



**PACKED**

Expertisecentrum Digitaal Erfgoed

# **Persistent identification: supporting digital humanities**

*Abstract for DHBenelux 2014*

Auteur: Alina Saenko

PACKED vzw - Expertisecentrum Digitaal Erfgoed  
Rue Delaunoestraat 58 bus 23  
B-1080 Brussel  
Belgium

e: [alina@packed.be](mailto:alina@packed.be)  
t: +32 (0)2 217 14 05  
t: +32 471 344 691  
w [www.packed.be](http://www.packed.be)

**Keywords:** Persistent identification, open data, semantic web, support of the digital humanities research, bottom-up approach, practical implementation, exchange and enrichment of data.

In September 2013 PACKED vzw initiated a project on *persistent identification* of collection data from ten art collections in Belgium: the Royal Museum of Fine Arts Antwerp (KMSKA), the Museum of Fine Arts in Ghent (MSK Gent), the Groeningemuseum Bruges, the Flemish Art Collection partnership (VKC), the Museum of Contemporary Art in Antwerp (M HKA), the Middelheim Museum, the Municipal Museum for Contemporary Art (S.M.A.K.), the Art Museum for Belgian Art in Ostend (Mu.ZEE), Lukasweb collection and the Art Collection of the Flemish Community.

The goal of the project is (1) to achieve persistent identification of works of art and the core identification metadata and (2) to research the possibilities for retrieval and enrichment. This way, the project endorses the concepts of *Open Data* and *Semantic Web* in the art and cultural heritage sector and enables heritage institutions to provoke new ideas for digital humanities research inside and outside their organisation.

*Semantic Web* projects, such as DBpedia and Europeana, often use a top-down approach when implementing persistent URI's, i.e. harvesting legacy data from institutions, integrating them in a database and identifying core entities by creating persistent URI's from scratch. The project initiated by PACKED vzw pursues a bottom-up approach, where museums use as much as possible existing metadata and self-controlled domain names to identify works of art, data records and digital representations, while including the persistence of these URI's in the operational strategy of the institutions. Using external authorities such as VIAF<sup>1</sup> and AAT<sup>2</sup> ensures the persistent identification of other entities in the identification metadata, such as agents and concept.

From a practical point of view, the project deals with the following research questions:

- What is the status as to the normalisation and identification of the collection data from the participating collections?
- How to distinguish between different data records and digital representations of the same work of art, but from different institutions?
- What is the sustainable syntax for a persistent URI?
- Which external authorities should be used to identify agents and concepts?
- Which tools do museums require to manage persistent URI's?
- To which extent does the use of persistent URI's improve the retrieval and enrichment of collection data?

From a technical point of view, the project explores the use of open source tools such as Open Refine, Solr, Lucene to normalise, integrate and index collection data and how these tools might become part of the toolbox of a collection manager.

The project aims at empowering museums to take responsibility for managing the online access to collection data and encourage them to step forward as an online publisher of their own knowledge base.

---

<sup>1</sup> <http://viaf.org>

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