

HERON VISUALISATION ENGINE.
VISUALISATION AND DISSEMINATION
OF SEMANTIC CULTURAL HERITAGE DATA

Over the past two decades, there has been a proliferation of software to create great 3D models of archaeological sites and objects, and there has been plenty of thought and discussion on data models for finds. The results of those exertions have been made public through institutional websites and specific portals, but now, a further step is necessary: the cultural heritage data and (meta)data need to be taken into the semantic web.

The *Heritage Online Visualisation Engine* (Heron VE) was born to do this. It provides tools for documenting, visualising and disseminating the semantic relations between sites, objects, documentation and narratives. Heron VE is scalable: it can be used to tell the story of a particular archaeological complex, but it can also illustrate relations between sites and objects which are widely separated in both time and space. With Heron VE, dissemination of (meta)data can take many different forms: Heron VE contains modules for presenting and reporting on data, but it can also be used to provide data only.

Heron VE was born in the early 1990s, when the author was attending the University in Utrecht. Working with both relational databases and SGML (e.g. TEI, the Text Encoding Initiative), the author developed a taste for flexible data models based on standards.

A significant step towards Heron VE was his work on *Digitale Kanalen – het water van Amsterdam* (Museumserver.nl 2001), a website recounting the story of water – highlighting, amongst other things, beer, shipping and flood defenses – in the history of Amsterdam. *Digitale Kanalen* combines digitized objects, images and videos from local archives and museums with several storylines created for the project. It is entirely metadata driven: everything a visitor sees is determined by metadata searches generated from either the graphical interface (flash) or hyperlinks.

The metadata model for *Digitale kanalen* is extremely simple. Objects and (fragments of) stories could be linked to themes, places and periods in time (expressed in a subset of the Dublin Core metadata standard).

Another contributing factor was the work done by the author on metadata standards for archaeology (for the Department of Archaeology and Building History of the city of 's-Hertogenbosch), archives and libraries, and an early interest in the semantic web. Here he was repeatedly confronted by the question of how to “translate” metadata from local systems, where context provides meaning and authorship is usually clear, to the World Wide Web, where these metadata may exist without any context, where specific languages

and character sets cannot be taken for granted and where it can only rely on automated processing to be found (as opposed to browsing by eye).

In the end, the World Wide Web consortium provided the solution as part of its work on RDF: semantic triples (*Semantic triple* 2016). Semantic triples capture the relationship (predicate) between two “things” (subject and object) in the form subject | predicate | object: Kainua 2017 (conference) | is held in | Bologna.

Below, one aspect of the historical town of Kainua has been captured in a triple as it is stored in Heron VE. NB: the examples below are simplified. For example, language codes for string data have been omitted, although they are present in the actual triples in Heron VE: Kainua | belongs to culture | Etruscan.

However, on the WWW text strings are a dangerous way to clearly identify things or relations. For example, in Dutch “Etruscan” would be “Etruskisch”, and where are we supposed to find out what the author of the triple meant exactly with “belongs to culture”? Therefore, our triple needs to be transformed into: <https://it.wikipedia.org/wiki/Kainua> | heronve:belongs_to_culture | <http://vocab.getty.edu/aat/300020471>.

Now, identifiers have been substituted for text. *Kainua* is now explained as “Kainua as described in the Wikipedia article”, the relation “belongs to culture” has been clarified as belonging to a namespace “heronve” (where we can direct any enquiries), and apparently, the triple uses “Etruscan” as defined by the Getty Art and Architecture Thesaurus.

For presentational purposes, a string of text and/or an image will of course be substituted for <https://it.wikipedia.org/wiki/Kainua>, but as long as everyone who wants to indicate that something is Etruscan uses the Getty-identifier, we can be sure that this information can be automatically linked. If several authority files are available – e.g. for people – the identifiers in the different authority files should link to each other.

Although authoritative lists of identifiers have become available for many things, a lot is still lacking. For example: the development of such an “authority file” for historical events is part of the work on Heron VE, and in the Netherlands, the author has just finished to work on an authority file for Dutch castles and country houses.

Heron VE distinguishes between five classes of “things”: object, actor, event, location and concept. Registering class could be done in the form of another triple (e.g. <https://it.wikipedia.org/wiki/Kainua> | heronve:has_class | object), but for Heron VE, we chose to create “Type-Strong” semantic nodes, where the class is already recorded as attribute of the “thing” itself (*IF4IT* 2017).

In some cases, a triple may have to be qualified. For <https://it.wikipedia.org/wiki/Kainua> | heronve:has_name | Kainua, we may want to add that this

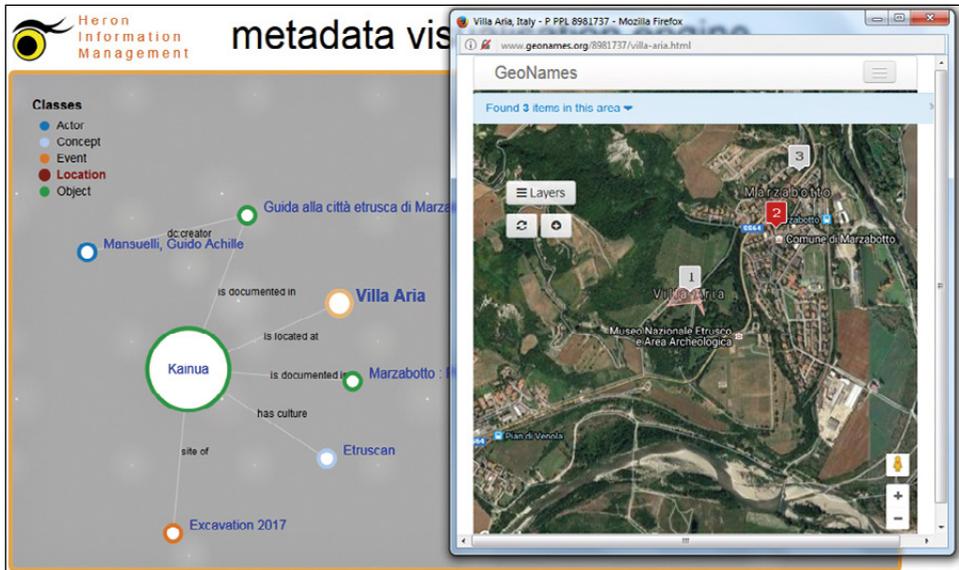


Fig. 1 – Heron VE collections called “locations”, which can be expanded.

name was only applied at a certain time: <triple ‘has name’> | heronve:valid_from_to | 550BCE-350BCE.

It might also be wise to register that the attribution of the name is not entirely certain: <triple ‘has name’> | heronve:certainty | possible.

Finally, in Heron VE, all triples can be annotated: <triple ‘certainty’> | heronve:has_annotation | “Name taken from locative in inscription found at the site”.

NB: it should be clear that those meta-triples regarding annotations and qualifications, however useful they are while working in Heron VE, are not part of the actual triples, and other environments with which triples are exchanged may not know what to do with these meta-triples.

With the arrival of the semantic triple, the basic problem of capturing metadata from many different sources in an unambiguous and entirely machine readable way has been solved. However, this does not mean that working with triples is easy. Triples are traditionally generated from data stored in other formats and then made available for processing by applications and specialized search engines. However, triples could be more: if managed carefully, they could be datasets in their own right, and on-line “editable” triple stores would be a great means of combining efforts of many contributors.

Heron VE is an attempt at such an editable triple store. It gets its triples either by transforming existing metadata records or by direct input. The

engine is able to transform metadata from a number of common standards (e.g. Dublin Core, MODS or the SPECTRUM data model) with no or minimal configuration. Mappings for other standards can be added with ease. For direct input, Heron VE provides a graphical interface.

Triples stored in Heron VE can be used in two different ways. On the one hand, it is possible to export them as building blocks for dedicated portals and services (the traditional approach mentioned above). For this, several serialization formats are supported (RDF/XML, Notation-3 (N3), Turtle, N-Triples, RDFa, and RDF/JSON). It is, however, also possible to browse, search and view the triples in the engine itself. As the V in the name suggests, this was one of the main reasons to create the engine to begin with. By default, all relations from a chosen “thing” to start with are shown, but they may be aggregated by class. For example, while viewing the concept “Etruscan”, the many connected places and people will not all be shown individually immediately: Heron VE will show collections called “people” and “locations”, which can then be expanded (Fig. 1). Filtering by class is possible, and the meta-triples for annotations and qualifications can be switched on or off. Finally, several reports are available, to provide information on either the triples regarding one “thing” or all triples per class or per authority file.

Heron VE is not a commercial development, although lessons learned from its development have been used by the author in projects for Heron LLP. Input and even collaboration is welcome. Developments can be followed at <http://heronim.co.uk/en/projects/ve/index.php>.

MARTIJN VAN DER KAAIJ

Heron Information Management LLP, The Netherlands
martijn@heronim.eu

WEB REFERENCES

- Semantic triple* 2016, https://en.wikipedia.org/wiki/Semantic_triple (accessed: 17/04/2017).
IF4IT 2017, *Type-Strong versus Type-Weak Semantic Relationships and Semantic Data Graphs* | *IF4IT* (<https://www.if4it.com/type-strong-versus-type-weak-semantic-relationships-semantic-data-graphs/>; accessed: 17/04/2017).
Museumserver.nl. 2001, *DIGITALE KANALEN* (<http://www.museumserver.nl/kanalen/>; accessed: 17/04/2017).

ABSTRACT

Over the past two decades, there has been a proliferation of software to create great 3D models of archaeological sites and objects, and there has been plenty of thought and discussion on data models for finds. The results of those efforts have been made public through institutional websites and specific portals, but now, a further step is necessary: the cultural heritage data and (meta)data need to be taken into the semantic web. The *Heritage Online Visualisation Engine* (further: Heron VE) provides tools for documenting, visualising and disseminating the semantic

relations between sites, objects, documentation and narratives. Heron VE is scalable: it can be used to tell the story of a particular archaeological complex, but it can also illustrate relations between sites and objects which are widely separated in both time and space. With Heron VE, dissemination of (meta)data can take many different forms: Heron VE contains modules for presenting and reporting on data, but it can also be used to provide data only, for example in several XML-formats or in N-triples. The designer of Heron VE has 25 years of experience with cultural heritage data in the field of archaeology, museums and libraries and has been working on the structured dissemination of cultural heritage data, first in the semantic web and now within the framework of linked data. This paper will illustrate the journey towards Heron VE, including considerations regarding the adoption, adaptation or rejection of existing data models and ontologies. It will also contain examples. These will mainly be based on data regarding sites of former castles and stately homes in the Netherlands, but it will become very clear that the Heron VE can be applied to many different cultural heritage datasets, including those regarding ancient urban areas.

