ICT-PSP Project no. 297158

EUROPEANAPHOTOGRAPHY

EUROPEAN Ancient PHOTOgraphic vintaGe repositoRies of digitAized Pictures of
Historical qualitY

Starting date: 1st February 2012

Ending date: 31st January 2015

Deliverable Number:	D 4.2
Title of the Deliverable:	Multilingual Mapping tool
Dissemination Level:	Public

Contractual Date of Delivery to EC:	Month 18
Actual Date of Delivery to EC:	July 2013

Project Coordinator

Company name : KU Leuven
Name of representative : Fred Truyen
Address : Blijde-Inkomststraat 21 B-3000 Leuven PB 3301
Phone number : +32 16 325005
E-mail : fred.truyen@kuleuven.be
Project WEB site address : http://www.europeana-photography.eu

Context

WP 4	Indexing and Multilingual Support
WP Leader	KMKG
Task 4.5	Metadata Multilingual Enhancement
Task Leader	KMKG
Dependencies	

Author(s)	Nacha Van Steen (KMKG)
Contributor(s)	
Reviewers	Nikolaos Simou (NTUA), David Inglesias (CRDI), Valentina Bachi (Promoter)
Approved by:	

History

Version	Date	Author	Comments
0.1	08/07/2013	KMKG - Nacha Van Steen	Finalized draft
0.2	24/07/2013	KMKG - Nacha Van Steen	
0.3	30/07/2013	KMKG - Nacha Van Steen	After internal review
1.0	30/7/2013	Nikolaos Simou (NTUA), David Inglesias (CRDI), Valentina Bachi (Promoter), Nacha Van Steen (KMKG)	Final reading

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

This deliverable accompanies the part of the MINT mapping tool that allows content partners to map their individual, local concepts to the central Europeana Photography vocabulary delivered in D4.1. This vocabulary can also be consulted online at http://bib.arts.kuleuven.be/photoVocabulary. Thanks to the work done by KULeuven for the publication of the vocabulary, and NTUA for the skosification of the vocabulary, it also meets the Linked Open Data requirements.[1]

The multilingual mapping tool has been developed by NTUA as a part of the general metadata mapping tool, so partners can work in two steps: first, the mapping of the local metadata schema to the LIDO metadata schema chosen in the course of the first year to serve as an intermediate (see D5.1); second, the mapping of the vocabulary and concepts contained in this metadata to the Europeana Photography vocabulary for purposes of enrichment and multilinguality. A practical manual can be found in point 3.2.2., allowing the partners to choose between three options: mapping of local concepts through MINT, attributing a constant value, or inserting the vocabulary's URIs in their local metadata.

The results of this enrichment process will be included in the tri-monthly narrative reports and progress reports, and be supported and supervised by a "metadata task force" in which the leaders of WP1, 3, 4 and 5 are represented. Further information on this point can be found in 3.3.1.

---

[1] http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/

2   INTRODUCTION

This document provides a description of the software developed for the multilingual enhancement of the metadata to be delivered by the Europeana Photography content partners to Europeana. It describes the context of the development, provides guidance on the use of the software and the added value to both the consortium and Europeana itself.

2.1   BACKGROUND

While the content to be provided to Europeana through the Europeana Photography consortium is fairly uniform in type (photographic materials from the invention of photography until the World Wars), there is a great diversity in subject matter, covering the entire range from local daily life to international public events.

The multilingual mapping tool serves as a way to connect local concepts with the central Europeana Photography vocabulary, as a way to augment and assure quality of the metadata delivered to Europeana, to translate keywords in use by the different partners into 12 languages and enrich the delivered metadata.

At the moment, this mapping tool serves mainly as a way to handle vocabulary mapping to the Europeana Photography vocabulary. However, it has been created in such a way that it allows different mappings, for instance to other available vocabularies.

2.2   ROLE OF THIS DELIVERABLE IN THE PROJECT

This deliverable is part of WP4, in close collaboration with WP5. The multilingual mapping tool is part of the overall MINT mapping tool, D5.2, and makes use of the Europeana Photography vocabulary, D4.1. The mapping tool provides the partners with the possibility for multilingual enrichment of their local metadata upon delivery to Europeana, as well as the uniformisation of their source metadata before delivery if necessary.

2.3   APPROACH

The multilingual mapping tool is part of the overall MINT mapping tool, D5.2, and makes use of the Europeana Photography vocabulary, D4.1. It will be used upon the completion of the mapping from the partners' local metadata schema to the LIDO intermediate metadata schema. At that time, a second mapping will take place, linking partners' local vocabulary to the central Europeana Photography vocabulary.

The results of this mapping and the ensuing enrichment will be collected through the tri-monthly narrative and progress reports. Support will be provided by NTUA and KMKG, as well as the metadata task force (3.3.1) that will come into being from M19 for coordination and support.

3 MULTILINGUAL MAPPING TOOL

3.1 DEVELOPMENT

The multilingual mapping tool has been developed by NTUA as part of the MINT mapping environment. This allows content providers to map not only their metadata schema's to the project's chosen intermediate, LIDO (Lightweight Information Describing Objects), see also D5.1, but at the same time complete a second mapping for the vocabulary.

The manual for the use of the MINT-tool has been provided to all content partners during the course of the project, and assistance is provided by NTUA, KMKG and the coordinators. The vocabulary to which each partner's own concepts will be mapped has been developed in accordance with the partners and was the subject of deliverable D 4.1. It is also available online for consultation and reuse at http://bib.arts.kuleuven.be/photoVocabulary. This deliverable will highlight the type of concepts that will be mapped through MINT, as well as provide a short manual focusing only on the vocabulary mapping. This manual will be based on the general MINT-manual.

For more information about the development of MINT as a whole, and the integration of a vocabulary mapping tool, please refer to D5.2.

3.2 USE

3.2.1 Mapping your metadata fields

During the content seminar, the consortium decided on the mandatory fields for all content partners. While most of these were inspired by the requirements set by Europeana, some mandatory fields were added specifically for the consortium based on standards for photographic cataloguing, to ensure a certain quality upon delivery of the metadata to Europeana, and to present a uniform, searchable collection to anyone interested. This uniformisation also forms the basis for annotation and/or enrichment through the use of an 12-language, hierarchically structured vocabulary that has been developed during the first year of the project, and was the subject of deliverable 4.1.

All content partners have received a training in the use of the MINT mapping tool during a seminar in March 2013, organized by NTUA. During this seminar, the partners have been trained in the general workings of the mapping tool, with a special focus on the mandatory fields as had been decided during the content seminar. At the same time, an introduction to vocabulary mapping was made, which we will expand upon in this deliverable.

For more information on the use of the MINT mapping tool, the mandatory fields for the consortium, the intermediate standard selected or the creation of the Europeana Photography, please read deliverables D 5.2, D2.1, D5.1 and D4.1, respectively.

3.2.2 Mapping your concepts

The first and most important faze of enrichment of the content partners' metadata is the aligning of their photographic techniques and the keywords describing their images with a central vocabulary, created for use in Europeana Photography. Within the project, this will be done for the techniques (LIDO intermediate element lido:eventMethod > lido:conceptID), the photographic practice (lido:eventDescriptionSet > lido:descriptiveNoteID), and the subjects/keywords (lido:subjectConcept > lido:conceptID), who have all been bookmarked in

MINT for ease of use. This vocabulary has been described in D4.1 and can be consulted online at http://bib.arts.kuleuven.be/photoVocabulary. While not strictly necessary for the digitization of photography or delivery to Europeana, this 12-language vocabulary[2] ensures that themes can be traced throughout the different collections. It provides translations for the keywords associated with photographic subjects and techniques, and allows better and more complete search results when searching Europeana for photographic content, providing a tool for research for multiple interested parties.[3]

Following is a practical guide to the use of the central Europeana Photography vocabulary during the enrichment process. It has been based on the MINT mapping tool manual[4], as described in D5.2, as well as the training given by NTUA.

> **Remark**: If the vocabulary of a provider is already skosified, he should follow option 1, and align their in-house vocabulary with the central vocabulary using the corresponding skos properties (broadMatch, narrowMatch, relatedMatch ,closeMAach, exactMatch,) using a tool like the Terminology Management Platform, developed within Linked Heritage.[5]
>
> If the in-house vocabulary is skosified and published according to Linked Open Data principles[6], partners can use option 3, but they would still have to align their terms to the central Europeana Photography vocabulary.

3.2.2.1   Option 1: Mapping of local concepts through MINT

Described as "value mapping" in the manual, this mapping takes place after the metadata mapping has been completed. It is a way to create uniformity in a content partner's own data, and through the use of a central vocabulary, in the data delivered by all content partners.

First, map the relevant metadata field from your dataset to the corresponding LIDO intermediate field by dragging and dropping it in the mapping editor. [7] The LIDO fields with thesauri are clearly marked *unmapped (thesaurus)*. You will see a green *V* sign appear to the left of the LIDO field label.

[2] Negotiations are underway to add a 13th language, Chinese.

[3] Van Steen, N. 2013. *WP4 factsheet*.

[4] Simou, N., Drosopoulos, N. 2012. *D 5.2 The MINT mapping tool.*

5 http://www.linkedheritage.eu/index.php?en/1/home

[6] http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/

[7] In all examples, we will focus on the mapping of photographic techniques. The same logic applies to all other vocabulary mappings.
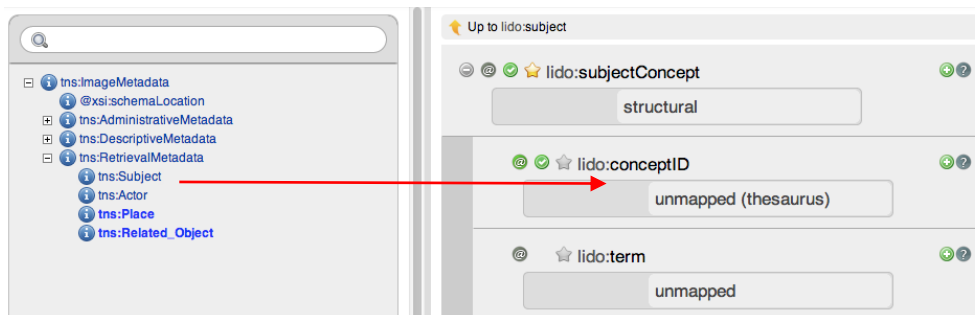
Mapping the techniques from the local metadata to the LIDO intermediate in the mapping editor.

Once the metadata field, containing keywords to be mapped to the central vocabulary, has been mapped to the corresponding LIDO intermediate field,
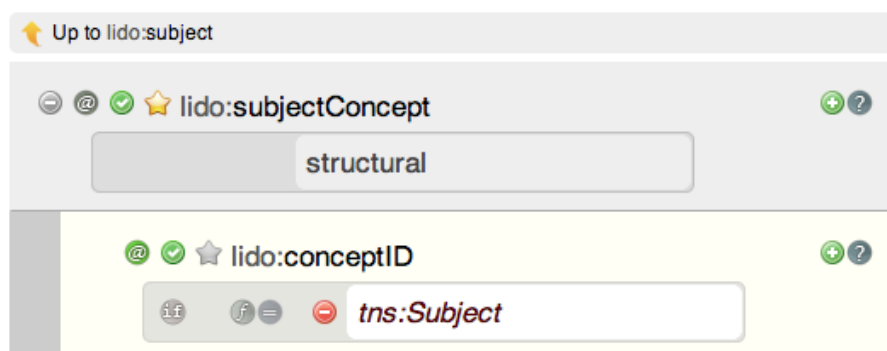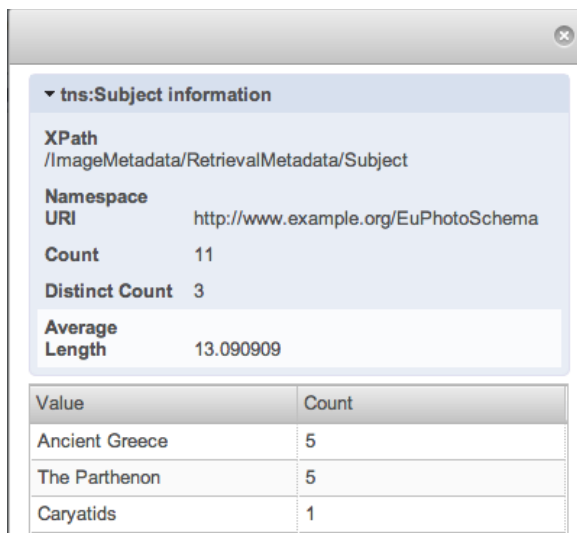

Element Subject from local metadata is mapped to the corresponding LIDO Element.

you can perform a value mapping by clicking the ⊜ icon. A new pane will open on the right hand side of the screen, the value mapping editor.
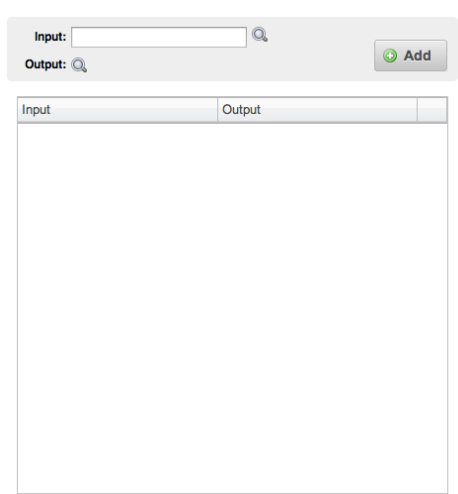
The Value Mapping Editor. On the bottom, some value mappings have already been made.
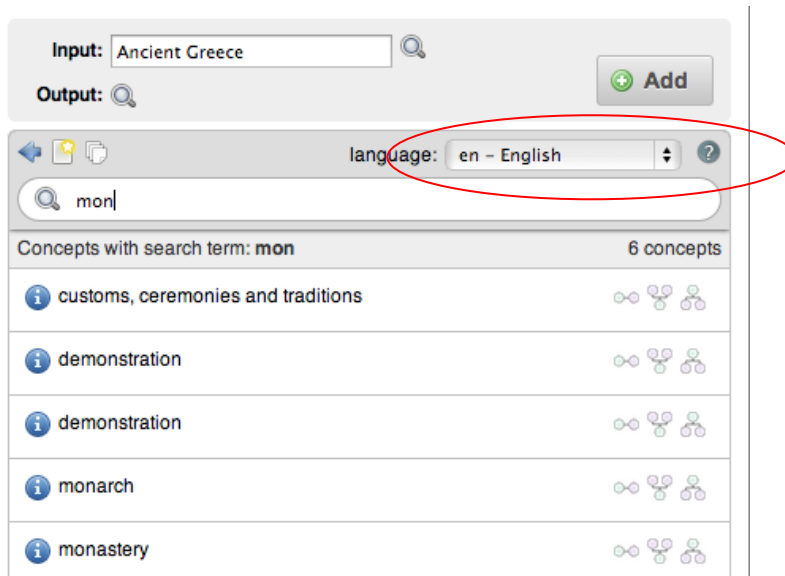
On the top of the pane you specify the input value of the selected element or you can browse values by clicking on 🔍. This will give you a list of concepts available in this specific local metadata field, for which you can create a mapping to the central vocabulary; simply click the desired concept for it to appear in the *Input* field.

The input-side of the value mapping: on the left the local metadata, on the right how often any term appears in the uploaded dataset. Simply click on a local concept for it to appear in the input *field.*

In the same way you can specify the desired output value: click the 🔍 next to output to open the Europeana Photography vocabulary. All languages are available for selection on the top right of the screen.

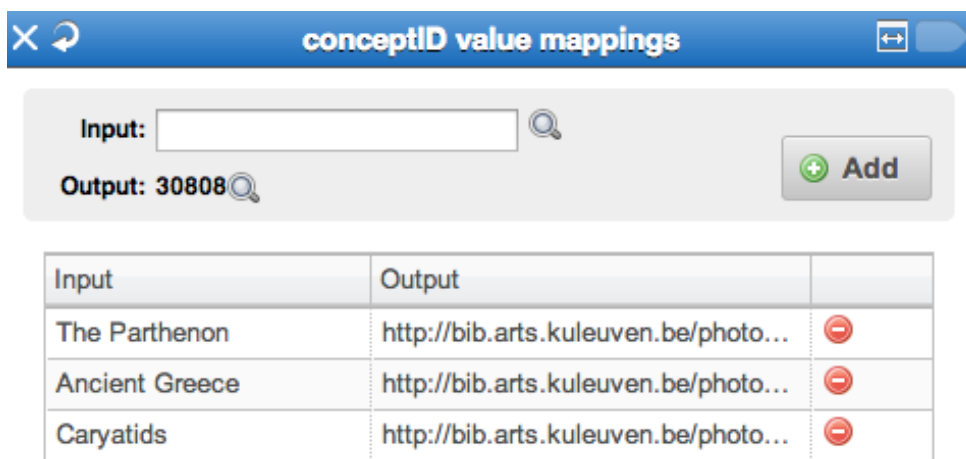*Browsing through the vocabulary in the value mapping editor*

Several methods are available to search for the desired Europeana Photography concept; always choose the most specific concept possible, to ensure quality of the published metadata:

- **Using search terms**: simply type a concept of part of a concept in the reference bar. All result will be shown underneath. This search is case sensitive, so be sure to use capitals where necessary. To check its place in the hierarchy, run through the hierarchy as described below.

- **View all concepts**: by clicking the 🗏 button, all concepts will be shown below in alphabetical order. Here again, run through the hierarchy to check the concept's

location in the vocabulary. To return to the top concepts of the vocabulary, click the ⬜ button.

- **Run through the hierarchy**: since the vocabulary has been hierarchically structured, you can also find a concept by going up and down the different levels. To go down one level in the hierarchy, click the narrower concepts button 🔻. To go back up one level, click the broader concepts button 🔺. The Europeana Photography vocabulary does not use related concepts. If no narrower/broader concepts have been found, return to the previous view with ◀ or go back to the top concepts with ⬜.
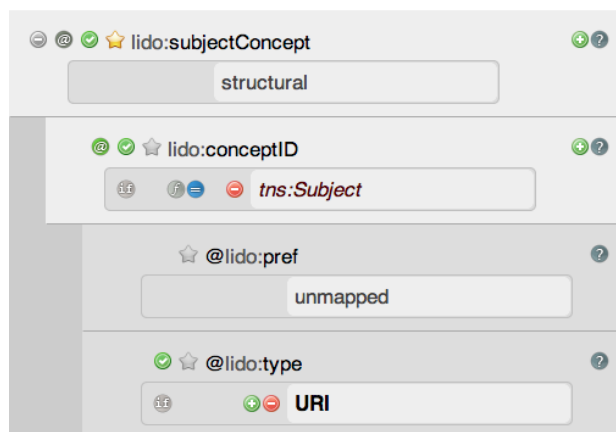
After you've selected the input from your own data, and the output from the central vocabulary, add the value mapping by clicking 🟢 Add and the mapping you've entered appears in the text area below. You can remove a value mapping by clicking on the 🔴 button on the right of it.

Result after mapping the local terminologies to the central vocabulary

Do not forget: when using the Europeana Photography vocabulary, fill in the corresponding lido:type as *URI.* To do this, open the attributes of the LIDO intermediate field with the @, and double click in the field next to lido:type. Type in *URI* and hit enter to save.

Setting the lido:type

3.2.2.2 Option 2: attributing a constant value

If the local metadata to be mapped to one of the three vocabularies is very limited or absent, it may be more interesting to attribute a constant value to all records in a given dataset. The workflow will be similar to the workflow described above, with one important difference: all of your records will have the same concept for this field!

In our example from 3.2.2.1, we have only 1 photographic technique. We can choose to follow option 1 (recommended in case more datasets will be uploaded), or use a standard value.

To attribute a constant value, double click in the field next to the LIDO intermediate label where you want to set a standard value. Do not map any of your local metadata to this field. The vocabulary mapping editor will open to the right, and the field in use will be highlighted with a dotted line.

On the left the mapping editor, with the selected field highlighted with a dotted line; on the right the vocabulary mapping editor from which the standard value will be chosen

The vocabulary will open in a new screen on the right, where you can browse in the same way as described in 3.2.2.1. Click on the required concept, and it will be added as a constant value. In our example, we will set color photography. It will appear in blue in the mapping editor.
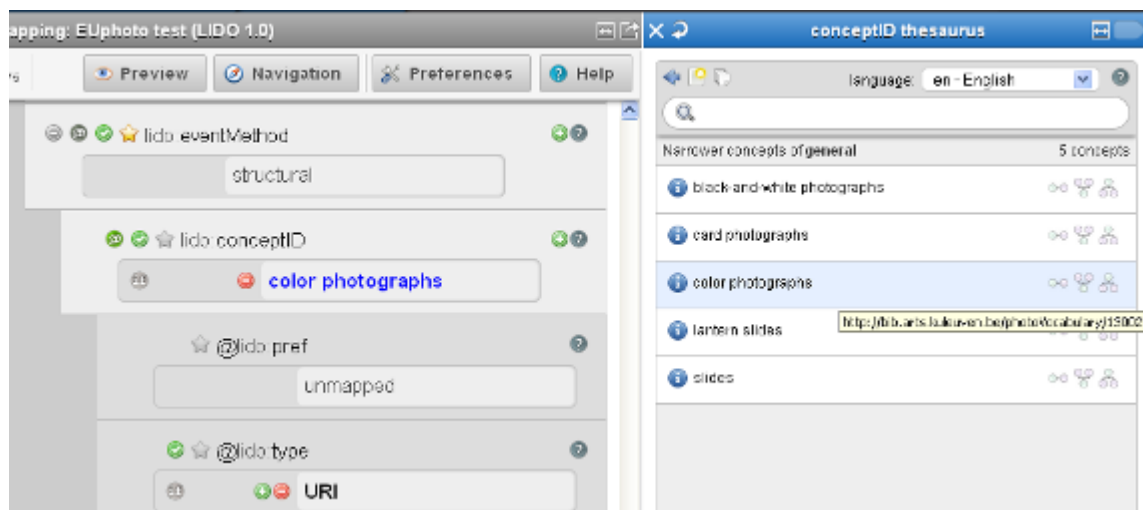
On the left, the mapping editor with the chosen standard value in blue, and the type URI*; on the right, the vocabulary mapping editor.*

Close the vocabulary mapping editor with the ☒. **Do not forget**: when using the Europeana Photography vocabulary, fill in the corresponding lido:type as *URI.* To do this, open the

attributes of the LIDO intermediate field with the @, and double click in the field next to lido:type. Type in *URI* and hit enter to save.

3.2.2.3 Option 3: inserting the URIs in your local metadata

This option requires making changes in your local metadata before uploading your dataset in MINT.

To use this option, you add the URIs from the EuropeanaPhoto vocabulary, as to be consulted on http://bib.arts.kuleuven.be/photoVocabulary in your local metadata, and export it. After upload in MINT, simply map the field for which you have chosen this option to the corresponding LIDO intermediate field by dragging and dropping it in the right location. No further mapping is required.
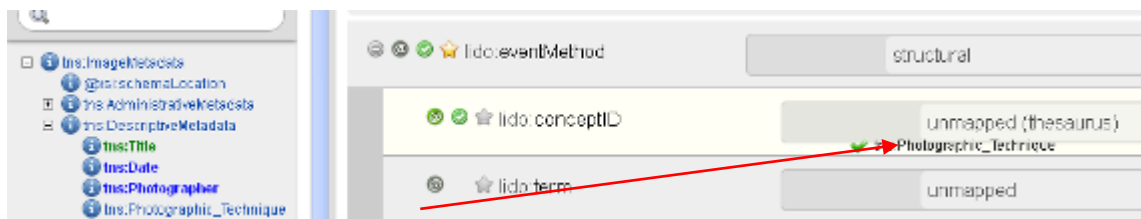
3.3 FUTURE

3.3.1 Metadata task force

At the plenary meeting in September 2013, the Metadata task force will be introduced to the consortium. This task force will serve as an extra way to keep production and delivery of metadata on track, as well as providing support to the partners.

This task force will consist of the leaders of WP1, 3, 4 and 5. Tasks include:

- WP1: monitoring and reporting on metadata production and delivery through the MINT mapping tool; monitoring and reporting on overall progress and intervene if needed; contact with Europeana for the delivery of metadata for publication, and verification of this publication.

- WP3: quality control, ensuring the availability of complete metadata upon delivery to Europeana.

- WP4: integration and support of the vocabularies into the produced metadata

- WP5: technical support using the MINT mapping tool

3.3.2 Possible future collaborations

The Europeana Photography consortium is currently investigating the possibilities of collaborating with the *Daguerreobase* project. For example, a possibility is that Europeana Photography provides the vocabulary created as part of WP4 and described in D4.1 for reuse in the project, and gets the opportunity to further enhance and enrich the vocabulary by allowing mapping of their partners' local concepts to the Europeana Photography vocabulary.

The existing collaboration with Linked Heritage will be deepened, the first tests for the use of the Linked Heritage's Terminology Management Platform are ongoing.

4 CONCLUSION

4.1 RESULTS

The multilingual mapping tool has been taken into production and is functioning as planned. The first tests have been concluded and provided the expected results, namely the enrichment and translation of partners' local keywords through the use of the central Europeana Photography vocabulary. Further results will be monitored by the metadata task force, and reported in the narrative and progress reports.

4.2 IMPACT

The multilingual mapping tool ensures an augmentation of the quality of the metadata delivered to Europeana, mainly through the provision of multilingual data, the uniformisation of themes and concepts, and therefore the enhancement of search results. This means that the whole of the Europeana Photography content will be easier to consult by the majority of interested parties. Furthermore, the interest shown by projects as Dagguereobase and Linked Heritage means possibilities for collaboration and eventually, the possibility of adding more languages to the 12-language vocabulary.