Learned Love

Proceedings of the Emblem Project Utrecht Conference on Dutch Love Emblems and the Internet (November 2006)

Edited by Els Stronks and Peter Boot, assisted by Dagmar Stiebral

DANS Symposium Publications 1 The Hague, 2007

Contents

The Dutch love emblem on the Internet: an introduction – Els Stronks and Peter Boot 1

PART 1 THE DUTCH LOVE EMBLEM

Creator of the earliest collection of love emblems? – Alison Saunders 13 Commonplaces of Catholic love – Arnoud Visser 33 Encoding the emblematic tradition of love – Marc van Vaeck 49 Churches as indicators of a larger phenomenon – Els Stronks 73 The Spanish epigrams in Vaenius's *Amoris divini emblemata* – Sagrario López Poza 93 Love emblems and a web of intertextuality – Jan Bloemendal 111 The *Ambacht van Cupido* from 1615 in Wroclaw (Poland) – Stefan Kiedron and Joanna Skubisz 119 Investing in your relationship – Arie Jan Gelderblom 131

The love emblem applied – Peter Boot 143

PART 2 THE DIGITISATION OF THE EMBLEM

The Emblem Project Utrecht as a knowledge site – Els Stronks 151 Traditional editorial standards and the digital edition – Edward Vanhoutte 157 The technical backbone of the Emblem Project Utrecht – Johan Tilstra 175 Digitising Dutch love emblems – Peter M. Daly 183 Setting the emblem schema to work – Thomas Stäcker 201 Mesotext. Framing and exploring annotations – Peter Boot 211

Colour plates

The technical backbone of the Emblem Project Utrecht

Johan Tilstra, De Tekstverwerkers/Emblem Project Utrecht

This paper is concerned with the technical backbone of the Emblem Project Utrecht (EPU). The backbone – the computer hardware and software – allows for the realisation of the basic aim of the project: the digitisation of about 25 Dutch love emblem books, from Daniël Heinsius's *Quaeris quid sit amor* (circa 1601) to Jan Suderman's *De godlievende ziel* (1724). The second aim of this project is no less important than the first: to reflect, based on practical experience, on the consequences of the digitisation of historical literary material. These consequences range from changes in how this material can be made accessible to, presumably, fundamental changes in scholarly research.

In a digitisation project like the EPU, the technical details of the digitisation process are tightly intertwined with the consequences of that process for the scholar using the digitised editions of the emblem books. These consequences in particular, and the broader consequences of digitisation of historical material in general, cannot be surveyed without a certain insight into the technical details behind digitisation projects. In this paper I will go into the technical setup used for the EPU, trying to explain some of the design choices made and elaborate on the practical and scholarly consequences of those choices.

Encoding seventeenth-century emblem books

At the start of the project, Peter Boot (co-creator of the project) argued for the use of public standards in order to ensure the long-term accessibility and usability of the digital material created by the EPU. In 'Accessing emblems using XML: Digitisation Practice at the Emblems Project Utrecht' (Boot 2004), Boot elaborates on the encoding techniques used for the EPU: XML for the syntax of the encoding of tags, and the *TEI Guidelines* for the vocabulary of the tags and their implementation. Using a few well-chosen examples, Boot describes how in a plain-text file, containing the transcription of one emblem in one of the emblem books of the EPU corpus, specific labels are added to specific pieces of information. These labels, or 'tags', follow the XML-syntax. They are therefore enclosed by angle brackets ('<' and '>'), have a specific, descriptive name and may contain so-called 'attributes'. A single piece of information enclosed by such tags constitutes a single XML-element, as shown below:

<title>Amoris divini et humani antipathia</title>

XML supports the nesting of elements. Consider the line group (starting with

the **<1g>**-tag and ending with the **</1g>**-tag) consisting of multiple lines (enclosed by the **<1>**- and **</1>**-tags), and the XML encoding of the emblem text will look like this:

What the encoding looks like – e.g. the angle brackets, the opening and closing tag combination – is governed by the XML specification. What is not specified by XML, is *which* particular tags can be used. This choice is up to the user. This is why XML is called a metalanguage: it is a language for describing nothing more than the syntax of an encoding language – the *vocabulary* of the encoding language is outside of its scope.

One can easily see that in that freedom there is potential chaos. In the example above, for the encoding of a group of lines there is no real argument for using the <lg>-tag over, for instance, the (hypothetical) <linegroup>-tag. And, while we are at it, why not come up with the better readable <LineGroup>-tag? Indeed, one could devise a complete set of custom tags, designed specifically for a single, one-time-only digitisation project. However, such a set probably would not be of much use for anything or anyone else, and would thus raise all sorts of compatibility issues.

This is where the *Guidelines* from the Text Encoding Initiative Consortium $(TEI-C)^1$ come in place. The TEI-C proposes a vocabulary for the digitisation of literary and linguistic material, together with extensive guidelines on how to use and implement it. The TEI-C's goals are to provide a vocabulary that is extensive enough to be of use to the majority – if not all – of digitisation projects in the humanities. In other words: with the TEI vocabulary, we are able to use the same 'words' to designate our segments of transcribed textual material – we are able to speak the same language. With such a common language across multiple humanities digitisation projects, we have passed the first hurdle (isolation), and are moving towards interoperability and interconnectivity.

Splitting the emblem

In digitisation projects the systematical approach that the computer demands bestows a certain level of rigidity or formality on the age-old practice of text annota-

```
1 http://www.tei-c.org/
```

tion – enriching text by adding 'metatext'. In *A Companion to Digital Humanities*, Allen H. Renear provides a clear overview of the emergence and the advantages of the practice of adding tags to the constituents of a text – the so-called 'descriptive markup approach'. These advantages are so many and so diverse (Renear provides an extensive list) that some people have suggested that, as Renear writes, 'it was not simply a handy way of working with text, but that it was rather in some sense deeply, profoundly, *correct*'. It leads to the particularly cogent model of text as an 'Ordered Hierarchy of Content Objects':

The model in question postulates that text consists of objects of a certain sort, structured in a certain way. The nature of the objects is best suggested by example and contrast. They are chapters, sections, paragraphs, titles, extracts, equations, examples, acts, scenes, stage directions, stanzas, (verse) lines, and so on. But they are not things like pages, columns, (typographical) lines, font shifts, vertical spacing, horizontal spacing, and so on.²

Equal pragmatism was needed for the digitisation of the EPU emblems: what details form an integral part of an emblem (according to current emblem theory, that is), and thus need to be encoded? What details, on the other hand, are of less or even no interest to emblem scholars? Renear continues, providing a provisional, practical standpoint:

'The objects indicated by descriptive markup have an intrinsic direct connection with the intellectual content of the text; they are the underlying 'logical' objects, components that get their identity directly from their role in carrying out and organizing communicative intention. The structural arrangement of these 'content objects' seems to be hierarchical – they nest in one another without overlap. Finally, they obviously also have a linear order as well: if a section contains three paragraphs, the first paragraph precedes the second, which in turn precedes the third'.³ (Renear 2004)

The EPU more or less followed this practice. All of the emblem books in its corpus are encoded using XML and the TEI vocabulary. The XML/TEI-combination provides for independence of the underlying technical platform used. Its widespread use and active user base combine the greatest number of possible uses with the best chances of having the encoded data survive the inevitable technological advances that time will bring. With the 'descriptive markup approach' also comes the perspective of texts – emblem books, in the case of the EPU – as the above mentioned 'ordered hierarchies of content objects'. Each emblem book is divided into texts (marked by 'text'-tags) – one text for each emblem. Further refinement of the tagging is done by classifying all discernible emblem elements (its *motto, pictura, subscriptio*, quotations, et cetera) and tagging each element as such ('div', for 'division'). The TEI guidelines provide the universal vocabulary for these specific elements. The data file that follows from the tagging process contains the

² Renear 2004.

³ Renear 2004.

original transcription of the emblem book, enhanced with the XML-tags that surround the pieces of text that form the individual emblems and its mottoes, *picturae*, et cetera. Thus, eventually it contains a hierarchy of XML-elements that represents our classification of the structural elements of the emblems in the original emblem book.

Though the perspective of emblem books as hierarchical structured codices does closely relate to the technical handling of the data in the EPU, it still remains to be seen whether emblem books are really only ordered hierarchies of content objects, and emblems are not more than just 'objects of a certain sort, structured in a certain way? Nevertheless, the strict standard for the XML-tags and, following from that, the separation from the 'language' of the emblem book (the text) and the 'metalanguage' (the XML-tags), allows for a certain level of machine readability of the encoded books. Without the XML-encoding, to a computer the sequence of characters in the transcribed text would be nothing more than just that. Having a computer to 'understand' that sequence of characters would require highly specialised artificial intelligence trying to make sense of the text - more or less like a human being would. With the XML-encoding, we are providing the computer with information on our (human) understanding, interpretation and classification of the structural elements - the hierarchy of content objects - of the emblem books and emblems. Clearly, this further opens the door to computer assisted emblem research.

Separating content from presentation

In day-to-day life, the data files containing the transcribed and encoded emblem books – in short, the XML-files – form the technical core of the EPU. Together, they are nothing more than just a set of files on a web server, one for each emblem book. These files are hosted on a server at the Utrecht University. From the beginning of the project up to the present day, work has been done – and still is being done – by the EPU editors to update, correct, revise, and enhance this core set of data files. Over time, transcription errors have been corrected and metadata like explanations, translations, links to parallels and sources and bibliographical references have been added. In their current state, the data files allow us – among other things – to present on the project's website⁴ faithful representations of all emblem books in the projects corpus.

The website is, of now, the most manifest presentation of the proceedings of the EPU. It is built using a so-called 'web development framework' that allows us, in abstract wording, to 'separate content from presentation'. In practice, it boils down to a rather simple setup. The web pages that make up the website are not simple HTML-pages stored on the server and waiting to be sent to a user requesting them. Instead, the HTML that a visitor of the EPU website is presented with in his/her web browser, is generated dynamically out of the central set of XML-files.

⁴ http://emblems.let.uu.nl/

This means that upon a request for a certain web page, specific XML-elements (the elements that are needed for the web page the user requested) are retrieved from the central data files, and transformed into HTML 'in real time'.

The software facilitating this setup is the web development framework Cocoon⁵, an open source software project from the Apache Software Foundation.⁶ The transformation from XML to HTML is done using so-called XSL-stylesheets. These stylesheets describe, in detail and under the governance of Cocoon, the process of extracting specific XML-elements, bundling them and subsequently transforming them into HTML.

Dynamically creating the web pages that visitors of the EPU website see has a number of advantages. The first is that changes to the underlying XML-files – fixing a transcription error, adding a hyperlink, et cetera – are immediately visible on the website. This alleviates the editors of the burden of constantly having to update the website trough technical procedures. The editors can immediately check every change. Every improvement can immediately be seen by the users of the website.

A second advantage is that the technical difficulties with current presentation formats like HTML (think about the differences between the browsers Mozilla Firefox and Microsoft Internet Explorer in the rendering of it), will not trickle down to the level of the encoding of the emblem books. Whereas the XML-standard is open and, at this moment, the best possible guarantee for ensuring the durability and persistence of digital data, every HTML-website that hopes to attract an audience suffers from Microsoft's woefully inadequate adherence to standards. Every website that wishes to be of service to web browsers that handle HTML in a well-ordered and predictable way, as well as to Microsoft Internet Explorer, has to deal with numerous banal, technical incompatibilities. With the setup used in the EPU the 'data layer' (the content: the transcribed emblem books, encoded in XML) is separated from the 'presentational layer' (the HTML website). These issues are strictly confined to the transformation process from the former into the latter.

More perspectives

The separation of the actual content from the presentation provides us with another advantage. With Cocoon, it is easy to create multiple, parallel transformation paths. Up till now, the EPU website is the most manifest presentation of the XML-encoded emblem books. It will probably stay that way for a long time to come. However, Cocoon is perfectly able to serve not only HTML, but other formats as well. The transformation from XML to HTML (executed, as we have seen, upon a user requesting a certain page from the website) can be augmented by more, parallel 'transformation paths'. Without too much effort, Cocoon can be set up to transform the underlying XML into, for instance, pdf-files; upon a user request, and in real time.

⁵ http://cocoon.apache.org

⁶ http://www.apache.org/

As pdf-documents are generally better suited for printing than HTML-pages, work is currently being done to add downloadable pdf-documents to every emblem on the website, as well as allowing users to download each of the emblem books in the corpus in pdf-format. Before the actual download, users will be able to specify what the pdf-document should look like; for instance, whether the original or the modernised spelling should be used. The request is being generated on the server and downloaded to the user's computer, where it can be viewed, stored and/or printed out.

Though allowing users to download pdf-files for each emblem is useful enough, the setup that the EPU website uses – employing Cocoon, thus allowing the above mentioned multiple, parallel transformation paths – provides a fundamental 'openness' of the underlying data. The website should be understood as only one of many possible 'perspectives' on the underlying data. Many more are possible and, indeed, currently in place or under development. The Open Archive Initiative⁷ will be able to query the EPU files for its own 'metadata harvesting' function, while Mnemosyne,⁸ a semantic web company that has contributed to the indexing of pictorial motives in the emblems, already extracts data on those motives from the EPU. All these methods of accessing the EPU data run in parallel, each without disturbing or interfering with the other. While the website is and probably will remain the most used method of 'accessing' the data files of the EPU, it will be clear by now that access to the emblems is not restricted to what is shown in HTML.

Future developments

The EPU was specifically set up to allow for yet unrecognized uses of the encoded data. Through the use of open source software and the open standard that XML is, it is hoped (though, of course, not guaranteed) that the encoded material will remain – at the very least – accessible in one form or another even after the project as it is now might have ended. Open access to the underlying data gives emblem (and other) scholars the possibilities to pursue research not yet thought of. Furthermore, the setup of the EPU is particularly suited for the incorporation of new tools like EDITOR and SANE⁹ that promise to allow users to add their own annotations to the encoded material. Although allowing users to enhance and perhaps even modify the original 'tagging' presents a myriad of security and integrity issues, it is a logical next step in the opening up of the underlying data.

Reflecting on these future developments would not have been possible without the actual hands-on experience the EPU provided us with. Clearly, the initial project goals have been met. The practice of digitising a corpus of 25 emblem books allowed us to reflect on the consequences of the digitisation of historical

⁷ http://www.openarchives.org/

⁸ http://mnemosyne.org/

⁹ EDITOR and its successor SANE are being developed at the Huygens Institute, http://www. huygensinstituut.knaw.nl

material. Furthermore, the EPU's digitised material forms a valuable addition to the evergrowing body of digitised historical material. Finally, the EPU promises to be a persuasive catalyst for further digitisation projects.

Bibliography

- Boot, P. (2004) 'Accessing emblems using XML: Digitisation Practice at the Emblem Project Utrecht' in S. López Poza (ed.), *Florilegio de estudios de emblemática: actas del VI Congreso Internacional de Emblemática de The Society for Emblem Studies = A florilegium of studies on emblematics: proceedings of the 6th International Conference of The Society for Emblem Studies: A Coruña, 2002.* Ferrol: Sociedad de Cultura Valle Inclán, 191-197.
- Renear, A.H. (2004) 'Text Encoding' in S. Schreibman, R. Siemens and J. Unsworth (eds), *A Companion to Digital Humanities*. Oxford: Blackwell. Also available as http://www.digitalhumanities.org/companion (last accessed June 6, 2007).