Army Museum collection and the semantic web

The ‘Allied Collections’ project

Everyone is online these days. We share knowledge, information, pictures and videos. Museums are also promoting themselves on the internet in many different ways. Museums are aware of the fact that they should actively communicate with the world outside and become part of the knowledge chain by making their collections available online. This is why the Army Museum has a Hyves page, it uploads photos to Flickr and Facebook and videos to YouTube and it is experimenting with Twitter as well. The eMuseum continues to grow in importance, serving as an interactive knowledge platform as well as providing such services as an online organiser for planning a visit. The museum not only presents itself on its general website www.legermuseum.nl, but also reaches out to new users by making its collection and knowledge accessible in several other places, e.g. in portals and databases. It is important, therefore, that the collection and knowledge on the collection are opened up in such a way that they can be found.

The project

In 2010 the Army Museum launched a project that aimed to offer end users an attractive way to access and search various integrated electronic collections. The name of the project is ‘Allied Collections’, for which Trezorix provides the technical support. The project is subsidised by the government as part of its *Digitaliseren met Beleid* digitisation programme.

This project stems from the fact that the museum has invested a great deal of time and money in digitising large parts of its collection over the last few years. This has given the museum an electronic goldmine of resources which can provide online visitors with a complete overview of Dutch military history.

The project brings together and opens up nine collections:

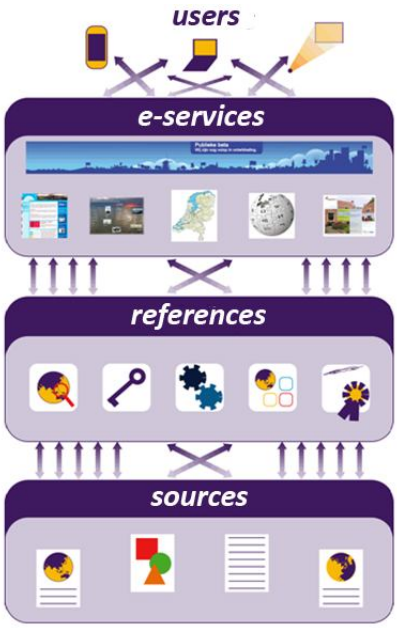
- The so-called Hoefer (the first director of the Army Museum in 1913) catalogue with approximately 11,000 cards with entries, descriptions, and often also depictions of military terms from the beginning of the previous century.
- The 1861 Landolt military dictionary with approximately 3,600 entries and descriptions.

- The Glossarium Armorum, an illustrated thesaurus with approximately 700 terms and depictions of armour (defensive weapons) in several languages.
- Articles from the Army Museum's Armamentaria year book about various military-historical subjects and the museum's collection.
- The famous 'Visser collection' of 175 historical hand-guns.
- A thesaurus on uniforms with links to parts of the collection with more than 900 entries and scope notes.
- 9,000 prints of uniforms of various armed services with descriptions from the collection of books on uniforms from the museum's military history library.
- A thesaurus of the Dutch armed services.
- Military vehicles with 90 descriptions and depictions.
- Official orders before 1940, approximately 1,200.

Until recently, only a small portion of this digitised material was available online. It was also not very easy to visualise the links between the various collections. The Allied Collections project changes this by collating data that is relevant to findability in an online findability layer.

The architecture

This findability layer is part of an infrastructure that is based on what is known as RNA architecture. RNA stands for Reference Networks Architecture, and is a web-based architecture that seeks to link sources of knowledge to or via the web and make the various types of content in those sources findable in a simple and straightforward manner.



Picture 1 shows the RNA architecture according to the Cultural Heritage Agency. The Army Museum has applied this set-up for constructing an extensive semantic network for military cultural heritage.

Three layers have been differentiated:

There is a *data layer* at the bottom, which consists of databases, file systems, web pages and other sources. This data is available via the internet. This layer represents the information offering.

The top layer is the one with *applications* or e-*services* that want to utilise the information offering in the bottom layer. Demand is determined in this application layer. End users use the applications to access the data.

A third layer between these two layers links them together. This layer contains the *network of references* and consists of linked metadata from the data layer and hierarchically arranged structures in which this metadata is given a place. This is the actual findability layer or reference layer. It serves as a sort of broker between the demand and supply sides, by linking high-quality access functionality to a detailed overview of the available data.

The working method

The Teylers Museum in Haarlem, the Netherlands' oldest museum, carried out a similar project. In that context, the challenge facing the Army Museum was aptly described by Marco Streefkerk of Digital Heritage Netherlands: 'With the problem of five collections, each with their own knowledge domain, laid down in knowledge systems with differing data structures and managed in collection administration systems with different software and suppliers, the museum represents a miniature version of the heritage universe. Traditionally,

we resolve that by exporting all the different metadata to one uniform format (generally XML) and converting the content (mapping) to the greatest common denominator, mostly a variation of Dublin Core. The museum has skipped this traditional approach and chosen directly for the semantic web.'

The same approach applies to the Army Museums' nine collections. To collate this miscellaneous data in the findability layer, a number of things have to happen. In short, data from the collection that is important for findability in the middle layer is brought together in the form of small sets called *content items*. A content item can refer to objects from the collection, to people, types of material, historical periods, etc. These content items are laid down in a computer readable format (RDF) and linked together via their properties. A content item of the 'book'-type can be linked via the property 'has author' to a particular person, i.e. the author.

Clean and standard data is required to be able to connect the content items. Museum staff have spent a great deal of time on the project aligning the data. Collaborative projects in the field of terminology are underway with other museums such as the Rijksmuseum in Amsterdam and the Royal Museum of the Armed Forces and of Military History in Brussels (Belgium).

Language technology

Standard tools such as MS Excel were used during the project as well as more bespoke specialised tools based on language technology, for example.
Language technology was used because the museum has many old text materials and catalogues with fairly diverse differences in language and spelling. *Natural language processing* was used to actually be able to efficiently integrate this content into the findability layer.

Facet determination

The museum is also introducing a new technology called facet determination to respond to the growing demand for military genealogy. This will enable visitors to independently identify uniforms in family portraits in a straightforward manner by excluding the characteristics in uniforms and armed services and in doing so determine in which of the armed services their ancestors served.

The perspective

The Army Museum is currently designing a separate website to host the 'Allied Collections', www.alliedcollection.org, which includes new technologies such as social media, where users collaborate via the internet. Visitors can tag objects with information or add comments and can also collect and save favourite objects together.
At this point in time, the website still only contains the Army Museum's collection, but this should change in the future. The technology is ready for the future. The Army Museum's data can easily be exchanged with other institutes. Given the multilingual terms employed, some collections lend themselves perfectly for comparison with and linking to collections from foreign heritage institutes.

This project is a *best practice* project for setting up a flexible online information environment that other organisations can easily link into and for integrating miscellaneous collections. By offering the collection in these new ways, the museum is seeking to prepare for the future. As Dirk Houtgraaf, Section Head of Knowledge Exchange at the Cultural Heritage Agency so nicely put it: 'We will soon no longer be a museum with a network, but a network with a museum'. The digital offering of knowledge and information on the collection, i.e. the museum's electronic platform, will become more and more important for the museums' future existence.

If, after reading this article, you still have questions or are interested in possibly participating in this project, please contact Annet Ruseler, Head of Collection Information, Army Museum, Delft, the Netherlands (a.ruseler@legermuseum.nl).

Annet Ruseler, Army Museum, Delft
Hans Nederbragt, Trezorix, Delft