

INDEXING AND MULTILINGUAL SUPPORT

by Nacha Van Steen - Koninklijke Musea voor Kunst en Geschiedenis (KMG)



While all content for EuropeanaPhotography is similar in technique and date, there is an enormous variety in subject matter, ranging from nature photography to press photography from all over Europe. The variety is substantial in most of the content partners' collections, so it stands to reason that the consortium as a whole had to find a way to combine metadata from all partners in order to simplify searches, enrich search results and enscribe the consortium's metadata within the whole of Europeana. This was the task of WP4, dedicated to the enrichment of the local cataloguing metadata, as well as to providing support for multilingual metadata. We've chosen to work on the enrichment of local cataloguing metadata by creating a multi-faceted, hierarchically structured, 12-language vocabulary, covering a large number of subjects, photographic types and photographic techniques. It is based on internationally recognized standards and utilizes the knowledge and expertise present in the content partners' institutions.



Survey of language features

Content providers delivering records to Europeana are required to provide a minimum set of metadata (EDM mandatory elements). EuropeanaPhotography however intends to enrich Europeana by delivering additional good quality metadata by using the LIDO metadata schema as the intermediate format. The survey of language features, conducted in the beginning of the project, aimed at mapping the available metadata in the partners' collections, the availability of the Europeana mandatory fields, as well as export methods and use of a standardized vocabulary.

The results of the survey helped to create a specific, keyword based vocabulary central to the enrichment process, as well as to define the mandatory fields for the EuropeanaPhotography consortium.

The EuropeanaPhotography vocabulary¹

For terminologies in the cultural field, very few standards exist. The Getty Institute has several thesauri, partly translated, and several local, very specific, thesauri and vocabularies do exist, but EuropeanaPhotography tries to capture something at the same time rather trivial, but very pertinent to the project: *what exactly am I looking at?* Using a faceted, hierarchically structured shared thesaurus allows us to address this question in a uniform, transparent way that can be adapted to every partner's need.

The vocabulary defined within this project is a key element of the semantic enrichment of the partners' metadata. While not strictly necessary for the digitization of photographs or delivery to Europeana, this vocabulary 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. The EuropeanaPhotography consortium has chosen to use a modular structure for the thesaurus to keep a good overview regarding the concepts and their arrangement in the hierarchical, relational structure. This structure has been agreed upon at the content seminar in M3, and refined throughout the development of the vocabulary. It consists of a subset of photographic techniques, one of photographic practices (why has a photo been taken?), and a series of keywords in a large number of fields. This vocabulary has been developed based on national and international standards, such as AAT², IPTC headers³ or the Sepiades vocabulary⁴, although a large part is inspired by the collections and themes to be contributed by the partners. This vocabulary has been completed (March 2014) in no less than 15 languages: English (as the pivotal language), French, Dutch, German, Italian, Polish, Danish, Bulgarian, Slovak, Lithuanian, Catalan, Spanish, Chinese, Ukrainian and Russian, and consists of over 500 concepts. Contacts with organizations all over the world, possibly leading to the addition of more languages (e.g. Hebrew translations from The Israel Museum) are currently ongoing. The vocabulary will be used in the annotation, translation and semantic enrichment process taking place at the time of mapping local source metadata schema's to the LIDO metadata schema used for EuropeanaPhotography.

Multilingual enhancement and enrichment

Throughout the remainder of the project period, the vocabulary will be used for semantic enrichment in the description of the photographs delivered to Europeana, by adding the keywords in the mapping stage through the MINT-tool. It will allow users to consult all collections at once with the use of keywords in their own language, and will unify the content from the different partners along thematic lines. Furthermore, it will serve as a way of creating thematic views of the content to be delivered on the EuropeanaPhotography website for dissemination. Enrichment will also take place by promoting the use of GeoNames as a way to unify place names in the content partner's data. They will be able to map, again through MINT, their local place names to the publically available GeoNames site, where maps can be consulted, as well as translations of place names in a large variety of languages. Furthermore, a controlled thesaurus of persons names will be compiled by the project partners and implemented in the project's enrichment process.

Collaborations

EuropeanaPhotography actively seeks to collaborate with other projects for enrichment, augmentation of quality and sustainability, and the sharing of expertise. In this light collaborations with Linked Heritage, EuroPhoto as well as the Digitizing Contemporary Art project have been set up.

¹ <http://www.europeana-photography.eu/index.php?en/115/deliverables>

² <http://www.getty.edu/research/tools/vocabularies/aat/>

³ <http://www.iptc.org/site/Home>

⁴ <http://www.loc.gov/marc/specifications/speccharmac8.html>

Project coordinator
Institute for Cultural Studies – KU Leuven
Fred Truyen
fred.truyen@arts.kuleuven.be

Technical coordinator
Promoter s.r.l.
Antonella Fresa
fresa@promoter.it



www.europeana-photography.eu
www.digitalmeetsculture.net/heritage-showcases/europeana-photography/

DIGITAL CULTURE
www.digitalmeetsculture.net