

ErfgoedLimburg.be
Integrated cultural heritage information

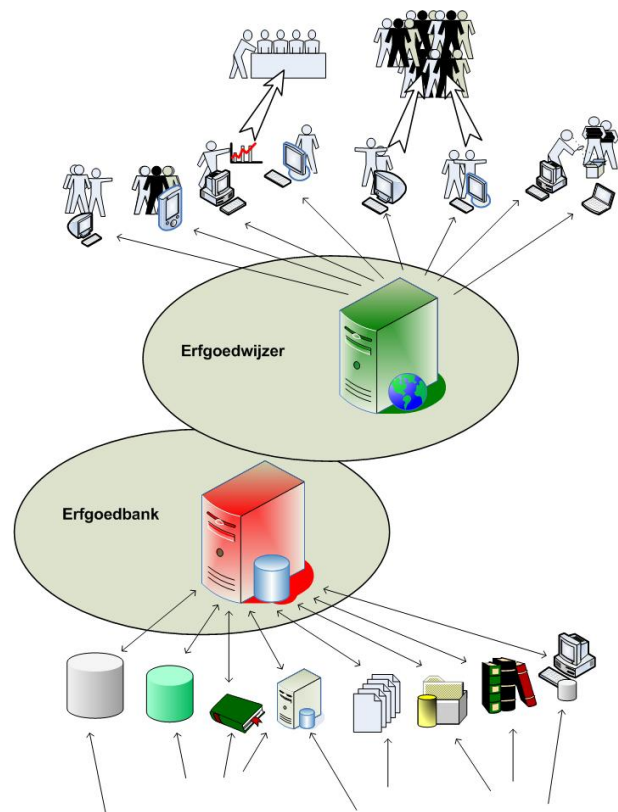
*Author:
Jef Malliet*

CIDOC06
GOTHENBURG
S W E D E N

Brief description of the project

ErfgoedLimburg.be will collect existing information about cultural heritage from a wide variety of inventories (the source databases). The scope aims at all categories of heritage: movable and immovable as well as immaterial heritage, in and from or related to the province of Limburg in Belgium. This information will be merged in a central repository ('*erfgoedbank*' – heritage bank). It will be jointly made available for public consultation, mainly through an Internet web interface ('*erfgoedwijzer*' – heritage index).

Relevant associations can be identified among the described heritage objects, across the source databases, and these will be made explicit through a conversion process. The available information will thus be enhanced, presenting the heritage objects within their context, and reconstructing their cultural historical background. This feature makes the project unique, but it is also one of its major challenges.



Collaboration

The project is implemented by the PCCE – Provincial Centre for Cultural Heritage of the Province of Limburg. The other institutions of the Province administration that are involved with cultural heritage are main partners in the project. Other museums and heritage collections in Limburg will become partners as they participate in the project with their inventories. Important information is available in inventories with a wider scope, such as the inventory of immovable heritage of the Flemish government, and the database of IRPA - the

Royal Institute for the Study and Conservation of Belgium's Artistic Heritage. It is planned to include in ErfgoedLimburg.be also parts of these inventories that are relevant for Limburg.

Target audience

It is anticipated that this kind of information and especially the added value from the collection from so many sources will be of interest to a wide variety of users, for a broad range of purposes. Therefore, the target audience is very broad:

- policy makers and heritage managers,
- academics and technicians who study and preserve cultural heritage,
- the education and tourism sectors, who use cultural heritage as a resource for their own activity and products,
- the general public, interested in learning about their history and environment.

A particular target group is the heritage sector itself, which contributes to the project, and can use it in support of its own activities. Benefits derive as much from the participation and exchange of experience as from the product itself.

Plan and status

The project has started in June 2005. It is financed mainly by the province of Limburg, with contributions from the European Union and from the Flemish government. The construction is planned over three years. Each year roughly coincides with one implementation phase.

- The first year is dedicated to building the *erfgoedbank*, the central repository. This will be done through the conversion of three pilot inventories, all managed by institutions of the province: the inventory of an ethnographic open-air museum, that of an archaeological museum, and a set of about 120 inventories from churches collected in a single database.
- The *erfgoedwijzer*, the public user interface, will be developed in the second year. At the same time three more inventories will be added to the repository.
- In the third year the whole system will be consolidated, and three or four more inventories will be added.

After this, ErfgoedLimburg.be will be operational and ready to accept information from any other inventories that are or become available. It will also be open to expansion beyond the province of Limburg.

Constraints

The source databases are inventories of museums, collections, immovable heritage, archives, libraries, archaeological sites, landscapes, etc. Their heterogeneity presents a number of difficulties:

- they have different contexts,
- they are structured differently,
- they use diverging terminology,
- they have distinct purposes,
- they are made with a variety of systems.

ErfgoedLimburg.be has no intention of replacing the source inventories, nor of imposing or unifying solutions for any of these aspects in the source databases. Their managers should be able to continue their work as they find most suitable. The information from the source databases will be copied and converted to the repository datastructure and authority lists.

Requirements and standards

The end product should have excellent usability and connectivity for information exchange. To support the requirements of the front-end, there are fortunately good tools and standards available:

- it must have a simple interface, easily understandable by non-technical users: the Internet www;
- it must have excellent technical interoperability: XML, RDF, OWL technologies, developed for the 'semantic web';
- it must be presented in a simple data structure, applicable to a wide range of heritage categories: Dublin Core;

- it must use widely agreed reference terminology and a suitable taxonomic hierarchy: the Getty AAT - Art and Architecture Thesaurus, and its translation into Dutch, AAT-Ned;
- it must have a solid semantic structure: CIDOC-CRM.

Field mapping

Conversion from the many different source database formats to the front-end formats is a complex task. Especially the adaptation to the CIDOC-CRM requires considerable interpretation, because the purpose and structure of the source databases have entirely different scopes. It would be very hard to maintain good consistency in the conversion from so many different sources. Therefore we decided to use an interchange format, which we called *Spil* (pivot), to mediate during the conversion. The conversion process can then be split in two more convenient phases: a conversion from the source databases to the *Spil*, and then a conversion from the *Spil* to the CIDOC-CRM. Other formats can easily be derived from the *Spil* or the CIDOC-CRM as appropriate. Conversion from the source to the *Spil* should remain quite linear, limited to a rather simple mapping of fields. Most of the interpretation involved in the conversion would then be in the second phase, which needs to be designed only once.

The format of the *Spil* must be compatible with the variety of contexts of the source information, without significant loss of information. If this can not be achieved with a single structure, then several data models for different kinds of heritage objects must be included. The structure must be similar to that of the source databases, so they can be easily mapped. We selected the data structure of Spectrum (the UK standard for museum documentation) as the basis for the *Spil* structure. (*note*: the Getty CDWA would have been an equally valid alternative, but it is less known among the museum community in Belgium.) Naturally, Spectrum was made for documentation of museum objects, within the context of a single museum. To make it suitable for ErfgoedLimburg.be, where information converges from many museums and contexts, some fields had to be added, and the meaning of some ‘units of information’ had to be slightly adapted. Those ‘units of information’ from Spectrum that are typically for collection administration, were not retained because they are outside the scope of

our database. So far our pilot inventories are from museums or museum-like collections. More verifications and adaptations will be required when we will start working with other inventories, such as for immovable heritage.

In order to facilitate the second phase of the conversion, from the *Spil* to CIDOC-CRM, good structural accuracy is required. To achieve this, we modelled the *Spil* structure using UML class diagrams (Unified Modelling Language). (*note*: while this paper is being written, the second phase of the conversion has not been implemented yet.)

Terminology mapping

All source databases use their own authority lists for controlling data input. When merging information from these sources, a major effort is required for the convergence of the terminology used in many different contexts. Instead of comparing the authorities one by one, we chose to compare them all with one thesaurus. The Getty AAT is a superb tool for this kind of job. It is intended to be a context-independent thesaurus for a wide range of heritage categories, and it is very accurate, mainly due to the systematic description of terms in scope notes.

We do not want to impose the use of the AAT on collection administrators for their inventories. However, we will ask them to relate their own terms to the relevant AAT terms, and to locate them within the AAT structure. We plan to use the AAT with its Dutch translation as the main supporting authority for the search tools.

Not all facets are covered in the AAT, but for those that are missing, there are no resources available that are readily suitable for usage in ErfgoedLimburg.be. The authority lists for these will be built on the basis of the lists currently used in the source inventories. Special attention is required for those subjects that are the main connectors for relationships among described objects: persons and organisations, places, time periods.

Quality improvement

The processes of mapping the field structures and the terminology require a thorough analysis of each of the source inventories. In most cases, it appears to be the first time that such an “audit” is performed. All kinds of inconsistencies and errors emerge, which were unexpected at first sight. The reasons are:

- Inventories are primarily considered an administrative requirement, and are currently not much used beyond that;
- Often they are built with special project resources, like subsidies, with the employment of temporary personnel; thus, many people have usually been involved in compiling the information;
- Experience helps to improve the registration process, causing changes in the interpretation of fields and terms;
- The pressure to complete the inventories is very high, and does not allow for looking back and adjusting the previously recorded information, such as for corrections or adaptations to changed interpretations.

These intrinsic issues must be addressed within the source inventories.

There are also a number of quality aspects related to the project, intended as publication of the inventory information. The inventories are normally made for use within the institution and the context of the collection. Terminology and concepts have specific meanings, which become ambiguous when confronted with other contexts.

Thus, the merging process meets many unexpected difficulties and requires much work within the source inventories. This work, however, will provoke a considerable quality jump for those inventories and will result in a much better and complete usage of the investments made for registration and documentation.

Organisational issues

The project is generating much interest from the heritage sector in the Province of Limburg, and beyond. The organisational structure provides a platform for partners to participate to

some extent in the process of building the repository. This makes people talk and think a lot about documentation and inventories, and the partners can learn from each other's experience. It will be an extra challenge to organise and maintain the network of heritage actors, but it is expected that the benefits to the heritage sector, and in particular regarding heritage information, will be very large as well.

The scale factor is very important in this context: it is fundamental that all interested partners can participate and feel part of the project, even the smallest ones. The province is therefore the appropriate scale to do this kind of project: a larger scale would exclude the smaller partners, a smaller scale would not generate sufficient information to make the repository and index work properly.