



## DELIVERABLE

**Project Acronym:** LoCloud  
**Grant Agreement number:** 325099  
**Project Title:** Local content in a Europeana cloud

### Report accompanying D2.4: Metadata preparation toolkit

**Version: 1**

**Authors: Eleni Iskou, Nasos Drosopoulos, Nikolaos Simou, Kostas Sismanis, Dimitris Gavrilis, Vassilis Tzouvaras**

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	x
C	Confidential, only for members of the consortium and the Commission Services	

## Revision History

Revision	Date	Author	Organisation	Description
0.1	10/06/2014	Eleni Iskou	NTUA	Deliverable outline
0.2	17/06/2014	Kostas Sismanis	NTUA	More publication and metadata reports
0.3	18/06/2014	Nikolaos Simou	NTUA	Interoperability Services
0.4	19/06/2014	Nasos Drosopoulos	NTUA	Annotation Group edit
0.5	25/06/2014	Vassilis Tzouvaras	NTUA	Executive summary
0.6	28/06/2014	Eleni Iskou	NTUA	First Draft
0.7	02/07/2014	Nikolaos Simou	NTUA	Updated according to review comments
0.8	03/07/2014	Eleni Iskou	NTUA	Second Draft
0.9	14/07/2014	Dimitris Gavrilis	DCU	Added section for common errors
1.0	15/07/2014	Vassilis Tzouvaras	NTUA	Final Draft

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This deliverable contains original unpublished work except where clearly indicated otherwise.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

## Contents

1	Executive summary .....	2
2	Metadata Preparation through MINT .....	3
3	Metadata upload service .....	5
4	Metadata statistics .....	8
5	Metadata mapping – transformation .....	10
6	Annotation – Group edit .....	12
7	Metadata reports .....	14
8	MINT to More Publication .....	18
9	Common errors & misconceptions .....	20
10	Conclusion .....	22
	Appendix 1: Definition of terms.....	23

## 1 Executive summary

This report accompanies and describes the contents of D2.4 metadata preparation toolkit that is available at :

<http://mint-projects.image.ntua.gr/locloud/>

It is important to mention at this point that this report does not describe MINT's functionality. The interested reader is referred to the online documentation that can be found at:

<http://mint.image.ntua.gr/mint2/>

and in D2.2 and D2.3, which are available at: <http://www.locloud.eu/Resources/Deliverables>.

The main objective of this report is to guide content providers in learning how to prepare their metadata for efficient publication to Europeana.

In the metadata preparation process, providers are checking if metadata datasets:

- ⤴ are well formed (i.e. encoding)
- ⤴ are in the correct format (i.e. xml, csv)
- ⤴ can be harvested using the OAI-PMH protocol
- ⤴ conform to a known metadata schema (e.g. CARARE, lido, EDM)
- ⤴ contain all mandatory elements, (isShowAt, title, unique ID)
- ⤴ contain valid links to the Digitized Cultural Heritage Objects

In Locloud we provide a set of tools and services that will facilitate providers to check and improve the level of quality and thus readiness of the metadata for publication at the Europeana portal.

In Locloud, providers are using MINT to ingest metadata to MORE. In this deliverable we emphasize on the set of tools and services that exist within MINT and can be used by providers to prepare their metadata. These set of tools are

- ⤴ Metadata upload
- ⤴ Metadata statistics.
- ⤴ Datasets reports
- ⤴ Mapping, Annotation and Group edit
- ⤴ Metadata transformation and validation

Additional information about the metadata preparation process including screencasts video can be found in <http://support.locloud.eu/>.

## 2 Metadata Preparation through MINT

In order to explain the usefulness of the MINT services in preparing the metadata, we provide a real world example.

A LoCloud content provider wants to provide metadata to Europeana using the MINT ingestion tool. The metadata are stored in a relational database (e.g. postgres sql). There are three ways available to ingest metadata to MINT.

1. setup an OAI-PMH repository,
2. export the metadata in XML formatting
3. export the metadata in CSV format.

The preferred way is the first one but it hides many technical difficulties – since the setup of a server that will satisfy the OAI-PMH specification is not a trivial task<sup>1</sup>. The second way hides also technical difficulties because of the necessary transformation of related data in a structural way while it exports metadata in a W3C metadata standard. The third way is the easiest but can become tricky as CSV is not a standard for exchanging metadata on the web and inconsistencies may occur.

In the first case, after setting up the repository, the first step is to check the OAI URL using MINT's metadata upload service and then to start the harvesting process. The MINT reporting tool will report if the harvesting process has been completed successfully.

In the second case, providers must check if the produced XMLs are valid with the correct encoding. This service is provided in MINT's metadata upload. In addition to the XML validation service, providers can check if the metadata conforms to a known metadata (LIDO, CARARE, EDM). In this case, providers can omit the mapping phase and can proceed straight to publish metadata to MORE and Europeana.

In the third case, the produced CSV will be upload to MINT where its validity and encoding are checked and then is transformed in XML for further processing.

Having uploaded the metadata dataset into MINT, the next step is to check the quality of the metadata. In the metadata statistics module, providers can view the values of the elements, information about the count, the distinct count, the frequency and finally can check the validity of the URL of the isShownAt and isShownBy elements. Providers must also check if all the elements required by the project are present in the imported metadata.

The next step is to transform the metadata to the one of the metadata standards provided by the project that are CARARE 2.0.2, CARARE 2.0.3, CARARE 2.0.4, LIDO and EDM. In case that not all mandatory elements are present or if the URLs are missing or if there is not any then providers can add them during the mapping process or they need to correct their metadata at their site and re-import them in MINT.

The annotation/group edit service can be used to add and/or edit or group edit existing records of the dataset. This service is available after metadata are mapped into one of the available metadata standards in MINT (CARARE, Lido, EDM).

---

<sup>1</sup> <http://www.openarchives.org/OAI/openarchivesprotocol.html>

## D2.4: Metadata preparation toolkit

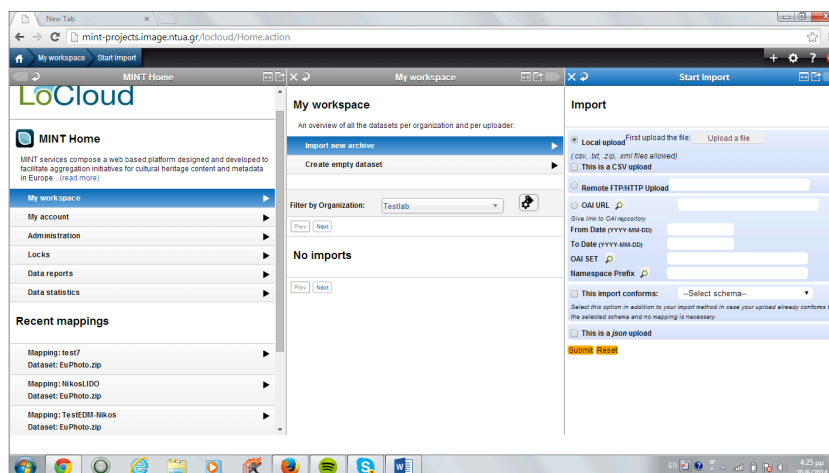
In addition to the above services, providers can check the progress using the data reports functionality. This service provides a visual way to present how many items have been mapped, transformed and published in MINT.

Additional information about the metadata preparation process including screencasts video can be found in <http://support.locloud.eu/>.

The rest of the deliverable contains all the details for preparing metadata using the MINT services. In particular, section 2 describes the metadata upload service. The third section presents the metadata statistics service. The following section outlines the metadata mapping and transformation service while the one follows the annotation and group edit service. Section 7 presents the data reports and finally we provide some information about the MINT MORE publication process. Section 9 presents some common errors that providers make and which prevent them from publishing content to Europeana.

### 3 Metadata upload service

Data providers, can import their metadata to the MINT ingestion platform by using common data delivery protocols, such as OAI-PMH, HTTP and FTP. More precisely, by selecting the button “import new archive” from “My workspace- MINT’s start page”, they can select from a range of protocols as shown in the figure below.



**Figure 1: MINT's metadata import functionality**

By selecting the HTTP upload, content providers can import data from their web browser while the FTP/HTTP option, allows them to import files from a remote FTP/HTTP location.

In addition to the aforementioned protocols content providers can also import data using an OAI-PMH URL. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository interoperability. Data Providers are repositories that expose structured metadata via OAI-PMH. Service Providers then make OAI-PMH service requests to harvest that metadata. OAI-PMH is a set of six verbs or services that are invoked within HTTP.

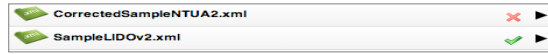
When a data provider uses an OAI-PMH URL he has to fill in the base URL and then click on the button next to it, in order to check its validity. As a valid OAI-PMH URL is considered a URL that serves the metadata according to the OAI-PMH specification<sup>2</sup>. Once the validity of the URL is confirmed the set and the prefix can be selected from a drop down list that is populated by the given OAI URL (server).

It is important to mention at this point that the metadata upload functionality of MINT is size depended – this means that the bigger an upload is the longer it takes to be imported. Especially for the OAI-PMH harvesting functionality that permits the upload of really very big datasets providers are highly recommended to plan their metadata upload a day before the day they intend to use on of the metadata toolkit services.

<sup>2</sup> <http://www.openarchives.org/OAI/openarchivesprotocol.html>

### Metadata formats validation services

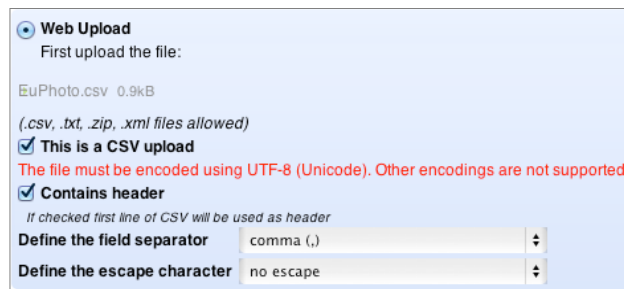
Data providers can import to the MINT ingestion platform XML, Comma Separated Values<sup>3</sup> (CSV) files or zip files containing the XML and CSV files. The validity of the uploaded metadata is checked according to the format used. If the import is valid a green tick appears next to it while if it is invalid a red X appears as shown in the figure below.



**Figure 2: Metadata format validation service**

In detail:

- CSV files are valid if the number of values per line - using the delimiter specified by the user - are equal to those of the header specified. Furthermore the encoding of the CSV files has to be UTF-8<sup>4</sup> ensuring in that way that all the characters will be parsed and displayed correctly.



**Figure 3: CSV upload**

- XML files are valid if they are well formed. A well-formed document in XML is a document that adheres to the syntax rules specified by the XML 1.0 specification in that it must satisfy both physical and logical structures<sup>5</sup>.
- A ZIP file containing XML files is valid if all the XML files contained are well formed.

It is important to mention at this point that for the XML and ZIP imports an additional validation service is offered by the MINT ingestion platform. Data providers can select the metadata schema (CARARE 2.0.0, CARARE 2.0.1, CARARE 2.0.3, EDM and LIDO) their import conforms to. Once the schema is set by the user the imported XML files are validated according to the schema.

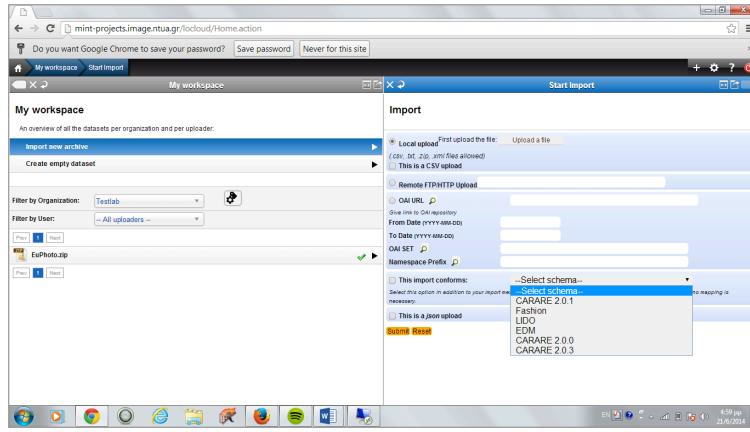
<sup>3</sup> [http://en.wikipedia.org/wiki/Comma-separated\\_values](http://en.wikipedia.org/wiki/Comma-separated_values)

<sup>4</sup> [http://en.wikipedia.org/wiki/Character\\_encoding](http://en.wikipedia.org/wiki/Character_encoding)

<sup>5</sup> "XML: The Well-formed Document". The UK Web Design Company. Retrieved 11 August 2013.



## D2.4: Metadata preparation toolkit



**Figure 4: Metadata schema validation service**

## 4 Metadata statistics

Data statistics is another service of the MINT ingestion platform. It provides to users detailed information of the imported metadata. By selecting “data statistics” they can see all the metadata elements of the imported dataset and by clicking on an element they can view its values. In the following screenshot we see the title element. In similar manner we can view the values of all metadata elements. Moreover, additional information is provided i.e. XPath, Namespace URI, count, Distinct count, Average length. In the example below the screenshot count and distinct count have same values, which means that this element has unique values.

The screenshot displays the 'Dataset Statistics' for 'EuPhoto.zip'. The main table lists metadata elements with their respective counts and distinct counts. The 'tms:Title' element is selected, and its details are shown in the right-hand panel. The details panel includes the XPath, Namespace URI, Count (5), Distinct Count (5), and Average Length (15.8). Below the details, a list of values for the title element is shown, including 'Caryatids', 'Parthenon', 'Parthenon East Pediment', 'Parthenon South Peristyle', and 'The Parthenon'.

Xpath	Count	Distinct
/ImageMetadata/DescriptiveMetadata/Title	5	5
/ImageMetadata/Location	5	1
/ImageMetadata/AdministrativeMetadata	5	5
/ImageMetadata/Identifier	5	5
/ImageMetadata/MetadataURI	5	5
/ImageMetadata/PhotoURI	5	5
/ImageMetadata/Organisation	5	1
/ImageMetadata/DescriptiveMetadata	5	5
/ImageMetadata/Title	5	5
/ImageMetadata/Date	5	1
/ImageMetadata/Photographer	5	1
/ImageMetadata/Photographic...	5	1
/ImageMetadata/Location	5	1
/ImageMetadata/Photographic...	5	1
/ImageMetadata/Material	5	1
/ImageMetadata/Description	5	5
/ImageMetadata/Copyright	5	1
/ImageMetadata/Dimensions	5	4

Figure 5: Dataset statistics

Dataset statistics assists the providers to check the mandatory elements that are defined with in the metadata target schema, to evaluate the quality of the imported metadata, to avoid duplicates. How rich are the imported metadata. Additionally they can check their diversity, for instance if the number of distinct count” is much smaller than the number of “count”, then the imported metadata are poor.

Possible errors may be indicated by observing the length of the element. The length shows how the metadata are extended or small. The users can observe the length to perceive errors, for example if the length of the description is very small then the description does not correspond to correct element

In the next example below we see the element subject, Count and distinct count have not same values, which means that this element has 3 unique values

## D2.4: Metadata preparation toolkit

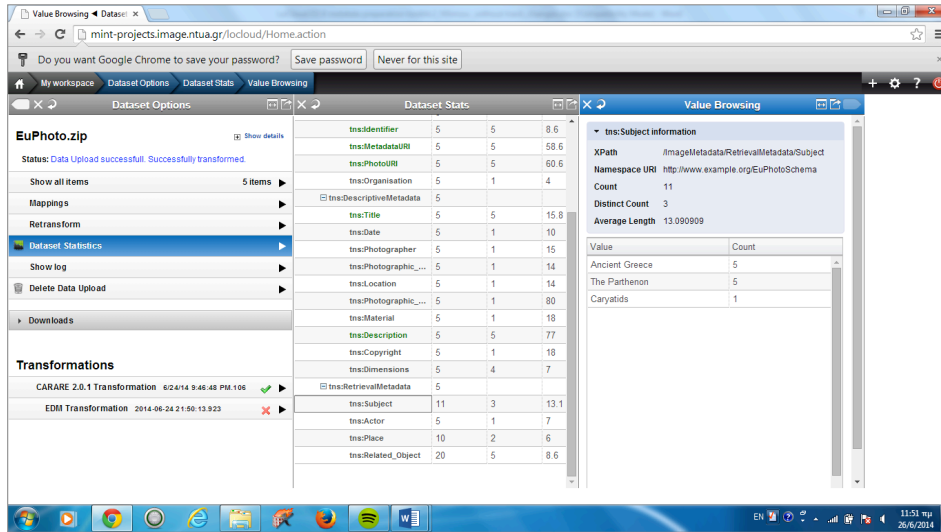
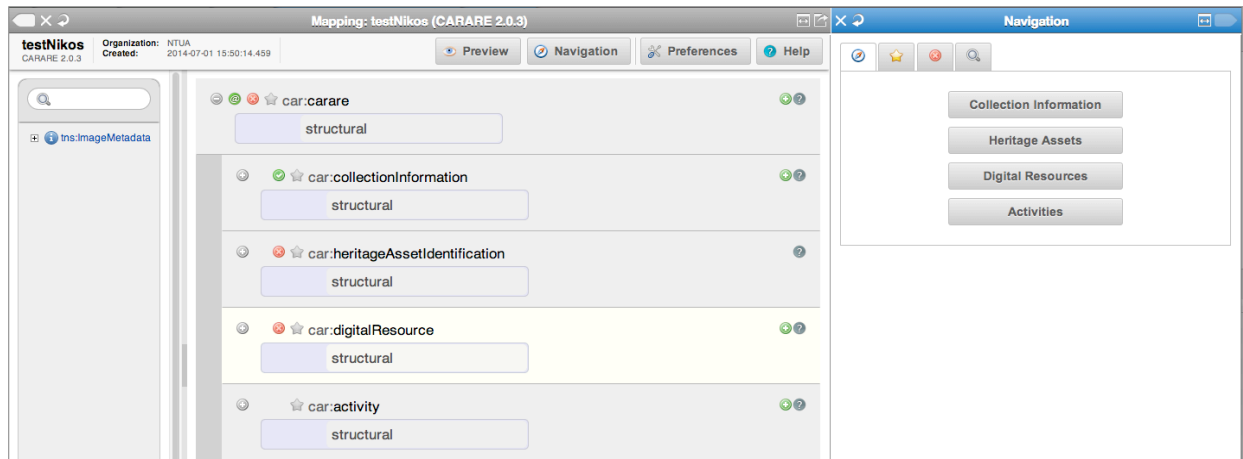


Figure 6: Dataset statistics

## 5 Metadata mapping – transformation

Metadata mapping is the crucial step of the ingestion procedure. It formalizes the notion of a metadata crosswalk, hiding the technical details and permitting semantic equivalences to emerge as the centrepiece. It involves a user-friendly graphical environment (*Figure 8* shows an example mapping opened in the editor) where interoperability is achieved by guiding users in the creation of mappings between input and target elements. LoCloud content providers can use this service for transforming their proprietary metadata to one of the metadata standards selected by the project and provided by the MINT tool that are CARARE 2.0.2., CARARE 2.0.3, CARARE 2.0.4. LIDO and EDM.

The user's mapping actions are expressed through XSLT style sheets, i.e. a well-formed XML document conforming to the namespaces in XML recommendation. XSLT style sheets are stored and can be applied to any user data, exported and published as a well-defined, machine understandable crosswalk and, shared with other users to act as template for their mapping needs.



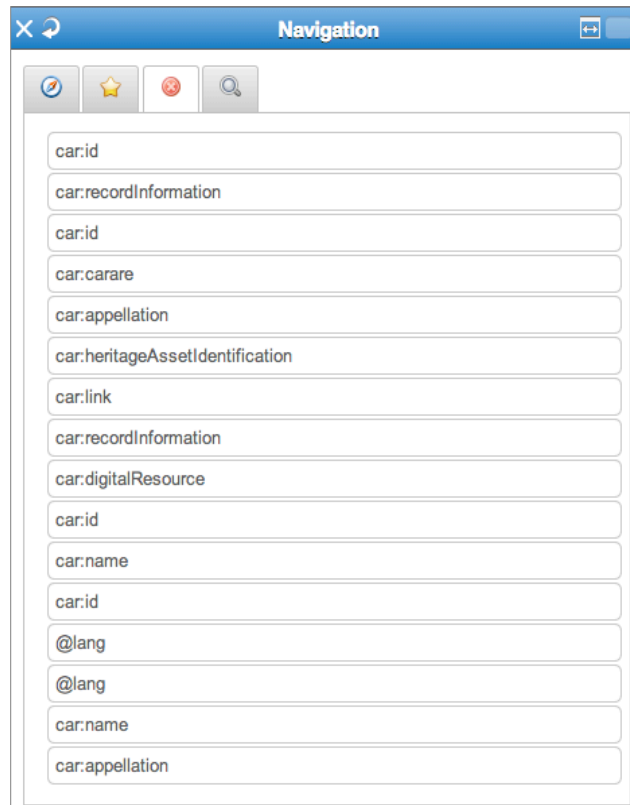
**Figure 7: Screenshot of the mapping editor**

The structure that corresponds to a user's specific import is visualized in the mapping interface as an interactive tree that appears on the left hand side of the editor. The tree represents the snapshot of the XML schema that is used as input for the mapping process. The user is able to navigate and access element statistics for the specific import while the set of elements that have to be mapped can be limited to those that are actually populated. The aim is to accelerate the actual work, especially for the non-expert user, and to help overcome expected inconsistencies between schema declaration and actual usage.

On the right hand side, buttons correspond to high-level elements of the selected target schema and are used to access their corresponding sub-elements. These are visualized on the middle part of the screen as a tree structure of embedded boxes, representing the internal structure of the complex element. The user is able to interact with this structure by clicking to collapse and expand every embedded box that represents an element, along with all relevant information (attributes, annotations) defined in the XML schema document. To perform an actual (one to one) mapping between the input and the target schema, a user has to simply drag a source element from the left and drop it on the respective target in the middle.

The user interface of the mapping editor is schema aware regarding the target data model and enables or restricts certain operations accordingly, based on constraints for elements in the target XSD. For example:

- when an element can be repeated then an appropriate button appears to indicate and implement its duplication.
- When an element of the target schema is mandatory and it is not mapped then it appears in red and also under error panel of the navigation pane (Figure 8)



**Figure 8: MINT's report on missing mandatory elements**

Several advanced mapping features of the language are accessible to the user through actions on the interface, including:

- String manipulation functions for input elements;  
This allows providers to use part of their metadata values e.g. the first or last 4 characters
- m-1 mappings with the option between concatenation and element repetition;
- Structural element mappings;
- Constant or controlled value assignment; This allows providers to add values to mandatory or recommended by the project elements that do not manage to upload
- Conditional mappings (with a complex condition editor);
- Value mappings editor (for input and target element value lists).

## 6 Annotation – Group edit

### Metadata authoring & editing services

A set of new services for authoring and editing metadata have been introduced in MINT. An annotator extends the basic metadata management functionalities with the ability to create, delete and modify records - XML records, CSV file rows and, for RDF models, resource records serialized in RDF/XML - in a dataset. A group edit service allows for massive 'tagging' of records using elements of the schema. Values can be free-text or retrieved from potentially associated lists – enumerations in the XSD or thesauri serialized in SKOS.

### Metadata Annotator

As CH metadata continue to evolve, it becomes more important for non-technical content providers to be able to contribute more effectively at all phases of the metadata publication cycle. The annotator, together with the ability to manage (create & delete records) datasets facilitate the deployment of more flexible aggregation workflows. It enables users to edit structured XML content, rendered in an intuitive data entry environment based on MINT's mapping tool UI components. This gives non-technical users the ability to create structured content including semantics and structure definitions based on XSD. The interface can switch between a simple view, in which each record is rendered fully as a tree and, a complex view where pre-defined bookmarks allow the grouping of specific groups of elements in a way that is more intuitive for the user. The preview interfaces present the metadata structure as it will appear in the target delivery format(s). In addition, records can also be rendered in HTML with an appropriate template for the metadata schema.

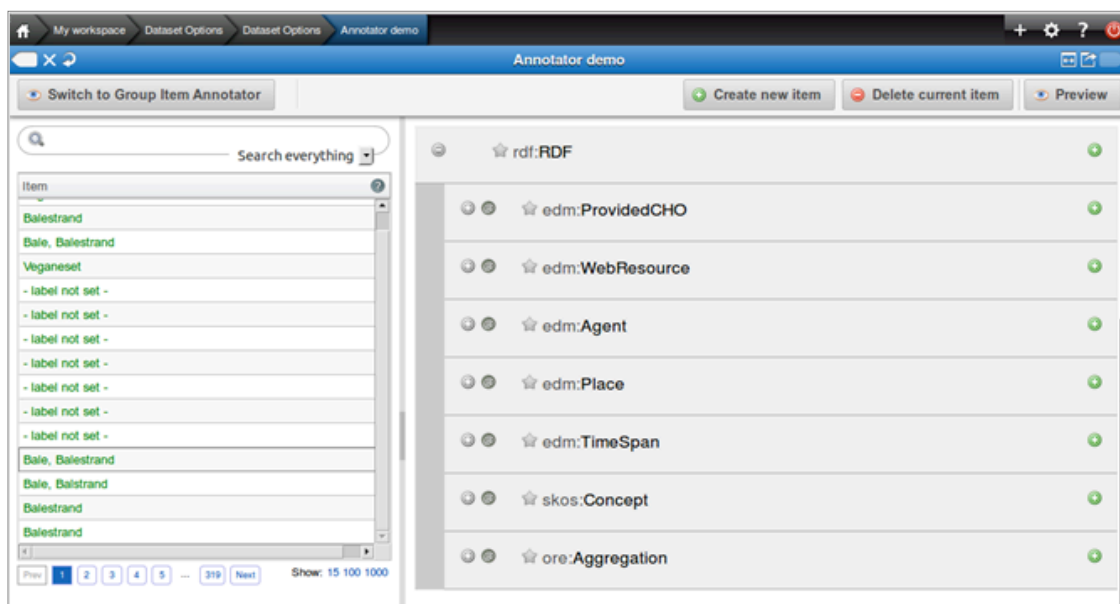


Figure 9: Annotator

### Group Edit

This prototype service is deployed to enable cleaning and enrichment operations based on filtering datasets. The search capability (full-text or in specific elements/XPaths) is used to define groups of records based on specific criteria while allowing to deselect and exclude

## D2.4: Metadata preparation toolkit

records from the result set. Based on this group generation, a number of operations are supported - on a schema level for the addition, deletion and update of elements with specific values (free text or controlled) and, on a data level for the application of conditional edits, string manipulation or numerical functions on already existing values. The resulting operations for the transformation of the dataset are queued for processing and a new version is created. Current developments include a more elaborate versioning system for subsequent 'edits' of a dataset and the ability to connect to external repositories for the discovery of potential resources of interest

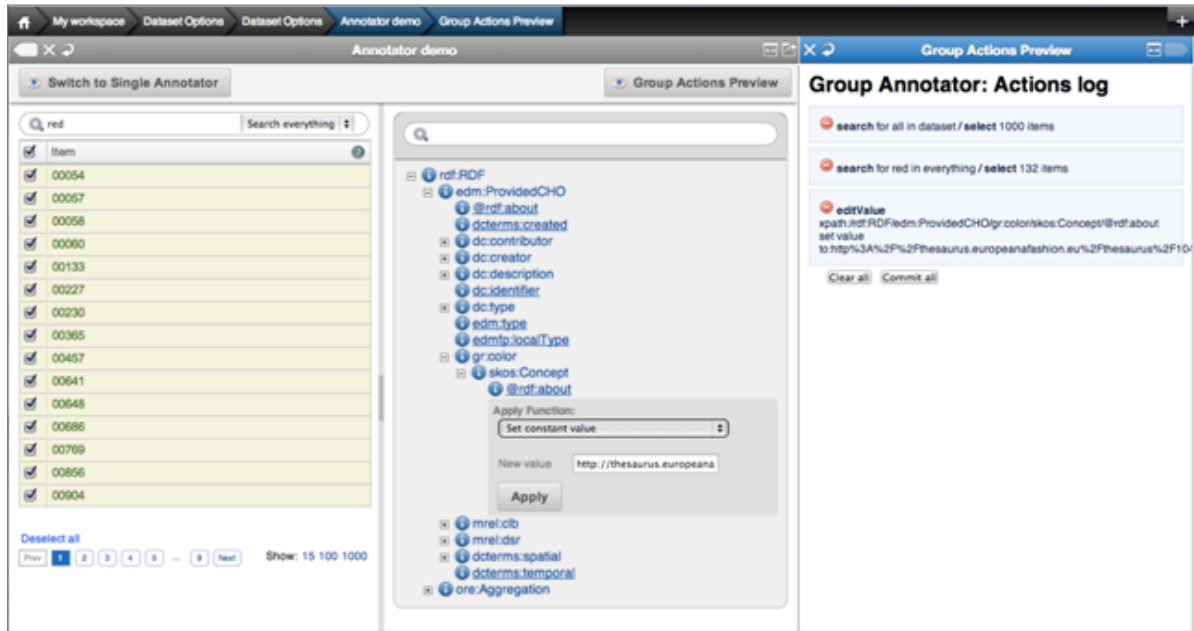


Figure 10: Group Edit

## 7 Metadata reports

The Metadata reports goal is to present in an intuitively manner a web presentation of the history of an Organization's actions made through Mint including:

1. Dataset Imports
2. Dataset Transformations
3. Dataset Publications
4. An overview of an Organization's progress. Including the presentation of all corresponding dataset item counts (imported, transformed, published), along with the Organization Project's specific, published items goal number.
5. An overall overview of all the Project's organizations imported, transformed or published and project goal items number.

Within the Mint interface, a user will find the Data Statistics button at the bottom of the initial Home list. Clicking it, brings a new window with available options as shown in the figure 10:

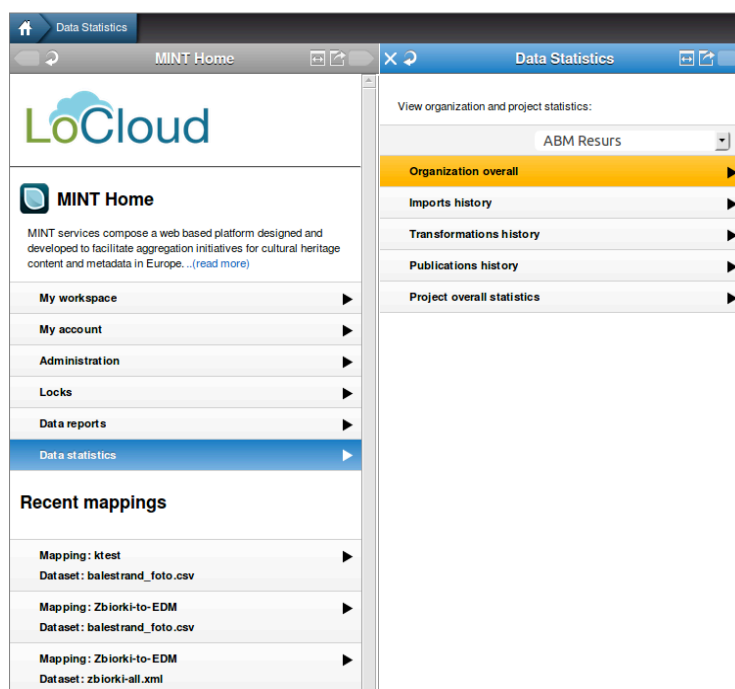


Figure 10: Data statistics

### Organization Overall

Selecting an organization at the drop down list and clicking on “Organization overall” brings a new window. Containing a table, with the total number of the organization's imported, transformed and published items. We can also see the number of published items per schema of interest depending on the project and set at the project's configured properties. Also we can see the organization's publication items goal number set within the project. An example screen is shown at Figure 11: Organization Overall.



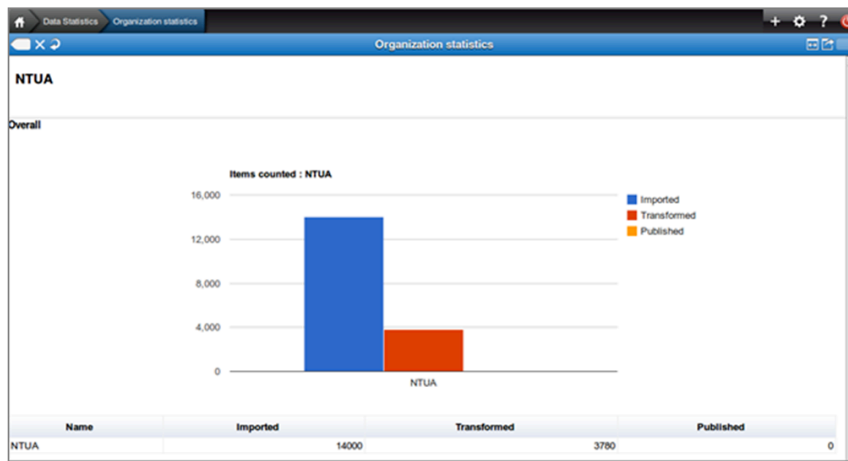


Figure 11: Organization Overall

### Organization Imports History

Within Data Statistics menu, by selecting an organization, and then clicking on Imports history brings up in a new window a table. Each table row represents a specific Data Import made by this organization, as show in Figure 12: Data Imports, containing details about the:

- ⤴ Name of the upload
- ⤴ Date of Import
- ⤴ Number of items contained.

Name	Date	Items	Invalid Items
balestrand_foto.csv	Jun 19, 2014, 7:53:00 PM	4077	4077
balestrand_foto.csv	Jun 19, 2014, 7:35:49 PM	4077	0
zbiorki-all.xml	Jun 17, 2014, 2:01:15 PM	1204	0
zbiorki-all.xml	Jun 17, 2014, 12:38:14 PM	1204	0
http://tbc.pionier.net.pl/zbiorki/oaipmh-repository.xml - DigitalArchiveOfMemorabilia	Mar 10, 2014, 3:51:49 PM	1186	0
http://kulturalvsdata.se/oaicat-europeana/OAIHandler/ - raa_bbrb	Feb 26, 2014, 4:04:48 PM	1849	0
MOMU.tgz	Feb 17, 2014, 8:09:34 PM	243	0

Figure 12: Data Imports

### Organization Transformations

By selecting an organization, and then clicking on Transformations history, brings up in a new window a table. Each table row represents a specific Transformation made by this organization, as show in Figure 13: Transformations, containing information about it like:

Parent Dataset Name, Date of Transformation, Item counts of in / valid items, Target Schema and the Mapping used. Finally in a chart we can see the number of items transformed per schema of interest.

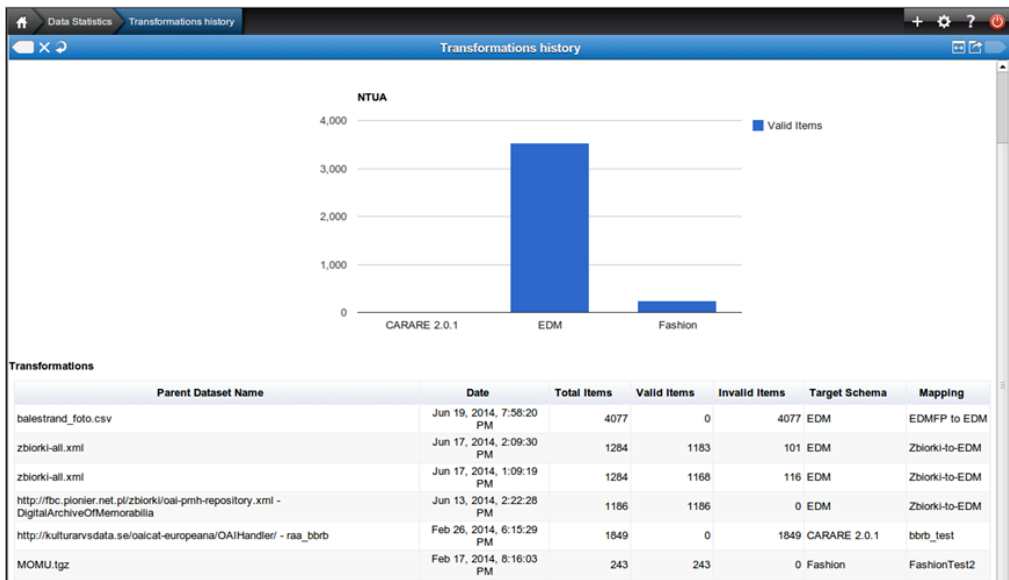


Figure 13: Transformations

### Organization Publications

By selecting an organization, and then clicking on Publications history, brings up in a new window a table. Each table row represents a specific Publication made by this organization, as show in Figure 14: Publications, containing information about it like:

Parent Dataset Name, Date of Publication, Item counts of in / valid items, Target Schema and the Mapping used.



Figure 14: Publications

### Project Overall

In this screen (Figure 14: Project Overall) we can see overall detail for all the project organizations together in one table that includes columns with the:

- ⤴ Imported items number
- ⤴ Transformed items number
- ⤴ Published items number

## D2.4: Metadata preparation toolkit

- Published per schema items number
- Project published goal item number (Aggregation and Goal)

All tables in the Data Statistics can be sorted by clicking on the header of the specific column of choice. For example someone can sort the table by clicking on Published Items.

All tables can also be exported by selecting the table with the mouse and copying it to a spreadsheet. This allows users to create their own reports.

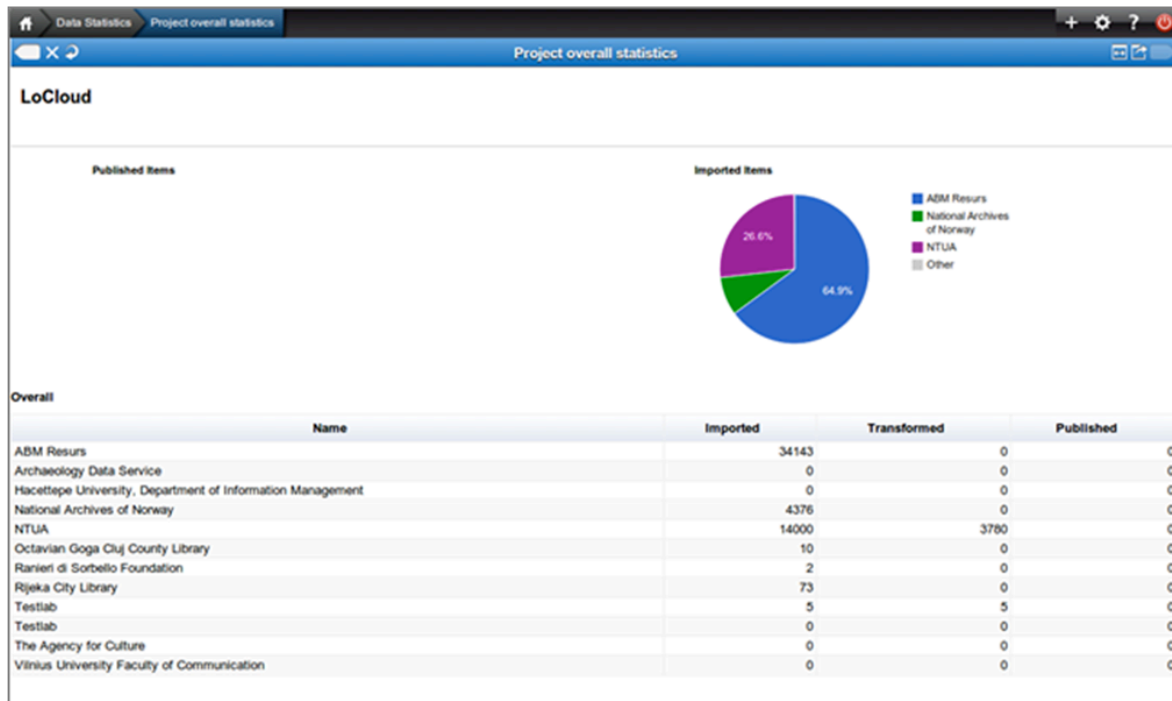


Figure 15: Project Overall

## 8 MINT to More Publication

In this section we present the MINT to MORE publication process.

The MoRE Publication process goal is to collect an Organization's items, that qualify for publication, in a data file that is transferred to the MoRE2 repository.

Within the MINT interface, a user by clicking on a particular dataset in his workspace list, can see a list of options as shown in Figure 16: Dataset Options.

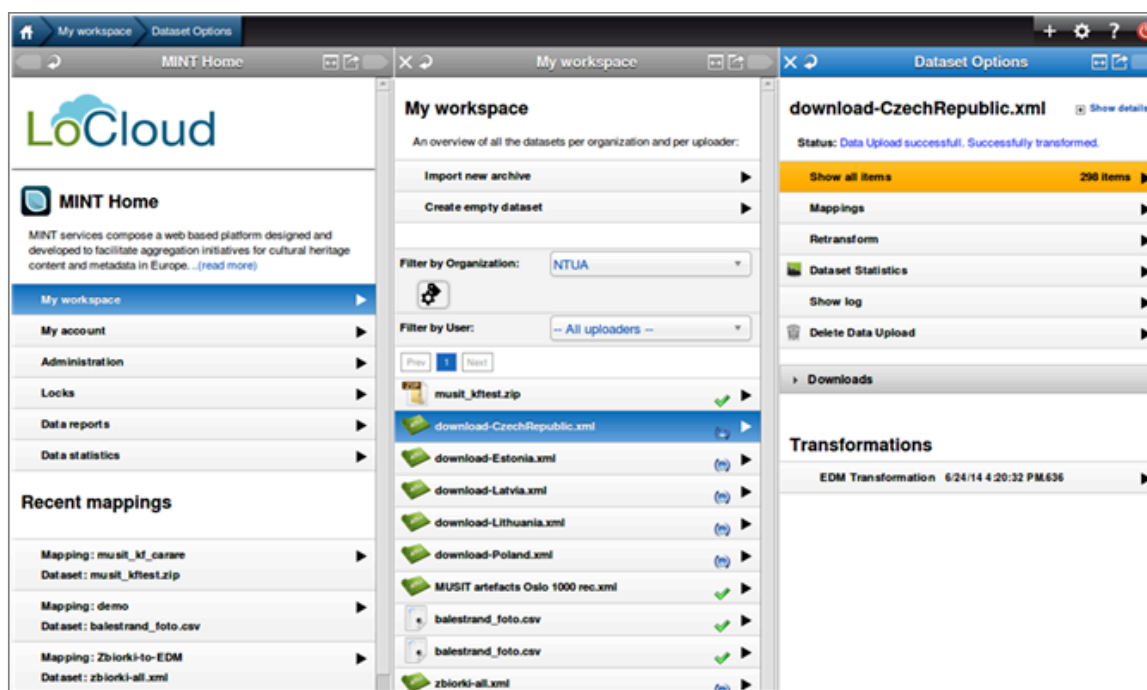


Figure 16: Dataset Options

When user has transformed successfully a dataset into the Publication Schema, will have access to a new button called Publish to More, appearing in the dataset options as show in Figure 17: Publish to MoRE2.

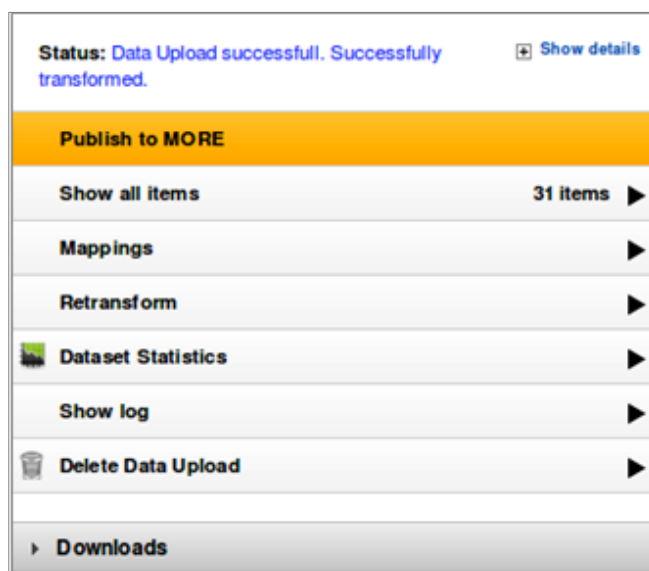
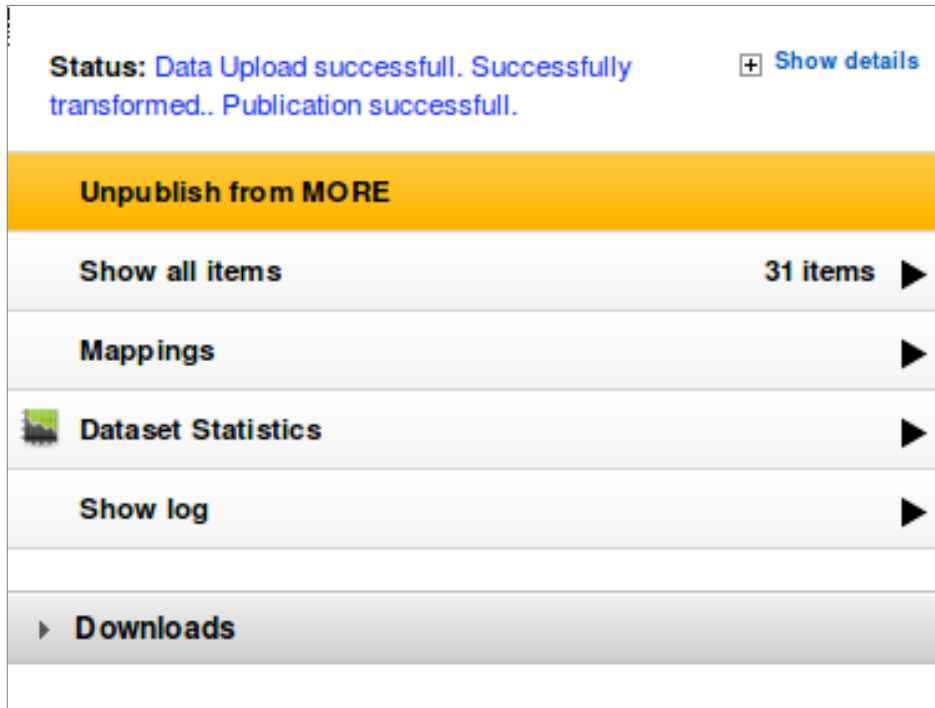


Figure 17: Publish to MoRE2

In a similar way the option to Unpublish a particular dataset from the MoRE2 repository will be available for all datasets already published as show in Figure 18: Unpublish fromMoRE2



**Figure 18: Unpublish fromMoRE2**

### Technical details

The publication to MoRE repository consists of :

- ⤴ Iterating all published datasets of the organization who is publishing, gathering all previously published items.
- ⤴ These items, along with the new published dataset 's items are added in a .tar file.
- ⤴ Every item is placed in a different directory along with an xml containing meta information about the item.
- ⤴ A URL is created that will be used by the MoRE2 repository to download the publication file.
- ⤴ The URL is announced to the MoRE2 repository so the transfer can be initiated.
- ⤴ While the publication process is executing the user sees a Publication in progress message in the dataset options.

In a similar way, when unpublishing a dataset the same process is repeated by excluding the items of the particular dataset.

## 9 Common errors & misconceptions

In this section a list of common errors and misconceptions are presented, which prevent the acceptance of metadata by Europeana.

### Native identifiers

Each metadata record must have a unique identity; represented by a unique identifier that accompanies every record. If the metadata are stored in an information system (of any kind) such as: a digital library, a repository, an access database, an SQL database, there is always an identifier (e.g. the primary key of the record's table). If the metadata are stored in an excel spread sheet, it is always good to create a column that represent the identifier and assign values to it. It is recommended that the identifier values contain some information about the provider and the collection. For example: **mymuseum:coins:1453**. The identifiers are extremely useful to track each record when submitting.

### Landing page URLs

Each record should be associated with a unique URI that points to an html page that describes the digital object. This is not only mandatory but also of major important because this is the URL that Europeana will use to point to your collection. Each digital object must have its own unique page (not one page for all the objects of a collection).

### Object thumbnail

Europeana requires that each digital object comes with a thumbnail image. This image will be used to present the object in the Europeana library portal. Although there are automated tools that can extract this from the object's landing page, it is always better to explicitly provide this information through a URI.

### Language attributes

Many metadata elements are encoded in multiple languages. For example a title and a description could be exported both in the native language of the provider and in English. It is most important to always include the language attribute when preparing the metadata. This will result in less ambiguous metadata and will allow for better handling / mapping using automated tools. If it is possible, always use some standard for language encoding (as defined by the metadata schema you are using).

### Spatial information

When including spatial information and more specifically coordinates, always bear in mind that Europeana requires WGS84 encoding. Although there are services in the aggregation process that can handle almost any coordinate system (transform it automatically to WGS84), is its most important to include the coordinate reference system that you use. This greatly simplifies the aggregation process and leads to better results.

### Rights statements

## D2.4: Metadata preparation toolkit

Every metadata record must be accompanied by a rights statement. Although content providers usually have their own rights statements and policies, Europeana asks for a rights statement to be selected a specific list (<http://pro.europeana.eu/available-rights-statements>) for the Europeana Rights element. It is always good to spend some time to study this list and align/map it to yours. This will allow you to define the correct rights statements during the metadata preparation phase so that the aggregation process runs smoothly.

## 10 Conclusion

In this deliverable, we presented how the MINT tool can be used by content providers to prepare their metadata for efficient publication to Europeana.

In particular, we illustrated how the metadata upload, metadata statistics, group-edit and annotation services can be used to improve the quality of the metadata. Using the metadata upload service, providers can check the validity of the XMLs or CSV and check the OAI URL if this exists. Using the metadata statistics can view the values inside the XMLs, check the links to the digital objects and inspect for unique IDs. Using the annotation and group-edit service can add missing values or even add entire items if required. Finally, using the metadata report can monitor the progress of the ingestion process.



**Appendix 1: Definition of terms**

OAI-PMH	Open Archives Initiative / Protocol for Metadata Harvesting
LIDO	Lightweight Information Describing Objects
EDM	Europeana Data Model
IsShownAt	EDM Elements
IsShownBy	EDM Elements
XML	Extensible Markup Language
CSV	Comma Separated Values
HTTP	Hypertext Transfer Protocol
FPT	File Transfer Protocol
UTF8	UCS Transformation Format—8-bit
URI	uniform resource identifier
XPATH	XML Path Language,
XSD	XML Schema
RDF	Resource Description Framework
XSLT	Extensible Stylesheet Language Transformations