Irene Scaturro (CTA-UNIROMA)

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise.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.

Catalogue:

Title	Content and Metadata, Selection, Aggregation and Augmentation
Identifier.de	DE4.2.2
Identifier.ISBN	
Creators	Pierfrancesco Bellini, Paolo Nesi Theofilos Malis, Natasa Sofou, Irene Scaturro
Subject	Content and Metadata Selection and aggregation
Description	Content and Metadata Selection and aggregation
Keywords	ECLAP content, ECLAP aggregation
Publisher	ECLAP
Date	05/07/2013
Format	Document
Type	PDF or DOC
Language	EN

Citation Guidelines

Author(s) name Surname, Deliverable number, Deliverable title, ECLAP Project, DD/MM/YY, URL: univocally determined on <http://www.eclap.eu>

ECLAP Copyright Notice

Depending on the document's declaration of accessibility on the title page, the following notices apply:

- the document is Public, it is available under the Creative Commons license: Attribution-NonCommercial-NoDerivs 3.0 Unported. This license permits non-commercial sharing and remixing of this work, so long as attribution is given.

For more information on this license, you can visit , <http://creativecommons.org/licenses/by-nc-nd/3.0/>



Please note that:

- You can become affiliated with ECLAP. This will give you access to a great amount of knowledge, information related to ECLAP services, content and tools. If you are interested please contact ECLAP coordinator Paolo Nesi at info@eclap.eu. Once affiliated with ECLAP you will have the possibility of using the ECLAP for your organisation.
- You can contribute to the improvement of ECLAP by sending your contribution to ECLAP coordinator Paolo Nesi at info@eclap.eu
- You can attend ECLAP meetings that are open to public, for additional information see www.eclap.eu or contact ECLAP coordinator Paolo Nesi at info@eclap.eu

TABLE OF CONTENTS

1	EXECUTIVE SUMMARY AND REPORT SCOPE	4
2	METADATA DESCRIPTORS INTEROPERABILITY.....	5
2.1	DEFINITIONS	5
2.2	RELATED DOCUMENTS	6
2.3	METADATA CROSSWALKS IN ECLAP	6
2.4	ECLAP MAPPING SCHEMA.....	7
2.5	MAPPING FROM CONTENT PROVIDERS SCHEMATA TO ECLAP SCHEMA	7
2.6	MAPPING STATISTICS.....	11
2.7	MAPPING FROM ECLAP SCHEMA TO EDM	13
2.8	DELIVERY METHOD TO EUROPEANA OFFICE	17
2.9	METADATA INTEROPERABILITY BEST PRACTICES	17
2.9.1	Derivation	18
2.9.2	Application Profiles	18
2.9.3	Crosswalks.....	19
2.9.4	Switching-across.....	20
2.9.5	Metadata Framework and Metadata Registry	20
2.9.6	A Metadata Repository Based on the Open Archives Initiative (OAI) Protocol	21
2.9.7	Aggregation	21
2.9.8	Element-based and Value-based Crosswalking Services.....	22
3	CONTENT SELECTION AND AGGREGATION	22
3.1	FEATURED CONTENT SECTION, CONTENT PARTNERS’ WALL, TOP RATED PAGE	23
3.2	SEARCH FILTERS AND LOCATION MAP	23
3.3	PLAYLISTS AND COLLECTIONS.....	24
3.4	SOCIAL GRAPH.....	25
3.5	MYSTORYPLAYER	26
4	CONTENT ENRICHMENT ON ECLAP SERVICE PORTAL, AUGMENTATION AND RICH MEDIA	28
4.1	LINKING TO GEONAMES DATASET.....	29
4.2	DATE EXTRACTION AND TIMELINE	30
4.3	NAMES IDENTIFICATION.....	30
4.4	ECLAP LINKED OPEN DATA SERVICE.....	31
4.4.1	ECLAP Semantic model.....	32
4.4.2	ECLAP LOD service	34
4.4.3	Content description	35
4.4.4	Taxonomy description	35
4.4.5	Annotations description	36
4.4.6	User description	37
4.5	SOCIAL GRAPH	37
5	BIBLIOGRAPHY	39
6	GLOSSARY	42
7	APPENDIX – COMPLETE MAPPINGS TO THE ECLAP SCHEMA	43
8	APPENDIX – COMPLETE ECLAP SCHEMA	47
8.1	GENERAL INFORMATION	47
8.2	DIGITAL RESOURCE INFORMATION.....	47
8.3	IPR INFORMATION	48
8.4	GEOSPATIAL INFORMATION.....	48
8.5	PERFORMING ARTS METADATA	48
8.6	DC – DUBLIN CORE METADATA.....	54
8.7	DCTERMS – DUBLIN CORE TERMS METADATA	58
8.8	TAXONOMY CLASSIFICATION	62
8.9	ECLAP GROUPS	62
8.10	ECLAP AGGREGATIONS.....	62

1 Executive summary and report scope

This report has the duty of summarising the activity performed in ECLAP regarding the Content and Metadata Selection, Aggregation and Augmentation. In the context of ECLAP, this means to talk about:

- Content description in terms of metadata and interoperability, metadata ingestion and mapping, metadata export towards Europeana
- Content Selection and contextualization, adoption and usage of content enrichment tools from the side of content providers
- Content Enrichment tool and technical aspects and results obtained.

As regards content description (see Section 2), the ECLAP project focussed on WP4 Content Provision and Augmentation involved the selection and delivery of content and metadata for a wide range of user communities as well as the definition of the harvesting metadata schema and its semantic mappings to a spectrum of commonly used standards. The major task has been to create the mappings between the descriptive and multimedia metadata, to ensure the semantic interoperability inside the ECLAP system and more importantly with the Europeana system. To this end, the report provides description of the metadata ingestion processes; ECLAP metadata schema; interoperability is achieved in the ECLAP domain; mapping statistics from content providers to the ECLAP portal; mapping of the elements of ECLAP to the Europeana Data Model; overview of the procedures and tools that are deployed within ECLAP, in order to establish interoperability between ECLAP's metadata and the Europeana repository in EDM, and metadata interoperability best practices adopted in order to ensure uniform access to the ECLAP and Europeana content.

The results obtained provided the evidence that the ECLAP consortium has worked well, integrating and defining the needed tools, providing support to the content providers, collecting content and metadata and finally providing the planned content to Europeana. The ECLAP consortium successfully addressed the complexity of the heterogeneous metadata mapping and also addressed the problems due to the lack of a standard for performing art metadata. ECLAP poses the basis for the further definition of a standard and related tools its management and exploitation.

For the Content selection and contextualization, Section 3 describes the activity performed on content for its selection, enrichment and aggregation. To aggregate content also means to contextualize a single resource in relation to other elements or resources: on ECLAP, content can be compared and contextualized at a transnational level, as the aggregation activities which can be performed on the portal allow content providers to relate their own items to those provided by other partners. For these reasons, the wide range of possibilities offered by ECLAP to further enrich and contextualize the resources and connecting them through users' annotations and tools such as Mystoryplayer and the Social Graph are particularly useful for the purposes of research and education. Both these tools have been selected by Europeana and included into the Europeana pro web pages.

For the technical aspects and tools for content enrichment, ECLAP provided a range of metadata enrichment activities (based on AXCP media grid). The typical metadata enrichments performed by ECLAP can be the addition of technical descriptors of source files, the addition of more languages, the geo localization passing from location named into metadata and descriptors to formal GPS position, the production of QR codes for museum inspection and linkage (see it as augmented reality first step), the content aggregation, the addition of comments and tags, the association of taxonomical classification, the establishing of connections with dbPedia open data, the addition of a formal IPR license descriptor, the association of univocal date and time, the association of any kind of identifiers, the production of LOD model, the content aggregations, etc. As a result, the content is described in terms of metadata based on the so called ECLAP semantic model which is much richer than the ECLAP ingestion model.

2 Metadata Descriptors Interoperability

WP4 (Content Provision and Augmentation) involves the selection and delivery of content and metadata for a wide range of user communities as well as the definition of the harvesting metadata schema and its semantic mappings to a spectrum of commonly used standards. This deliverable reports the results of Task 2 (Metadata/descriptors interoperability maps) of WP4. Its main purpose is to provide all the necessary mappings between ECLAP and Europeana. In addition, part of this task is to create the mappings between the descriptive and multimedia metadata. This task will ensure the semantic interoperability inside the ECLAP system and more importantly with the Europeana system. This report focuses on the interoperability of metadata standards and classifications, between ECLAP and Europeana.

The report is divided in the following Sections:

- Section 2.3 describes the metadata ingestion processes taking place in the ECLAP portal:
 - A) When content is contributed to ECLAP project, the corresponding metadata are transformed into a new form under the predefined ECLAP schema.
 - B) In order to be available to the Europeana portal the same content is transformed from ECLAP schema to EDM and then published to Europeana.
- Section 2.4 provides a general summarization of the ECLAP metadata schema, i.e. the information required for the documentation of performing arts content, as those decided by the providers. A more detailed description of it is available in Section 0 in the Appendix.
- In order to illustrate how the interoperability is achieved in the ECLAP domain, Section 2.5 demonstrates a mapping procedure from content providers' schemata to the ECLAP schema. Section 2.6 provides mapping statistics from content providers to the ECLAP portal.
- The mapping of the elements of ECLAP to the Europeana Data Model is presented in Section 2.7. Moreover in Section 2.8 there is an overview of the procedures and tools that are deployed within ECLAP, in order to establish interoperability between ECLAP's metadata and the Europeana repository.
- Finally, Section 2.9 presents the metadata interoperability best practices adopted in order to ensure uniform access to the ECLAP and Europeana content.

2.1 Definitions

Metadata

Metadata is a modern term for the bibliographic information that libraries traditionally entered into their catalogs or databases, or registration information about collections that museums have entered into their systems; however the term metadata is most commonly used to refer to descriptive information about digital resources [1].

Metadata Schema

A metadata schema is simply a set of elements with a precise semantic definition, optionally connected by some structure [3]. The semantics of a schema is defined by the meanings of its elements. Schemas usually define the names of elements together with their semantics and optionally content rules that define how content values must be formulated (e.g., capitalization, allowable content values). For the encoding of the elements and their values, a schema can define syntax rules [4].

Metadata Repository

A Metadata repository is a database created to gather, store, and distribute contextual information about metadata. This contextual information of metadata include meaning and content, policies that govern, technical attributes, specifications that transform, and programs that manipulate [5].

Metadata Ingestion

Integration of metadata from heterogeneous sources is a major issue when connecting cultural institutions to digital library networks. Uniform access to metadata is impeded by the structural and semantic heterogeneities of the metadata and metadata schemes used in the source systems. The term metadata ingestion refers to the integration of metadata from heterogeneous sources into a single repository.

Metadata Interoperability

According to [2] “Interoperability is the ability of multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality”.

2.2 Related Documents

This part of the document is closely related to the Deliverable 4.3 “Metadata Descriptors Interoperability”.
<http://www.eclap.eu/115119>

2.3 Metadata Crosswalks in ECLAP

Metadata crosswalks facilitate the interoperability and exchange of metadata. As defined in Section 2.1, a crosswalk is a mapping of the elements, semantics, and syntax from one metadata schema to those of another. Within ECLAP, a crosswalk can be considered as follows:

Content providers use their preferable standard for their metadata representation. When content is contributed to ECLAP project, the corresponding metadata need to be transformed into a new form under the predefined ECLAP schema. The available metadata (regardless the standard they conform to) are in xml format. In order to transform their metadata, content providers use the ECLAP Metadata Ingestion Service portal (MINT), and map their metadata to the ECLAP schema by defining an XSLT that is used in the mapping phase as depicted in Figure 2-1. Although, all metadata are uniformly stored in ECLAP under ECLAP schema, in order to be available through Europeana portal they first need to be transformed from ECLAP schema to EDM and then published to Europeana. Thus a new mapping takes places, transforming metadata from ECLAP schema to EDM as depicted in Figure 2-2. For more details on the aforementioned hierarchy and process the reader can refer to D3.1- Infrastructure: ingestion and processing content and metadata and D4.2.1 Content and Metadata, Selection Aggregation and Augmentation.

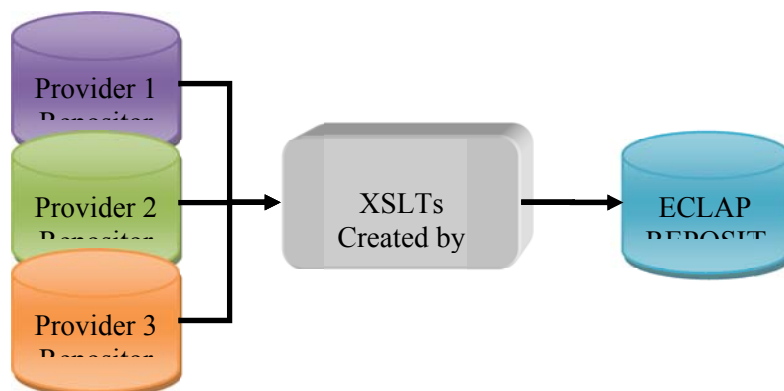


Figure 2-1: Metadata ingestion from multiple metadata repositories to the ECLAP repository.

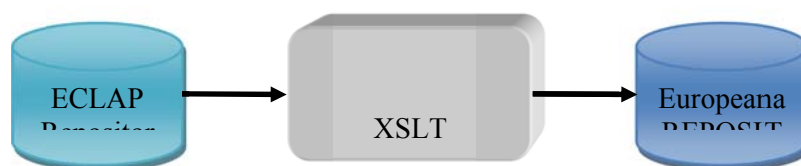


Figure 2-2: Metadata ingestion into the Europeana repository.

2.4 ECLAP Mapping Schema

In this section a brief description of the current ECLAP metadata schema is presented. The complete description is reported in Appendix.

The metadata schema is divided in the following parts:

- **General information** about the content, it reports the internal identifiers like the axoid, the drupal nid, the provider name and id and the url to the content on ECLAP portal
- **Information about the digital resource**, it reports information on the resource format (audio, video, image, document, 3d, crossmedia, etc.), in case of video content if it is available in MD and HD resolutions, if the resource is available on smartphones, the url to the preview gif (animated for video), etc.
- **Information on IPR**, on intellectual property rights for the digital resource, it contains the Europeana rights URL, if the content is public or private, the title and description of the IPR model associated with the content.
- **GeoSpatial information about the resource**, reports the GPS position associated with the content as well as the extent of the area surrounding the position.
- **Dublin Core metadata** (DC and DCTERMS), reports Dublin Core metadata describing the content.
- **Performing Arts metadata**, reports the metadata specific for performing arts:
 - FirstPerformance Date, Location, City and Country, indicates where and when the premiere of the piece depicted in the digital resource was performed.
 - Performance Date, Location, City and Country, indicates where and when the performance depicted in the digital resource was performed.
 - PerformingArtType, for the type of performing art (e.g. theatre, dance, etc.)
 - PlotSummary, summary of the plot
 - PerformingArtsGroup, name of the theatre or dance company or musical group (e.g. Momix)
 - Cast, name/names of cast member
 - PerformersAndCrew, name/names of performers and crews of the performance
 - Professional, people involved in the performance indicating which role each person had in the performance (e.g. Actor, director, set designer etc.).
 - Object, object used in the performance
 - PieceRecord, script of the play
 - Genre
 - HistoricalPeriod
 - ArtisticMovementAndActingStyle, Artistic movement and acting styles in which the work can be categorized (e.g. Classicism, Dada, Epic, Expressionism, etc.)
 - RecordingDate, date when the recording was made
- **Taxonomy based classifications**, reports information about the taxonomy terms associated with the content, for each term is reported the label in every language, the term id, and the id of the top term for the hierarchy and the path from the term to the top term.
- **ECLAP Groups**, reports the ECLAP Groups to which the content is associated with.
- **ECLAP Aggregations** (Collections & Playlists), for collections and playlists reports the identifiers (axoids) of the content in the collection/playlist, for playlists items it is also eventually present the startTime/endTime/duration indications.

2.5 Mapping from Content Providers Schemata to ECLAP Schema

This section demonstrates an example of how interoperability of metadata is achieved when different institutes contribute metadata corresponding to different schemata. As we have already stated,

interoperability is achieved via crosswalking, i.e. the corresponding contributed metadata are transformed into a new form under the predefined ECLAP schema.

Our example focuses on two different institutions contributing metadata: the *Dario Fo & Franca Rame Archive* and the *Ljubljana museum*. Suppose for example that the *Dario Fo & Franca Rame Archive* contributes to ECLAP a PDF file with the following metadata: the *document's title* is “Grasso è bello!”, its *creators* are “Franca Rame and Dario Fo”, its *subject* is “testo”. The entire record is depicted on Table 2-1 in its XML format. On the other hand the Ljubljana museum contributes a JPEG image file with the following metadata: its *title* is “Costumes, Leo Kulaš (No. 19), L'impresario di Smirne”, its *Slovenian title* is “Kostumi, Leo Kulaš (št.19), L'impresario di Smirne”, its *creator* is “Leo Kulaš archive”. The entire record is depicted on Table 2-2 in its XML format. It is obviously that the two institutions adopt different schemata in order to describe their data.

Table 2-1: A metadata record provided by the Dario Fo & Franca Rame Archive

```
<Export>
  <Record>
    <Filename>content/archivioforame/PARL/2008/TEST/55500/55500.pdf</Filename>
    <Title>Grasso è bello!</Title>
    <Creator>Franca Rame, Dario Fo</Creator>
    <Subject>testo</Subject>
    <Description>Dattiloscritto dell'atto unico "Grasso è bello!" di Franca Rame, Dario Fo e
alcune idee di Jacopo Fo. Revisione di Franca Rame aggiornata al 2008.</Description>
    <Format>pdf</Format>
    <Language>it</Language>
    <Coverage>Milano, Coordinates: Lat: 45 28 00 N degrees minutes Lat: 45.4667 decimal
degrees Long: 009 12 00 E degrees minutes Long: 9.2000 decimal degrees </Coverage>
    <Rights>CTFR</Rights>
    <Extent>55pag</Extent>
    <ProviderID>CTFR</ProviderID>
    <ProviderContentID>55500</ProviderContentID>
    <AggregationID>PARL</AggregationID>
  </Record>
</Export>
```

Table 2-2: A metadata record provided by the Ljubljana museum

```
<metadata>
  <record>
    <dc.Title>Costumes, Leo Kulaš (No. 19), L'impresario di Smirne</dc.Title>
    <dc.Title_slo>Kostumi, Leo Kulaš (št.19), L'impresario di Smirne</dc.Title_slo>
    <dc.Creator>Leo Kulaš archive</dc.Creator>
    <dc.Publisher>Produced by: HNK Zagreb</dc.Publisher>
    <dc.Description>Costume photo from the performance L'impresario di Smirne, written by
Carlo Goldoni, directed by Paolo Magelli</dc.Description>
    <dc.Subject>theatre performance, costume design</dc.Subject>
    <dc.Rights>For information about the rights holder please contact Muzeum, Institute of Art
Production, Mediation and Publishing Ljubljana.</dc.Rights>
    <dc.Type>Image</dc.Type>
    <dc.Language>en</dc.Language>
    <dc.IsPartOf>Leo Kulaš</dc.IsPartOf>
    <ProviderContentID>MLJ2346_LKulas_Impresaria</ProviderContentID>
  </record>
</metadata>
```



```

<ContentLocation>In_process/MLJ2346_LKulas_Impresaria.jpg</ContentLocation>
<IPRModel>MUZEUM_all_CC-BY-NC-ND </IPRModel>
<Public>Yes</Public>
<FirstPerformanceCountry>Croatia</FirstPerformanceCountry>
<FirstPerformanceDate>1993</FirstPerformanceDate>
<ProfessionalCostume_designer>Leo Kulaš</ProfessionalCostume_designer>
<ProfessionalDirector>Paolo Magelli</ProfessionalDirector>
<ProfessionalPlaywright>Carlo Goldoni</ProfessionalPlaywright>
<PerformingArtType>Theatre</PerformingArtType>
<HistoricalPeriod>XX Century</HistoricalPeriod>
    
```

</record>

</metadata>

In order to make these two different structured metadata records uniformly available based on the ECLAP metadata schema, two different mappings (crosswalks) were constructed.

The *CTFR-mapping-export* was constructed by the *Dario Fo & Franca Rame Archive* in order to map its content to the ECLAP portal. In Table 2-3 we see the summary of this mapping (the complete mapping can be found in Table 7-1 in the Appendix). For example, based on the two first rows of this mapping the */DC/subject* field of the ECLAP schema corresponds to the */Export/Record/Subject* field of the initial schema, while the */DC/subject/@lang* of the ECLAP schema is statically mapped to *it* (Italian) for every record in the *Dario Fo & Franca Rame Archive*. Similarly the */DC/title* field of the ECLAP schema corresponds to the */Export/Record/Title* field of the initial schema, while the */DC/title/@lang* of the ECLAP schema is statically mapped to *it* (Italian) for every record in the initial schema. In Table 2-4 we see the XML corresponding source and target metadata records for the fields of subject, title, subject language and title language.

The *Muzeum_new_v06a_mapping* was constructed by the *Ljubljana museum* in order to map its content to the ECLAP portal. In Table 2-5 we see the summary of this mapping (the complete mapping can be found in Table 7-2 in the Appendix). For example, based on the two first rows of this mapping the */DC/subject* field of the ECLAP schema corresponds to the */metadata/record/dc.Subject* field of the initial schema, while the */DC/subject/@lang* of the ECLAP schema is statically mapped to *en* (English) for every record in the *Ljubljana museum* archive. Similarly the */DC/title* field of the ECLAP schema corresponds to the */metadata/record/dc.Title* field of the initial schema, while the */DC/title/@lang* of the ECLAP schema is statically mapped to *en* (English) for every record in the initial schema. In Table 2-6 we see the XML corresponding source and target metadata records for the fields of subject, title, subject language and title language.

Table 2-3: The CTFR-mapping-export summary

Source	Target
/Export/Record/Subject	/DC/subject
It	/DC/subject/@lang
/Export/Record/Title	/DC/title
It	/DC/title/@lang
...	

Table 2-4: Mapping from the source XML presented in Table 2-1 to the ECLAP XML

Source
<pre><Export> <Record> ... <Subject>testo</Subject> <Title>Grasso è bello!</Title> </Record> </Export></pre>
Target
<pre><eclap:ECLAPObjectWrap xmlns:eclap="http://www.eclap.eu/ECLAPSchemaV0" xmlns:dc="http://purl.org/dc/elements/1.1/" ...> <eclap:ECLAPObject> ... <eclap:DC> ... <dc:subject xml:lang="it">testo</dc:subject> <dc:title xml:lang="it">Grasso è bello!</dc:title> </eclap:DC> </eclap:ECLAPObject> </eclap:ECLAPObjectWrap></pre>

Table 2-5: The MUZEUM_new_v06a_mapping summary

Source	Target
/metadata/record/dc.Subject	/DC/subject
en	/DC/subject/@lang
/metadata/record/dc.Title	/DC/title
en	/DC/title/@lang
...	

Table 2-6: Mapping from the source XML presented in Table 2-2 to the ECLAP XML

Source
<pre><metadata> <record> <dc.Title>Costumes, Leo Kulaš (No. 19), L'impresario di Smirne</dc.Title> <dc.Subject>theatre performance, costume design</dc.Subject> </record> </metadata></pre>
Target
<pre><eclap:ECLAPObjectWrap xmlns:eclap="http://www.eclap.eu/ECLAPSchemaV0" xmlns:dc="http://purl.org/dc/elements/1.1/" ...> <eclap:ECLAPObject></pre>

```
...
<eclap:DC>
...
<dc:subject xml:lang="en">theatre performance, costume design</dc:subject>
<dc:title xml:lang="en">Costumes, Leo Kulaš (No. 19), L'impresario di Smirne</dc:title>
</eclap:DC>
</eclap:ECLAPObject>
</eclap:ECLAPObjectWrap>
```

It is evident from Table 2-4 and Table 2-6 that after the transformation takes place the metadata originating from the Dario Fo & Franca Rame Archive and the metadata originating from the Ljubljana museum are uniformly represented with respect to the ECLAP schema.

The complete transformed XML metadata records corresponding to the metadata records depicted in Table 2-1 and Table 2-2 are presented in Table 7-3 and Table 7-4 respectively.

2.6 Mapping Statistics

This paragraph contains the progress report of the production and delivery of metadata to the ECLAP project. The purpose of this paragraph is to ensure partners are able to create metadata compliant with the ECLAP standard, develop a relationship with their national aggregators and make necessary amendments to their internal systems to deliver content to ECLAP. The partners themselves manage the process of creating and transforming their metadata.

The statistics examine three different parameters for each partner:

- Imports: Each import is a set of metadata describing performing arts objects. An import is a file uploaded to the MINT Mapping Tool -and through it to the ECLAP portal- with a variety of methods such as HTTP upload, FTP upload, remote FTP/HTTP upload and OAI-PMH V2 protocol. The file related to each import is either an XML, or CSV, or a ZIP file containing XML and CSV files.
- Items: This parameter refers to the performing arts items described by the set of metadata contained in an import.
- Transformations: Each set of items conforms to an XML schema and should be transformed to the corresponding ECLAP schema via the mapping created by the MINT tool. The number of transformations refers to the number of items that are validly mapped to the ECLAP schema.

For example, a user uploads an xml file (1 import), containing metadata for 123 performing arts items (123 items) from which only 120 are validly mapped to the ECLAP schema (120 transformations).

Figure 2-3 shows the number metadata imports contributed in years 2011, 2012 and 2013. The total number of imports is 277 for the years 2011, 2012 and 332 for the year 2013.

Figure 2-4 shows the number items examined in years 2011, 2012 and 2013. The total number of items is 136400 for the years 2011, 2012 and 156116 for the year 2013.

Figure 2-5 shows the number metadata records produced in years 2011, 2012 and 2013. The total number of imports is 101058 for the years 2011, 2012 and 120599 for the year 2013.

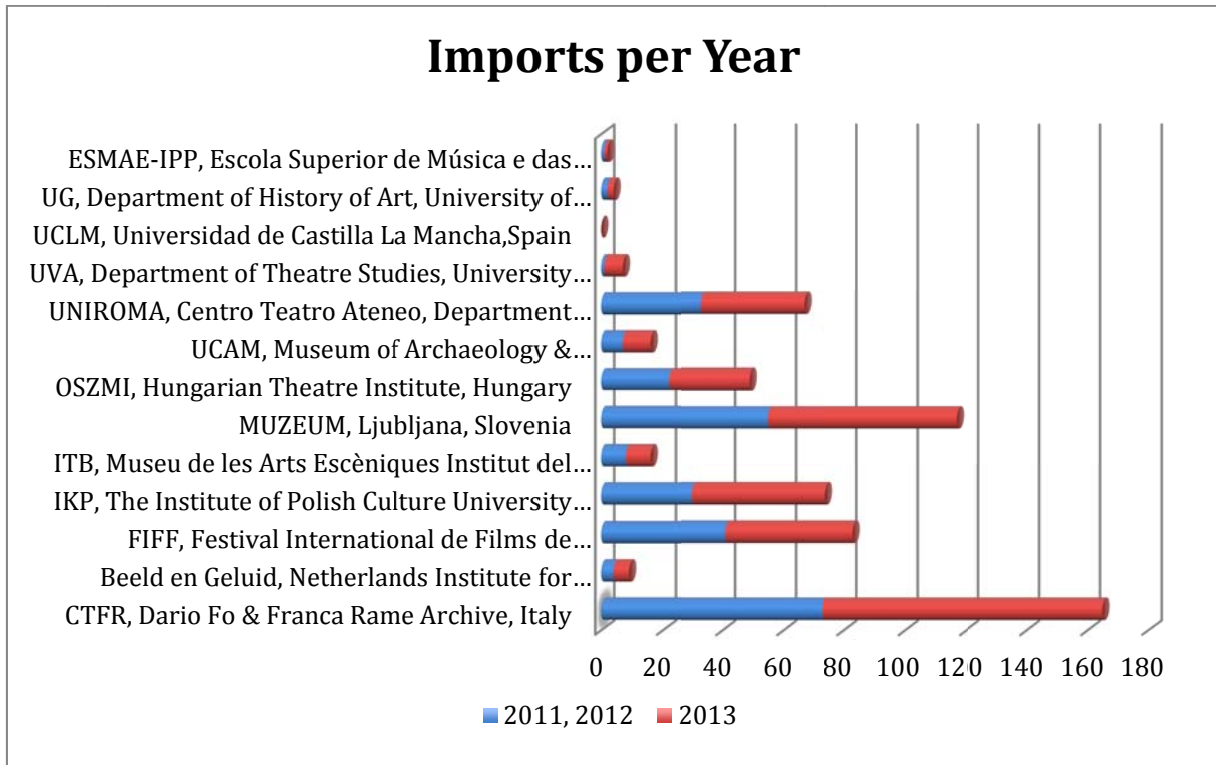


Figure 2-3: Metadata Imports per Year, up to the end of May 2013 (ONLY data passed via MINT an additional 20-25% passed directly on ECLAP service portal)

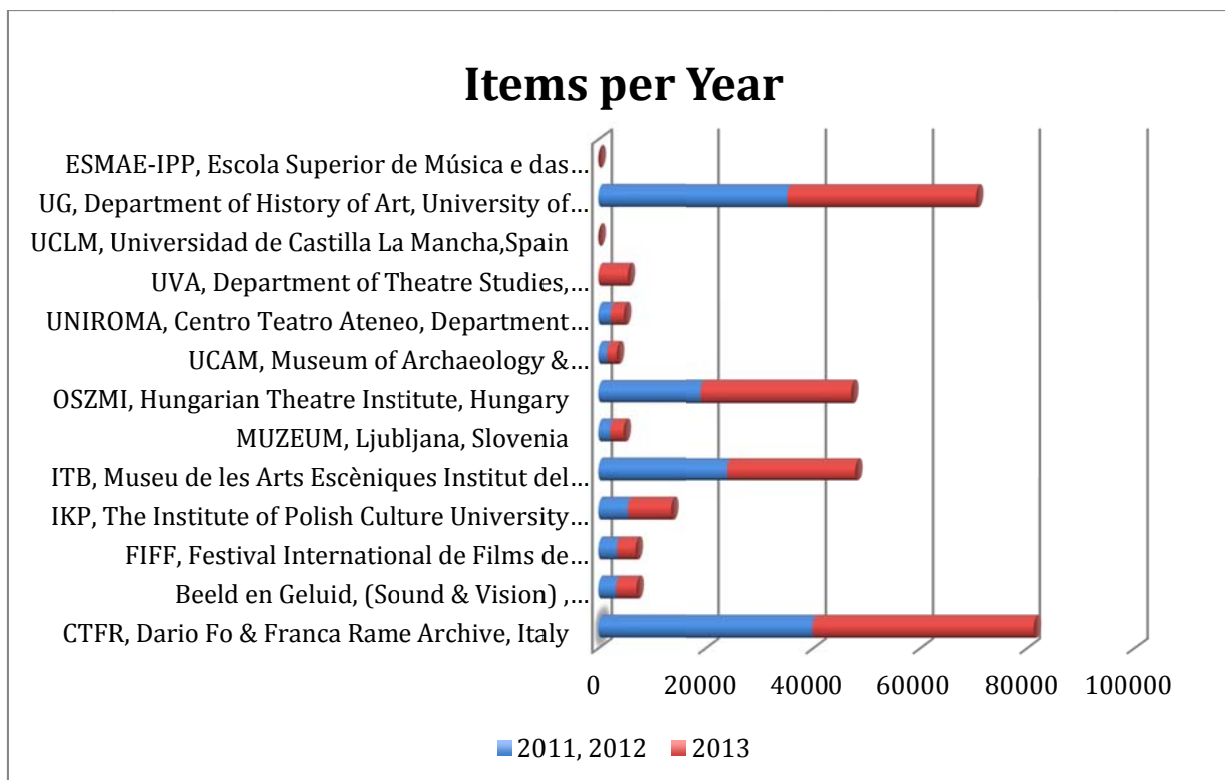


Figure 2-4: Items per Year, up to the end of May 2013 (ONLY data passed via MINT an additional 20-25% passed directly on ECLAP service portal)

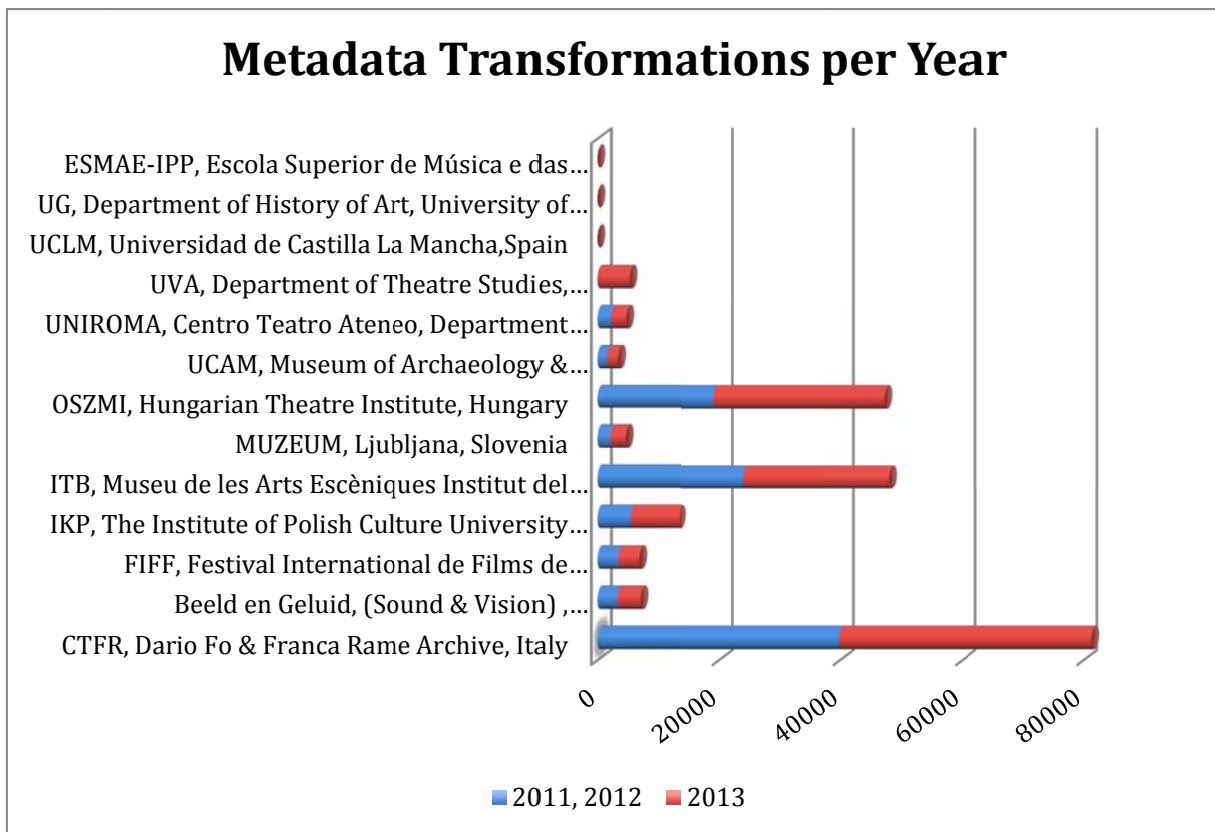


Figure 2-5: Metadata Transformations per Year, up to the end of May 2013 (ONLY data passed via MINT an additional 20-25% passed directly on ECLAP service portal)

It should be noted that the previously presented figures refer only to metadata mapped to ECLAP via the MINT tool and not the total available metadata in ECLAP. Content Providers can also provide and exploit services for providing:

- Single content/item and metadata: can be uploaded in one shot via the upload on ECLAP Social Service Portal by using the ECLAP Upload Service and web page. Once uploaded the Edit Object functionality/action allows the user to: edit metadata, upload a different icon/thumbnaïl, update the digital file, upload a new version of the digital file, update/change the classification, and define the IPR model via the IPR Wizard, etc.
- Multiple/massive content/items and metadata: This can be performed in one shot by using the ECLAP Upload Service and web page. The user can select in the file selection dialog several files at the same time.

2.7 Mapping from ECLAP Schema to EDM

In this section, it is reported how the ECLAP metadata is mapped to EDM (this process and the mapping has been developed by DSI, UNIFI) using an Object centric perspective (the only one that now Europeana ingestion supports), it have to be noted that in the material to be provided to Europeana in many cases it does not represent strictly an Object (like a book, a painting, a sculpture, ...) while often it represents an event happened in the past, the performance.

In the following is reported how the ECLAP metadata are mapped to EDM elements. The Dublin core elements (dc and dcterms) are mapped directly to the ProvidedCHO elements while the PerformingArts metadata are mapped to DublinCore elements where possible, also the taxonomy associations are mapped to DublinCore depending on the top hierarchy element (Subject is mapped to dc:subject, PerformingArtType to dc:type, HistoricalPeriod to dterm:temporal, etc.).

For each ECLAP content is provided: one ProvidedCHO element, one WebResource element representing the ECLAP portal page showing the content and one Aggregation element aggregating the two preceding ones. The @ sign indicate an attribute of the element.

edm:ProvidedCHO	
@rdf:about	axoid
dcterms:*	All ECLAP dcterms fields
dcterms:issued	“<PerfArts.FirstPerformance.Date> (first performance)”
dcterms:issued	PerfArts.Performance.Date
dcterms:spatial	PerfArts.Performance.Place
dcterms:spatial	PerfArts.Performance.City
dcterms:spatial	PerfArts.Performance.Country
dcterms:temporal	PerfArts.HistoricalPeriod
dcterms:temporal	
@rdf:resource	“http://www.eclap.eu/Classification/HistoricalPeriod/<histPeriodId>”
dcterms:references	PerfArts.PieceRecord
dc:*	All ECLAP dc fields
dc:date	PerfArts.RecordingDate
dc:creator	PerfArts.PerformingArtsGroup
dc:contributor	“<PerfArts.Professional> (<role>)”
dc:contributor	PerfArts.Cast
dc:contributor	PerfArts.PerformersAndCrew
dc:contributor	PerfArts.PersonRecord
dc:contributor	PerfArts.ProductionRecord
dc:description	PerfArts.PlotSummary
dc:description	PerfArts.Object
dc:subject	
@rdf:resource	“http://www.eclap.eu/Classification/Subject/<subjId>”
dc:subject	PerfArts.Genre
dc:subject	
@rdf:resource	“http://www.eclap.eu/Classification/Genre/<genreId>”
dc:type	PerfArts.PerformingArtsType
dc:type	
@rdf:resource	“http://www.eclap.eu/Classification/PerformingArtType/<paTypeId>”
dc:type	PerfArts.ArtisticMovementAndActingStyle
dc:type	
@rdf:resource	“http://www.eclap.eu/Classification/ArtisticMovementAndActingStyle/<amasId>”
edm:type	based on Resource.Format (video, audio, image, document)

The histPeriodId, subjId, genreId, paTypeId, amasId are the ids of the terms in the ECLAP taxonomy to which the content is associated with. The SKOS taxonomy defining the concepts used are provided to europeana using a specific file.

edm:WebResource	
@rdf:about	“http://www.eclap.eu/europeana/<axoid>”
edm:rights	IPR.EuropeanaRightsUrl
ore:Aggregation	
@rdf:about	“<axoid>:aggregation”
edm:aggregatedCHO	axoid
edm:dataProvider	eclap:ProviderName
edm:provider	“ECLAP, e-library of Performing Arts”
edm:rights	IPR.EuropeanaRightsUrl

edm:isShownAt “http://www.eclap.eu/europeana/<axoid>”
edm:object eclap:Preview

This mapping should be enhanced by enriching the metadata with associations with Places, TimeSpans, Agents thus integrating the text metadata with an association with an rdf resource coming from linked open data initiatives or well known authority files as VIAF for person names, GeoNames for places etc.

The following is an example of mapping the metadata of an Image from the Dario Fo and Franca Rame Archive.

The source metadata is:

```
<eclap:Content axoid="urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d">
  <eclap:url>http://www.eclap.eu/drupal?q=home&axoid=urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d</eclap:url>
  <eclap:nid>35923</eclap:nid>
  <eclap:Version>2</eclap:Version>
  <eclap:InsertUpdateTime>2011-08-06T00:36:26</eclap:InsertUpdateTime>
  <eclap:ProviderId>CTFR</eclap:ProviderId>
  <eclap:ProviderName>Dario Fo & Franca Rame Archive</eclap:ProviderName>
  <eclap:DefaultMetadataLanguage>it</eclap:DefaultMetadataLanguage>
  <eclap:Resource>
    <eclap:Format>image</eclap:Format>
    <eclap:Type>image</eclap:Type>
    <eclap:Width>597</eclap:Width>
    <eclap:Height>800</eclap:Height>
    <eclap:Extension>.jpg</eclap:Extension>
  </eclap:Resource>
  <eclap:Platforms>
    <eclap:AvlForPDA>yes</eclap:AvlForPDA>
    <eclap:AvlForIPhone>yes</eclap:AvlForIPhone>
    <eclap:AvlForPC>yes</eclap:AvlForPC>
  </eclap:Platforms>
  <eclap:IPR>
    <eclap:IsPublic>yes</eclap:IsPublic>
    <eclap:IPRTitle>CTFR IPR</eclap:IPRTitle>
    <eclap:IPRDescription>CTFR model</eclap:IPRDescription>
    <eclap:EuropeanaRightsUrl>http://www.europeana.eu/rights/rr-f/</eclap:EuropeanaRightsUrl>
    <eclap:LicenseUrl>http://bpnet.eclap.eu/drupal/?q=node/2862</eclap:LicenseUrl>
  </eclap:IPR>
  <eclap:Preview>http://www.eclap.eu/gif/urn_axmedis_00000_obj_36b2407e-0ca0-4f44-892b-ebf254118a2d.gif</eclap:Preview>
  <eclap:DublinCoreMetadata>
    <dc:description xml:lang="it">Tavole a colori di Dario Fo.</dc:description>
    <dc:terms:extent xml:lang="it">1 pagine</dc:terms:extent>
    <dc:rights xml:lang="it">Archivio Dario Fo e Franca Rame, alcuni diritti riservati</dc:rights>
    <dc:title xml:lang="it">Sant' Ambrogio - 2005 Testo teatrale di Dario Fo. 32/116</dc:title>
    <dc:type xml:lang="it">Disegni</dc:type>
    <dc:date xml:lang="it">2005</dc:date>
    <dc:language xml:lang="it">it</dc:language>
  </eclap:DublinCoreMetadata>
  <eclap:PerformingArtsMetadata xml:lang="it">
    <eclap:Performance>
      <eclap:Country>Italia</eclap:Country>
    </eclap:Performance>
  </eclap:PerformingArtsMetadata>
</eclap:Content>
```

```
</eclap:PerformingArtsMetadata>
<eclap:Classification>
  <eclap:term id="504" vid="5" root="664" path="664">
    <eclap:label xml:lang="it">Teatro</eclap:label>
    <eclap:label xml:lang="en">Theatre</eclap:label>
    <eclap:label xml:lang="da">Teater</eclap:label>
    <eclap:label xml:lang="nl">Theater</eclap:label>
    <eclap:label xml:lang="fr">Théâtre</eclap:label>
    <eclap:label xml:lang="de">Theater</eclap:label>
    <eclap:label xml:lang="el">Θέατρο</eclap:label>
    <eclap:label xml:lang="hu">Színház</eclap:label>
    <eclap:label xml:lang="pl">Teatr</eclap:label>
    <eclap:label xml:lang="pt">Teatro</eclap:label>
    <eclap:label xml:lang="es">Teatro</eclap:label>
    <eclap:label xml:lang="ca">Teatre</eclap:label>
    <eclap:label xml:lang="sl">Gledališče</eclap:label>
  </eclap:term>
</eclap:Classification>
<eclap:Group id="2862">
  <eclap:label xml:lang="en">Dario Fo & Franca Rame Archive</eclap:label>
</eclap:Group>
</eclap:Content>
```

That is mapped to EDM as:

```
<rdf:RDF ...>
  <edm:ProvidedCHO rdf:about="urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d">
    <dcterms:extent xml:lang="it">1 pagine</dcterms:extent>
    <dcterms:spatial>Italia</dcterms:spatial>
    <dc:description xml:lang="it">Tavole a colori di Dario Fo.</dc:description>
    <dc:rights xml:lang="it">Archivio Dario Fo e Franca Rame, alcuni diritti riservati</dc:rights>
    <dc:title xml:lang="it">Sant' Ambrogio - 2005 Testo teatrale di Dario Fo. 32/116</dc:title>
    <dc:type xml:lang="it">Disegni</dc:type>
    <dc:date xml:lang="it">2005</dc:date>
    <dc:language xml:lang="it">it</dc:language>
    <dc:type rdf:resource="http://www.eclap.eu/Classification/PerformingArtsType/504"/>
    <edm:type>IMAGE</edm:type>
  </edm:ProvidedCHO>

  <edm:WebResource rdf:about="http://www.eclap.eu/europeana/urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d">
    <edm:rights>http://www.europeana.eu/rights/rr-f/</edm:rights>
  </edm:WebResource>

  <ore:Aggregation rdf:about="urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d:aggregation">
    <edm:aggregatedCHO rdf:resource="urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d"/>
    <edm:dataProvider>Dario Fo & Franca Rame Archive</edm:dataProvider>
    <edm:provider>ECLAP, e-library for Performing Arts</edm:provider>
    <edm:rights>http://www.europeana.eu/rights/rr-f/</edm:rights>
    <edm:isShownAt rdf:resource="http://www.eclap.eu/europeana/urn:axmedis:00000:obj:36b2407e-0ca0-4f44-892b-ebf254118a2d"/>
    <edm:object rdf:resource="http://www.eclap.eu/gif/urn_axmedis_00000_obj_36b2407e-0ca0-4f44-892b-ebf254118a2d.gif"/>
```



```
</ore:Aggregation>
<skos:Concept rdf:about="http://www.eclap.eu/Classification/PerformingArtsType/504">
  <skos:prefLabel xml:lang="it">Teatro</skos:prefLabel>
  <skos:prefLabel xml:lang="en">Theatre</skos:prefLabel>
  <skos:prefLabel xml:lang="da">Teater</skos:prefLabel>
  <skos:prefLabel xml:lang="nl">Theater</skos:prefLabel>
  <skos:prefLabel xml:lang="fr">Théâtre</skos:prefLabel>
  <skos:prefLabel xml:lang="de">Theater</skos:prefLabel>
  <skos:prefLabel xml:lang="el">Θέατρο</skos:prefLabel>
  <skos:prefLabel xml:lang="hu">Színház</skos:prefLabel>
  <skos:prefLabel xml:lang="pl">Teatr</skos:prefLabel>
  <skos:prefLabel xml:lang="pt">Teatro</skos:prefLabel>
  <skos:prefLabel xml:lang="es">Teatro</skos:prefLabel>
  <skos:prefLabel xml:lang="ca">Teatre</skos:prefLabel>
  <skos:prefLabel xml:lang="sl">Gledališče</skos:prefLabel>
</skos:Concept>
</rdf:RDF>
```

2.8 Delivery Method to Europeana Office

The delivery of metadata to the Europeana Office it is done using an OAI-PMH server that publish the metadata using the ECLAP schema. An XSLT has been developed to transform the source XML to EDM compliant schema; this XSLT is used at Europeana to transform records to EDM.

For each partner a specific OAI set has been created allowing updating only the relevant records.

The ingestion and production of records for Europeana follows the following steps:

1. Partners upload XML with metadata to the Metadata Mapping Portal that is mapped and then published for ECLAP portal
2. When a publish notification is received the metadata records are ingested in the ECLAP portal
3. If content is set to be present on the ECLAP FTP, it is checked if it is present and files that are not present are reported to the user.
4. If no errors are found the ingestion of content starts.
5. After 1 week from the ingestion/content upload the content is checked for the minimal metadata needed by Europeana. In case it is ok the content workflow state transit from Uploaded to Under-Approval otherwise an email is sent to the user.
6. When content is Under-approval it can be manually published or automatically published for Europeana (depends on the choice of the partner)
7. Regularly content that is marked published for Europeana is published on the OAI-PMH server. In this phase it is also checked if the preview icon with the minimum requirements for Europeana is present.

2.9 Metadata Interoperability Best Practices

In the context of the ECLAP/EUROPEANA initiative there exist two phases of metadata ingestions, as presented in Section 2.3 (Figure 2-1 and Figure 2-2). Initially metadata from various providers are ingested to the ECLAP repository, while during the second phase of ingestion ECLAP metadata are made available to the EUROPEANA repository.

One of the most important principles in metadata ingestion is interoperability. In the ECLAP context interoperability can be interpreted as the ability to provide uniform access to ingested metadata from different sources. In this section we discuss the best practices applied to ingest proprietary metadata into the ECLAP digital library network and to process ECLAP metadata in terms of search and retrieval.

This paragraph is based on the study on metadata interoperability best practices presented in [2] by L. M. Chan and M. L. Zeng. According to [2] these practices are performed on three different levels:

- Schema level: Efforts are focused on the elements of the schemas, being independent of any applications. The results usually appear as derived element sets or encoded schemas, crosswalks, application profiles, and element registries.
- Record level: Efforts are intended to integrate the metadata records through the mapping of the elements according to the semantic meanings of these elements. Common results include converted records and new records resulting from combining values of existing records.
- Repository level: With harvested or integrated records from varying sources, efforts at this level focus on mapping value strings associated with particular elements (e.g., terms associated with subject or format elements). The results enable cross-collection searching.

In this paragraph we focus on the best practices adopted by ECLAP in order to achieve metadata interoperability:

2.9.1 Derivation

In this approach, a new schema is derived from an existing one. In a collection of digital databases where different components have different needs and different requirements regarding description details, an existing complex schema may be used as the source or model from which new and simpler individual schemas may be derived. Specific derivation methods include adaptation, modification, expansion, partial adaptation, translation, etc. In each case, the new schema is dependent on the source schema.

According to Section 2.4, the ECLAP metadata schema (as defined by DSI UNIFI) is divided in different parts as depicted in Figure 2-4. The Dublin Core metadata part (DC and DCTERMS) reports Dublin Core metadata describing the content. Therefore the ECLAP schema adopts elements from the Dublin Core schema and may be considered its derivation.

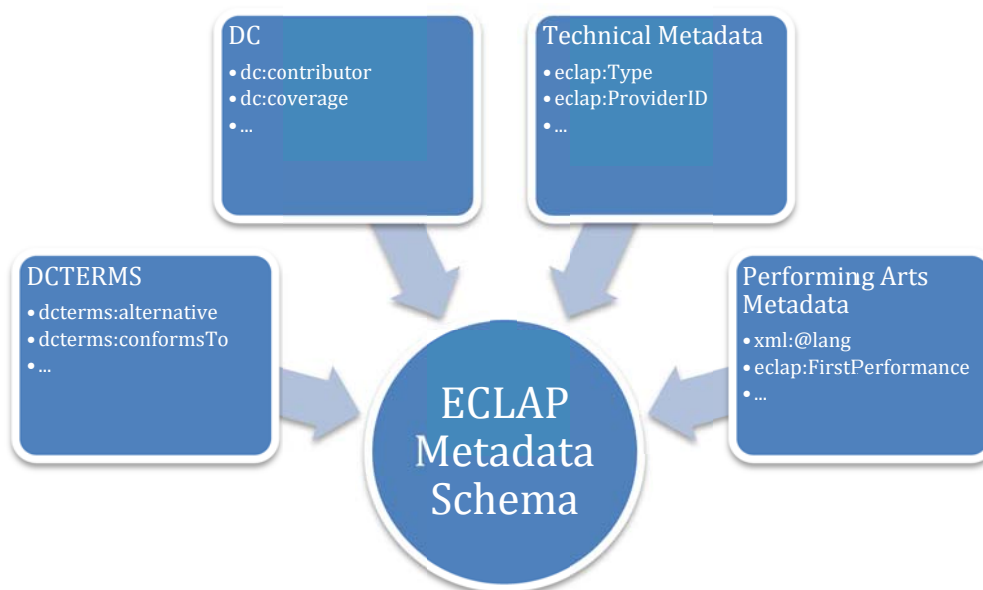


Figure 2-4: ECLAP Metadata Schema

2.9.2 Application Profiles

Even within a particular information community, there are different user requirements and special local needs. The details provided in a particular schema may not meet the needs of all user groups. Therefore, based on the notion that metadata standards are necessarily localized and optimized for specific contents, the concept of application profiles, a typical approach to accommodating individual needs, emerged [6]. While a particular existing schema or schemas are used as the basis for description in a particular digital library or

repository, individual needs are met through a set of specific application guidelines or policies established for a particular interest or user group.

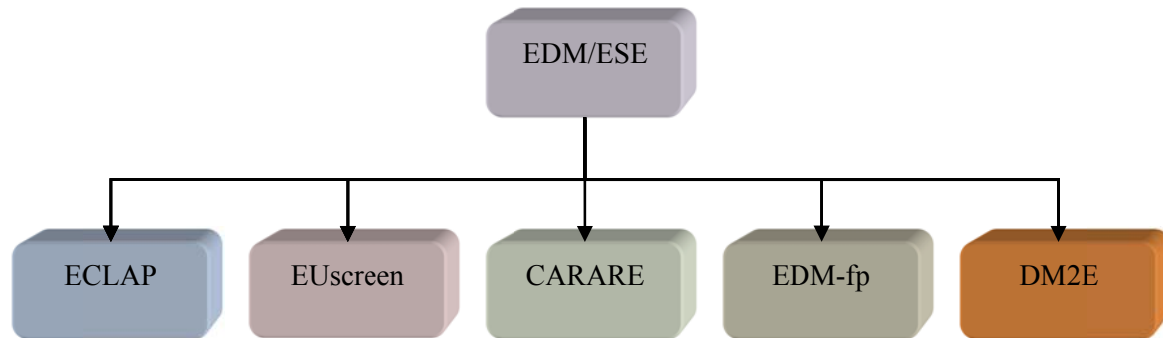


Figure 2-5: Application profiles of the Europeana Data Model and Europeana Semantic Elements Specification

According to Figure 2-5, the ECLAP, EUscreen, CARARE, EDM-fp, and DM2E schemas may be considered application profiles of the EDM and ESE schemas specifically designed to meet the needs of different user communities. The EUscreen schema is designed to represent metadata related to videos, stills, texts and audio from European broadcasters and audiovisual archives. The CARARE schema is designed to establish a service that will make digital content for Europe's unique archaeological monuments and historic sites interoperable with Europeana. The EDM-fp schema is an adaptation of the EDM schema adapted to the needs of the fashion sector. Finally the DM2E schema was developed for use within the Digital Humanities community.

2.9.3 Crosswalks

A crosswalk is “a mapping of the elements, semantics, and syntax from one metadata scheme to those of another” [7]. Currently, crosswalks are by far the most commonly used method to enable interoperability between and among metadata schemas. This method begins with independent metadata schemas. Attempts are made to map or create crosswalks between equivalent or comparable metadata terms. The mechanism used in crosswalks is usually a chart or table that represents the semantic mapping of data elements in one data standard (source) to those in another standard (target) based on the similarity of function or meaning of the elements [8].

In the ECLAP context there exist two levels of crosswalking. On the first level, as depicted in Figure 2-6 each partner contributing content and metadata to the ECLAP initiative has its metadata content organized based on a different schema. Suppose that we have three different providers contributing content organized based on the metadata schemas A, B, and C. Based on these schemas the metadata instances A.1, A.2 (corresponding to schema A) and B.1, B.2 (corresponding to schemas B and C) were created. Based on three different mappings (crosswalkings) mA, mB, and mC these metadata instances are transformed to the ECLAP.1, ECLAP.2, ECLAP.3, and ECLAP.4 metadata instances that comply with the ECLAP metadata ingestion schema. The metadata mapper adopted has been MINT, while the ECLAP metadata ingestion model is the one provided by ECLAP portal. These metadata can be uniformly accessed from the ECLAP portal as performing art content.

On the second level of crosswalking, the metadata instances ECLAP.1, ECLAP.2, ECLAP.3, and ECLAP.4 are transformed to the metadata instances EUROPEANA.1, EUROPEANA.2, EUROPEANA.3, and EUROPEANA.4 based on the mapping mE. This second crosswalk is performed directly by the ECLAP service portal directly without the usage of a metadata mapper.

The new metadata instances comply with the EUROPEANA schema and can be uniformly accessed from the EUROPEANA portal along with other digitized items from various domains. It should be noted that:

- the MINT tool of NTUA has been used to execute a part of mapping of the first crosswalk. NTUA has been responsible for set up and maintenance of MINT and support partner in its usage.

- most of the maps have been directly developed by ECLAP partners, SI has supervised their development and implementation with particular care to those of CTRF as stated into the contract.
- ECLAP ingestion model has been developed by DSI and used as a reference for ingestion
- ECLAP in addition is enriching the ECLAP semantic model for ingestion adopted for the first layer of crosswalking to reach the real ECLAP semantic model with a larger metadata semantic model. This activity is performed by ECLAP back office and tools to add: annotations, comments, geolocations, vip naming via dbpedia, collections, playlists, score, technical metadata, etc. etc. (activities and models developed by DSI UNIFI).
- ECLAP full semantic model has been developed by DSI and used as a source for creating the EDM by generation exploiting the ECLAP tools directly developed by DSI also.

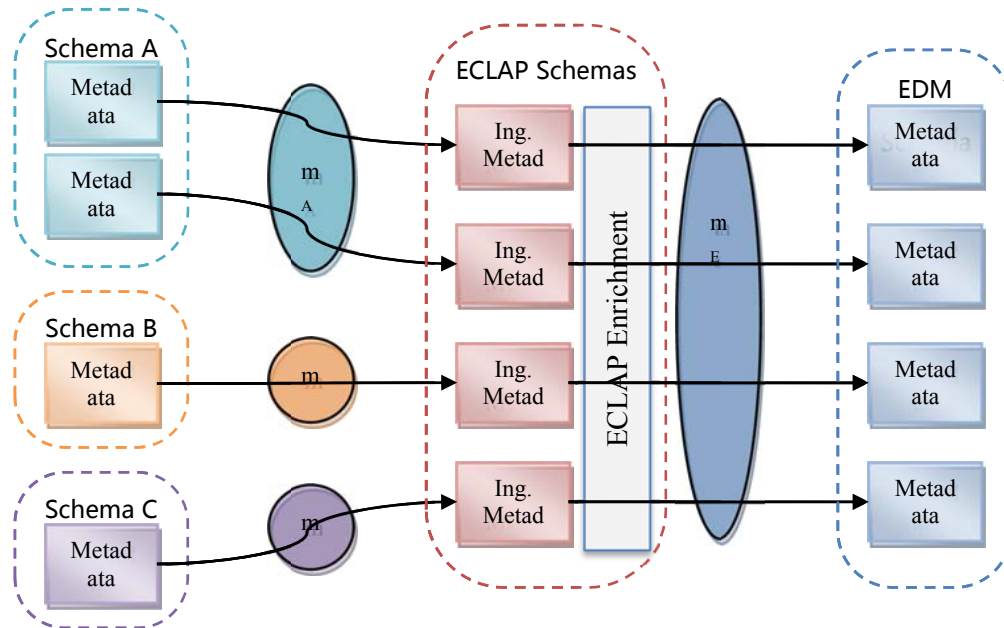


Figure 2-6: Crosswalking example in the ECLAP context.

2.9.4 Switching-across

Crosswalking among multiple schemas is not only extremely tedious and labour intensive but also requires enormous intellectual effort. For example, a one-way crosswalk requires one mapping process ($A \rightarrow B$), and a two-way crosswalk requires two mapping processes ($A \rightarrow B$ and $B \rightarrow A$). The process becomes more and more cumbersome when more schemas are involved. A four-schema crosswalk would require twelve (or six pairs of) mapping processes. As a result, using a switching schema (new or existing) to channel crosswalking among multiple schemas has become a well-accepted solution.

2.9.5 Metadata Framework and Metadata Registry

A framework can be considered as a skeleton upon which various objects are integrated for a given solution. The need for a metadata framework is best demonstrated by emerging digital preservation efforts. While many organizations have developed metadata for digital preservation in support of their own activities, such efforts have been conducted largely in isolation, lacking any substantial degree of cross-organizational coordination. It becomes obvious that a metadata framework is needed to represent a consensus of leading experts and practitioners and could be readily applied to a broad range of such activities [9].

The EUROPEANA metadata schemas ESE and EDM are metadata frameworks for providing a single access point to millions of books, paintings, films, museum objects and archival records that have been digitized throughout Europe. Thus Europeana is an authoritative source of information coming from European cultural and scientific institutions. Moreover the ECLAP metadata schema is a framework for providing a single access point for performing arts content. Contributions to Standardization bodies have been prepared and

provided to the corresponding bodies. See ECLAP DE4.3.1 Performing Arts Metadata vs Standardization, <http://www.eclap.eu/115121>.

The purpose of a metadata registry is fairly straightforward: to collect data regarding metadata schemas. Because the reuse of existing metadata terms is essential to achieving interoperability among metadata element sets, the identification of existing terms becomes a prerequisite step in any new metadata schema development process. Thus the presence of a metadata registry application “promotes the wider adoption, standardization and interoperability of metadata by facilitating its discovery, and reuse, across diverse disciplines and communities of practice” [10].

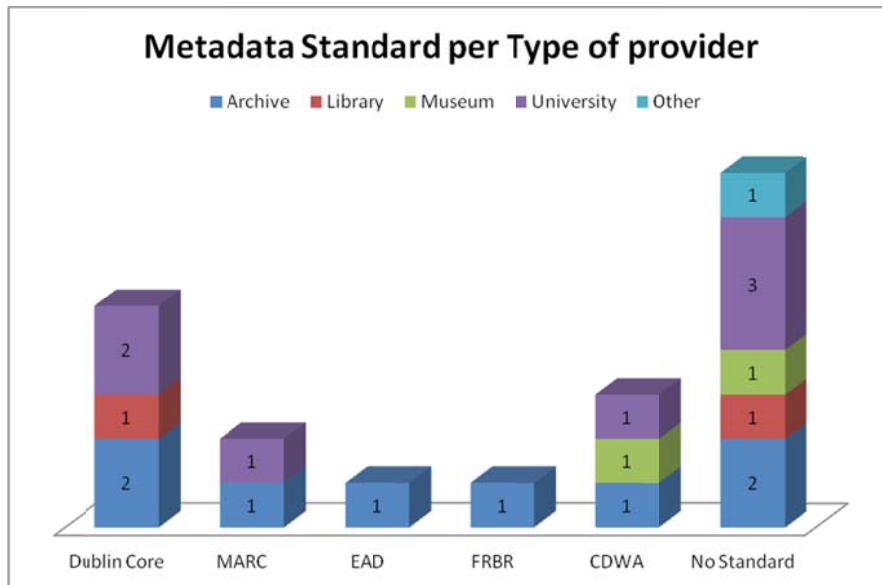


Figure 2-8: Metadata standards adopted by ECLAP content providers.

In the ECLAP domain, a metadata registry was created based on the results and the analysis of the provided questionnaires and samples of the most commonly used metadata standards among content providers. Dublin Core, MARC, EAD, CDWA and a simplified customization of FRBR are the mostly adopted metadata schemas as depicted in Figure 2- while the majority of content providers did not use any standard for their metadata.

The work performed on standards and ECLAP metadata highlighted that an urgent action is needed to converge to widely shared metadata standard for performing arts. ECLAP posed the basis for this work as highlighted in Deliverable <http://www.eclap.eu/115121> and in the results received from the standard bodies.

2.9.6 A Metadata Repository Based on the Open Archives Initiative (OAI) Protocol

OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is a protocol developed by the Open Archives Initiative. It is used to harvest (or collect) the metadata descriptions of the records in an archive so that services can be built using metadata from many archives.

The ECLAP Repository has facilitated the construction of an automated "ingestion" system, based on the Open Archives Initiative Protocol for Metadata Harvesting (<http://www.eclap.eu/115121>). Through this protocol, metadata flows into the Metadata Repository with a minimum of ongoing human intervention.

In ECLAP two different OAI-PMH servers have been set up. One, located in Athena, for the ingestion of metadata in output of the MINT tool and accessed by the ECLAP Service Portal. A second located in Florence at ECLAP Portal premises to export the metadata to Europeana.

2.9.7 Aggregation

As described, the ECLAP Metadata Repository employs an automated ingestion system based on OAI-PMH, whereby metadata flows into the Metadata Repository with a minimum of ongoing human intervention. The ECLAP, from this perspective, functions essentially as a metadata aggregator.

Four categories of problems that limit the usefulness of metadata have been identified by: (1) Missing data: elements not present, (2) Incorrect data: values not conforming to proper usage, (3) Confusing data: embedded html tags, improper separation of multiple elements, etc., and (4) Insufficient data: no indication of controlled vocabularies, formats, etc. It is possible that these problems could be partially eliminated through a process called aggregation in a metadata repository. The notion behind this process is that each metadata record contains a series of statements about a particular resource, and therefore metadata from different sources can be aggregated to build a more complete profile of that resource.

ECLAP allows to aggregate and to relate metadata originating from different content providers. A detailed report on content aggregation is presented in the corresponding part of this document.

A second source of aggregation in ECLAP has been performed by the ECLAP tools to create: collections, playlists and annotations. Some of the semantic information have been used to enrich the metadata and provide to Europeana richer EDM models.

2.9.8 Element-based and Value-based Crosswalking Services

While presently crosswalks have paved a way to the relatively effective exchange and sharing of schema and data, there is a further need for effective crosswalks to solve the everyday problem of ensuring consistency in large databases that are built of records from multiple sources. According to [11] there is a need for robust systems that can handle validation, enhancement, and multiple character encodings and allow human guidance of the translation process. It is stated that usable crosswalks must have the following characteristics:

- A set of mappings between metadata standards that is endorsed by a stakeholder community.
- A machine-processable encoding.
- A well-defined relationship to source and target metadata standards, which must make reference to particular versions and syntactic encodings.

The tools used by ECLAP allowed value based crosswalking services and enrichment by enabling a set of conditional restriction related to the metadata values.

3 Content selection and aggregation

The dramatic increase of the quantity of cultural heritage resources available online is challenging our usual ways of studying and accessing culture [14]. Cultural heritage institutions are facing the necessity to guarantee the retrieval, valorisation and, eventually, the re-usability of their Non Book Materials and texts.

Within ECLAP, the possibility of aggregating content - pertaining to different media type, institutions, performing arts and nations - in significant clusters, represents an opportunity for performing arts institutions to increase the accessibility to their resources. For example, each ECLAP partner, through its own wall page, can present users with an overview of the collections hosted by ECLAP and enhance the most precious and rare resources through the Featured Content page or, as embedded content, within the presentation page.

To aggregate content also means to contextualize a single resource in relation to other elements or resources: on ECLAP, content can be compared and contextualized at a transnational level, as the aggregation activities which can be performed on the portal allow content providers to relate their own items to those provided by other partners. In addition to this, digital objects can be compared with the content already accessible through Europeana, since, whenever a user chooses to visualize a resource on the ECLAP portal, some related objects from Europeana are shown as embedded thumbnails provided with provenance information.

It has to be noted that, in the case of the traditional performing arts forms, historians looking back at the past cannot directly analyze the object of their studies (*i.e.* the performance) as it may happen, for example, in the field of fine arts, but need to collect documents of different nature (*e.g.*: letters, sketches, biographies, articles, interviews, etc.) and type (texts, images, video and audio recording) to get an idea of what the event performed in the *hic et nunc* might have been [15]. Besides, scholars do not just reconstruct the *mise-en-scène*, but also investigate, comparing different documents, the artistic process and the coeval socio-political, economic and cultural circumstances in which they took place [16]. Also students are usually encouraged to

apply the same method to study the performing arts phenomena, especially when they are approaching the subject of their studies from a theoretical perspective.

For these reasons, the wide range of possibilities offered by ECLAP to further enrich and contextualize the resources and connecting them through users' annotations and tools such as Mystoryplayer <http://www.eclap.eu/3748> and the Social Graph <http://www.eclap.eu/116088> are particularly useful for the purposes of research and education, which proved to be the most representative area ECLAP end-users belong to [17]. Both these tools have been selected by Europeana and included into the Europeana pro web pages. Other categories of users, such as the performing arts lovers and professionals, will benefit from the aggregations obtained through the tools, as they can be useful to orient those users who are not domain experts to browse through the ECLAP environment, while professionals will be able to deepen their knowledge or to transform the ECLAP cultural repository in a source to draw material to produce new creations.

3.1 Featured Content Section, Content Partners' Wall, Top Rated Page

Some of the content provided by partners is unique or particularly rare: in order to enhance such resources and guarantee their retrieval, they are clustered both in the Featured Content section and in each partner's Wall Page. The first solution displays an overall selection of the best of the ECLAP resources as aggregated content, the latter clusters best items pertaining to the same provider and relates them to their "mother" institution thus contextualizing them in respect to the collection they belong to.

Each content partner has a group, which is accessible through the Content Providers menu: groups have their own blog, a forum and different pages to display their items. The first page the user visualizes when choosing to explore a group is the Wall Page, that can be customized by each content provider and allows institutions to present their history, activities and mission. Most of the ECLAP partners also enriched this page by describing their collections and contextualizing them, both from a cultural and an historical point of view. Best items of each content provider are displayed as embedded videos, images and texts.

Through the Featured Content section and the Wall Pages, users can easily retrieve what domain experts consider as valuable documents for the performing arts history, while a third option, represented by the Top Rated section, makes it also possible to browse that content users contributed to aggregate with their positive evaluation and comments.

3.2 Search Filters and Location Map

Content aggregation can be performed using the ECLAP Search Filters and the Classification, through the following categories:

- Content Language
- Format
- Genre
- Group
- Historical Period
- Management and Organisation
- Metadata Language
- Movements and Styles
- Performing Arts
- Professionals
- Published by
- Subject

Filters and Classification allow users to aggregate items and contextualize them in relation, for example, to a specific genre or historical period, while the document type and format can be selected both through the filters and through an additional search button. The classification, the specific performing arts metadata elements and the filtered search make it possible to draw cross-cultural comparisons, for example, to realize that the same genre may occur in different time span depending on the country in which the event took place or to see how the same subject is addressed in different geographical areas.

Geographic information on the resources can be further analyzed through a geolocation tool which links items to the Location Map in order to aggregate and contextualize content in regard to the place it was

created <http://www.eclap.eu/drupal/?q=home/location> . The map is not just showing the single item provenance (fig. 3-1), but also a cluster of all the other resources related to the same area. Through the zoom, it is also possible to get a global overview of the geographic provenance of the ECLAP content.

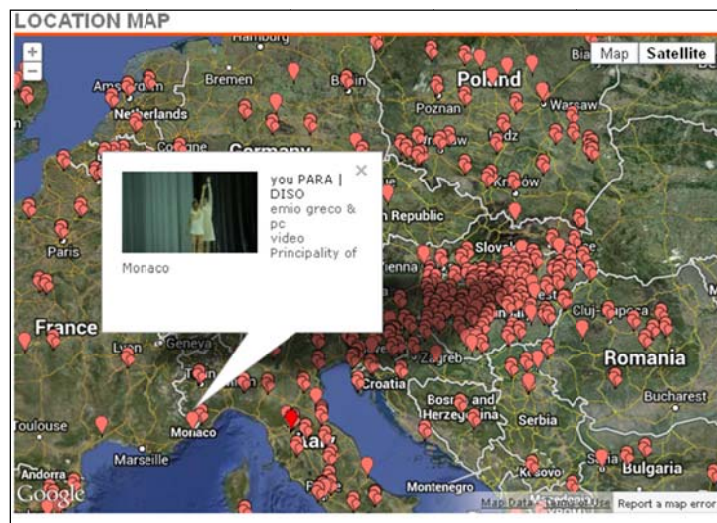


Figure 3-1: Location Map

3.3 Playlists and Collections

Another way to aggregate content is the creation of playlists and collections <http://www.eclap.eu/3738> .

Playlists give users the possibility to play series of multimedia items in a specific order, defined during the editing phase. The user may choose to insert in the playlist not an entire video or audio, but just the fraction of a video, audio or an image for a certain time interval (fig. 3-2). Once the editing phase is completed, the user's playlist can be saved and metadata records can be added, in order to store it in the database as a digital object and to make it retrievable both for its creator and other users [6].

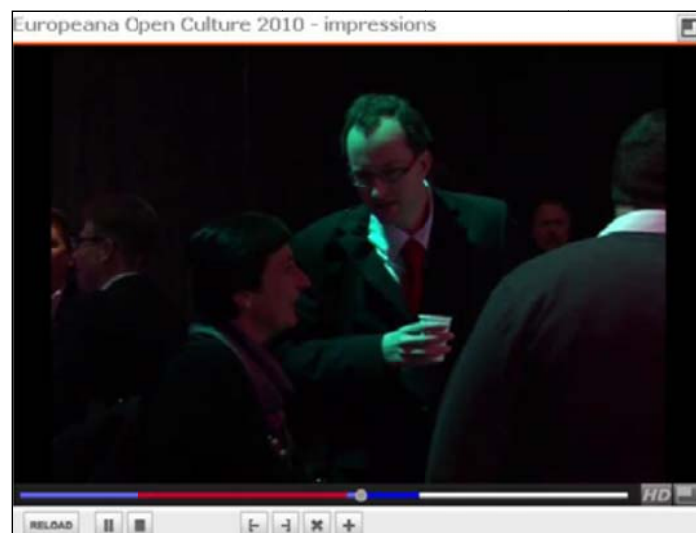


Figure 3-2: Player loaded in order to choose a media segment for the playlist

Collections are sets of objects: audio, video, images, texts, playlists, zip files, pdf documents, etc., grouped together on the basis of the user's purposes. Collections can be monographic or built as a first step to collect the content needed to prepare a lesson for the e-learning environment (LMS). Collections may be published on the portal or saved by the user as private content, visible only to their creator, if published, collections are indexed as new content and become visible to all registered users [19]. This process is functional for leisure

and social activities, but especially for educational purposes, as both private and published collections can be automatically exported on the LMS area of ECLAP. Teachers can therefore organize their presentations for classes and then publish them, thus allowing their colleagues to reuse learning objects that have been already shaped to illustrate a specific topic or point out a certain aspect of a subject. From this point of view, ECLAP can operate as a learning content repository supplying teachers and students with items already aggregated from a critical point of view.

The possibility to gather items in playlists that can then be included in collections is of great importance to assemble resources belonging to the same series such as photographs, letters or video clips of a same event. For example, shootings of the same moment of a show taken by different angles can be grouped within a playlist, while the sequence of the different parts of the show can be reconstructed by organizing these playlists in chronological order within a collection. Transnational collections related to a specific performance or artist can also be shaped clustering diverse document types associated to the same event or person but provided by different institutions.



Figure 3.3: Block to organize collections

3.4 Social Graph

Through the Social Graph it is possible to get an immediate visual representation of all the items and events connected to an object through all kind of different relationships (fig. 3-4) <http://www.eclap.eu/116088>. This interactive application aggregates related objects and collections, dynamically showing each connection through arcs and declaring the type of relationship through nodes. The user can decide which kind of clusters and relationships to explore by ticking boxes which can open or close the existing connections built on the basis of people (Creator, Colleagues, Administrators, Groups, Writer, Group Member), events (Annotations, Publications, Comments, User's favourites, Taxonomy) or related content (Collections, Object List, Related Objects). This way the item is rapidly contextualized in respect to other items, but also to the taxonomy, annotations and comments. Moreover, it is possible to view all the events related to the ECLAP lifecycle of the resource, for example to know who published it and to which institution it belongs to. The basic idea for this navigation tool is to make the browsing experience more intuitive and associative, helping the user to find new connections within the content aggregation and with other users¹.

¹ For further details please visit <http://www.eclap.eu/116088> and paragraph 4.6 of this deliverable.

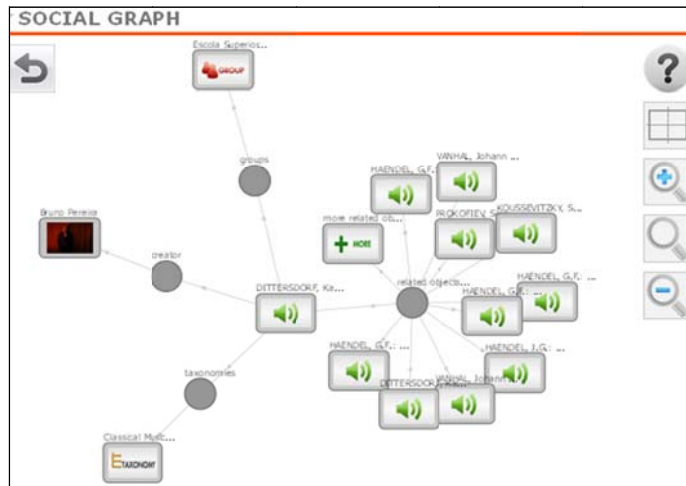


Figure 3.4 – Social graph

3.5 MyStoryPlayer

MyStoryPlayer is a tool that allows users to annotate multimedia objects and draw temporal connections by organizing items on a multiple timeline <http://www.eclap.eu/3748>. In MyStoryPlayer items do not need to be played one after another, but can coexist on the same page simultaneously, integrating each other and thus giving the user the possibility to layer content in a more complex structure than that allowed by playlists and collections (fig. 3-5).

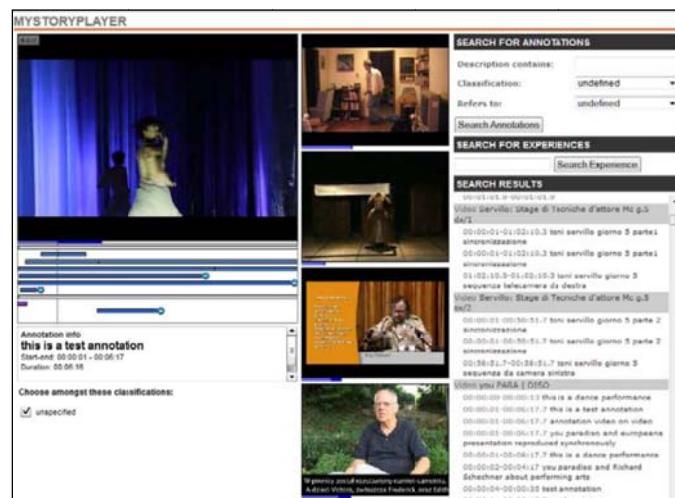


Figure 3-5 – MyStoryPlayer interface

This tool also allows synchronizing audiovisual objects, which is of capital importance in the case of the multicamera shootings of performances, rehearsals, seminars and acting classes provided by ECLAP. When all the three angles perspectives (frontal, left and right side) caught by the camera are aggregated and played simultaneously, it is possible, for example, to analyse the acting technique of an actor by viewing on the same page and in the same moment what he is doing with his right and left profiles and with his whole body. As for the case of playlists and collections, to cluster items in a structured way is useful to reconstruct events (fig. 3-6) and to prepare presentations or multimedia essays on a certain subjects. In the educational environment, MyStoryPlayer allows teachers to prepare dynamic lessons showing images, texts, audio and

video recordings, moreover, annotations can be added on the timeline and displayed during a specific time interval.

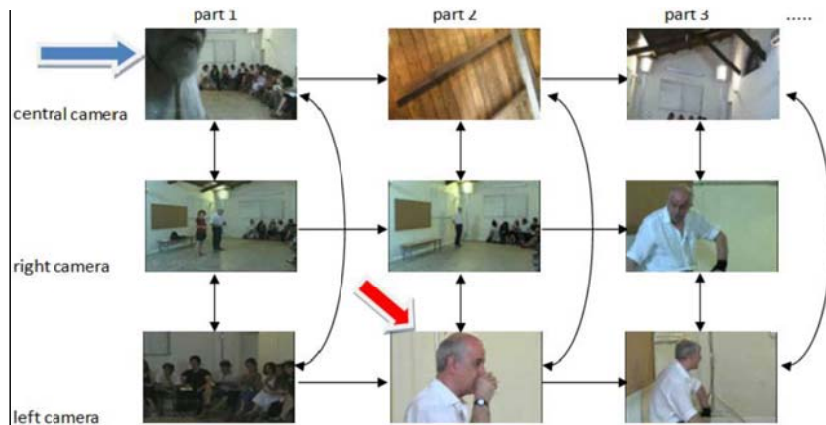


Figure 3-6: Reconstruction of an event through MyStoryPlayer

MyStoryPlayer do not only allows multiple contemporary views but also multiple choices on what should be enhanced and put on the foreground: each time the same structured content is replayed, users can view in different ways, performing a new experience, which can be saved and shared with other users, thus giving teachers the opportunity to reuse and enrich their colleagues' learning objects and students to access the didactic content.

The semantic connections can be visualized through the Social Graph, while from a specific page it is possible to explore the network of relationships among audiovisual content defined by using MyStoryPlayer, as shown in the figure below (fig. 3-7), where two different views are available: by clicking on them, users can access the corresponding interactive page, where they can drag and drop a video to see its relationships with the others [20].

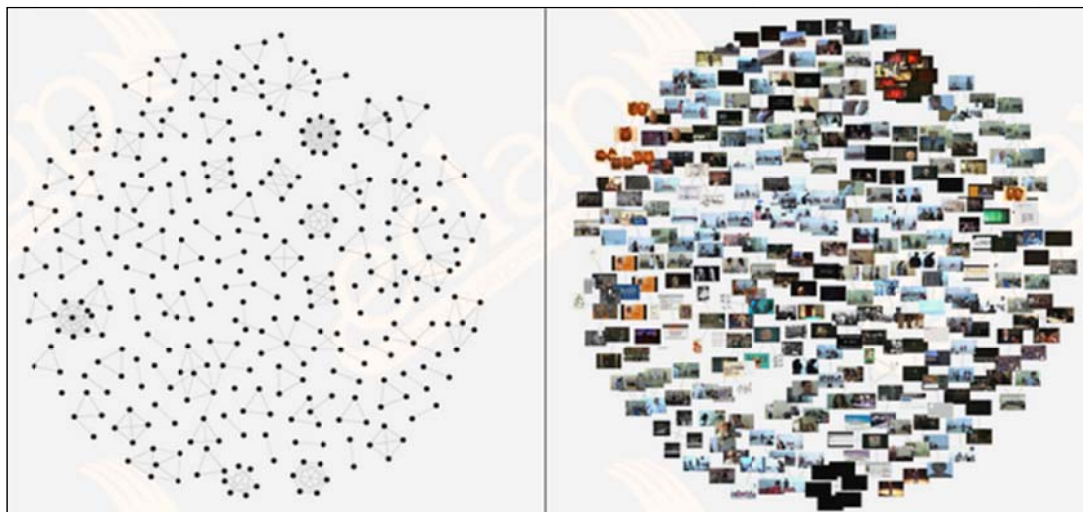


Figure 3-7: Relationships among audiovisual content

4 Content Enrichment on ECLAP service Portal, augmentation and rich media

ECLAP provides services and tools for automated content ingestion, adaptation, metadata ingestion and editing, semantic information extraction, indexing and distribution by exploiting the most innovative and consolidated technologies [34], [35], [36], [37], [38], [39], [40], [41]. The content ingestion may start taking metadata and content files from any kind of archive and/or database or by providing them via FTP and/or web based utilities. Once the metadata area ingested, an intelligent content processing back office is capable of collecting and automatically repurposing content for distribution via pc and mobiles, coping with more than 500 digital file formats.

ECLAP supports the institutions in all their activities: metadata selection and mapping, content ingestion, to the definition and management of permissions and licenses on contents, and finally managing their users on ECLAP services. For details about the ECLAP ingestion process and facilities see: <http://www.eclap.eu/3739>

According to ECLAP workflow, the obtained metadata are sent to Europeana only after that the metadata have been enriched and linked to a reachable digital resource and when the IPR details have been finalized, with needed quality level. The IPR management and the assignment of access restrictions is a way to enable the increment of possible available content on the internet. Permissions as IPR models can be enforced on content by each ECLAP institution (content owner), by using the IPR Wizard tool [38]. An ECLAP IPR Model can be associated with each single content or collection. The IPR model has been derived from the work performed on MPEG-21. This means that access rules are imposed to restrict and regulate the content access taking into account: content format (video, audio, document, etc.), actions/rights (play, download, stream, embed, etc.), device (PC, mobile, mobile application), users' type (private, public, educational, etc.), location (nationality, university, ..), resolution (HD, high quality, medium, low, etc.) [37]. This model for content distribution with IPR management, is associated with a strong legal model as Terms of Use and privacy policy (see them on the portal). This allowed ECLAP to have in two years of content ingestion and distribution only one IPR resolution to be managed, that have been solved in a couple of emails, despite to the high number of newly published and accessed content items, and content providers involved, coming from different nationalities. For details about the IPR model and tools see: <http://www.eclap.eu/4021>

The ECLAP content management performs a wide range of metadata enrichment activities (based on AXCP media grid [42]). The typical **metadata enrichments** performed by ECLAP can be the addition of technical descriptors of source files, the addition of more languages, the geo localization passing from location named into metadata and descriptors to formal GPS position, the production of QR codes for museum inspection and linkage (see it as augmented reality first step), the content aggregation, the addition of comments and tags, the association of taxonomical classification, the establishing of connections with dbPedia open data, the addition of a formal IPR license descriptor, the association of univocal date and time, the association of an UUID (permitting the management of any kind of identifiers that may be available for the single content element such as: ISBN, ISAN, ISMN, private coding IDs, etc.), the production of LOD model, the content aggregations, etc. [37]. As a result, the content is described in terms of metadata based on the so called ECLAP semantic model which is much richer than the ECLAP ingestion model [43]. For the ECLAP metadata ingestion model: <http://www.eclap.eu/40876>.

Therefore, the activity of **content publishing to Europeana** and as **Open Data** is very simple for the institutions since the content can be automatically ingested in several different manners, processed and automatically adapted for format, metadata and IPR, to arrive to be published as full content information on ECLAP service portal, while a subset of the metadata are provided to Europeana according to the EDM, Europeana Data Model. The enriched metadata are also made available in different formats (as LOD, OAI-PMH, DC, etc.) for the former institutions (content providers) and for massive diffusion and promotion. Thus, the content published on ECLAP is widely indexed on all major engines and can be accessed from any kind of device, from PC, TV and mobiles. The ECLAP service also allows to update and review the extended metadata, for example to make corrections, add more details and/or additional languages, create links with

other sources, etc., and automatically provides the updates towards Europeana, LOD, ECLAP, mobile, etc. For more details about the ECLAP semantic model see ECLAP deliverable on page support: <http://www.eclap.eu/3609>

The ECLAP content model on WEB also allows the embedding (Copy HTML) of ECLAP content in third parties portals. The institutions may exploit ECLAP as streaming service for the audio-visual content by integrating in their lighter servers the provided HTML code of the ECLAP players. For details about citation and embedding: <http://www.eclap.eu/40876>

4.1 Linking to Geonames dataset

In order to link content record with Geonames, the data fields containing geographical information were analyzed to find matches with names that are present in the Geonames dataset. The fields considered are the (first)performance place, city and country. Since exact matching does not produce sufficient results the matching was done using full text search of the metadata field over the geographical names, the results have been filtered requiring that the words of the matched name have to be present in the metadata field. Moreover, when the country field is identified for the identification of city or place the search is limited to names of that country.

An analysis was performed on the values matched for the (First) Performance City, Country and Place as reported in Table 3-1. For each of the distinct values of the fields were evaluated if the match with the geoname entity was correct, wrong or missing. It can be seen that there is a high precision for City and Country and quite low for the Place. While there is a high recall for the City and quite low for the Country.

	values				
	precision	recall	#good	#wrong	#missing
City	94,74%	96,12%	198	11	8
Country	98,72%	52,74%	77	1	69
Place	65,77%	--	684	356	?

No evaluation was done for the missing places for the high number of places to be evaluated.

The high number of missing for the country is due to the fact that are present countries names with 2 letter ISO codes that are not present in the geonames names indexed. This can be improved considering also this possibility. While the high number of wrong for the Place values is due to places matched that do not agree with the city, in general the places do not provide values that are present in the geonames database when the city and/or country are provided. However there are records that have in the Place field also City and Country information and for these values the matching was good (precision 82,19%).

Since many objects share the same value for the Place, City and Country in Table 3-2 is reported the same evaluation considering the number of objects having the value that was considered good, wrong or missing.

	objects				
	precision	recall	#good	#wrong	#missing
City	98,08%	98,83%	23708	463	280
Country	99,54%	16,08%	7354	34	38368
Place	54,04%	--	4962	4220	?

This leads to similar results the number of repeated values increases the precision for City and Country while decreasing it for Places and increasing the recall for City and decreasing it for Country.

4.2 Date extraction and timeline

A module was developed to normalize the data information that is available in all the different fields that are present like dc:date, dcterms:issued, but also present in other fields that are non necessarily related to dates like dc:description and in fields like PerformanceDate, FirstPerformanceDate, RecordingDate, etc.

There are cases that the date is ambiguous like 02/03/2010, is it March 2nd or February 3rd? To disambiguate the date we made the hypothesis that the same provider uses the same date format for all the dates and when the date is ambiguous the most frequent way of representing dates by the provider is used to interpret the date.

The date information that is extracted is used to build an interactive viewer of the timeline, a prototype is available at <http://www.eclap.eu/timeline> as it can be seen in Figure 4-1. The user can zoom on a date double clicking on it or selecting a bar can see for the date selected in which kind of fields that date is used. The user can also check/uncheck which kinds of fields she is interested in for example in Figure 4-2 is reported the case of performance dates.

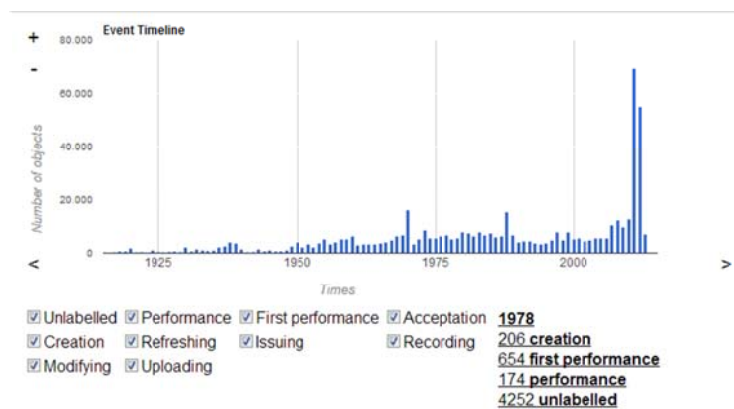


Figure 4-1: Timeline viewer prototype

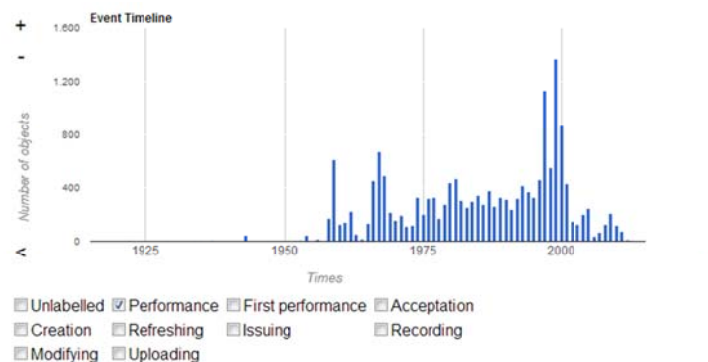


Figure 4-2: Timeline for performance dates

4.3 Names identification

Another enrichment developed is the identification of names in the metadata fields. The problem of Named Entity Recognition in text is a well-known problem and it can be solved with a variety of solutions ranging from simple grammars to machine learning. However there is also the problem of identifying synonyms and reviewing the names automatically identified.

The solution developed is made of three phases

1. The metadata fields are analyzed to identify names using an application based on the GATE framework [21] and in particular the ANNIE plugins with some additional Jape rules. The names

- identified are store in the db and associated with the metadata field referring to it. A special text analyzer for the Italian developed by DSI in the past has been exploited.
2. A web application is used to review the names identified allowing the user to blacklist wrongly identified names and to aggregate more names under the same name telling the system that they are synonyms of a ‘master’ name. In this phase the system suggests possible synonyms for a selected name (see Figure 4-3).
 3. The names that are identified as ‘root’ names are searched on dbpedia and among the ECLAP users to possibly associate metadata fields to such entities.

The three phases are run in parallel allowing the user to not wait for the elaboration of the huge number of metadata fields to start working on blacklisting wrong names or associating synonyms.

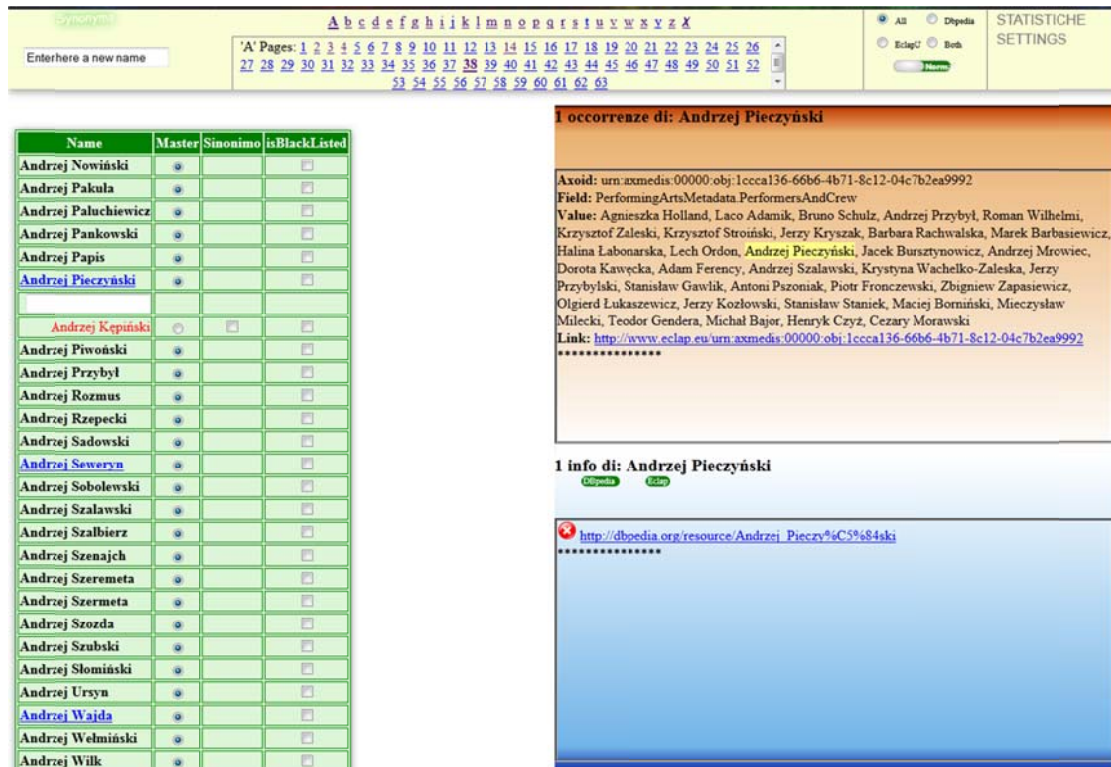


Figure 4-3: Phase 2 web application

4.4 ECLAP Linked Open Data Service

Linked Data is a technique for data publishing that uses common Web technologies to connect related data and make them accessible on the Web. It is based on identifying resources with HTTP Uniform Resource Identifiers (URI), and, using standards like the Resource Description Framework (RDF) to provide data about these resources and to connect them to other resources on the Web [22].

Regarding the description of resources the best practices for publishing linked data suggest to reuse vocabularies already available. Reuse can be done by using classes and properties as they are or by creating a specific vocabulary and defining sub classes and sub properties of the ones already defined.

Some well known basic vocabularies are present:

- Dublin Core (<http://purl.org/dc/terms/>) for the description of human-created arti-facts,
- Friend of a Friend (<http://xmlns.com/foaf/1.0/>) for the description of persons, organization and relations among them,
- Creative Commons (<http://creativecommons.org/ns#>) for the representation of legal information about works,
- Basic Geo Vocabulary (http://www.w3.org/2003/01/geo/wgs84_pos#) for basic properties for the representation of geographical coordinates.

In the field of performing arts, there are some contributions, not all the aspects are covered by a single vocabulary. The Music Ontology [23] used by BBC programmes and music [24] and DBtune covers only the music related information, the Linked Movie Database has a vocabulary specific for the film domain and other ontologies like DBpedia [25] and Freebase are quite generic.

Another relevant aspect is the description of annotations of multimedia content, the Annotea project [26] was one of the first to adopt semantic web technologies for annotations but it was originally designed for annotations of web sites and therefore offers limited capabilities for annotating multimedia objects. The LEMO annotation framework [27] built on top of Annotea model supports annotations of media fragments [28]. Recently the Open Annotation Collaboration [29] model has been proposed and it is designed for the use as linked data.

4.4.1 ECLAP Semantic model

In Figure 4-4, the general ECLAP semantic model is shown where almost all the mayor entities managed by the system are reported.

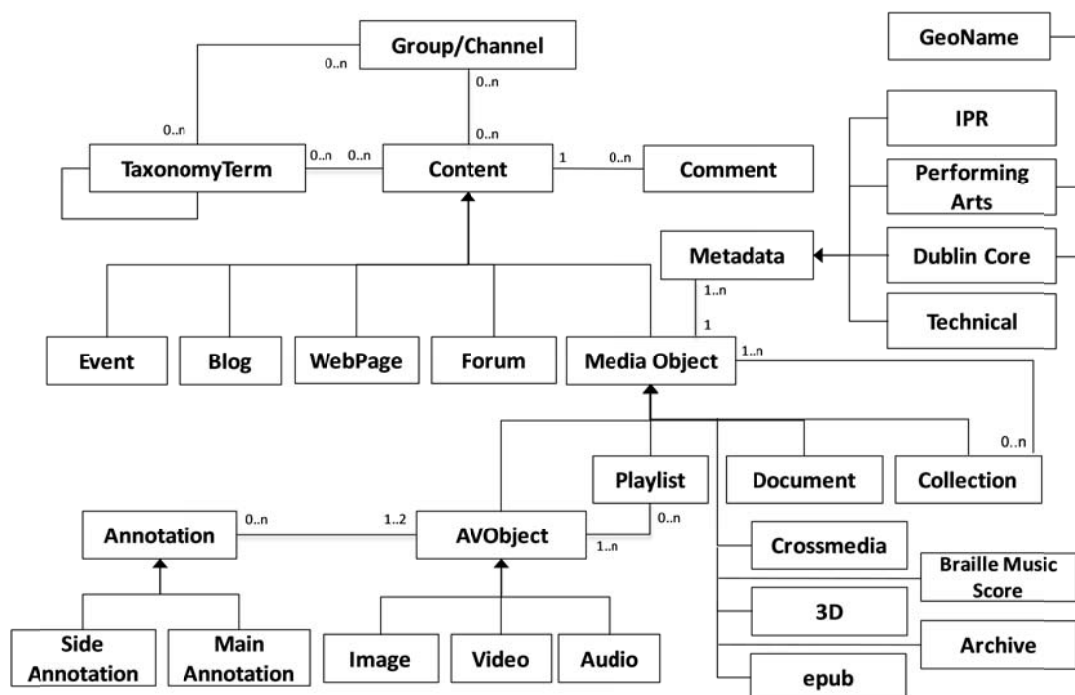


Figure 4-4: ECLAP Semantic model

The *Content* elements that represent the content managed by the portal are associated with *Groups/Channels* providing the content (each ECLAP content provider has a group). Moreover each content can have *Comments* and can be associated with terms taken from a taxonomy. *Content* is specialized in *Events*, *Blogs*, *WebPage*, *Forum* and *Media Objects*. Blogs, WebPages and Forums are used for providing news, general unstructured information and for stimulating users discussions on specific topics while *Media Objects* represent the multimedia content and their aggregations that are accessible from the ECLAP portal. The *Media Objects* are specialized in *AVObjects* (Image, Video, Audio) that can be used in *Annotations* and in *Playlists*. *Playlists* aggregate *AVObjects* in a sequence allowing to use even a fragment of the Audio/Video. *Collections* aggregate a set of *Media Objects* and in this case can include also *Documents*, *Playlists* and also other *Collections*. *Annotations* are used to associate a textual description with an audio visual object or to its fragment, moreover it is possible to associate with another audio visual object (or its fragment) allowing to link two audio visual contents or even different parts of the same content. This kind of annotations are presented to the user via the MyStoryPlayer tool [30], [31].

In Figure 4-5, the relations among *Users* and other entities are depicted. A user may be a member of one or more groups and he/she can be the group administrator. Each content is provided by a user and each *Media*

Object may be marked as favorite (similar to the facebook “Like”) by an user, moreover a group administrator can mark the content as to be featured on the featured object list on the portal. *Comments* and *Annotations* are linked to the user that created them. Finally users are linked with other users with the ‘knows’ relation that builds the classical ‘social graph’ and each user can specify the topic of interests among the taxonomy terms.

Each *Media Object* is associated with different sets of metadata, the general *Dublin Core* metadata (e.g. title, subject, type, description), the *Technical* metadata related to the content and its management (e.g. audio/video duration, workflow type), the metadata per IPR management and specific metadata for performing arts information (e.g., performance place, performance date, performing arts type, etc.). To be noted that some of the Dublin core and performing arts metadata elements (e.g. coverage, spatial, performance place, performance city and country) can be associated with GeoNames entities to allow to link to the GeoNames linked data service.

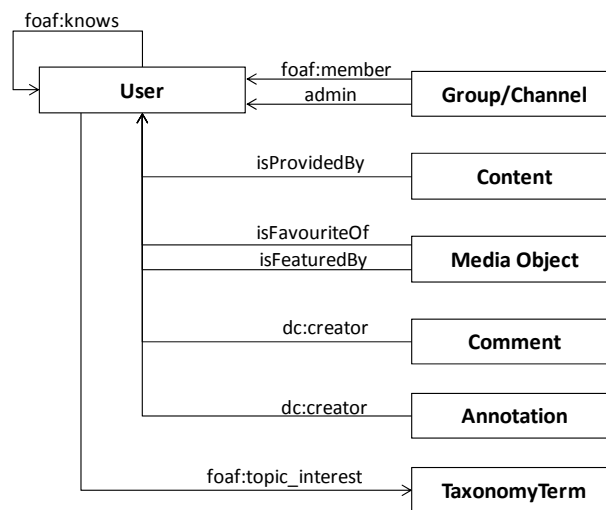


Figure 4-5: Relations of users with other entities

In Table 4-1, some of the properties specifically defined for performing arts are reported. These properties are defined as specialization of Dublin core properties and they were identified analyzing the metadata schemas used by ECLAP partners as well as schemas used by other projects and metadata standards. Among the properties identified information about the performance depicted in the resource (place, city, country and date), the premiere of the performance (place, city, country and date), the contributors to the creation of the performance each with the specific role (actor, dancer, light designer, hairdresser) is present.

Table 4-7. Some performing arts metadata properties

<i>property</i>	<i>sub property of</i>	<i>description</i>
performancePlace	dcterms:spatial	theatre or venue of the performance
performanceCity	dcterms:spatial	city of the performance
performanceCountry	dcterms:spatial	country of the performance
performanceDate	dcterms:issued	date of the performance
firstPerformancePlace	dcterms:spatial	venue of the premier of the performance
firstPerformanceCity	dcterms:spatial	city of the premier
firstPerformanceCountry	dcterms:spatial	country of the premier
firstPerformanceDate	dcterms:issued	date of the premier
plotSummary	dcterms:abstract	summary of the plot
performingArtsProfessional	dcterms:contributor	person involved in the performance realization
dancer	performingArtsProfessional	
actor	performingArtsProfessional	
director	performingArtsProfessional	
...		

A complete description of the ECLAP metadata fields is reported in [32].

4.4.2 ECLAP LOD service

The ECLAP portal allows to access RDF descriptions of digital resources that are available on it using specific URIs, the RDF description of the resource is provided in case of a LOD enabled browser otherwise the standard web browsers are redirected to the usual HTML page with a human readable description <http://www.eclap.eu/103998> . Also the taxonomy terms used to classify content are accessible using LOD as well as the content annotations that relate them, the groups to which the content is bound (e.g., the group of the provider) and the ECLAP users.

The URIs currently supported are in the forms:

- <http://www.eclap.eu/resource/object/<axoid>>
- <http://www.eclap.eu/resource/term/<tid>>
- <http://www.eclap.eu/resource/annotation/<aid>>
- <http://www.eclap.eu/resource/group/<gid>>
- <http://www.eclap.eu/resource/user/<uid>>

where *<axoid>* is the identifier assigned to the content when uploaded (e.g., urn:axmedis:00000:obj:04e0caef-b33b-4f4a-ba50-a80d96766192), *<tid>* is the identifier of the vocabulary term (e.g., 501 for Dance), *<aid>* is the identifier assigned to the annotation, *<gid>* is the identifier of the group (e.g., 3160 for the Development group) and *<uid>* is the identifier of the user (e.g., 1 is the portal administrator).

The following links are present among:

- the content and the vocabulary terms,
- a content and the aggregated content (e.g., collection, playlist) containing it,
- the content and the groups that are used to provide the content (each ECLAP content provider has a group),
- a content and the annotations that describe it,
- the users and content, groups and annotations,
- the content and the GeoNames vocabulary for the places where performances were held, they are provided as a result of an enrichment made on the metadata.

In Figure 4-6, an example of how a content is related with vocabulary/taxonomy terms, collections and annotations is reported. For the description of the entities a specific ontology has been designed, this ontology is available as a linked data. All URIs used for properties and classes are dereferenceable and point to the ontology description (e.g., <http://www.eclap.eu/schema/eclap/performancePlace>) both as RDF and human readable documentation in HTML.

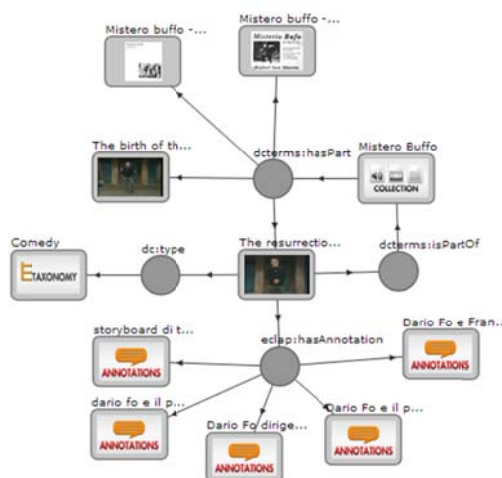


Figure 4-6: Example relation among a content with collections, taxonomy terms and annotations

4.4.3 Content description

Each content is described using RDF, the fields that are already Dublin Core terms in the ECLAP model are provided as they are, while the specific fields for ECLAP are provided by using specific properties (e.g., *eclap:performancePlace*) that are declared refinements of more generic properties taken from standard schemas (e.g., *dcterms:spatial*).

The relations with the vocabulary are provided by using specific properties (e.g., *eclap:genre* for the terms of the genre hierarchy) linking the LOD URIs to the terms. Also these properties are declared as sub properties of Dublin Core terms.

The relations with other aggregated content like collections are provided using *dcterms:isPartOf* and *dcterms:hasPart* properties. Relations with the group of the content provider that is giving the content are offered by specific properties, *eclap:isProvidedBy* and *eclap:provides* (both sub properties of *dc:relation*). These relations allows to link all the content, in particular it can be useful for crawlers allowing them to harvest all the content of a provider. The following is an example:

```
<rdf:RDF ...>
  <rdf:Description rdf:about="http://www.eclap.eu/resource/object/urn:axmedis:00000:obj:04..">
    <dc:title xml:lang="en">you PARA | DISO</dc:title>
    <dc:description xml:lang="en">In July 2010 Emio Greco and Pieter C. Scholten presented their performance "you
  PARA | DISO" at Salle Garnier de l'Opéra de Monte-Carlo. You PARA | DISO is the last performance around Dante's
  Divina Commedia. ... </dc:description>
    <dc:publisher xml:lang="en">iTheatre</dc:publisher>
    <dc:subject xml:lang="en">dance</dc:subject>
    <dc:format>video</dc:format>
    <dc:rights xml:lang="en">erik lint</dc:rights>
    <dc:rights xml:lang="en">emo greco & pc</dc:rights>
    <dc:creator xml:lang="en">emio greco & pc</dc:creator>
    <dc:creator xml:lang="en">erik lint</dc:creator>
    <eclap:performancePlace>Salle Garnier de l'Opéra de Monte-Carlo</eclap:performancePlace>
    <eclap:performanceCity>Monte-Carlo</eclap:performanceCity>
    <eclap:performanceCountry>Monaco</eclap:performanceCountry>
    <eclap:performanceCountry rdf:resource="http://sws.geonames.org/2993457/" />
    <eclap:performanceDate>July 2010</eclap:performanceDate>
    <eclap:choreographer>Emio Greco</eclap:choreographer>
    <eclap:choreographer>Pieter C. Scholten</eclap:choreographer>
    <eclap:historicalPeriod
      rdf:resource="http://www.eclap.eu/resource/term/567" />
    <eclap:performingArtsType
      rdf:resource="http://www.eclap.eu/resource/term/501" />
    <eclap:managementAndOrganization
      rdf:resource="http://www.eclap.eu/resource/term/514" />
    <dcterms:isPartOf
      rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:00000..." />
    <eclap:hasAnnotation
      rdf:resource="http://www.eclap.eu/resource/annotation/SideAnnotation_130..." />
    ...
  </rdf:Description>
</rdf:RDF>
```

4.4.4 Taxonomy description

ECLAP provides six thesauri of terms for the classification of content (for a total of 231 terms):

- Subject (e.g., Teaching, Philosophy, Multiculture)
- Genre (e.g., Comedy, Comic, Drama)
- Historical period (e.g., Contemporary, Classical, XX Century)
- Movement and style (e.g., Experimental, Theatre of the absurd)
- Performing arts type (e.g., Dance, Ballet, Music, Rock, Theatre, Noh)
- Management and organization (e.g., Performance, Choreography)

Each term in the thesauri is described using SKOS [33], the relations among the concepts are provided using the *broader/narrower* properties, and each term is described with multilingual labels in 13 different languages. Moreover, each term is linked with all the content items that use that term using a specific *isSubjectOf* property. The following is an example:

```
<rdf:RDF>
  <skos:Concept rdf:about="http://www.eclap.eu/resource/term/501">
    <skos:prefLabel xml:lang="it">Danza</skos:prefLabel>
    <skos:prefLabel xml:lang="en">Dance</skos:prefLabel>
    ...
    <skos:broader rdf:resource="http://www.eclap.eu/resource/term/664"/>
    <skos:narrower rdf:resource="http://www.eclap.eu/resource/term/540"/>
    <skos:narrower rdf:resource="http://www.eclap.eu/resource/term/539"/>
    <skos:narrower rdf:resource="http://www.eclap.eu/resource/term/507"/>
    <skos:narrower rdf:resource="http://www.eclap.eu/resource/term/506"/>
    <eclap:isSubjectOf
      rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:000... "/>
    <eclap:isSubjectOf
      rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:000... "/>
    ...
  </skos:Concept>
</rdf:RDF>
```

4.4.5 Annotations description

The annotations are used to describe the whole content or its fragment associating it with it a textual description and with another content or fragment. The annotations can be also associated with a classification term (e.g. scene, gesture, character).

The annotations are described using the Open Annotation Collaboration ontology (<http://www.openannotation.org/>) that is currently in working draft, the *hasTarget* property refers to the object being annotated, the *FragmentSelector* class is used to specify the temporal fragment of the annotated resource that is subject to the annotation and the *hasBody* property refers to the annotation body that can be the reference to another content or a text description. The *annotatedBy* property is used to relate the annotation to the user that created it and the *annotatedAt* indicates when the annotation was created. The following is an example:

```
<rdf:RDF xmlns:oa="http://www.w3.org/ns/oa#"
  xmlns:cnt="http://www.w3.org/2011/content#">
  <oa:Annotation
    rdf:about="http://www.eclap.eu/resource/annotation/SideAnnotation_13010...">
    <oa:hasTarget>
      <oa:SpecificResource>
        <oa:hasSource
          rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:0.."/>
        <oa:hasSelector>
          <oa:FragmentSelector>
            <rdf:value>t=npt:10,60</rdf:value>
            <dcterms:conformsTo rdf:resource="http://www.w3.org/TR/media-frags/" />
          </oa:FragmentSelector>
        </oa:hasSelector>
      </oa:SpecificResource>
    </oa:hasTarget>
    <oa:hasBody
      rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:..."/>
    <oa:hasBody>
      <cnt:ContentAsText>
        <cnt:chars>this is an annotation</cnt:chars>
        <dc:format>text/plain</dc:format>
      </cnt:ContentAsText>
    </oa:hasBody>
    <oa:annotatedBy rdf:resource="http://www.eclap.eu/resource/user/1"/>
    <oa:annotatedAt>2013-02-28T20:00:00</oa:annotatedAt>
  </oa:Annotation>
</rdf:RDF>
```

4.4.6 User description

Considering the privacy implication of publishing personal information about the user, minimal personal user information is provided, namely only the nickname is provided. However other relations are available such as:

- the ‘knows’ relation that connect with ‘friends’ users,
- the featured content,
- the favorite content,
- the uploaded content,
- the annotations created,
- the subscribed groups,
- the taxonomy terms of interest to the user.

The following is an example of the description of an user:

```
<rdf:RDF ...>
  <foaf:Person rdf:about="http://www.eclap.eu/resource/user/45">
    <foaf:nick>bellini</foaf:nick>
    <foaf:knows rdf:resource="http://www.eclap.eu/resource/user/1" />
    <foaf:topic_interest rdf:resource="http://www.eclap.eu/resource/term/501" />
    <eclap:isMemberOf rdf:resource="http://www.eclap.eu/resource/group/3160" />
    <eclap:isAdminOf rdf:resource="http://www.eclap.eu/resource/group/3160" />
    <eclap:createdAnnotation
      rdf:resource="http://www.eclap.eu/resource/annotation/SideAnnotation..." />
    <eclap:hasFavorite
      rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:000..." />
    <eclap:hasFeatured
      rdf:resource="http://www.eclap.eu/resource/object/urn:axmedis:000..." />
  </foaf:Person>
</rdf:RDF>
```

The property *isMemberOf* is the inverse of the *foaf:member* property and the *createdAnnotation* property is the inverse of *oa:annotatedBy*. The *hasFavourite* property is defined as a sub property of *foaf:interest*

4.5 Social graph

The ECLAP portal allows to display and to navigate the relations present among the entities managed by the portal <http://www.eclap.eu/116088> . The ‘Social Graph’ of a Media Object is shown when a content is played and the graph related to the user is shown at user login. This graph is a simplification of the information that is available via Linked Data and the terminology used for relations is not always the same used in LOD.

The graph is made of two kind of nodes, rectangular shaped nodes represent entities (content, terms, users, etc.) while circular shaped nodes represent relations. Directed edges connect an entity node to a relation node and a relation node to an entity node. Regarding the user interactions the user is able to:

- **Expand** an entity node with its relations adding them to the graph
- **Focus** on an entity, in this case the graph is cleared and only the focused node is shown with its relations
- **Open** the page associated with the node
- **use the Back button** to go back to previous states of the graph (e.g. after a focus)
- **Zoom/Pan** the view
- **Hide/show** types of relations to reduce the complexity of the graph

A special node is the 'More' node that is presented when in a relation are present many nodes (e.g., the content associated with a group). In this case, providing all nodes could be infeasible thus a limited number of nodes is provided and a 'more' node is added to the relation. Clicking on it other nodes are added to the relation in a way similar to classical pagination used to present long lists in HTML.

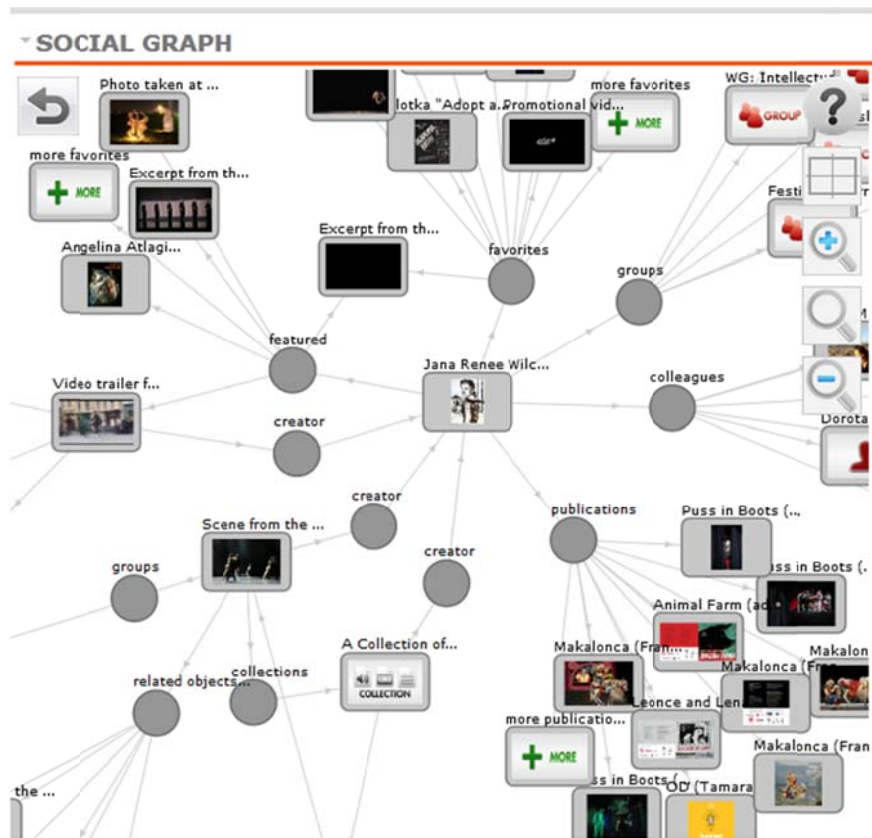


Figure 4-7: The ECLAP Social Graph

In Figure 4-7, an example of ECLAP social graph of a content is shown after expanding some nodes.

The relations visualized by the social graph are:

- *MediaObject* → **creator** → *User* (who has uploaded the object)
- *MediaObject* → **groups** → *Group, ...* (the groups where the object is associated)
- *MediaObject* → **collections** → *Collection, ...* (the collections where the object is present)
- *MediaObject* → **places** → *Place, ...* (the geonames places referred by the object)
- *MediaObject* → **taxonomies** → *TaxonomyTerm, ...* (the taxonomy terms associated with the object)
- *MediaObject* → **annotations** → *Annotation, ...* (the annotations associated with the object)
- *MediaObject* → **comments** → *Comment, ...* (the comments associated with the object)
- *MediaObject* → **related objects** → *MediaObject, ...* (the objects that are ‘similar’ to an object)

- *User* → **featured** → *MediaObject, ...* (the objects that were featured by the user)
- *User* → **publications** → *MediaObject, ...* (the objects uploaded by the user)
- *User* → **colleagues** → *User, ...* (the other user that are colleagues of the user)
- *User* → **groups** → *Group, ...* (the groups subscribed by the user)

- *Group* → **members** → *User, ...* (the members of the group)
- *Group* → **administrators** → *User, ...* (the administrators of the group)

- *Place* → **formed** → *Place, ...* (the other geonames that are part of a geoname)
- *Place* → **belong** → *Place* (the geoname contains another geoname)
- *Place* → **objects** → *MediaObject, ...* (the objects associated with the geoname place)

- *TaxonomyTerm* → **broader** → *TaxonomyTerm* (the term that is broader than another term)
- *TaxonomyTerm* → **narrower** → *TaxonomyTerm, ...* (the terms that are narrower than the term)

- *TaxonomyTerm* → **objects** → *MediaObject*, ... (the objects that are associated with a term)
- *Collection* → **objects** → *MediaObject*, ... (the objects that are part of the collection)
- *Comment* → **writer** → *User* (the user that wrote the comment)
- *Annotation* → **writer** → *User* (the user that created the annotation)

The Social Graph is also presented in the Europeana ThoughtLab page on new ways of searching and browsing (<http://pro.europeana.eu/web/guest/thoughtlab/new-ways-of-searching-and-browsing#SocialGraph>).

We have analysed how the users interact with the social graph. We have found that from the content visualizations coming from 16,031 different IP addresses 740 interacted also with the social graph (expand, more, focus, open) that is about 4.6%.

The most requested operation is to open a node (36%), then to explore a node (33%) and then to see the more related content (12%), the focus operation is at about 10% on 3030 operations requested since the social graph was activated (2013-01-29) until the end of June 2013. In these six months there were an average of 19 interactions per day with a maximum of 77 interactions (2013-04-12) and an average of 3.3 interactions per IP address per day with a maximum of 15 interactions per IP per day.

Figure 4-8 reports in more details the distribution of the interactions among the different types of actions.

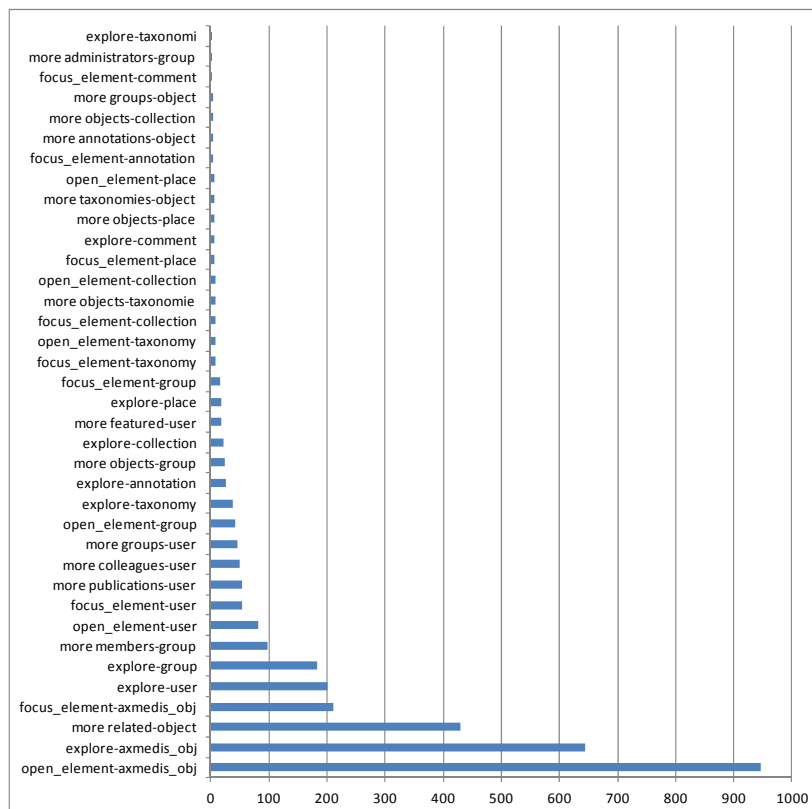


Figure 4-8: distribution of user interactions on social graph

5 Bibliography

- [1] C. D. P. M. W. Group and others, “Dublin Core Metadata Best Practices,” Athabasca University Experts Team, “Canadian Culture Online (CCO) Policy: Metadata Strategy and Development of a Matrix of Metadata Elements” (Athabasca, Alberta, Canada: Athabasca University, 2004). www.canadianheritage.gc.ca/progs/pcce-ccop/reana/sm-ms, 2006.

- [2] L. M. Chan and M. L. Zeng, “Metadata Interoperability and Standardization-A Study of Methodology Part I: Achieving Interoperability at the Schema Level,” *D-Lib magazine*, vol. 12, no. 6, 2006.
- [3] E. Rahm and P. A. Bernstein, “A survey of approaches to automatic schema matching,” *the VLDB Journal*, vol. 10, no. 4, pp. 334–350, 2001.
- [4] B. Haslhofer and W. Klas, “A survey of techniques for achieving metadata interoperability,” *ACM Computing Surveys (CSUR)*, vol. 42, no. 2, p. 7, 2010.
- [5] L. T. Moss and S. Atre, *Business intelligence roadmap: the complete project lifecycle for decision-support applications*. Addison-Wesley Professional, 2003.
- [6] P. Johnston, “Metadata and interoperability in a complex world,” *Ariadne*, vol. 37, 2003.
- [7] N. Press, “Understanding metadata,” *National Information Standards*, p. 20, 2004.
- [8] A. Gilliland-Swetland, M. Bacca, and T. Gill, “Introduction to metadata: Pathways to digital information,” *Getty Information Institute*, 2000.
- [9] B. Lavoie, R. Dale, M. Alexander, K. Bradley, M. May, R. Guenther, B. Hruley, C. Lupovici, O. Rieger, D. Sergeant, and others, “Preservation Metadata and the OAIS Information Model. A Metadata Framework to Support the Preservation of Digital Objects,” *OCLC/RLG Working Group on Preservation Metadata*, 2002.
- [10] S. L. Weibel and T. Koch, “The Dublin core metadata initiative,” *D-lib magazine*, vol. 6, no. 12, pp. 1082–9873, 2000.
- [11] C. J. Godby, J. A. Young, and E. Childress, “A repository of metadata crosswalks,” 2004.
- [12] O. A. Initiative and C. Lagoze, *The Open Archives Initiative Protocol for Metadata Harvesting. The Initiative*, 2002.
- [13] M. Doerr, S. Gradmann, S. Henniecke, A. Isaac, C. Meghini, and H. van de Sompel, “The Europeana Data Model (EDM),” *Proceedings of IFLA*, 2010.
- [14] L. Manovich, “An Introduction to Cultural Analytics”, in *Changing Sceneries Changing Role – Part IV. Keeping your Best Content and Metadata, Selected Papers, from the FIAT Media Management Seminar 15-16 May 2008 Stockholm, Hilversum: Beeld and Geluid*, 2010.
- [15] F. Marotti, “Note di metodo per lo studio del teatro della regia”, *Biblioteca Teatrale* 8: 33-36, 1973.
- [16] M. De Marinis, “Capire il teatro. Lineamenti di una nuova teatrologia” *Rome: Bulzoni*, 1999.
- [17] J. Shurmann, P. Eversmann, N. Mitolo, P. Nesi, L. Baltussen, M. Borelli, I. Scaturro, “DE 6.2.1 ECLAP Specific Services”, 2012, available at <http://www.eclap.eu>
- [18] V. Valentini, “Dopo l’avanguardia: la performance della crisi”, in B. Marranta, *American Performance 1975-2005*, V. Valentina (eds), *Rome: Bulzoni*, 2006.
- [19] P. Bellini, I. Bruno, D. Cenni, A. Fuzier, P. Nesi, M. Paolucci, M. Serena, “MyStoryPlayer, Collections, Playlists: Content aggregator tools in ECLAP Project”, 2011, available at <http://www.eclap.eu/>
- [20] P. Bellini, P. Nesi, S. Marco, “MyStoryPlayer for ECLAP. Non Linear Story Modeling and Navigation Tools based on Semantic Annotations. User Manual”, 2013, available at www.eclap.eu

- [21] Cunningham H.; Maynard D.; Bontcheva K.; Tablan V., "GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications", Proceeding of the 40th Anniversary Meeting of the Association for Computational Linguistics, 2002.
- [22] Bizer C., Heath T., Berners-Lee T.: Linked Data—the story so far. In: *Int. J Semantic Web Inf Syst* 5(3):1–22 (2009)
- [23] Raimond Y., Sandler M.B.: A Web of musical information. In: Bello JP, Chew E, Turnbull D (eds) *ISMIR*, pp 263–268 (2008)
- [24] Kobilarov G., Scott T., Raimond Y., Oliver S., Sizemore C., Smethurst M., Bizer C., Lee R.: Media meets Semantic Web—how the BBC uses DBpedia and Linked Data to make connections. In: *Proceedings of the 6th European Semantic Web conference*. Springer, Berlin, pp 723–737 (2009)
- [25] Bizer C., Lehmann J., Kobilarov G., Auer S., Becker C., Cyganiak R., Hellmann S.: DBpedia — a crystallization point for the Web of Data. *J Web Sem* 7(3):154–165 (2009)
- [26] Kahan J., Koivunen M.R.: Annotea: an open RDF infrastructure for shared Web annotations. In: *WWW '01: proceedings of the 10th international conference on World Wide Web*. ACM Press, New York, pp 623–632 (2001)
- [27] Haslhofer B., Jochum W., King R., Sadilek C., Schellner K.: The LEMO annotation framework: weaving multimedia annotations with the Web. *Int J Digit Libr* 10(1):15–32 (2009)
- [28] W3C: Media fragments URI 1.0. W3C Media Fragments Working Group URL <http://www.w3.org/TR/media-frag/> (2009)
- [29] Haslhofer B., Simon R., Sanderson R., Van de Sompel H.: The Open Annotation Collaboration (OAC) Model. In: *Proc. of the 2011 Workshop on Multimedia on the Web (MMWEB '11)*. IEEE Computer Society, Washington, DC, USA, pp 5-9 (2011)
- [30] Bellini P., Nesi P., Paolucci M., Serena M.: Models and tools for content aggregation and audiovisual cross annotation synchronization. In: *Proc. of 2011 IEEE Int. Symp. on Multimedia*, pp 210-215 (2011)
- [31] Bellini P., Nesi P., Serena M.: Mystoryplayer: Semantic Audio Visual Annotation And Navigation Tool. In: *Proc. of the 17th Int. Conf. on Distributed Multimedia Systems* (2011)
- [32] ECLAP DE4.3 Metadata Descriptors Interoperability, <http://www.eclap.eu/urn:axmedis:00000:obj:6d1e8e22-55d4-4082-9486-f37a3752704a>
- [33] W3C Semantic Web Deployment Group: SKOS simple knowledge organization system reference. URL <http://www.w3.org/TR/2009/REC-skos-reference-20090818/> (2009)
- [34] P. Bellini, I. Bruno, D. Cenni, P. Nesi, M. Paolucci, M. Serena, "A New Generation Digital Content Service for Cultural Heritage Institutions", *Proc. of the ECLAP 2013 conference, 2nd International Conference on Information Technologies for Performing Arts, Media Access and Entertainment*, Springer Verlag LNCS, 2013. <http://www.eclap.eu/drupal/?q=it/node/119325>
- [35] E. Bellini, P. Nesi, "Metadata Quality assessment tool for Open Access Cultural Heritage institutional repositories", *Proc. of the ECLAP 2013 conference, 2nd International Conference on Information Technologies for Performing Arts, Media Access and Entertainment*, Springer Verlag LNCS, 2013. <http://www.eclap.eu/drupal/?q=it/node/119325>
- [36] Pierfrancesco Bellini, P. Nesi, "A Linked Open Data service for performing arts", *Proc. of the ECLAP 2013 conference, 2nd International Conference on Information Technologies for Performing Arts, Media Access and Entertainment*, Springer Verlag LNCS, 2013. <http://www.eclap.eu/drupal/?q=it/node/119325>

- [37] P. Bellini, I. Bruno, P. Nesi, "A workflow model and architecture for content and metadata management based on grid computing", Proc. of the ECLAP 2013 conference, 2nd International Conference on Information Technologies for Performing Arts, Media Access and Entertainment, Springer Verlag LNCS, 2013. <http://www.eclap.eu/drupal/?q=it/node/119325>
- [38] P. Bellini, P. Nesi, M. Paolucci, "IPR Management Models for Cultural Heritage on ECLAP Best Practice Network", Proc. of the IEEE International Conference on Comunicazion, IEEE ICC 2013, Workshop "Beyond Social Networks: Collective Awareness", 9-13 June 2013, Budapest, Hungary, IEEE Press.
- [39] P. Bellini, I. Bruno, P. Nesi, M. Paolucci, "A Static and Dynamic Recommendations System for Best Practice Networks", Proc. of the Human-Computer Interaction Conference, 21-26- July 2013, Nevada, USA, Press by Springer Verlag LNCS.
- [40] P. Bellini, D. Cenni, P. Nesi, "On the Effectiveness and Optimization of Information Retrieval for Cross Media Content", Proceeding of the KDIR 2012 is part of IC3K 2012, International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 4-7 October 2012, Barcelona, Spain.
- [41] P. Bellini, D. Cenni, P. Nesi, "Indexing and Searching Cross Media Content in a Social Network", Proceeding of the ECLAP 2012 International Conference on Information Technologies for Performing Arts, Media Access and Entertainment", Florence University Press, May 2012, ISBN: 978-88-6655-125-6
- [42] P. Bellini, I. Bruno, D. Cenni, P. Nesi, "Micro grids for scalable media computing and intelligence on distributed scenarios", IEEE Multimedia, IEEE Computer Soc. 2012.
- [43] I. Bruno, M. Paolucci, P. Bellini, N. Mitolo, "DE3.3.2 infrastructure Content and Metadata Processing and Semantification" <http://www.eclap.eu/urn:axmedis:00000:obj:a722b357-4644-4076-a259-0cbc6260ad74>

6 Glossary

AMICO	Art Museum Image Consortium
CDWA	Categories for the Description of Works of Art
CRM	Conceptual Reference Model
DC	Dublin Core
EAD	Encoded Archival Description
EDM	Europeana Data Model
ESE	Europeana Semantic Elements Specification
FRBR	Functional Requirements for Bibliographic Records
IMS	Instructional Management Systems
IPTC	International Press Telecommunications Council
LIDO	Lightweight Information Describing Objects
MARC21	Machine-Readable Cataloguing
METS	Metadata Encoding and Transmission Standard
MODS	Metadata Object Description Schema
MPEG	Moving Pictures Expert Group
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OWL	Ontology Web Language
OWL DL	Ontology Web Language Description Logics
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SKOS	Simple Knowledge Organisation System
SPECTRUM	Standard ProcEdures for CollecTions Recording Used in Museums
VRA	Visual Resources Association
XML	Extensible Markup Language

7 Appendix – Complete Mappings to the ECLAP schema

Table 7-1: The CTFR-mapping-export summary

Source	Target
/Export/Record/Subject	/DC/subject
it	/DC/subject/@lang
it	/DC/description/@lang
it	/DC/title/@lang
it	/DC/type/@lang
/Export/Record/ProviderContentID	/TechnicalMetadata/ProviderContentID
/Export/Record/ProductionRecord	/PerformingArtsMetadata/ProductionRecord
/Export/Record/Coverage	/DC/coverage
/Export/Record/TableOfContents	/DCTERMS/tableOfContents
/Export/Record/PerformanceCountry	/PerformingArtsMetadata/Performance/Country
/Export/Record/Cast	/PerformingArtsMetadata/Cast
it	/PerformingArtsMetadata/@lang
/Export/Record/Provenance	/DCTERMS/provenance
/Export/Record/PersonRecord	/PerformingArtsMetadata/PersonRecord
/Export/Record/RecordingDate	/PerformingArtsMetadata/RecordingDate
/Export/Record/Extent	/DCTERMS/extent
/Export/Record/HistoricalPeriod	/PerformingArtsMetadata/HistoricalPeriod
/Export/Record/Creator	/DC/creator
/Export/Record/PerformingArtType	/PerformingArtsMetadata/PerformingArtType
/Export/Record/PerformingArtsGroup	/PerformingArtsMetadata/PerformingArtsGroup
BASIC_CONTENT	/TechnicalMetadata/Type
/Export/Record/Created	/DCTERMS/created
/Export/Record/Medium	/DCTERMS/medium
/Export/Record/PieceRecord	/PerformingArtsMetadata/PieceRecord
/Export/Record/Rights	/DC/rights
/Export/Record/Language	/DC/language
ctfr-ipr-model-1	/TechnicalMetadata/IPRModelID
/Export/Record/Title	/DC/title
/Export/Record/PerformanceCity	/PerformingArtsMetadata/Performance/City
/Export/Record/AggregationID	/TechnicalMetadata/AggregationID
/Export/Record/Relation	/DC/relation
/Export/Record/HasVersion	/DCTERMS/hasVersion
/Export/Record/Date	/DC/date
/Export/Record/Source	/DC/source
/Export/Record/Issued	/DCTERMS/issued
/Export/Record/FirstPerformanceDate	/PerformingArtsMetadata/FirstPerformance/Date
/Export/Record/PerformancePlace	/PerformingArtsMetadata/Performance/Place
/Export/Record/Type	/DC/type
/Export/Record/Title	/TechnicalMetadata/AggregationName
/Export/Record/Genre	/PerformingArtsMetadata/Genre
/Export/Record/Filename	/TechnicalMetadata/ProviderContentUrl
/Export/Record/ArtisticMovementAndActingStyle	/PerformingArtsMetadata/ArtisticMovementAndActingStyle
/Export/Record/Professional	/PerformingArtsMetadata/PerformersAndCrew
/Export/Record/Format	/DC/format

/Export/Record/Contributor	/DC/contributor
/Export/Record/Publisher	/DC/publisher
/Export/Record/FirstPerformancePlace	/PerformingArtsMetadata/FirstPerformance/Place
it	/DC/rights/@lang
/Export/Record/IsFormatOf	/DCTERMS/isFormatOf
/Export/Record/IsPartOf	/DCTERMS/isPartOf
/Export/Record/IsVersionOf	/DCTERMS/isVersionOf
CTFR	/TechnicalMetadata/ProviderID
/Export/Record/FirstPerformanceCountry	/PerformingArtsMetadata/FirstPerformance/Country
/Export/Record/PlotSummary	/PerformingArtsMetadata/PlotSummary
/Export/Record/Description	/DC/description
/Export/Record/PerformanceDate	/PerformingArtsMetadata/Performance/Date
it	/DCTERMS/tableOfContents/@lang
/Export/Record/FirstPerformanceCity	/PerformingArtsMetadata/FirstPerformance/City

Table 7-2: The MUZEUM_new_v06a_mapping summary

Source	Target
/metadata/record/ManagementAndOrganization	/PerformingArtsMetadata/ManagementAndOrganization
/metadata/record/dc.Subject	/DC/subject
en	/DC/subject/@lang
en	/DC/description/@lang
en	/DC/title/@lang
en	/DC/type/@lang
/metadata/record/ProviderContentID	/TechnicalMetadata/ProviderContentID
/metadata/record/Cast	/PerformingArtsMetadata/Cast
/metadata/record/PerformanceCountry	/PerformingArtsMetadata/Performance/Country
en	/PerformingArtsMetadata/@lang
/metadata/record/dc.Alternative	/DCTERMS/alternative
/metadata/record/HistoricalPeriod	/PerformingArtsMetadata/HistoricalPeriod
en	/DCTERMS/isPartOf/@lang
/metadata/record/dc.Creator	/DC/creator
/metadata/record/PerformingArtType	/PerformingArtsMetadata/PerformingArtType
BASIC_CONTENT	/TechnicalMetadata/Type
/metadata/record/dc.References	/DCTERMS/references
/metadata/record/dc.Rights	/DC/rights
/metadata/record/dc.Language	/DC/language
/metadata/record/IPRModel	/TechnicalMetadata/IPRModelID
/metadata/record/dc.Title	/DC/title
/metadata/record/PerformanceCity	/PerformingArtsMetadata/Performance/City
/metadata/record/ProfessionalActor	/PerformingArtsMetadata/Professional
en	/DCTERMS/references/@lang
/metadata/record/FirstPerformanceDate	/PerformingArtsMetadata/FirstPerformance/Date
/metadata/record/PerformancePlace	/PerformingArtsMetadata/Performance/Place
/metadata/record/dc.Type	/DC/type
/metadata/record/ContentLocation	/TechnicalMetadata/ProviderContentUrl
/metadata/record/ArtisticMovementAndActingStyle	/PerformingArtsMetadata/ArtisticMovementAndActingStyle
/metadata/record/dc.Contributor	/DC/contributor
en	/DC/publisher/@lang
/metadata/record/dc.Publisher	/DC/publisher

/metadata/record/FirstPerformancePlace en	/PerformingArtsMetadata/FirstPerformance/Place /DC/rights/@lang
/metadata/record/dc.IsPartOf slo	/DCTERMS/isPartOf /DCTERMS/alternative/@lang
MUZEUM	/TechnicalMetadata/ProviderID
/metadata/record/FirstPerformanceCountry	/PerformingArtsMetadata/FirstPerformance/Country
/metadata/record/PlotSummary	/PerformingArtsMetadata/PlotSummary
/metadata/record/dc.Description	/DC/description
/metadata/record/PerformanceDate en	/PerformingArtsMetadata/Performance/Date /DC/contributor/@lang
en	/DC/creator/@lang
/metadata/record/FirstPerformanceCity Actor	/PerformingArtsMetadata/FirstPerformance/City /PerformingArtsMetadata/Professional/@role

Table 7-3: Transformation of metadata records on Table 2-1 with respect to the ECLAP schema

```
<?xml version="1.0" encoding="UTF-8"?>
<eclap:ECLAPObjectWrap xmlns:eclap="http://www.eclap.eu/ECLAPSchemaV0"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xalan="http://xml.apache.org/xalan">
  <eclap:ECLAPObject>
    <eclap:DC>
      <dc:coverage>Milano, Coordinates: Lat: 45 28 00 N degrees minutes Lat: 45.4667 decimal degrees Long:
009 12 00 E degrees minutes Long: 9.2000 decimal degrees </dc:coverage>
      <dc:creator>Franca Rame, Dario Fo</dc:creator>
      <dc:description xml:lang="it">Dattiloscritto dell'atto unico "Grasso è bello!" di Franca Rame, Dario Fo e
alcune idee di Jacopo Fo. Revisione di Franca Rame aggiornata al 2008.</dc:description>
      <dc:format>pdf</dc:format>
      <dc:language>it</dc:language>
      <dc:rights xml:lang="it">CTFR</dc:rights>
      <dc:subject xml:lang="it">testo</dc:subject>
      <dc:title xml:lang="it">Grasso è bello!</dc:title>
    </eclap:DC>
    <eclap:DCTERMS>
      <dcterms:extent>55 pag</dcterms:extent>
    </eclap:DCTERMS>
    <eclap:TechnicalMetadata>
      <eclap:Type>BASIC_CONTENT</eclap:Type>
      <eclap:ProviderID>CTFR</eclap:ProviderID>
      <eclap:ProviderContentID>55500</eclap:ProviderContentID>
      <eclap:ProviderContentUrl>content/archivioforame/PARL/2008/TEST/55500/55500.pdf</eclap:ProviderCo
ntentUrl>
      <eclap:AggregationID>PARL</eclap:AggregationID>
      <eclap:AggregationName>Grasso è bello!</eclap:AggregationName>
      <eclap:IPRModelID>ctfr-ipr-model-1</eclap:IPRModelID>
    </eclap:TechnicalMetadata>
    <eclap:PerformingArtsMetadata xml:lang="it">
      <eclap:FirstPerformance/>
      <eclap:Performance/>
    </eclap:PerformingArtsMetadata>
  </eclap:ECLAPObject>
</eclap:ECLAPObjectWrap>
```

Table 7-4: Transformation of metadata records on Table 2-2 with respect to the ECLAP schema

```
<?xml version="1.0" encoding="UTF-8"?>
<eclap:ECLAPObjectWrap xmlns:eclap="http://www.eclap.eu/ECLAPSchemaV0"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xalan="http://xml.apache.org/xalan">
<eclap:ECLAPObject>
<eclap:DC>
<dc:creator xml:lang="en">Leo Kulaš archive</dc:creator>
<dc:description xml:lang="en">Costume photo from the performance L'impresario di Smirne, written by
Carlo Goldoni, directed by Paolo Magelli</dc:description>
<dc:language>en</dc:language>
<dc:publisher xml:lang="en">Produced by: HNK Zagreb</dc:publisher>
<dc:rights xml:lang="en">For information about the rights holder please contact Muzeum, Institute of Art
Production, Mediation and Publishing Ljubljana.</dc:rights>
<dc:subject xml:lang="en">theatre performance, costume design</dc:subject>
<dc:title xml:lang="slo">Kostumi, Leo Kulaš (št.19), L'impresario di Smirne</dc:title>
<dc:title xml:lang="en">Costumes, Leo Kulaš (No. 19), L'impresario di Smirne</dc:title>
<dc:type xml:lang="en">Image</dc:type>
</eclap:DC>
<eclap:DCTERMS>
<dcterms:isPartOf xml:lang="en">Leo Kulaš</dcterms:isPartOf>
</eclap:DCTERMS>
<eclap:TechnicalMetadata>
<eclap:Type>BASIC_CONTENT</eclap:Type>
<eclap:ProviderID>MUZEUM</eclap:ProviderID>
<eclap:ProviderContentID>MLJ2346_LKulas_Impresaria</eclap:ProviderContentID>
<eclap:ProviderContentUrl>In_process/MLJ2346_LKulas_Impresaria.jpg</eclap:ProviderContentUrl>
<eclap:IPRModelID>MUZEUM_all_CC-BY-NC-ND </eclap:IPRModelID>
</eclap:TechnicalMetadata>
<eclap:PerformingArtsMetadata xml:lang="en">
<eclap:FirstPerformance>
<eclap:Country>Croatia</eclap:Country>
<eclap>Date>1993</eclap>Date>
</eclap:FirstPerformance>
<eclap:Performance/>
<eclap:Professional role="Costume_designer">Leo Kulaš</eclap:Professional>
<eclap:Professional role="Director">Paolo Magelli</eclap:Professional>
<eclap:Professional role="Playwright">Carlo Goldoni</eclap:Professional>
<eclap:PerformingArtType>Theatre</eclap:PerformingArtType>
<eclap:HistoricalPeriod>XX Century</eclap:HistoricalPeriod>
</eclap:PerformingArtsMetadata>
</eclap:ECLAPObject>
</eclap:ECLAPObjectWrap>
```

8 Appendix – Complete ECLAP Schema

8.1 General information

axoid

unique identifier identifying the content on the portal, the id is based on UUID identification

url

The url on the portal where the content is available it is of the form:
http://www.eclap.eu/drupal?q=home&axoid=<axoid>

nid

drupal node id identifier, id associated by drupal to the content.

Version

Version number of the content, it is incremented when an update is performed

InsertUpdateTime

Date and time when the content was uploaded or when it was updated. The date is in the format YYYY-MM-DDThh:mm:ss.

ProviderId

ECLAP provider acronym used to identify the provider

ProviderName

Complete name of the provider

DefaultMetadataLanguage

default language used for the description of the resource, it should be a 2 letter ISO language code.

8.2 Digital resource information

Format

The resource format it can be “audio, video, document, image, crossmedia, 3d, archive, tool, playlist, collection”

Type

Specifies better the resource format for crossmedia (html, flash) and document (document, epub, pdf, excel, slide, braille music)

Width

Width of the frame for image or video

Height

Height of the frame for image or video

Duration

Duration of audio or video resource in the form “hh:mm:ss.mm”

AvlMDVideo

Is “yes” if a medium definition video resource is available

AvlHDVideo

Is “yes” if a high definition video resource is available

Extension

The file extension for the digital resource.

Preview

Url used for the content preview, for video it is an animated gif.

AvlForPDA

Is “yes” if a version for WindowsMobile 6.5 has been produced

AvlForIPhone

Is “yes” if the content can be used on iPhone or in general on Smart Phones (iOS, Android, WindowsPhone7)

AvlForPC

Is “yes” if the content can be used on a Personal Computer (Windows/MAC/Linux)

8.3 IPR information

IsPublic

Is “yes” if the IPR model associated with the content is public

IPRTitle

The title given to the IPR model associated with the content

IPRDescription

The description given to the IPR model associated with the content

EuropeanaRightsUrl

The Europeana Url given to the IPR model associated with the content

LicenseUrl

The license url given to the IPR model associated with the content

8.4 GeoSpatial Information

This section can provide a set of GPS coordinates

Latitude

The decimal representation of the latitude

Longitude

The decimal representation of the longitude

Radius

The radius in meters of the area where the content is “active”

8.5 Performing arts metadata

In this section are reported the metadata specific for performing arts.

FirstPerformance Place

Name of the theatre or venue where the performance taken place for the first time.

Examples	“Théâtre des Bouffes du Nord”
Count	0..1

(1, 0..1, 0..many, 1..many)	
Notes	the first performance is the première, therefore its “place”, might not correspond with the place in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held at: “Théâtre des Bouffes du Nord”, but what we are looking at on the ECLAP portal might be a video of the performance held months later - while the show was touring – at “The Globe Theatre”
Refinement of	dcterms:spatial

FirstPerformance City

Name of the city where the first performance taken place.

Examples	“Paris”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “City”, might not correspond with the city in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held in: “Paris”, but what we are looking at on the ECLAP portal might be a video of the performance held months later - while the show was touring – in “London”
Refinement of	dcterms:spatial

FirstPerformance Country

Name of the country where the first performance taken place

Examples	“France”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “Country”, might not correspond with the country in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held in: “France”, but what we are looking at on the ECLAP portal might be a video of the performance held months later - while the show was touring – in “England”.
Refinement of	dcterms:spatial

FirstPerformance Date

Date of the first performance

Examples	“2000-11-20”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “date”, might not correspond with the date in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held in: “2000-11-20”, but what we are looking at on the ECLAP portal might be a video of the performance held months later, in “2001-04-05”
Refinement of	DCTerms.issued

Performance Place

Name of the theatre or venue where the shown performance taken place

Examples	“The Globe Theatre”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “place”, might not correspond with the place in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held at: “Théâtre des Bouffes du Nord”, but what we are looking at on the ECLAP portal might be a video of the performance held months later - while the show was touring – at “The Globe Theatre”
Refinement of	dcterms:spatial

Performance City

Name of the city where the shown performance taken place.

Examples	“London”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “City”, might not correspond with the city in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held in: “Paris”, but what we are looking at on the ECLAP portal might be a video of the performance held months later - while the show was touring – in “London”
Refinement of	dcterms:spatial

Performance Country

Name of the country where the shown performance taken place

Examples	“England”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “Country”, might not correspond with the country in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held in: “France”, but what we are looking at on the ECLAP portal might be a video of the performance held months later - while the show was touring – in “England”.
Refinement of	dcterms:spatial

Performance Date

Date of the shown performance

Examples	“2001-04-05”
Count (1, 0..1, 0..many, 1..many)	0..1
Notes	the first performance is the première, therefore its “date”, might not correspond with the date in which the show was recorded. For example: the opening night of “The Tragedy of Hamlet” directed by P. Brook might be held in:

	“2000-11-20”, but what we are looking at on the ECLAP portal might be a video of the performance held months later, in “2001-04-05”
Refinement of	dc:terms:issued

PerformingArtsGroup

Name of the theatre or dance company or musical group (if present)

Examples	“Momix”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	None
Refinement of	dc:creator

PlotSummary

Summary of the plot

Examples	“Prince Hamlet mourns both his father's death and his mother, Queen Gertrude's remarriage to Claudius. The ghost of Hamlet's father appears to him and tells him that Claudius has poisoned him: Hamlet swears revenge, etc.”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	None
Refinement of	dc:description

Cast

Name/Names of a member of the cast.

Examples	“Ryszard Cieślak, Rena Mirecka, Antoni Jahołkowski, Mieczysław Janowski, Maja Komorowska, Stanislaw Scierski”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	Use this element only if the Professional elements cannot be used, as the case of a cast written in a single text that cannot be easily split in all the different professional people
Refinement of	dc:contributor

PerformersAndCrew

Name/Names of a performers and crew of a performance.

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	Use this element only if the Professional elements cannot be used, as the case of a performers written in a single text that cannot be easily split in all the different professional people
Refinement of	dc:contributor

Professional

A list of the people involved in the performance indicating which role each person had in the performance (eg. Actor, director, set designer etc.). It includes all the information listed in a playbill, such as the artistic cast of the show and the technicians, but also the names of the troupe which recorded the performance (eg. Cameraman, Director of Photography, etc.). Possible roles are:

- Acrobat
- Actor
- Adaptator
- Architect
- Assistant_director
- Casting
- Choreographer
- Clown
- Composer
- Concept_originator
- Costume_designer
- Critic
- Dancer
- Director
- Dramaturge
- Hairdresser
- Light_designer
- Make-up_artist
- Marketing_manager
- Mask_designer
- Mime
- Musician
- Patron
- Performer
- Playwright
- Producer
- Puppet_designer
- Scenographer
- Seamster
- Set_builder
- Set_designer
- Singer
- Sound_designer
- Stage_manager
- Technician
- Theatre_manager
- Theoretician
- Translator
- Other

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	None
Refinement of	dc:contributor

Object

Objects used in the performance, (i.e. Sets, Costumes, Props, Programs, Prints, Drawings,...)

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	
Refinement of	DC.description

Genre

The genre in which the work can be categorized (i.e. Ballet, Butho, Commedia dell'Arte, Drama, Feast Flamenco, etc)

Examples	“Tragedy”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	we will work on a shared vocabulary for this
Refinement of	dc:subject

PerformingArtType

Type of performing art present in the content.

Examples	“theatre”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	identified in WP4 as cinema, dance, music, theatre, performance art
Refinement of	dc:type

HistoricalPeriod

Historical period the topic of the resource refers to.

Examples	“XV century”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	None
Refinement of	dcterms:temporal

ArtisticMovementAndActingStyle

Artistic movement and acting styles in which the work can be categorized (e.g. Classicism, Dada, Epic, Expressionism, etc.)

Examples	“Futurism”
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	we will work on a shared vocabulary for this
Refinement of	dcterms:type

ManagementAndOrganization

management and organization...

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	Deprecated

RecordingDate

Date of creation of the digital object,

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	Use this element in case what it is recorded is not a public performance (e.g. an interview) otherwise use the Performance Date
Refinement of	dc:date

PersonRecord

Credits for the audio or video recording

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	Deprecated, use Professional elements
Refinement of	dc:contributor

PieceRecord

Credits for the text or image. The meaning of this field is a bit complex.... The text we are dealing with in this field is the script of the play. We intend this field to be filled out with the original title of the performance (eg. Medea) - which might differ from the title of the item (eg. Photo of Medea_2) - and with the name of the person who wrote the script. The records pertaining to the novel or the literary work which inspired the script should be mapped in the field "reference" instead; the field "reference" should also include the title of the novel and its author(s).

Examples	Title: Il Principe Costante; scenario: Jerzy Grotoski; adaptation: Julius Slowacki
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	None
Refinement of	dcterms:references

ProductionRecord

Credits of the production team. the name of the producer(s) and of other people involved in the organization.

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Notes	Deprectaed, use the Professional element with the appropriate role
Refinement of	dc:contributor

8.6 DC – dublin core metadata

This section contains information about the dublin core metadata to be associated with the content to be ingested:

title

The name given to the resource. Typically, a Title will be a name by which the resource is formally known. The title of the original analog or born digital object. The title should be significant.

Examples	"Romeo and Juliet"
Count (1, 0..1, 0..many, 1..many)	1..many
Language	Mandatory
Notes	None

creator

An entity primarily responsible for making the content of the resource. Examples of a Creator include a person, an organization, or a service. Typically the name of the Creator should be used to indicate the entity. In ECLAP, the name of Partner uploading is kept automatically in a separate field. This is the name of the creator of the original analog or born digital object. . This field should be used only to indicate the creator of the work of art (usually the director for a performance, the author if we are dealing with a book, the composer if we are uploading a script and so on). Often, in devised work, the creator might be the whole company or the actors might collaborate with the director. Nevertheless I guess we need to set a rule to be applied to every situation, so that I would consider actors and other artistic figures as contributors and eventually explain in the field "description" if their role as creator of the performance was capital.

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Language	Optional
Notes	None

subject

The topic of the content of the resource. Typically, a Subject will be expressed as keywords or key phrases or classification codes that describe the topic of the resource. Recommended best practice is to select a value from your own classification scheme. This is the subject of the original analog or born digital object.

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Language	Mandatory
Notes	None

description

An account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content. A description of the original analog or born digital object.

Examples	...
Count (1, 0..1, 0..many, 1..many)	0..many
Language	Mandatory
Notes	None

publisher

The entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity. In ECLAP, the name of Partner that has provided the content is automatically tracked and stored in a different field. The name of the publisher of the original analog or born digital object.

Examples	In case of a performance review the name of the newspaper where the review was published
Count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
Notes	None

contributor

An entity responsible for making contributions to the content of the resource. Examples of a Contributor include a person, an organization or a service. Typically, the name of a Contributor should be used to indicate the entity. In most cases, the authors of a document are listed here. The name of contributors to the original analog or born digital object. This could be a person, an organisation or a service.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

date

A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [Date and Time Formats, W3C Note, <http://www.w3.org/TR/NOTE-datetime>] and follows the YYYY-MM-DD format. If the full date is unknown, month and year (YYYY-MM) or just year (YYYY) may be used. Many other schemes are possible, but if used, they may not be easily interpreted by users or software. Use for a significant date in the life of the original analog or born digital object. Use `dcterms:temporal` (or `dc:coverage`) if the date is associated with the topic of the resource.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

type

The nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMIType vocabulary <http://dublincore.org/documents/dcmi-type-vocabulary/>). To describe the physical or digital manifestation of the resource, use the FORMAT element. The type of the original analog or born digital object as recorded by the content holder, this element typically includes values such as photograph, painting, sculpture etc.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Mandatory
notes	None

format

The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Examples of dimensions include size and duration. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [<http://www.iana.org/assignments/media-types/>] defining computer media formats). The unqualified element includes file format, physical medium or dimensions of the original and/or digital object. Use this element for the file format of the digital object or born digital originals. Internet Media Types [MIME] are highly recommended (<http://www.iana.org/assignments/media-types/>). Use of the more specific elements `dcterms:extent` (dimensions) and `dcterms:medium` (physical medium) is preferred where appropriate.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

identifier

An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Examples of formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN). This is the identifier for the original analog or born digital object.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional

notes	None
--------------	------

source

A Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system. In general, include in this area information about a resource that is related intellectually to the described resource but does not fit easily into a Relation element. In ECLAP, this value should be the URL or the filename of the original resource. The file uploaded and the URL provided in the upload form are tracked automatically in different fields. This element can be used for several different types of source that are related to the object (such as reference sources). The name of the content holder should no longer be recorded here as a new element.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

language

A language of the resource. Use ISO 639 two letter language tags (it, en, fr, de, el, ...) Use this element for the language of textual objects and also where there is a language aspect to other objects e.g. sound recordings, posters, newspapers etc). If there is no language aspect to the digital object (e.g. a photograph), please ignore this element. This element is not for the language of the metadata of a resource, which may be described in xml:lang attribute. In case the digital object presents more languages, use more language elements, one for each language.

examples	en, it, fr, de, el, hu, es, ca
count (1, 0..1, 0..many, 1..many)	0..many
language	No
notes	None

relation

A reference to a related resource. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system. This is information about resources that are related to the original analog or born digital object.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

coverage

The extent or scope of the content of the resource. Coverage will typically include spatial location (a place name or geographic co-ordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [Getty Thesaurus of Geographic Names, <http://www.getty.edu/research/tools/vocabulary/tgn/>]). Where appropriate, named places or time periods should be used in preference to numeric identifiers such as sets of co-ordinates or date ranges. Coverage is the unqualified spatial or temporal coverage of the original analog or born digital object. Use of the more specific dcterms:spatial and dcterms:temporal elements is preferred where possible.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional

notes	None
--------------	------

rights

Information about rights held in and over the resource. Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource. This is a free text element and should be used for information about intellectual property rights or access arrangements for the digital object that is additional to the controlled value provided in europeana:rights.

examples	“All rights reserved”
count (1, 0..1, 0..many, 1..many)	0..many
language	Mandatory
notes	None

8.7 DCTERMS – dublin core terms metadata

In this section are reported the dcterms elements that are supported, that are the ones supported by Europeana.

alternative

An alternative name given to the resource. Typically, an Alternative title will be a name by which the resource is alternatively referred and it is different from the formal Title. Any alternative title by which the original analog or born digital object is known. This can include abbreviations or translations of the title.

Examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Mandatory
notes	None

tableOfContents

A list of subunits of the resource. A list of the units within the original analog or born digital resource object.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Mandatory
notes	None

created

Date of creation of the resource. This is the date when the original analog or born digital object was created.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

issued

Date of formal issuance (e.g., publication) of the resource. The date when the original analog or born digital object was issued or published.

examples	...
count	0..many

(1, 0..1, 0..many, 1..many)	
language	Optional
notes	None

extent

The size or duration of the resource. Refinement of format. Size or duration of the digital object and the original object may be recorded.

examples	“30 pages”, “01:15:20”
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

medium

The material or physical carrier of the resource. Refinement of dc:format.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

isVersionOf

A related resource of which the described resource is a version, edition, or adaptation. Changes in version imply substantive changes in content rather than differences in format. Refinement of dc:relation. See also dcterms:hasVersion.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

hasVersion

A related resource that is a version, edition, or adaptation of the described resource. Changes in version imply substantive changes in content rather than differences in format. Refinement of dc:relation. See also dcterms:isVersionOf. Use dcterms:hasFormat for differences in format.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

isReplacedBy

A related resource that supplants, displaces, or supersedes the described resource.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

replaces

A related resource that is supplanted, displaced, or superseded by the described resource.

examples	...
-----------------	-----

count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

isRequiredBy

A related resource that requires the described resource to support its function, delivery, or coherence.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

requires

A related resource that is required by the described resource to support its function, delivery, or coherence.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

isPartOf

Is Part Of - A related resource in which the described resource is physically or logically included. Use for the name of the collection which the digital object is part of.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

hasPart

A related resource that is included either physically or logically in the described resource. Refinement of dc:relation. See also dcterms:isPartOf.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

isReferencedBy

Is Referenced By: A related resource that references, cites, or otherwise points to the described resource.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

references

A related resource that is referenced, cited, or otherwise pointed to by the described resource.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many

language	Optional
notes	None

isFormatOf

A related resource that is substantially the same as the described resource, but in another format. Refinement of dc:relation. See also dcterms:hasFormat.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

hasFormat

A related resource that is substantially the same as the pre-existing described resource, but in another format. Refinement of dc:relation. See also dcterms:isFormatOf. Use dcterms:hasVersion for differences in version.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

conformsTo

An established standard to which the described resource conforms. Refinement of dc:relation. The names of standards that the digital object (digitized or born digital) complies with and which are useful for the use of the object.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

spatial

Spatial characteristics of the resource. Information about the spatial characteristics of the original analog or born digital object, i.e. what the resource represents or depicts in terms of space. This may be a named place, a location, a spatial coordinate or a named administrative entity.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

temporal

Temporal characteristics of the resource. The temporal characteristics of the original analog or born digital object i.e. what the resource is about or depicts in terms of time. This may be a period, date or date range.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

provenance

A statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity, and interpretation. The statement may include a description of any changes successive custodians made to the resource. This relates to the ownership and custody of the original analog or born digital object.

examples	...
count (1, 0..1, 0..many, 1..many)	0..many
language	Optional
notes	None

8.8 Taxonomy Classification

For each drupal taxonomy term associated with the content it is reported:

label

The label of the term in each available language

id

attribute with the drupal id for the term

root

attribute with the id of the root term where the term is a descendent.

vid

attribute with the id of the vocabulary of the term

path

attribute with the term ids separated by spaces from the root to the term (e.g. "664 668")

8.9 ECLAP Groups

For each drupal og group associated with the content it is reported:

label

the label of the group

id

attribute with the id of the group

8.10 ECLAP Aggregations

If the content is a playlist or a collection are reported the set of axoids that belong to the playlist/collection.

For playlists more information is provided for each content in the playlist:

For audio and video:

startTime

attribute with the time instant in seconds from the audio/video start representing the time in the resource to start resource Audio/Visual rendering, if omitted the resource start time is intended

endTime

attribute with the time instant in seconds from the audio/video start representing the time in the resource to end the resource Audio/Visual rendering, if omitted the resource end time is intended

For images:

duration

the duration in seconds of the image display