



UNICUM: a Portal to Dutch Academic Heritage

Henriette Reerink

University Library Amsterdam, Netherlands,
henriette.reerink@uva.nl

Abstract

The UNICUM development project, commissioned by the Academic Heritage Foundation (SAE), is being carried out by the five classic Dutch universities in 2010–11. UNICUM, short for ‘University Collections and University Museums’, has received a national government grant to create a digital portal to Dutch academic heritage. The portal will present both academic archives and museum and library collections. Images, collection metadata and items can be found on one site. The UNICUM idea is inspired by the Online Archive California.

The project is important because it crosses the traditional sector boundaries between museums, libraries and archives, it creates awareness of the opportunities this cross-sectoral approach offers, and it retains the context of — and the relation between — objects within collections as a whole. Moreover, the joint effort brings to light the importance of creating metadata according to international standards to stimulate re-use and exchange of content. In addition, UNICUM has to be regarded as a technical project in which multi-level descriptions will be presented and browsed in a structured way (collections linked to objects and archives linked to separate documents). After the project is finished, the focus will shift to generating content.

UNICUM aims to create structured and integrated access to academic heritage by:

- using international standards (CCO, CDWA Lite, Dublin Core) to stimulate exchange of metadata
- examining the potential of EAD as an exchange standard for (non-archival) collections and for structuring related items
- determining a common method and creating an input module for registration at collection level
- formulating ‘Best Practice Guidelines’ for registration of material at collection and item level

- choosing available and established thesauri which comprise all aspects of the future content
- developing an integrated format for thematic and highlight descriptions
- focusing on the interrelation between items on the one hand and collections as a whole on the other, and vice versa, an interrelation which tells the story of and gives meaning to cultural heritage
- harvesting of the content of the aggregation by Europeana

Key Words: metadata; cross-sectoral; academic heritage; international standards; portal; harvesting

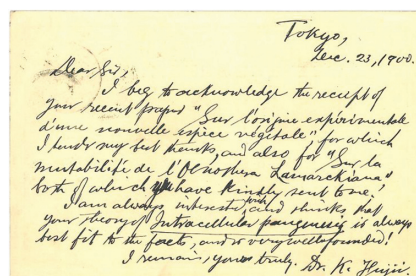
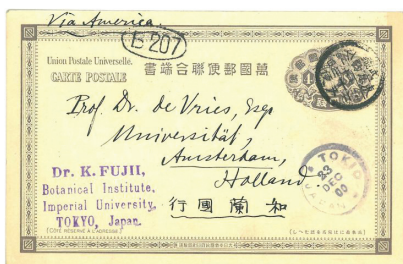
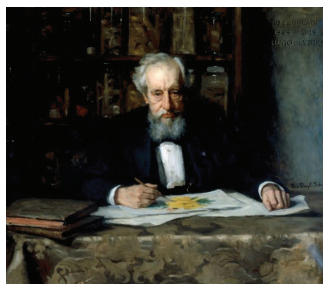
Imagine you are interested in the Dutch biologist Hugo de Vries and his teaching methods. You search for his name in the UNICUM portal, and several hits from different Dutch universities are returned, among them a rough description of his archive, a collection of his educational botanical wall charts (Figure 1), the Dutch translation of Darwin's book *The Origin of Species*, a painted portrait of Hugo de Vries (Figure 2), letters addressed to him by a colleague in Tokyo (Figures 3a,3b), and a collection of papers relating to the Botanical Garden in Amsterdam. If the various collection managers have registered and uploaded their materials, these treasures will appear on your screen and give you an insight into the man Hugo de Vries, his times, and the context in which he worked.

Introduction

UNICUM, which stands for University Collections and University Museums, is intended to create a portal for academic heritage of the Dutch universities. Academic heritage comprises those collections which have historically grown and have come into being to meet the educational and research purposes of the universities. Examples are, for instance, historic microscopes, anatomical models or photographs of university buildings. In addition, faculty archives, paintings belonging to universities, as well as rare book collections are part of the academic collections.

Because of this diversity, the UNICUM portal will present both academic archives and museum and library collections. Crossing the boundaries that traditionally separate museums, libraries and archives is one specific aim of the project. Another feature is that the context of and the relationship between

Figs. 1, 2, 3a, 3b:



objects within the collections as a whole is retained by linking single items to collection descriptions.

As such UNICUM is a metadata project that intends to convert the partners' metadata to common international standards. It is also a technical project in

which multi-level descriptions of a great variety of collections are connected to items and their images.

The Online Archive of California, which was conceived in 2002 and has expanded ever since, inspired the idea for UNICUM. In the Online Archive of California more than 200 Californian cultural heritage institutions present their material at both collection and item level. In the UNICUM project, we will start out with five institutions. Our project was commissioned by the SAE, the Dutch Academic Heritage Foundation, which at the time represented the five classic Dutch universities: Utrecht and Groningen are represented by their university museums and Leiden, Delft and Amsterdam by their university libraries. Recently four more universities have joined the SAE — Maastricht, Eindhoven, Wageningen and the Free University of Amsterdam.

In 2010 the Academic Heritage Foundation received a national grant for innovative development projects. The Dutch Ministry of Education, Culture and Science expects the UNICUM project to contribute to a national portal to Dutch cultural heritage. In addition UNICUM is to disseminate expertise on metadata at collection and item level, as well as expertise on technical infrastructures across the different heritage sectors. The UNICUM portal will be ready in the Spring of 2012.

The project demonstrates that museums and libraries have different types of expertise to bring to the process: libraries tend to have more hands-on experience with information technology and with applying international standards, whereas museums are experts in presenting their material. Both sectors can benefit from the interaction.

The Digital Production Centre (DPC) of the University of Amsterdam is responsible for the technical infrastructure of the project. DPC uses open, international standards and open source software. In the coming years, DPC will host the portal and technically maintain it.

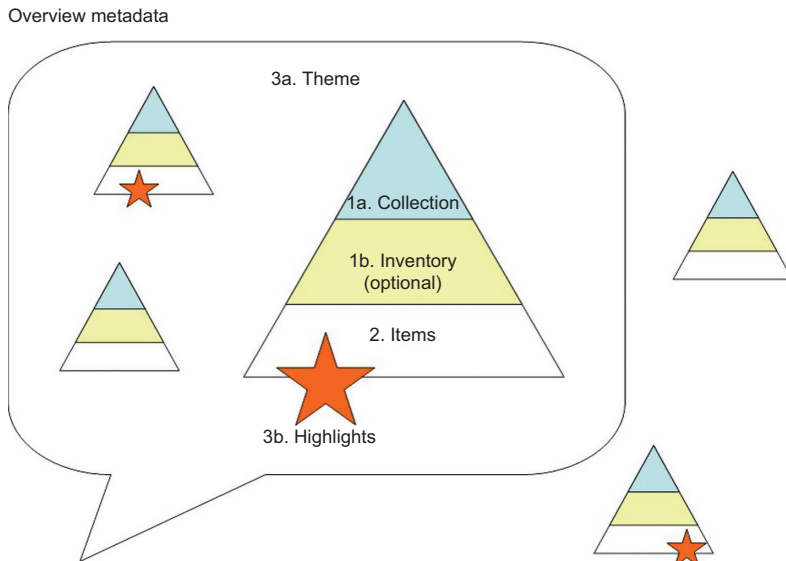
The UNICUM portal is designed to serve the interested general public, the researcher, as well as the collection manager himself. At this stage we are concentrating on developing the project as a whole, and not so much on delivering content or on creating a website with the newest features.

The Portal: Collections, Items and Stories

Our portal will be built around four related components (Figure 4):

- 1a. collection descriptions (museum and library collections)
- 1b. with/without inventories (archives)
2. item descriptions (museum, library and archival objects)
3. stories (3a. themes and 3b. highlights)
4. images.

Fig. 4:



Crossing the boundaries that traditionally separate museums, libraries and archives will be one particular feature of the project. Linking collections with items is another major challenge. Since all universities are faced with time lags in describing their holdings, collection descriptions have been introduced into the portal out of practical necessity. A subsidiary advantage of using collection descriptions is that, with time, items may

be linked to collection descriptions, thus enriching the data with context information.

Collections

Analogous to the Online Archive of California, we use the EAD format to describe both museum and library collections. Best Practice Guidelines for collection registration will be one of the deliverables of the project.

The archival EAD standard is known for its multi-leveled complexity, but DPC has made available a specially designed input module which simplifies the process. This is a tool which can transform the delivered content into EAD xml. Institutions are supplied with a login to make use of the input tool through the Internet. Universities already working with EAD can supply the DPC with their own generated xml and do not have to make use of the input module.

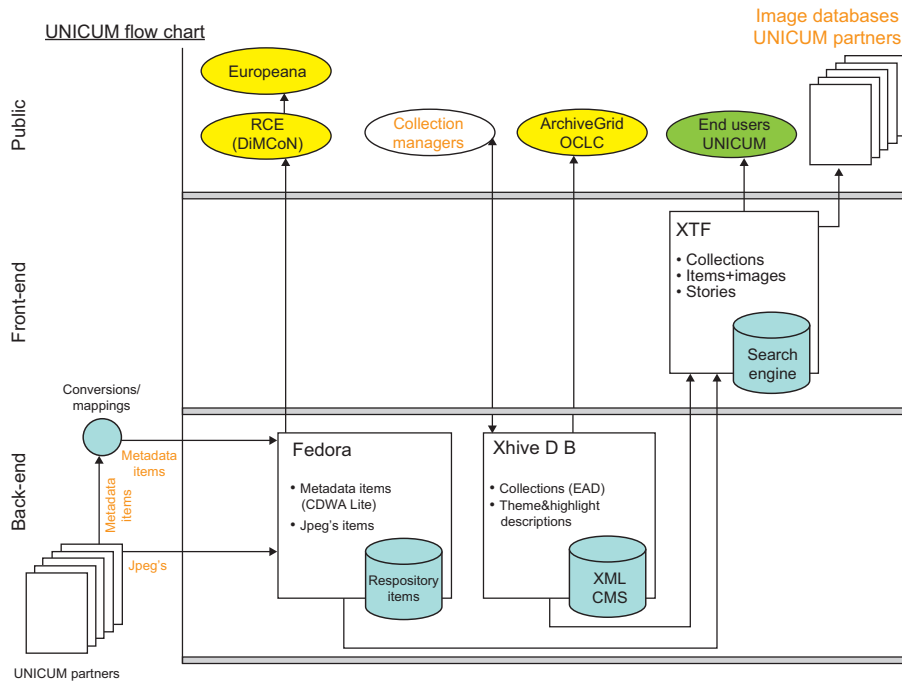
The museums in the project are content with the EAD format, and have already described some of their collections this way. They actually requested a similar format for describing their items, but that was beyond the scope of this project.

Inventories

Multi-layered archives can be very well described by means of the inventory levels offered by EAD. It is possible to link various archival series to a basic, upper-level description, such as, for example 'correspondence of Hugo de Vries and Elizabeth Day Palmer, 1904–1933' to the entry of Hugo de Vries' archive. An online input module for inventories will also be developed within the project. It is expected that only experienced EAD cataloguers will make use of this facility — or they will deliver the xml directly to the DPC.

Single items make up complete collections or archives and, in the end, constitute the core of the portal because of their images. The next section deals with those items.

Fig. 5: UNICUM flow chart.



Items

As shown in Figure 5, the UNICUM partners will recurrently supply DPC with metadata exports of their items to be uploaded into the UNICUM aggregation. This is already quite a challenge, as the UNICUM library partners are the only ones who more or less work with international metadata standards. Standardisation facilitates mapping to the UNICUM aggregation.

The museums have never described their holdings according to international standards. This hampers international exchanges of data.

In some cases organisations did not even use one universal method within their own organization, various collections having been described in mark-

edly different ways. The pitfalls of describing varied materials are well known. It must be emphasized that *consistency* is a key factor here. Even when items have been catalogued incorrectly, as long as this has been done consistently, the errors can be easily corrected. Below are two examples of UNICUM situations where mistakes resulted in data loss:

- At some point in time, a university museum transferred its metadata to another database system. This was not done as carefully as necessary and all distinguished elements were placed in one or two fields in the new database. At the time, no-one was aware of the consequences, and there were no back-up copies. As a result, years of work were lost.
- When trying to obtain a dump of metadata from a university museum collection, we came across a file published on the Internet that contained the required metadata which differed considerably from the file of metadata extracted from the database. The html file was much richer in data than the database file. It turned out that records were updated in the static html file, instead of in the database.

Within UNICUM we recommend that our partners use the metadata content standard of the Visual Resource Association CCO, which stands for Cataloguing Cultural Objects. The project does not *impose* standards, but tries to persuade the UNICUM partners by highlighting their benefits. The data structure chosen for UNICUM is the recommended data structure of CCO: CDWA Lite of the Getty Museum. All metadata of partner institutions is converted or mapped to this format, the native DPC format for UNICUM. A language field was added to this format, since language is not used as a distinguishing criterion in the museum world, whereas it is an essential prerequisite in the library and archive domains.

Stories

To enliven the portal's future website, the possibility has been created to write stories about special themes or objects. These stories are produced by means of an input form which can contain a limited number of words. In this way universities can work together to create thematic profiles of Dutch academic heritage.

Images

As for the images, large thumbnails of the items are presented to the user in UNICUM. These are linked to the image databases of the universities which own the items. To see the complete picture instead of the thumbnail, the user is directed to the particular website of the owning institution. This also applies to composite objects, such as books, which will only be presented as a single thumbnail within the portal.

Web-based image databases are usually not used by smaller museums. UNICUM offers them the possibility to increase the international visibility of their holdings in a relatively easy and instant way.

Thesauri

The site can be searched by both word and by attributed keywords. Whereas the two university libraries (Amsterdam and Leiden) make use of controlled vocabularies and (inter)national thesauri, most museums work with lists of Dutch key words of their own design. Obviously, such lists are not conducive to international data exchange in these days of globalisation. In UNICUM we plead the case of controlled vocabulary or thesauri to mend this problem. Within UNICUM we chose to make use of the AAI and the NBC, the Dutch Basic Classification which classifies according to academic discipline. Both thesauri contain English and Dutch key words.

International Exchange: ArchiveGrid and Europeana

The collection descriptions in the EAD format will be periodically sent to ArchiveGrid, the OCLC database of archival collection descriptions. For this purpose the abstracts of these descriptions have been translated into English, and the keywords also are submitted in their English version.

As mentioned earlier, the images and metadata of the partner universities are delivered to DPC by means of periodic exports. At the moment, it is not yet possible for either the museums to be harvested or for DPC to harvest data from UNICUM partners. We hope to tackle this issue in a future project.

The object metadata and thumbnails of the aggregated items will be harvested by Europeana. Presently, Europeana is not yet able to process metadata at the collection level. The ongoing European ApeNet project aims at contributing multi-levelled archival descriptions to Europeana.

Europeana does not do business with individual institutions or new portals, but with national aggregators. In the Netherlands, the Ministry of Education, Culture and Science distributed the cultural heritage sector roughly into three Europeana aggregators according to types of material:

1. RCE, the Dutch National Cultural Heritage Service for museum material
2. The National Library of the Netherlands for text material
3. The Netherlands Institute for Sound and Vision for audiovisual material.

UNICUM will deal with the RCE, since the majority of our material can be marked as museum material and, not in the least, because the RCE is already applying a well-functioning tool to convert the UNICUM metadata to the Europeana format. This tool was developed by Delving, the software company which wrote the Europeana software. This, of course, is a major advantage, since Delving knows all the technical ins and outs of Europeana.

The Delving SIP-Creator, as the tool is called, is an open source conversion tool which can be used and adjusted by anyone interested according to his or her needs. If the input is according to international standards, as is the case within UNICUM, a sustainable mapping can easily be created. And if it is not, the tool creates practical out-of-the-box conversions to Europeana.

Issues that still need to be solved concern: copyright on the items, the interval during which UNICUM will be harvested by the RCE, and questions of data ownership. These issues should be solved by a set of contractual agreements, one between the Academic Heritage Foundation and Europeana, and one with the RCE. In addition, there is the risk of uploading the same content to Europeana more than once. Europeana has established a working group to deal with this issue. Lastly, each institution has to decide by which aggrega-

tor it wants to be harvested, because, obviously, institutions will only want to pay for this one time.

Added Value?

What is the added value of the UNICUM portal? One may wonder whether it would not be more practical if the five UNICUM partners would deal directly with the RCE or the National Library. That might be the case if becoming part of Europeana were the only goal of the partners. However, the merits of using standards with regards to metadata are immediately tangible within UNICUM, where collections and items are linked with the use of the EAD format for collection descriptions.

By commissioning this portal the Dutch Academic Heritage Foundation can stimulate the presentation of academic heritage and use UNICUM to create a distinct profile for itself. The portal may also work for collection managers to fine tune their collections. But, more importantly, the university museums who have joined UNICUM will not have to initiate the project by themselves, as we can do it collectively. And that is what we wanted from the start: cooperation to meet the challenges and opportunities of globalisation.

Acknowledgement

With thanks to Klara Ashley Bonder for editing the text.

Web Sites Referred to in the Text

UNICUM <http://academiccollections.wordpress.com/>

Online Archive of California <http://www.oac.cdlib.org/>

SAE (Stichting Academisch Erfgoed), Dutch Foundation for Academic Heritage
<http://www.academischerfgoed.nl/>

Leiden University Library <http://www.library.leiden.edu/>

Universiteitsmuseum Utrecht <http://www.uu.nl/NL/universiteitsmuseum/Pages/default.aspx>

Groningen University Museum <http://www.rug.nl/museum/index>
Amsterdam University Library <http://cf.uba.uva.nl/nl/>
Delft Library <http://www.library.tudelft.nl/>
Digital Production Centre (DPC) http://www.uba.uva.nl/digital_production_centre/home.cfm
Open source software used by DPC (XTF) <http://xtf.cdlib.org/>
EAD <http://www.loc.gov/ead/>
CCO <http://cco.vrafoundation.org/>
CDWA Lite http://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.html
AAT <http://www.aat-ned.nl/>
NBC <http://www.kb.nl/vak/basis/bc04.pdf>
ArchiveGrid <http://archivegrid.org/web/index.jsp>
Europeana <http://www.europeana.eu/portal/>
ApeNet <http://www.apenet.eu/>
Three Dutch Europeana aggregators <http://digitalecollectie.nl/>
RCE <http://www.cultureelerfgoed.nl/>
The National Library of the Netherlands <http://www.kb.nl/index-en.html>
The Netherlands Institute for Sound and Vision <http://instituut.beeldengeluid.nl/index.aspx?ChapterID=8532>
Delving <http://www.delving.eu/>
SIP-Creator <http://vimeo.com/19291418>