Mesotext

Mesotext
Digitised Emblems, Modelled Annotations and
Humanities Scholarship

Mesotekst
Gedigitaliseerde emblemen, gemodelleerde annotaties en de
geesteswetenschappen
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit
Utrecht op gezag van de rector magnificus, prof.dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties in het
openbaar te verdedigen op vrijdag 20 november 2009 des
middags te 12.45 uur

door

Peter Boot

geboren op 25 september 1961
te Alphen aan den Rijn

Promotoren: Prof.dr. E.M.P. van Gemert

                        Prof.dr. P.W.M. Wackers

Co-promotoren: Dr. A.J. Gelderblom

                        Dr. F. Wiering

Table of Contents

Part I. Context of this Study

1   Introduction

1.1   Subject

The central thesis of this book is that digital annotations make it possible to create live connections between the primary literature that scholars study, and the secondary literature that they produce. It will show how digital tools can help scholars annotate the sources that they study, and help them conceptualise the issues that interest them. These typed annotations can facilitate entry into and exploration of primary texts, and can provide the supporting arguments for the articles and studies that scholars write about primary texts. I will call these typed annotations *mesotext*, and the book converges towards the chapter that introduces the concept and explores its usefulness in the context of humanities scholarship.

Mesotext is the scholarly equivalent of the measurement data of science. And just as scientific data do not result from random measurements but originate from hypotheses based in theories, or models, scholarly annotation is implicitly or explicitly based on a model of the area under study. The model that informs a body of annotation is what makes it more than a collection of random observations. In according a central place to the notion of modelling, this book is much indebted to the view of humanities computing put forward by Willard McCarty (2005). Disciplined application of a model, which is what the computer helps us do, reveals the limitations of the model and the gap between the model and the phenomena it tries to account for. This is what helps us refine or reject the model, and thus helps us make progress.

The background of this book is digital emblem studies. The book's theses emerge from the application of annotation technology to emblems. The emblem, which will be introduced in chapter 2, was a genre in especially sixteenth and seventeenth century European literature. It typically consists of a motto, a picture (woodcut or engraving), and an epigram, often joined by one or more learned quotations and prose explanations. Emblem creators were heavy users of what we would now call 'cut and paste', and borrowed freely from the visual arts, classical literature and each other. Emblems could deal with a great variety of subjects, from love to religious instruction, from applied morality to politics. Emblem scholars are therefore to be found among historians of literature and the arts, but also among political historians, theologians, historians of mentality, and the like. The emblem has often been described as essential to the study of the early modern era. The emblem's use of multiple

media, its wide variety of subject matter and its many intertextual relations make emblem studies a very suitable area for experiments in humanities computing.

1.2   Field

If there should be one field that this thesis belongs to, it would be humanities computing. The practice of modelling that McCarty puts squarely at the heart of humanities computing is what is facilitated by the annotation tools that I will describe. Another of McCarty's characteristics of the field that quite fits this thesis is that the field is by nature experimental. Doing humanities computing is to be engaged in a process of trial and error. The following chapters will show examples of these. I have not tried to hide them, heeding John Unsworth's advice to the computing humanist, to 'record your every doubt and misstep, and aspire to be remembered for the ignorance which was uniquely yours, rather than for the common sense you helped to construct' (1997a).

Humanities computing is by nature an interdisciplinary field. From a disciplinary perspective, McCarty characterises it as a methodological commons, fed by concepts and methods borrowed from computer science and other disciplines, which are then applied to other fields, such as law or musicology or literature. The concept of annotation is important in a number of fields: fields that belong to or are close to literary studies, such as classical studies or textual scholarship, and fields further removed, such as library science, the social sciences or computer science. This book draws part of its inspiration from the way annotations are used in these other fields. In a sense, what this book does is to help bring the concept of annotation into the methodological commons of humanities computing, and to develop methods and tools for applying it.

The field that this book originated in was emblem studies. The original plan was to do a semiotic analysis of emblem literature and devise digital tools for storing, querying and exploring the analysis. Over the years, the emphasis has shifted away from content to methodology and technology. Still, the emblem was the field that the annotation tools were designed for, and the book contains two chapters of unadulterated emblem studies.

A necessary condition for what the book sets out to do is digitised primary material, and part II of this book is devoted to emblem digitisation. Different ideas exist about the best way to do emblem digitisation and there are correspondingly different practices. Some of these lend themselves more easily than others to an integration of digital emblem collections and the research that uses them. I will discuss the relevant features.

From another perspective, the subject matter of this book is the scholarly digital edition. The edition is an important subject, both because annotation has always been an indispensable ingredient of the scholarly edition, and because the digital edition is the vehicle of choice for the primary texts that

are the target of annotation. Annotation is one way of making an edition into more than merely a store of existing knowledge: the edition should be an environment for study and research and a receptacle for new knowledge. I will suggest ways to bring that about.

Finally, in addressing scholarly research environments, the book is of necessity also concerned with scholarly publishing. Mesotext makes no sense if it cannot be digitally explored, and for that exploration to be useful, the primary texts that it refers to and the secondary texts it supports should be available online and should be accessible. Moreover, unlike today's ubiquitous PDF documents, they should be interactive, addressable and expandable. We need to find ways to create live connections from our scholarship into the data that support it, and our present publishing system is not equipped to do that.

1.3   Audience

The book then was written with multiple audiences in mind. Those involved in emblem digitisation will probably be interested in part II, those that work in humanities computing will prefer part III, emblem scholars may like part IV, while the final part may again be mostly attractive to those that work in humanities computing. Section 1.4 provides a roadmap to the book.

I very much hope for readers that will look beyond their own fields. For students of humanities computing, I believe the emblem is an interesting phenomenon, because of its bimedial nature, its relations to the visual arts and its otherwise central location in Renaissance culture. But more importantly, I hope that emblem scholars, and others working in Renaissance arts and literature, will take note of and think about the issues dealt with in the chapters of this book that are not specifically emblematic.

This is important for two reasons, the first of which is that there is nothing like 'the epistemological power of self-limiting method' (McCarty 2005: 20) to help clarify one's ideas and learn about one's subject, whatever it is. The annotation and modelling tools described in part III and reflected on in part V of this book embody that self-limiting method and are therefore, as will be argued in chapter 13, useful to humanities scholars.

The second reason is that humanities scholars should want to contribute to the development of the tools and procedures of digital scholarship. As McCarty notes, 'the tendency has been to relate computer science and the humanities according to their separate concerns as these are imagined by those peripheral to humanities scholarship' (191). If as humanities scholars we want to be more than the 'welcoming conquered' (10) there is no alternative to getting involved into thinking about and designing the future tools of our trade.

1.4   Roadmap for Readers of this Book

This roadmap very briefly characterises the chapters of this book and their interconnections. Part I of the book is introductory. The present chapter provides an overview of its contents and describes some of the context in which it was written. Chapter 2, 'An Introduction to Emblem Studies', introduces the emblem and (digital) emblem studies. The chapter describes the history of emblem studies and discusses the potential role for digital technologies in emblem studies. Chapter 3, 'Annotating the Digital Edition', then introduces and defines the concept of annotation. It examines the way annotation facilities are currently implemented in editions and digital library environments, and looks at neighbouring fields for other applications of annotations.

Part II, 'Emblem Digitisation and the Emblem Project Utrecht', deals with a necessary precondition for digital scholarship: the availability of digitised source material, i.e. digital emblems. It contains a number of papers that discuss the Emblem Project Utrecht (EPU) and other emblem digitisation projects. Chapter 4, 'A Model for Digital Emblem Editions', opens with a brief look at the first beginnings of emblem digitisation, and then sketches a model that can be used to characterise emblem digitisation projects. It also discusses the suitability of XML for emblem digitisation. The next chapter, 'Digital Editing and Text Modelling: The Case of the Emblem Project Utrecht', looks at the role of modelling in designing a digital emblem site such as that of the EPU.

One of the books digitised within the EPU is Otto van Veen's *Amoris Divini Emblemata* (Antwerp 1615). This book of religious love emblems provided the material for the exploratory investigations reported on in Part III of the book, which discusses a number of approaches to digital (emblem) annotation. The first approach is represented by EDITOR, an annotation tool developed at the Huygens Institute. EDITOR facilitates the creation and manipulation of annotations to digital editions. One of the essential characteristics of EDITOR is its user defined annotation types consisting of multiple fields. The annotation types implement the notion of a model that informs the mesotext, as mentioned above. EDITOR is described in chapter 6 'Digital Edition Annotation using EDITOR'. The chapter explains the need for an annotation program that can be used in the context of XML-based editions. It lists requirements for a scholarly annotation facility and describes EDITOR's data model, functionality and interface. It considers a number of issues that arose during the program's development. It also presents the web facility for (graphical) display of EDITOR annotations and continues with a discussion of how EDITOR was put to use in an investigation into theatricality in emblems. Chapter 7, 'A SANE Approach to Annotation in the Digital Edition', discusses the issue of how to merge EDITOR-created annotations with a digital edition. The chapter proposes a wider annotation exchange infrastructure dubbed 'SANE', for 'Scholarly ANnotation Exchange'. SANE would consist of an annotation mark-up language (for storing

annotations) and an annotation request protocol that clients could use to request annotation information. The chapter shows a prototype of an edition requesting annotation information from a SANE server and embedding the returned information, including interactive graphics, in the edition's web pages.

A second annotation approach that I take here is ontological annotation. The model that informs the annotations in this case is formulated in terms of the Web Ontology Language, OWL.[1] Models formulated in OWL can be much more expressive than models formulated in EDITOR, as OWL classes can have subclasses, and can be related to other classes using properties. Using OWL, one can create a model of the world underlying a text, rather than just make statements about text fragments. Chapter 8, 'Decoding Emblem Semantics', presents the foundations for this approach. It offers a general-purpose semiotic model and an implementation of this model in terms of an ontology. A computer ontology is a digital representation of the types of objects, with their properties and relations, that exist in a domain. It implements this ontology using semantic web technology and shows how it might be applied. It also argues for the importance of creating annotations that are closely attached to the annotated text and that are based on an explicit model of the phenomena being annotated. Chapter 9, 'Creating a Metaphor Index', applies these ideas to the creation of an index of metaphor in *Amoris Divini Emblemata*. Based on a simple definition of metaphor it indexes part of the metaphoric material in the book and shows an application where the user can query the metaphor data from the text of the emblems, from a secondary text that discusses the metaphors, and from the metaphor model.

A third approach uses the Text Encoding Initiative's (TEI) facilities for annotation, as applied specifically to annotating parallel text structures. The annotation model is here formulated in terms of TEI feature structures. The expressive power of TEI feature structures, as used here, is roughly equivalent to EDITOR's annotation types. Chapter 10, 'Towards a TEI-based Encoding Scheme for the Annotation of Parallel Texts', first discusses requirements, specifically in the context of annotating parallel texts, such as translations. It proposes a TEI customisation that can be used in this context. It also introduces the concept of a TEI proxy document, to be used for annotation of texts of which a full TEI representation is not available.

Part IV presents two 'traditional' studies into *Amoris Divini Emblemata*. Chapter 11, 'A Mirror to the Eyes of the Mind' is about the metaphors that the book employs. It argues that the recurring metaphors (of landscape, light, and nature) contribute to making the book into a coherent whole. Chapter 12, 'Playing and Displaying Love', focuses on a number of theatrical aspects in *Amoris Divini Emblemata*. It argues that the protagonists of the emblem pictures, Divine Love and the human soul, in many respects resemble actors,

---

[1] Not a typo.

staging symbolic scenes in front of an audience. This suggests a way of looking at their gestures, body-language and utterances.

The last part of the book, part V, reflects on what has and what has not been accomplished. Chapter 13, 'Mesotext', makes the case for what in the opening of this chapter I described as the central thesis of the book, the claim that digital annotations allow us to create a connection between the primary literature that scholars study, and the secondary literature that they produce. Based on what has been presented in part II it will introduce the concept of mesotext and place the concept in the context of recent discussions about scholarly publishing. The last chapter, chapter 14, looks back at the annotation concepts, tools and technologies discussed in the book and tries to assess their potential. It considers some of the wider questions these tools raise. It claims there is a need for an experiment that will put to test the theses I argued for in this book.

Notes

This book unavoidably uses technical terminology and many acronyms. Rather than explain these in each chapter where they occur, I provide a glossary at the end of the book.

The thesis includes a number of chapters that report on the development of experimental annotation tools. For illustrative purposes, I have placed demo applications, source code and demo files on the web for most of these tools. All are accessible from http://peterboot.nl/thesis/. A small number of documents that reflect essential design decisions I include as appendices in this book.

All hyperlinks were checked in April 2009.

2   An Introduction to Emblem Studies

This book is a study into computational methods that will be helpful in accomplishing, documenting and querying a decoding of the semantics of the European emblem. Digitising emblem books and developing digital tools to support emblem scholarship require some understanding of the emblem and the scholarly interest in it. Accordingly, this chapter will introduce developments in emblem studies: emblems and emblem books in section 1, and the main interests motivating emblem study in section 2. Section 3 will introduce the subject of digital emblem studies.

2.1   The Emblem

This section will introduce the emblem and consider the difficulties scholars have encountered who wanted to define the emblem and delimit it from neighbouring genres. It will also mention some of the characteristics of national emblem traditions and some of the unifying forces that have kept these traditions together.
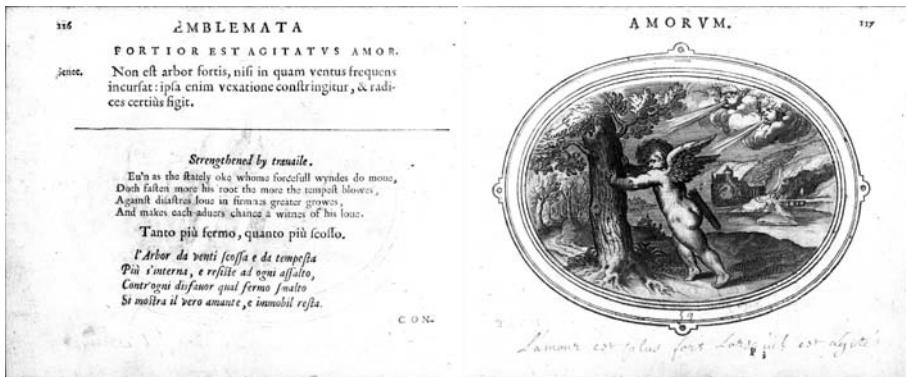
Figure 1 *Fortior est agitatus amor*. Otho van Veen, *Amorum Emblemata* (Antwerp 1608), pp. 116-17.

I will use Peter Daly's summary definition of the emblem: 'emblems are composed of symbolic pictures and words; a meaningful relationship between the two is intended; the manner of communication is connotative rather than

denotative' (1998: 8). A typical emblem would be *Fortior est agitatus amor* from Otto van Veen's *Amorum Emblemata* (Antwerp, 1608). See Figure 1.

The emblem consists of a motto, *Fortior est agitatus amor* ('Love is stronger when assailed'), a Latin quotation from Seneca, epigrams in English and in Italian, both of which have their own motto-like title, and a picture. The picture shows Cupid, leaning against a tree that is battered by winds. The English epigram states:

> Ev'n as the stately oke whome forcefull wyndes do move,
> Doth fasten more his root the more the tempest blowes,
> Against disastres love in firmnes greater growes,
> And makes each advers chance a witnes of his love.[2]

The emblem shows the process of the perfection of love under duress (Buschhoff 2004: 92). The first part of the epigram describes the metaphorical object, the tree, restating the Seneca quote;[3] the second part applies the metaphor to the subject of love. This two-part construction is very typical of emblem epigrams. The amorous interpretation of the tree is reinforced by the presence of Cupid in the picture. The sample emblem also shows a number of formal properties characteristic of the genre: the importance of Latin, the use of multiple languages, the reference to a classical author, the use of a single page or a book opening for each emblem – properties that are in no way mandatory, but do recur in quite a number of emblems and emblem books. And it is not just the formal properties of this emblem that it shares with others: the topos of the tree that is strengthened by storms was used by at least four other emblematists, e.g. by Juan de Horozco y Covarrubias (Henkel & Schöne 1996: 147-48). He compares the tree's strength to virtue rather than to love. The example shows in a nutshell how emblem writers could use the same commonplaces while not merely repeating each other.

The emblem genre is commonly held to have come into existence when in 1531 the Augsburg publisher Steyner added pictures to a collection of epigrams, the *Emblematum Liber* (Alciato 1531), by humanist and lawyer Andrea Alciato (Miedema 1968). Many of these epigrams described visual objects such as statues, and consequently they lent themselves very well to illustration. The collection became popular immediately, and was repeatedly reprinted and expanded. It was translated into a number of vernacular languages and became a subject for much learned commentary. The *Alciato at Glasgow* website (Alciato

---

[2] I modernised 'u' to 'v'.

[3] 'Non est arbor fortis, nisi in quam ventus frequens incursat' (No tree becomes rooted and sturdy unless many a wind assails it). From *De Providentia* 4, 16 (identified by Jan Bloemendal).

2007) makes available 22 of the many editions and shows the richness of this tradition.[4] Alciato's emblems gave birth to a genre that has been estimated to include 6500 editions (Daly 2002: 261), published in especially the sixteenth, seventeenth and eighteenth centuries.[5] Alciato's book was not the only impetus towards the new genre, however. Russell finds an 'emblematic process' at work in fifteenth century France, where works of allegory such as Petrarch's *Triumphs* were illustrated and condensed into what were essentially emblems (Russell 1985: 164-72). Sebastian Brant's *Narrenschiff* (1494) has been considered an emblem book (Clements 1960: 9). The proto-emblematic genre of the device, a combination of picture and motto that stated a personal ambition, developed during the fourteenth and fifteenth centuries (Russell 1985: 23-28).

It is notoriously difficult to give a precise definition of what constitutes an emblem (Daly 1998: 3). There is no single property that distinguishes emblems from non-emblem texts. The 'rule' that describes an emblem as consisting of picture, motto and subscription 'may actually be honoured more frequently in the breach than in the observance', as David Graham writes (2005: 131). Prefacing his attempt to distinguish form, structure and function in the emblem, Graham notes that structural heterogeneity is the rule, rather than the exception. It is not clear whether the 'remarkable degree of shared functionality' which he suggests may be hidden by the proliferation of forms, will also serve to distinguish the emblem from related or similar genres. Definitions in term of emblem structure 'reveal themselves to be quite inadequate when confronted with real French emblem books' (Russell 1985: 162). Earlier attempts to identify a uniquely emblematic distribution of functions over the emblem's structural components have failed, and must fail, as recently argued by Carsten-Peter Warncke, who in his discussion of Alciato notes: 'Alle diese heterogene Züge der Alciatschen Emblemveröffentlichungen wurden vorbildlich für die Gesamtentwicklung der Emblematik, weswegen es sich verbietet, Emblemdefinitionen an den Äußerlichkeiten der Emblemform festzumachen. So war es möglich dass die Emblematik sich inhaltlich und formal überaus unterschiedlich und variantenreich entwickeln konnte' (2005: 52).

John Manning devotes the introduction of his book *The Emblem* (2002) to the multiple forms the emblem may take. 'One cannot understate the variety as well as the pervasiveness of emblematic modes of thought during the [Renaissance and Baroque] period', he writes. Not all emblems have pictures, not all are enigmatic, not all are moralising. Manning derides the 'hunt for the fabulous Idealtypus' (21) of German scholarship. Nor can the emblem be described as a collection of either internally consistent sub-genres, such as the love emblem or the Jesuit emblem, or coherent national traditions. The best answer to the question what an emblem is, may be to recognise that '[p]art of the emblem's distinction - as was recognised early on - was its diversity', and to

---

[4] Höltgen counts 170 Alciato editions (1986: 26).
[5] Höltgen's estimate is considerably lower: 2000.

take the approach that '[e]verything that was called an emblem *was* indeed an emblem'.

A problem of this nominalist approach is that many emblem books did not identify themselves as such. Heckscher and Sherman, in their *Emblematic Variants* (1995), list words that emblem writers used on their title pages to characterise their books. It is highly insightful to see the very different semantic fields that were used: there are mirrors, garlands, thesauruses, symbols, anatomies, eagles, gardens, and truth sayers. This is obviously of great consequence if we want to have an idea of what emblem writers proposed to do and what they wanted their books to be. But what is important for us in this context is to note that only a quarter of the books used the word 'emblem' to describe themselves.

For the purpose of digitisation and digital analysis it is important to be aware of the multiple forms that the emblem could take. We do not require a definition that strictly separates emblems from non-emblems based on a single characteristic. It is important to be aware of the existence of neighbouring forms, such as the emblematic fable (P. J. Smith 2006); it is not that important to be able to decide what counts as an emblem and what does not.

The emblem can only be studied and understood when due consideration is given to its predecessors in form, content and way of thinking. I mention a number of these predecessors, following Daly's account (1998: 9-42).

Alciato based up to 50 of his epigrams on a collection of Greek epigrams known as the *Greek Anthology*, epigrams that prefigured the emblematic tradition in describing a visual object and drawing a conclusion from that description (Scholz 2002: 30). Beyond this particular debt, both Alciato and the wider emblem tradition owed a large debt to classical mythology, history and allegory. Figures from e.g. mythology would represent particular virtues, vices, or lines of conduct. One form in which the classical inheritance was transmitted consisted of collections of 'loci communes' that Renaissance pedagogy highly valued. Bath has stressed the importance of the compilations and *florilegia* for the emblem, and the emblem writers' contribution to that tradition (1994: 28-56). Collections of proverbs and adages, such as Erasmus' *Adagia*, were also to influence the emblem tradition.

Hieroglyphs provided an important source of motifs for emblem writers. They were thought to represent ancient wisdom and to have a somehow natural and inherent significance. The Renaissance understanding of hieroglyphs provided what Daly calls a 'symbolic mode of thought' for emblem writers (27). Many of the hieroglyphic signs had also found their way into medieval bestiaries, such as the *Physiologus* (e.g. the lion that sleeps with its eyes open or the traditional accounts of the phoenix and the pelican). Much of this animal lore would be used by emblem writers.

Another source of inspiration for emblematists was the impresa. The impresa, or device, consisted of a picture and brief motto, and was worn by noblemen and military officers. It was thought to express a personal aspiration, in contrast with the generally applicable truths expressed by the emblem. The impresa was especially fashionable in Italy, where it was popularised by the invading French armies around the turn of the sixteenth century, and Italian theories of the impresa deeply influenced contemporary thinking about the emblem. The most important of these theorists was Paolo Giovio, whose *Dialogo dell'imprese militari et amorose* (1555) formulated a number of conditions for the perfect device. Related elements of the semiotic matrix[6] of the Renaissance were the commemorative medals and the system of heraldry.

A last source of influence to be mentioned here is medieval nature symbolism and the allegorical reading of the Bible. Nature, seen as God's second book (Scholz 1993), was interpreted to have multiple allegorical readings. A lion could be Christ-like, but could also represent the devil, the heretic, or the Christian, dependent on the (supposed) property being foregrounded. This created a network of meanings that precludes any simple 1-1 relationship between symbolic sign and meaning.

What this briefest of overviews of the emblem's forerunners shows, is that, like any significant cultural phenomenon, the emblem had an intellectually and artistically diverse ancestry. This may help explain the variety of forms it took, both geographically and temporally. The refusal of the emblem genre to conform to any straitjacket, and the emblematists' willingness to incorporate diverse material ensured the liveliness and attractiveness of the genre. It is also a warning against any attempt to enforce simplifying models in emblem digitisation projects and projects for digital analysis of emblems.

During its history of almost three hundred years, in the various national cultures the emblem developed along different lines. Most of the larger studies in the emblem take this national perspective. Examples would be Adams (2003), Russell (1985, 1995) and Saunders (1988, 2000) for France, for Britain Bath (1994) and Höltgen (1986), for Germany Höpel (1987) and for the Low Countries Porteman (1977). I will sketch the developments in five European countries based on Daly's *Companion to emblem studies* (2008).[7]

The French were the earliest adopters of the emblem genre (Russell 2008): French printers printed a large number of Alciato editions, but from the 1530s also produced indigenous emblem books such as those by De la Perrière and Corrozet. The first completely religious emblem book, by Georgette de Montenay, was published in France in 1571. In the seventeenth century the French began to read foreign and translated emblems, often in polyglot editions. Seventeenth-century emblems were often religious, e.g. the collection of heart

---

[6] A concept borrowed by Scholz (2002: 16) from Thomas Greene (1982).
[7] I will discuss the emblem in the Low Countries at the end of this section.

emblems by Luzvic. In court circles, the device was very important. Devices also played an important part in interior decoration. Their design was complex and subject to many rules; it was said, probably in jest, that the design of a good device could represent a life's work. A specialised *Petite Académie* for the study of devices and related material was created in 1663. In the eighteenth century, French interest in the emblem was more or less over. In German-speaking regions, even though Alciato's book was first printed in Germany, interest took longer to take hold (Peil 2008). From the 1580s emblem books began to appear, often either in Latin or bilingual collections, usually with an ethical-moral direction. Early German emblems, e.g. in the *Politisches Schatzkästlein* (Meisner 1623-1631), often integrated a townscape in the emblem picture. Later, religious emblems became very important, both protestant and catholic. They were often published in illustrated collections of sermons. A specific German invention was the *emblema triplex*, a combination of three individual emblems in a coherent whole. The religious love and heart emblems from the Netherlands were often imitated. The presence of many courts in German-speaking regions meant there was also ample room for political, usually panegyric, emblems.

In Britain, unlike in these other countries, there never was a translation of Alciato (Silcox 2008), although Geffrey Whitney's *Choice of emblems* (1586), the earliest secular British emblem book, includes 87 imitations of Alciato. Another important secular work was Wither's *Collection of Emblemes, Ancient and Moderne* (1635), a collection based on plates made for Rollenhagen's *Nucleus emblematum selectissimorum* (1611-1613). The religious emblem in Britain started as vehemently anti-catholic, but developed into a largely non-sectarian meditative genre. The British emblem is unique in having had a revival in the nineteenth century. The Spanish emblem flowered somewhat later than elsewhere, in the second half of the seventeenth century (Bernat Vistarini & Cull 2008). It was more baroque than elsewhere, less influenced by humanism, and more by the mind-set of the Counter-Reformation. Spain had its own Alciato translations and a number of learned commentators. Among books with moral content, Sebastián de Covarrubias' *Emblemas Morales* (1610) stands out. Among the political emblems, Saavedra Fajardo's *Idea de un príncipe politico cristiano* (1640), dealing with the education of a Christian prince, became very influential all over Europe.

The emblem tradition of the Low Countries can be characterised by a number of highlights:[8] in the second half of the sixteenth century a series of translations of foreign emblem books, printed by Plantin in Antwerp; then in the first two decades of the seventeenth century the flowering of the (secular) love emblem, overtaken, from 1615 onwards, by its religious counterpart; a growing realism, in the emblems of Visscher's *Sinnepoppen* (1614) and Cats's *Proteus* (1618), in the second decade of the century; *Gulden Winckel* (1613) and

---

[8] About the Dutch emblem see Porteman (1977), Stronks (2008) and Porteman & Smits-Veldt (2008: esp. 43-45, 179-83, 248-52, 310-12, 483-91, 856-70)

Vorsteliicke warande der dieren (1617) by Vondel, a book of emblems and one of emblematic fables respectively, in the same period; and finally the emblems of Jan Luyken who on his own, around 1700, recapitulates nearly the whole of the Dutch emblem tradition: publishing *Duytse lier* (an emblematically organised collection of love poetry, 1671), two volumes of religious love emblems among which *Jezus en de ziel* (1678), and a number of works where objects from ordinary life (such as trades or household utensils) are emblematised, e.g. *Het menselyk bedryf* (1694).

For this thesis, the most important of these highlights is the love emblem, in both its secular and its religious variety.[9] The Emblem Project Utrecht digitised the most important works from the love emblem corpus. The love emblem was created by Daniel Heinsius and a circle of his humanist friends at Leiden university. *Quaeris quid sit amor* (1601) was the first specimen of the genre. *Quaeris* was followed by books of love emblems by Van Veen, Hooft and Cats. In Heinsius's book, love was presented in a Petrarchist vein, where the male lover suffers at the hand of his cruel mistress. Later the emphasis moved towards bourgeois courtship and marriage. One of the most characteristic features of the love emblem is the presence of Cupid in the emblem pictures. Cupid's presence provides a generic motif with amorous meaning, as we saw in the sample emblem above. Van Veen, also known as Vaenius, was the first to rework the emblems of love in a religious sense. In his *Amoris Divini Emblemata* (1615) Divine Love replaced Cupid, and showed the significance of symbolic objects and situations to the human soul, who had taken on the guise of a young girl. *Amoris Divini Emblemata* is the book that most of my examples will be taken from. It was to inspire Hugo's *Pia Desideria* (1624) and through that a multitude of other pious emblem books.

These national developments were held together by a number of common strands. Alciato's emblems and his imitators represented one such strand. The neo-Latin emblem of course transcended national boundaries. Manning views the emblem book as a manifestation of 'a shared European neo-Latin culture of festive celebration' (2002: 9). Among Latin emblems the Jesuit emblems have a prominent place (Dimler 2007): Jesuits were among the leading theorists of the genre, and also, more importantly, they used the emblem as a vehicle for education (Porteman 1996b). Thematically, emblems all over Europe often dealt with the same subjects: virtue, religious instruction, and politics. A 'new' emblem subject, such as love, could become widely popular all over Europe in a few years (Warncke 2004). One reason why emblematic fashions could quickly spread over the continent is the common practice of creating polyglot emblem books. Books would often use Latin and a single vernacular,[10] but they could also

---

[9] About the love emblem see Manning (2002: 166-84) and Stronks & Boot (2007b). About the secular love emblem see Porteman (1983) and Sebastián López (2001).
[10] 7 out of 22 books in the mentioned *Alciato in Glasgow* project do so.

include multiple vernaculars.[11] Emblem encyclopaedias published around the end of the seventeenth century might use eight languages for each emblem.

Emblems did not appear only in emblem books: '[t]he emblem (...) helped to shape virtually every form of verbal and visual communication during the sixteenth and seventeenth centuries' (Daly 1986: 160). Emblems were used in painting, in architecture, in the ephemeral arts, in interior decoration. 'No domestic or public space was left unfilled by some appropriate emblematic decoration' (Manning 2002: 25). This reuse of the emblem outside of the sphere of the book is one among several factors motivating emblem studies.

2.2   Emblem Studies

It is important to provide a short introduction to emblem scholarship, as the digital collections and tools that this book will discuss are meant to serve emblem scholars. The usefulness of these tools and collections is measured by the degree to which they meet the requirements of emblem scholarship. Therefore the subjects, procedures and preoccupations, trends and directions of emblem scholarship are very relevant to any discussion of digital tools and collections.

Any overview of emblem studies should begin with the theories of Mario Praz.[12] Praz' *Studies in Seventeenth-Century Imagery* (1975, first impression 1939), viewed the emblem in the context of mannerism, and described it as a conceit, principally the result of an exercise of wit, which succeeds in bringing together things apparently unrelated (a lover and a crocodile, to give an example from Van Veen). The result is an enigmatic form, directed at a learned public. This enigmatic character of the emblem was seen by Heckscher and Wirth as its defining characteristic: the emblem is essentially an enigma, posed by picture and motto and resolved by the subscription. In their influential encyclopedia article 'Emblem, Emblembuch' they write: 'Man hat es beim E[mblem] demnach mit einer Vereinigung vom Wort des Lemma mit dem Bild der Icon zu einem Rätsel zu tun, dessen Auflösung durch das Epigramm ermöglicht wird; zum Besonderen der Auflösung gehört, dass sie inhaltlich auf ein moralisches Ziel hin zugespitzt ist und in dieser Einseitigkeit und Übertreibung ein Mittel abgeben soll, die Wahrheit, die das E[mblem] enthält, zu verdeutlichen' (1959: col. 95).

Perhaps the most influential analysis of the emblem has been given by Albrecht Schöne in his *Emblematik und Drama im Zeitalter des Barock* (1993, first impression 1964). Schöne sees the emblem not as an exercise of wit but as an expression of meanings sanctioned by traditional bodies of knowledge

---

[11] Van Veen's *Amorum Emblemata* was simultaneously published in three trilingual editions, in a total of five languages (Porteman 1996a).
[12] This overview is based on among others Bannasch (2007: 18-31)

handed down from antiquity and the Middle Ages, such as mythology and bible exegesis. In Peter Daly's summary (1979: 68), Schöne's theory holds that:

> (1) the three-part emblem fulfils a double function of 'representation and interpretation';
> (2) the picture is characterised by 'potential facticity' and 'priority of idea';
> (3) the mode of thought is 'verbindlich', a development of medieval symbolism and exegesis.

That is to say: all three constituents (motto, picture and subscription) can partake in both representing an emblematic object or situation and interpreting that object. But as the picture is the only constituent that represents reality immediately,[13] the picture is where the process of interpretation begins and therefore it is logically prior to the text. The picture must represent (potential) reality, as the depicted object's inherent significance is what the subscription then proceeds to explain.[14] The emblem reveals God-given correspondences between objects. Only emblems whose functioning conforms to this model are real emblems; allegory, which in Schöne's view is simply a way of referring to abstract ideas, has no role in the emblem.

Daly has been Schöne's interpreter in the non-germanist world, mainly through his *Emblem theory: Recent German contributions to the characterization of the emblem genre* (1979). By and large, Daly accepts Schöne's analysis, but stresses that there are other modes of thought for the emblem than just the typological variety which Schöne elevated to the rank of an ideal type. The other modes of thought (i.e. 'the relationship between object and concept', 80) are hieroglyphic ('manifests itself in those emblems in which a strange or inorganic combination of individual motifs is assembled to represent a general notion', 82) and allegorical ('is seen in operation where the emblem writer employs an existing pictorial generalisation, such as a personification, to illustrate a general truth, which is not rendered specific in any way', 92). That is not to say that each emblem can be assigned to a single category: there are also mixed and double modes of thought. Schöne's demand that the pictured object should correspond to a (potential) reality should be relaxed for the allegorical and hieroglyphic emblems.

Michael Bath has proposed replacing the concept of potential facticity with the structuralist notion of the 'vraisemblable', 'the clichés and stereotypes of accepted knowledge through which particular representations become

---

[13] '… sie und nur sie repräsentiert ganz unmittelbar, nämlich auf anschaubare Weise …' (Schöne 1993: 26).
[14] '… was durch die emblematische subscriptio dann ausgelegt wird, indem diese die in der pictura beschlossene, über die res picta hinausweisende significatio offenbar macht' (ibid.).

plausible or intelligible' (1994: 6). The supposed facticity is also one of the main issues for Bernhard Scholz, one of the most vehement critics of Schöne. Instead of Schöne's ontological foundation for the emblem, Scholz uses a semiotic approach and proposes the concept of 'Referenzialisierbarkeit', the possibility of being assigned a referent object in reality. This 'referencialisibility' should be seen in the context of a 'Sinngehalt', a system of statements recognised as true within a certain community (2002: 257-58). Within a Sinngehalt that includes the traditional lore of medieval bestiaries, pelicans spill their blood over their young and lions sleep with their eyes open. Their occurrences in emblems are therefore referencialisible. Scholz uses the word 'Zeichenvergessenheit' to characterise Schöne's disregard for the emblem as a (constructed) sign (40, n. 79).

  Another aspect of Schöne's views that Scholz has criticised is the use of an emblematic ideal type to come to grips with the variety of emblematic genres and subgenres. Scholz considers Schöne's ideal type unhistorical and normative. What is needed is a clear distinction between text type ('Textsorte') and genre. The text type is a modern scholarly construct, the genre is a historic institution as perceived by contemporary readers and writers (2002: 271-302). Scholz then proceeds to give a four-level model for the analysis of the text type. Specific emblematic genres can be characterised by the choices made at each of these levels, and so can emblem theories. In a number of ground-breaking studies Scholz has applied his ideas to Roemer Visscher's *Sinnepoppen* (1979, 1984).

  In the Low Countries, without doubt the most prominent emblem scholar is Karel Porteman. His main contributions, among many others, are perhaps his introduction to Dutch emblem literature (1977), which includes a typology of emblem books, his edition of Hooft's love emblems (1983) and his study on the emblematic exhibitions held in Brussels Jesuit colleges (1996b). His successor in the Leuven chair, Marc van Vaeck, wrote a monumental dissertation about Van der Venne (1993) and published extensively about other aspects of the emblem.

  Utrecht University has a strong tradition in emblem research. Scholz did his main work on emblems while at Utrecht. William Heckscher worked at Utrecht as a professor of Iconology and inspired a generation of scholars with, among other things, his interest in the emblem (Vinken 1993). In the influential exhibition catalogue *Tot Lering en Vermaak* (To instruct and entertain, 1976) his successor Eddy de Jongh showed how Dutch seventeenth-century painters incorporated emblems in genre paintings. From the perspective of Dutch literature, interest in the emblem was shown by W.A.P. Smit's studies into the emblematic aspects of Vondel's dramas (1957). Van Gemert (1991) looks critically at Smit's and especially Schöne's views on emblem and theatre. Hans Luijten's edition of Cats's emblems was a Utrecht PhD-thesis (1996). The tradition is carried on by Arie Gelderblom, who edited Luyken's *Duytse lier*

(1996) and wrote extensively on his emblems (e.g. 2000) and Els Stronks, specializing in Dutch religious emblematics.[15]

2.3   Introducing Digital Emblem Studies

The emblem can assume many forms, it has many sources, it is related to many other cultural phenomena and it influenced the visual arts and other aspects of life. The interests of emblem scholars are correspondingly diverse. Some focus on emblems proper, others on the relationship between emblems and other genres; some are interested in national peculiarities, others on shared characteristics; some are interested in historical details, others prefer larger and theoretical questions.

The question of how emblem studies could profit from the work being done in humanities computing seems a natural one to ask. Apart from the more mundane applications of technology (word processing, electronic library catalogues, reference management), up to now the main application of technology has been in emblem digitisation, indexing and editing. In this book I will argue that digital technologies have applications in emblem studies beyond these 'preparatory' scholarly activities. The historic and interpretive research that the preparatory activities facilitate can also be supported by digital tools, and this book's main topic of interest is these tools' potential and significance.

Work in digital emblem studies could do a lot to facilitate the efforts of emblem scholars. It could indeed make available much more digital source material (emblems and related primary material). But it could also develop tools to support annotation of this material; to support the creation of online collaborative bibliographies; to support visualisation of the relations between books; to support manipulation of structured text; to support research into intertextuality in a corpus; to support the activity of comparison, both in direct interaction with the emblem scholar and 'on its own'; to support indexing, searching, manipulation and other forms of study of images; to connect to image banks; to support the use of multiple languages at multiple levels of linguistic expertise; or to assist in doing text analysis.

To my knowledge 'Digital emblem studies' is a field that has not been explicitly identified before. Digital emblem studies is emblem study that develops and employs digital technology. To a large extent this will be technology developed elsewhere, but put to good use for emblem analysis. To some extent, technology will be period-specific, for instance when the printed books of the early modern period are being analysed. In these cases, emblem scholars will want to collaborate with technologists and other students of the

---

[15] Together with Van Vaeck and PhD students Roggen and Dietz, she works on a project about the religious emblem tradition in the Low Countries in the light of Herman Hugo's *Pia Desideria*.

era. But for an analysis of specifically emblematic aspects of early modern material (the way text and word interact to form a meaningful whole) emblem scholars will have to take the lead – no one else will do it (Stronks 2007b: 153-54).

I will not speculate about a future in which some or all of this digital functionality is available. The body of the book will be devoted to annotation support (part II). In chapters 4 and 5 I will discuss a necessary condition for digital annotation, the electronic accessibility of emblem books. But the potential importance of an annotation facility may need explanation. In the next chapter, I will look at the concept of annotation and its significance for humanities scholarship.

3   Annotating the Digital Edition

3.1   Introduction

In this chapter I will introduce the concept of annotation, its present availability in digital editions, and its promises. Part III of the book will propose a number of annotation tools that implement this concept.

      Why annotation? Is annotation not something of secondary importance? I will argue it is not: section 3.4 in this chapter will show some of its promises, and a number of the following chapters will set out its potential significance. Historically, annotation has been very important in the development of scholarship. There were times when a great deal of scholarship consisted in glosses in the margins of other texts. In editorial work, annotation has always remained highly important, and auto-annotation (that is, placing footnotes in one's own texts) is ubiquitous. One of the functions of the present chapter is to create a view of annotation that is sufficiently wide to encompass the many forms it can take.

      What does this chapter contain then? Section 2 defines what I understand by 'annotation' and describes the kind of annotations to be discussed in this book. In section 3 I describe a number of annotation facilities presently used in online editions, while section 4 offers some perspectives on annotation from neighbouring fields or applications. Section 5 contains some final remarks.

3.2   Definition

In the scholarly edition, editors traditionally provide notes to give the reader the information that is necessary to grasp the author's meaning. This practice of annotation is described by e.g. Battestin (1981). For the purposes of this study I will use a wider sense of the word, where annotation subsumes note-taking and other activities that result in some form of commentary on texts. 'Annotation' can be defined from many angles. Most obviously, 'annotation' can refer to a process and to its result (MacMullen 2005). My interest will be in the annotation-as-result. For some, this result is primarily a physical thing (Brush *et al.* 2001): 'An annotation is a marking

made on a document at a particular place'. Others have focused on the content of the annotation (David & Robert 2006): '(…) a written, oral or graphic document usually attached to the host document'. Others again have stressed the relation between the annotated and the annotating resources (Bargeron *et al.* 2001): 'The (…) annotation object draws together a collection of anchors and encapsulates the annotation relationship among the resources they specify'.

For my purposes, I will define an annotation as a resource, brought to bear on another resource. The word 'resource' suggests an information science perspective, but annotations were around long before there was information science. I choose the word because it is sufficiently general to cover all things that may merit annotation: books, films, maps, but also objects like brain scans or the human genome. Similarly, annotation content is not limited to text; a sculpture representing Daphne can be considered an annotation to the *Metamorphoses*: it is a response to the literary work and in that sense an interpretation. The expression 'brought to bear on' makes clear that there are three ingredients in every annotation: (1) an *annotatum*, something that is being annotated, the annotation target, (2) an *annotans*, something that is predicated of the annotatum, the annotation content, and (3) an *annotator*, who brings the annotans to bear on the annotatum. The relation between annotation content and annotation target must be seen as meaningful in order for the annotation to make some kind of sense. An annotation is always made with an audience in mind, and later in this book we will examine how readers can use and explore annotations.

Annotations can be explanatory words in the margin of a text, or in footnotes, or collected together at the end of a work or even in separate volumes. Annotations can also be 'sub-textual', so to speak, like exclamation marks or underlinings. They can also be more than just text: a video clip of a scene from *Hamlet* can be thought of as an annotation to that scene. The assignment of a predefined set of categories (such as Iconclass) to a series of art works is also an act of annotation.

The list of purposes for which annotations are made is similarly wide. They can be made to help memorisation, to summarise, to assist understanding, to bookmark: all more or less private purposes. But annotation can also be used for very public purposes: to clarify a text, or an aspect of a text, or a theory about a text. One of the central tenets of this book is that annotations may also be *both* private and public, or begin life as a private annotation and end up being used to support scholarly argument. An annotation can be like a quotation in a Renaissance scholar's notebook: written down when he was a schoolboy and used in many publications later in life (Moss 1996). This is an example of the more general phenomenon that annotations can be repurposed. I will come back to this in section 3.4.

Not all annotation is motivated by scholarly purposes. People annotate train schedules and lists of holiday destinations, but these are out of scope for

this thesis. Very interesting but equally out of scope is the suggestion that in reading for pleasure of literary fiction, (digital) annotation may serve social exchange among readers (Chisik & Kaplan 2006). This book will be occupied with annotation of texts that are studied for scholarly purposes.

Usually, scholarly annotation will aim at the incremental creation of sharable knowledge. It requires some form of organisation for the annotations to be successfully managed and exchanged. I will employ the concepts of annotation types and sets for that purpose. An annotation set is a collection of annotations that are created or stored together for some purpose. A typed annotation is an annotation that makes explicit some of its meaning or purpose, using an annotation type that is repeatedly applied in the annotation set. In structured annotation the annotating resources share a recurring structure. This structure will facilitate processing (filtering, visualisation) that is necessary for helpful display of the annotations. In this book I will be concerned with sets of typed and structured annotations.

3.3   Online Annotation in Digital Editions

This section will discuss the present state of annotation facilities in the digital edition. What I am looking for is a presentation of annotation tools in the context of the text to be annotated, with the annotations displayed alongside the edited text. Out of scope are therefore external annotation tools that might be used in conjunction with the digital edition, such as Google Notebook, Diigo or Notefish.[16]

It should be possible to annotate not just at the text level, but at the level of text constituents as well – pages, paragraphs or other. Annotations may be private, shared or public. Adding annotations may require registration, perhaps open only to experts, or may be open to every visitor. The annotations may be subjected to a review process. They may consist of free text comments, of a simple tag, or have a more complex structure. They may be meant for the purpose of explanation, discussion, transcription, translation, or for other purposes. My main focus of interest is in annotation in scholarly text editions, but it is hard to define these terms ('scholarly', 'text', and 'edition') precisely, and there is no need to do so here. I will limit the overview to online interactive facilities, and so exclude annotation facilities available with CD-ROM editions.[17]

---

[16] http://www.google.com/notebook, http://www.diigo.com/, http://www.notefish.com/.

[17] This is not to say that there are no CD-ROM editions that have remarkable annotation facilities. Most CD-ROMs, like most websites, have no annotation features, yet there might still be more appealing annotation functionality available on CD-ROMs than on websites. To give an example from the emblem field: on the Wither and Rollenhagen CD-ROM produced by Studiolum (Daly & Young 2002) users can create notes, named bookmarks and multiple varieties of (named) highlighting. In the CD-ROM Bible published

Wikisource and Wikibooks

The first annotation platform that I want to discuss here is that provided by wiki software. The reason for this is not that wiki software should be especially suitable for serious annotation, but it is because wiki has captured the imagination of many. When I try to explain to non-scholars what annotation is and why it is important, the wiki is their point of reference – 'ah, so what you want is a kind of wiki!' And then I find myself explaining that, yes, I want a wiki, but what I really want is something that can do a little bit more than the free text that wikis facilitate.

A wiki is a publication system where multiple people can collaborate in publishing web pages (Leuf & Cunningham 2001). A simple markup-system, sometimes hidden by a wysiwyg editing system, allows non-technical persons to create and modify pages. Typically, wiki pages are edited collaboratively, allowing editors to extend earlier content of a page. The best known wiki is Wikipedia, but wiki systems are used as knowledge management systems in many organisations. In the field of digital humanities e.g., the TEI and Digital Classicist communities have their own wikis.

A second reason why I feel that a discussion of wiki software should open this section is the vitality and energy of the wiki community. Academic use of annotation facilities tends to be experimental and very careful and seems more worried by potential risks than excited about possibilities. The wiki community, so far predominantly non-academic, tends to act rather than to worry, and may help academics imagine the kind of environment that would help them do their work (Kuiper 2009).

Among the sister projects of Wikipedia are Wikisource and Wikibooks. Wikisource describes itself as 'an online library of free content publications collected and maintained by the community', Wikibooks is a 'community for creating a free library of educational textbooks that anyone can edit'. Both Wikisource and Wikibooks can accommodate editions with annotations. In the case of Wikisource, the primary object will be the edited text, to which the annotations are supplementary material. In the case of Wikibooks, the new intellectual content will be the annotation, to which the annotated text can be provided as reference material. The usual case in what Wikibooks names 'annotated books', however, is that the annotated text is only present as a

---

by Importantia (2004) readers can create verse-based commentaries, named bookmarks, and multiple named verse lists. This last facility is worthy of note, because it shows structure. Verse lists are not just text, they consist of entries that can be handled by the computer. This sort of structure is what the annotation tools to be discussed in this book should facilitate. However, with web access set to become ubiquitous, it seems reasonable to expect the CD-ROM to gradually disappear as an edition platform. Perhaps more importantly, the CD-ROM is a closed medium. Annotations on texts cannot be shared with students or (present or future) colleagues. From my perspective, the CD-ROM is ultimately a dead-end street.

reference.[18] Such is the case in, e.g., the annotations to *The Complete Peanuts*.[19] Annotations to a number of chapters of *The Brothers Karamazow*, on the contrary, are displayed next to the texts.[20]

On Wikisource, the focus is on providing editions of existing texts. These editions may be annotated, a mechanism that seems most popular in the German Wikisource site. Sometimes notes from an existing edition are included, as in the *Zimmerische Chronik*,[21] but often both transcription and notes are original, as with *Von abtuhung der Bylder*.[22]

Wiki software can be used outside of the realm of the Wikimedia foundation, the organisation that operates Wikipedia and its sister projects. There is an extensive wiki site devoted to the annotations of novelist Thomas Pynchon's works, as there are sites devoted to Terry Pratchett, to Eco's *The Mysterious Flame of Queen Loana*,[23] and no doubt to many other authors and works. As these sites do not contain the text being annotated, they are somewhat outside the scope of this section, but much of this work is serious and inspiring. It shows a willingness to create and share knowledge that is essential for social annotation to become a success. Academic use of the wiki platform is taking off in genomics (Marx 2009) and in the 'open notebook science' movement in chemistry (J.-C. Bradley 2008a). Humanities scholars should follow their lead.

Weblogging and CommentPress

Weblogs are publication systems that are typically used to publish series of short, diary-like articles, usually presented in reverse chronological order. Most weblog systems have a comment facility that readers may use to comment, communicate with the author, etcetera. While weblogs are not usually employed as an edition platform, weblogs are home to a significant volume of today's primary writing. Weblog comments themselves are, in a psychological sense, more 'primary' than the annotations one expects from a scholarly editor, in that they are often written on the spur of the moment and show personal involvement. Still, the ubiquity of the comment form, not just on blogs but on a plethora of news sites,[24] cannot but influence what we will expect of the future's digital edition.

---

[18] I base this statement on an inspection of mainly the English, Dutch and German versions of Wikibooks in February 2009.

[19] http://en.wikibooks.org/wiki/Annotations_of_The_Complete_Peanuts

[20] http://en.wikibooks.org/wiki/The_Grand_Inquisitor

[21] http://de.wikisource.org/wiki/Zimmerische_Chronik

[22] http://de.wikisource.org/wiki/Von_abtuhung_der_Bylder

[23] http://pynchonwiki.com/, http://wiki.lspace.org/wiki/Annotations, http://queenloana.wikispaces.com/.

[24] General purpose sites such as Topix (http://www.topix.com) or specialised sites like Slashdot (http://slashdot.org/).

A number of edition projects make use of weblog functionality, and some of these allow readers to annotate using the standard weblog annotation facilities. An example is the publication of George Orwell's diaries, one day at a time, by the Orwell Prize (s.a.). The entries are heavily commented by visitors, sometimes rather facetiously. The comments to Samuel Pepys' diaries (Gyford s.a.) are by and large more serious. The Pepys diaries are embedded within a larger site that houses an encyclopaedia and a section of in-depth articles on Pepys and his age, which users can contribute to.[25]

A weblog plugin created especially to facilitate paragraph-by-paragraph signed comments is CommentPress (Fitzpatrick 2007). CommentPress is a 'theme' for the popular blog tool WordPress. To each section and paragraph it adds an icon that readers may use in order to add comments. In the right sidebar CommentPress adds a display of recent comments and a menu showing other options for browsing comments, such as 'by commenter' and 'by section'. CommentPress has been used as a discussion platform, e.g. in the discussion of the Scholarly Publishing Office's *Ithaka* report (2007). It has been used in an educational setting for the annotation of Melville's *Pierre, or, The Ambiguities*.[26] Martin Luther King's *Letter from Birmingham Jail* has been published and analysed using the CommentPress platform.[27]

One of the noteworthy things about this is that the success of the WordPress weblog platform is what makes CommentPress a useful tool. If we had a shared publishing platform for digital editions, it would be worthwhile to develop annotation facilities for that platform. In the absence of such a platform, edition projects must either develop annotation tools themselves or forego annotation functionality.

Scholarly Edition Projects

Unlike the wiki or weblog approaches, the projects that I will discuss in this subsection function within a more traditional editorial background.

The Virtual Humanities Lab at Brown's University (Zafrin & Armstrong 2005) is a site where scholars of the Italian Renaissance collaborate on a number of important texts from the period.[28] The project conceives itself as a 'virtual editing house' where scholars will have the possibility to add semantic encoding and annotation to the edited text. 'Semantic encoding' includes marking references to persons, terms, and themes. Zafrin and Armstrong also mention

---

[25] Another example of an interactive weblog-edition: *Defoe's Review*, edited by Christopher Flynn (s.a.). A non-interactive example: the war diary of Virginie Loveling, edited by Bert Van Raemdonck (2004).

[26] Site no longer available.

[27] Site no longer available. The Wayback Machine at the Internet Archive provides proof that the page was once there: http://web.archive.org/web/20080130140520/http://www.futureofthebook.org/nateste arns-/letterfrombirminghamjail/.

[28] http://golf.services.brown.edu/projects/VHL/index.php

structural elements, quotations, rhetorical devices, etc. It should be possible to annotate text as well as other annotations, possibly even in a forum. Even though this is clearly an experimental site, it is traditional in the sense that annotations can only be made by established scholars with privileged access. The annotations they create are accessible to the public.

Another annotation system was developed for the Online Chopin Variorum Edition (OCVE) (J. Bradley & Vetch 2007).[29] The OCVE will house facsimiles of multiple editions and manuscripts of Chopin's piano works. Students of the edition can attach annotations to locations on the facsimile. Annotations can be assigned to categories and grouped into trees. This provides a measure of support for a process of discovery through exploratory annotation (J. Bradley 2003). Annotations can be either private or public. Like the system developed for the Virtual Humanities Lab, the OCVE's system is still at an experimental stage. Even though it was mainly used by the editors, it was developed with the edition's end users in mind. But, as in the VHL's case, the edition has not yet reached the stage of end user usability.

The publication platform Telma (Bertrand 2006) was developed by the French IRHT and the École Nationale des Chartes, with the specific goal of publishing medieval charters.[30] Telma contains limited facilities for annotation and bookmarking. The facilities are open to all registered users, and anyone can register. All annotations are private. The annotations are not shown with the annotated texts, but can only be accessed through a separate menu. There is no possibility for further organisation or categorisation of the notes.

Digital Library Projects

The number of scholarly edition projects that offer annotation facilities is surprisingly, and disappointingly, low. The digital library community however hardly does better, even though the appropriateness of annotation support in the digital library is uncontested (Gazan 2008; Waller 2003). A number of sites offer annotation facilities at document level,[31] which is not enough for our purposes. Digital collections that are created using ContentDM software contain a facility for bookmarking pages. The collected bookmarks are private.[32] In some of these collections, users can create notes to the digital objects. The notes are public, and there is no way one can ask for all the notes one has created. Arko *et al*. (2006) describe a system for shared annotation of items in the Digital Library for Earth System Education (DLESE). The DLESE system is mainly oriented towards recommending items to other users of the library. In the National

---

[29] http://www.ocve.org.uk/
[30] http://www.cn-telma.fr/
[31] Public 'reviews' in the Internet Archive (http://www.archive.org/details/texts), private labels and reviews in the 'My Library'-facility in Google Books (http://books.google.com/).
[32] See for an example the edition of Phillis Wheatley's *Poems on various subjects, religious and moral*, http://www.sc.edu/library/spcoll/wheatley/wheatleyp.html.

Library of Australia's newspaper digitisation program, users can tag articles and create notes. Notes can be private or shared. One can ask for a list of all the tags or comments one has created. Users can also help correct the OCR-ed text of the newspaper.[33]

A test version of the new system for Gallica, the French National Library's digital library, goes a little further, and offers bookmarking and tagging at page level.[34] The bookmarks can be grouped into folders. The tags can be used to retrieve the corresponding pages. This goes some way towards developing the digital library into a research tool. All bookmarks and tags are private however. They cannot be shared with other researchers.

The annotation system employed in the digital collections of the University of California (UCLA) is more powerful.[35] The reader can create proper notes, not just tags, that are shown next to the page. The annotations and bookmarks can be stored as virtual collections, that can be either private or public. It is even possible to allow others to edit the same collection. A researcher could use a virtual collection to store pointers to the locations in a book that are relevant to an article he or she has written. A teacher could add clarifications to a number of pages he has assigned to a class. A researcher might even use the facility for in-depth annotation of an entire work, although page level annotations may not be the best means for doing that.

Collaboratories

Depending on one's strictness in defining scholarly text editions, a number of other sites might be mentioned here. Among these are a number of 'collaboratories', sites that facilitate collaboration among scholars in the study of, usually, documents. Annotation is an essential facility for collaboration, although successful collaboration requires more. Examples are Collate, a film collaboratory (Frommholz *et al.* 2003) and IPSA, a collaboratory for the study of botanical illustrations (Agosti *et al.* 2003). No doubt one of the most ambitious collaboratories is HyperNietzsche (D'Iorio 2000). Here annotation is only an ingredient in a site facilitating genetic study of Nietzsche's manuscripts.[36]

A very remarkable site is Footnote.[37] Footnote contains historic material, mostly from the US National Archives and other agencies. Visitors to the site can add what is called 'annotations' – basically transcriptions of names, places, dates and other text, and free-form comments. They can create 'spotlights' – essentially a way of highlighting and commenting images or image

---

[33] http://ndpbeta.nla.gov.au/ndp/del/home.
[34] http://gallica2.bnf.fr/.
[35] See for an example the Canon Law Digital Edition at http://digital.library.ucla.edu/canonlaw/.
[36] HyperNietzsche is being replaced by NietzscheSource (http://www.hypernietzsche.org/).
[37] http://www.footnote.com/.

regions. References to spotlights can be embedded in external web pages. Another way for visitors to add information to the collection is to connect images. Connections can be typed: the system provides a number of standard types, but the visitor can create his or her own connection type, and can even further annotate the connection. Visitors can also upload new material. Perhaps the most interesting way in which visitors can contribute value is by the creation of 'story pages': pages that tell a story around images from the Footnote archives and possibly other, uploaded images. Footnote does many things that would be useful in digital editions. Footnote is, however, a commercial venture, and not all of the content on the site is freely accessible.

Another commercial collaboratory is Crossroads (Nunn 2008). Crossroads is software available in conjunction with Readex's Archive of Americana. Users can create projects that can be private or visible to others. Within projects, users can create bookmarks, highlights in documents, annotations and reading lists and decide whether they want to share these with other Crossroads users. It is clear Readex has developed important functionality. The fact that the annotations created within Crossroads cannot be accessed from outside makes it less interesting from a scholarly perspective.

Summing Up

It seems self-evident that an integrated annotation facility, even if the annotations are only accessible to the annotator himself, can be very useful. In spite of this, the number of digital editions, however one defines them, that offer annotation facilities is surprisingly low. Where sites provide annotation facilities, they often hide the annotations in separate windows, making them hard to access and use.

The question why annotation facilities are so rare is hard to answer. From the perspective of the researcher, who would entrust his research data to an unknown party, a certain amount of reserve may be understandable. The provider of the digital edition, however, should have no such qualms. No one seems to have argued that annotation facilities are pointless. The answer must be that budgets are short and annotation facilities feel like an 'extra', something nice to have but not essential when a deadline looms.

3.4   Perspectives on Annotation

The amount of annotation functionality available in digital editions at present is less than impressive. Thus there is ample opportunity to look again at what annotation might be and what annotation tools might accomplish in the domain of literary studies. I will accordingly describe a number of annotation tools and studies from other fields: media studies, linguistics, spatial hypertext, qualitative social science research, semantic web studies and digital library studies. A brief tour of some of the surprising uses of annotation outside the field of literary

studies may help prevent an unimaginative re-creation of the traditional edition's footnotes in the new medium. I have looked for ways to widen the scope of annotation with respect to purpose, form, structure, application and audience. At the end of this section I will look at what these 'foreign' perspectives, from possibly off-line tools, might suggest for the future of literary studies if conceived of as an on-line and interactive field of inquiry.

(a)   Annotation as Creating a Selection: Generated Video Documentaries

Bocconi *et al*. (2005) describe a project where annotations on video fragments facilitate the creation of small documentaries consisting of related fragments. The video fragments that the researchers use are taken from a collection of interviews with US citizens about the war in Afghanistan. The annotations express opinions, in the form of sentences consisting of a subject, a modifier and a predicate, such as 'military action (subject) never (modifier) effective (predicate)'. Using linguistic knowledge about relations between words (opposition, similarity, specialisation, generalisation), the collection of annotations is searched for statements related to the statements made in a given video fragment. From the resulting set of fragments, documentaries can be generated based on their rhetorical relationship (support, contradict, warrant etc.) to the given fragment. It is a tempting thought to use this technology in a study of emblem book organisation. Can we see patterns in the ways emblems are grouped? Do these patterns provide rhetorically convincing arguments? This example highlights three points of interest:

1. structured annotation is powerful. Because of the sentence structure of the annotations it is possible to find related statements and to specify the type of relation between fragments. This type is essential to the selection of fragments that are appropriate for inclusion in the generated documentaries;

2. annotation is a starting point for further processing. What this annotation processing makes clear is that not all annotations are meant to be read by humans only, and the design of annotation tools should be informed by that;

3. annotation facilitates (advanced) selection. Just like simple categorisations facilitate simple selections (which would be facilitated by e.g. tagging the interviews as pro- or anti-war), structured descriptions facilitate complex selections.

(b)   Intensive Linguistic Annotation

The most intensive use of annotation is probably made in linguistics. Several projects have worked on platforms for sharing and reusing annotations. One of the more successful was the EU-funded NITE project. Carletta *et al.* (2005, which reprints a paper circulated in 2000) sets out the project desiderata. It argues for the need for a platform that supports reuse of linguistic data, and that can be

used by linguists, psychologists and humanities scholars. The platform should be integrated, structured, but not monolithic. Support for specific annotation structures is important: 'Software developers attempting to write usable interfaces to the data need to understand the structure which users will be expecting to see. Data sets where this structure is not reflected in the data manipulation and display methods imposes an extra burden on the human because it requires that information to be carried in the head'. The annotation platform should use an inspectable file format. The format is provided by the XML recommendation, published in 1998 (Bray *et al.* 1998). Stand-off annotation avoids the limitation of a single hierarchy.

The resulting data model is introduced by Carletta *et al.* (2003), who also describe the software developed to handle linguistic annotation. Elements (annotations) have a type, have attributes, children (other elements), timing information and features (a named relation to another element, such as the antecedent of an anaphor). Data are stored in groups of heavily interlinked XML files. The software provides an API for loading, saving and manipulating the data. A query language is defined by Evert & Voormann (2002).

With the help of these facilities, layers of particularly dense annotation can be created, and, more importantly, queried efficiently. The queries can incorporate the hierarchical structures implied by many annotation types (e.g. syllables contained in words) but can also relate multiple independent annotation categories, such as prosodic annotation and part-of-speech-tagging (Gut *et al.* 2004).

Some of the interesting characteristics of this application of annotation are (and I continue the list from the previous example) :
4.   any individual annotation is irrelevant. Displaying one annotation would be pointless. It is only by querying the entire annotated corpus that one learns something. Facilitating retrieval is the only point of this type of annotation;
5.   again, formalised structure is necessary. Linguistic analysis requires placing elements in complex hierarchies that must have a representation that the computer knows to handle; and
6.   finally, all of these annotations are shared. They are created to be queried. In that respect, they are diametrically opposed to the exploratory annotations made possible by the next group of annotation tools.

(c)   Personal and Light-Weight Annotation Using Workspaces

Exploratory annotation is what VIKI and VKB (Visual Knowledge Builder) workbenches facilitate. The central idea in VIKI, and later VKB, has been the possibility to record and manipulate ad hoc, partial, or ambiguous interpretation. As Marshall wrote: 'I can change a document to red to signal to myself that it's important, or that I need to come back to it, or that it's about

Greece. It is these lightweight classifications that we need to reclaim in our digital library reading machines' (1998b: 7). For Marshall, that is the 'truly annotative' property of VIKI. In VIKI, users manipulate visual properties of objects (colour, location, border) in order to organise their knowledge of a domain. The ease with which these manipulations can be made contrasts with the cumbersome process of creating traditional hypertexts. This is especially important in situations where there is too much information or where circumstances evolve rapidly. VIKI was developed with the 'emergent nature of structure' (Marshall *et al.* 1994: 13) in mind. It avoids the need to commit to structure prematurely.

The task that VIKI is meant to support is interpretation, defined as 'that part of writing, collaborating, or thinking in which people collect the materials of interest and make sense of them in the light of their task and the background they bring to it' (1994: 15). Users create structure in the material by manipulating its appearance or by explicit grouping into types, collections and composite objects. VIKI software recognises some of the structures implicit in the groupings the user created and uses the recognised collections to suggest explicit structure to the user. It thus emulates the interpretive structures arising 'through filing, in conversation, in juxtapositions of documents on one's desktop, in notes and annotations, in markings and marginalia' (15).

From the point of view of literary annotation, a shortcoming of this system is that all annotation exists at object level. There is no way to attach information to a span of text within an object. The tools to be discussed in part III of this book do facilitate that. But there is much in VIKI that is worth noting: the exploratory tagging, the system-assisted creation of structure, the way the system exploits one's visual and spatial memory, the natural extension of one's pre-digital working methods. Bradley (2008b) records the many debts his thinking on annotation, reflected in his annotation tool Pliny, owes to VIKI.

Among the additions in VKB is more support for publication of workspaces (Shipman III *et al.* 2001). A number of presentational enhancements help workspace authors create more accessible workspaces. Links into workspaces should help build a web of what is known as 'spatial hypertext', and thus facilitate annotation sharing and repurposing.[38] It is not clear, however, whether the informal and exploratory nature of VKB's structures will lend themselves to wider sharing. Marshall (1998b) argued for the repurposability of annotations. In a different context, Marshall & Brush (2004) found a number of differences between personal and public annotation, which they explained arguing that '[u]sually personal annotations reflect unselfconscious reactions to reading material, while public or shared annotations on a document reflect specific communicative intent'. This suggests a fundamental difference between personal and public annotation.

---

[38] In contrast to the hypertext of today's web, that consists of nodes and links.

In this context, noticeable features of VIKI and VKB are:

7.   annotation can be exploratory, tentative;
8.   annotations are not necessarily expressed in words: they can use colours or geometric shapes and similar visual cues to assist understanding;
9.   annotation can help the process of interpretation.

(d)   Coding in Social Science Qualitative Data Analysis

Literary research is, to a large extent, a close investigation of texts. That is also what many social scientists do, when they analyse interviews, first person narratives and other texts. There is quite a number of computer programs that assist them in doing that analysis. Among these programs are Kwalitan, NVIVO and Atlas.ti.[39] Programs such as these may be relevant to the humanities, as remarked before by e.g. Bradley (2003) and Carletta *et al*. (2005).

The most important function of these programs is to assist in the coding of text fragments (Boeije 2005). The user creates short descriptive codes and assigns these to text fragments. Dependent on the research methodology, the codes may be determined beforehand or may be chosen based on an open-minded reading of texts. The codes can be organised into one or more code trees. Memos can be attached to codes, to texts and to fragments for discussion, explanation or questioning. Memos can sometimes be typed. Sometimes colours can be assigned to codes in order to help understand e.g. the thematic structure of texts, but also to facilitate scanning of texts for assigned codes. Besides these facilities for visual inspection, there are also simple and more complex retrieval and search facilities that help find text fragments based on assigned codes, content, and context. Context can to some extent be formalised in terms of text-level attributes: for interviews, these attributes could be used to specify age and sex of the interviewee, date of the interview, etc.

The coding of texts and the subsequent retrieval of texts based on these codings are steps in the formulation of a theory. Some programs also assist in the formulation of such theories, in the form of diagrams and networks that visualise the relations between codes or between the concepts or entities they may be thought to represent.

Codes and memos are annotations. It is interesting to see how the creation of exploratory but intensive annotations can help in the creation of detailed understanding. Whereas the VKB and VIKI workspaces supported what Marshall (1998a) would describe as extensive annotation, the tools I discuss here are geared towards intensive study of a number of texts. The purpose of such studies is to arrive at a theory or model of (some aspects of) these texts. One reason why social scientists use tools such as these, and humanities scholars (by and large) do not, may be that socials scientists are trained in

---

[39] http://www.kwalitan.nl/, http://www.qsrinternational.com/, http://www.atlasti.com/.

methodology. They are taught that there are valid ways of arriving at knowledge, and that research publications should demonstrate that the publication's conclusions were reached following a valid path. The programs discussed here help the researcher to follow such a path and help him/her in reporting about it, as the details of the analysis are easily accessed and retrieved. Humanities students (again, by and large) get no such training. Bradley (2008b) notes the dearth of publications about research methods in the humanities. As McCarty argues, the humanities often feel suspicious about explicit methodology (McCarty 2007).

Remarkable about this application of annotation is:
10.   it is geared towards the creation of a model. As in the workbenches mentioned in the previous example, initial annotation may be explorative, but here the aim is to arrive at a scholarly theory;
11.   to argue for a need of annotation tools such as these, is not just to argue for software support of current practice of work in the humanities, but also to argue for different practices, a subject to which I will return in the final chapter of this book.

(e)   Annotation as the Application of a Semantic Model

A good deal of annotation research has been done in semantic web studies. The aim of these studies is always to make web data machine-readable. A number of researchers have experimented with annotations that describe events or actions in images or stories. Schreiber *et al*. (2001) describe an experiment in photo annotation where the image content is annotated using, among other things, sentence templates consisting of agent, action, object, and setting (e.g., 'orange ape', 'eating', 'banana', 'at dawn'). Similarly Hyvonen *et al*. (2007) and Junnila *et al*. (2006) use thematic roles in the annotation of actions. Actions are related to events, that in turn are associated with a time and place. To model the development of a narrative, events are connected using a separate process ontology. The description can be expanded by associating themes and moods, and possibly other attributes. The authors use this model in the description both of scenes in the *Kalevala*, the Finnish national epic, and of cultural heritage artefacts such as paintings, photos and museum objects.

The objective for both groups of researchers is to facilitate searching. Simple descriptions using keywords are not powerful enough to limit search result sets in larger collections and they cannot facilitate retrieval of related objects, e.g. paintings that depict a given scene. The sentence templates or event-based descriptions relate a number of aspects that can then be used in focused queries. In both cases, the context is enhancing the accessibility of cultural heritage.

The event ontology provides a (simple) model for the description of events in the corpus. In the application that the articles describe, the knowledge that will be stored using the model will consist of presumably uncontested and

accepted facts. There is no reason, however, why the technology should not be used to formulate and store new research findings. Scholars interested in any aspect of a text, not just narrative, could create a model for the field they are interested in and use the model to annotate the texts they investigate. The annotations will then provide the basis for exploring and querying the text and the research findings. This will help both the researcher himself/herself and his readers in assessing the value of reported findings.

Noticeable about this type of annotation is:
12.   annotations can be based on complex ontologies and are used to make explicit structured semantic content of cultural artefacts;
13.   it supports search functionality for a wider audience than just fellow scholars.

(f)   Annotations Supporting Social Navigation

A number of researchers have investigated the value of readers' annotations in guiding (future) readers. Farzan & Brusilovsky (2005) created navigation hints based on annotations users had made in educational material. Users liked the annotation facilities and navigation was shown to be affected by the hints. Users also seemed to use their own annotations as a basis for revisiting pages. Shipman III *et al*. (2003) found evidence of correlation between passages annotated during preparation of a paper and passages cited in the paper. Bradshaw & Light (2007) found significant overlap in annotated passages among annotators, warranting development of a system for reading recommendations based on annotated passages.

Use of annotation for social navigation support is part of a wider trend. In a number of bookmark services users learn from each other's tagged bookmarks. Del.icio.us was the first service of this kind,[40] citeULike is a similar service targeted at academic papers. In a sense, when YouTube shows the 'most discussed' or 'most viewed' videos it does a similar thing: it uses earlier readers' (or viewers') choices in guiding new readers. But the phenomenon is really ubiquitous: on-line stores show book recommendations based on buying behaviour ('people who bought this book also bought ...'), LibraryThing has a list 'People with this book also have...', CNN lists 'stories you might be interested in based on past viewing', etcetera.

For scholarly work, recommendation features seem most suitable to secondary literature. Citations and citation numbers already are a form of recommendation indispensable in scholarship. References to literature collected by Zotero (a scholarly reference management tool, built as a Firefox extension) may in the future provide another popularity count (Cohen 2007). It certainly would not be impossible, however, to do similar things with scholarly

---

[40] Sites mentioned in this paragraph: http://delicious.com/, http://www.citeulike.org/, http://www.youtube.com/, http://www.librarything.com/, http://cnn.com/.

annotations to primary literature. One could guide readers to the most heavily annotated chapters in e.g. the *Divina Commedia*, or perhaps even visualise how these annotation patterns have changed over time. The focus of interest would then move away from the primary text to its reception in scholarship. More generally, the phenomenon of quotation can be seen as one book recommending another, or recommending certain passages in another book. Popular passages, in that sense, certainly deserve attention, and generated annotations that reflect that popularity would be a very useful research tool.

Some of the noticeable aspects of this final example are:
14.   annotations function as recommendations: they show interesting objects and locations, and it therefore makes sense to count them;
15.   annotation patterns can be used to create profiles of persons and possibly of research interests.

Discussion
The technologies and investigations that I have discussed here exemplify different approaches to annotation and have different functionalities. In literary studies they might support different phases of scholarship, or facilitate different types of scholarship. Scholarship as a social activity would benefit most from the 'social navigation' approach (f). The other approaches are not explicitly social, though all (except, possibly, VIKI) would result in annotation data that others should be able to inspect. The linguistic annotation (b) and semantic model (e) approaches result in structured annotation. This facilitates navigating and querying the data in ways that are still quite unusual in humanities scholarship. The video annotation tool adds sophisticated processing to generate documentaries. Here we move towards fields like automated storytelling. Structured annotation will be the main issue in the tools that I will discuss in part III of this book. Part V will examine what may happen when these structured annotations can be explored, or inspected in detail, by other researchers.

        Some aspects of these approaches will not come back in the rest of this book. One of these is the usefulness of automated creation of annotations. The cases of linguistic annotation (b) and semantic modelling (e) would certainly profit from (semi-)automatic annotation. Natural Language Processing (NLP) software would help in automated annotation and would also help in the further processing of annotations, as in (a). Another aspect that I will not come back to is the spatial organisation of annotations proposed by VIKI (c) and the social science QDA tools (d). The motivation for leaving these aspects out of this thesis is pragmatic: no doubt these are important aspects of annotation technology, but the book is full enough as it is. They are issues for future research.

3.5   Conclusion

This chapter introduced the concept of annotation and looked at its applicability in the context of the digital edition. It defined an annotation as a resource, seen as meaningfully related to another resource. It singled out sets of typed and structured annotations, which are the topic of interest for this book. It then surveyed the annotation facilities presently available in digital editions. The main result was that these facilities are conspicuously missing in most editions, or are still at an experimental stage where present. The most innovative and inspiring use of annotations is outside the scholarly domain, in the wiki context and in Footnote, or in a commercial product, as in Crossroads.

For inspiration I then turned to a number of annotation tools and studies further removed from the domain of literary studies. Annotation, it appears, can take many forms. It can be public or private, extensive or intensive, structured or unstructured, formal or informal. In some cases, structured annotation based on complex models is necessary, in other cases digital annotation can be limited to the equivalent of a series of exclamation marks in the margin. In some cases, annotations represent the growth of understanding, created in the very process of annotating. In other cases annotations apply existing categories to a new corpus in order to make it accessible to queries. And while it is not necessarily useful to share exclamation marks, it seems useful (even if not current practice) for fellow researchers to be able to access the coded files that underlie research publications in social science, and it is self-evident that linguistic annotations are created for shared use.

There is a marked contrast between the promises that annotation facilities hold, and the at best skimpy annotation facilities now actually available in digital editions. It would seem a worthwhile attempt to re-create annotation facilities that include some of these promises, in the context of the digital edition. In part III of this book I will present three different approaches to structured annotation in the digital edition. Not all of the possibilities that were mentioned in the previous section have been incorporated in the tools I will discuss. In the book's final chapter I will compare the promises sketched in this chapter and the results given in what follows.

Part II. Emblem Digitisation and the Emblem Project Utrecht

4   A Model for Digital Emblem Editions

The Emblem Project Utrecht (EPU) was created by Els Stronks and the present author, in 2002.[41] In the period 2003-2006 the project received funding from NWO, the Netherlands Organisation for Scientific Research. The project's aims were, on the one hand, digitisation of the core corpus of Dutch love emblem books, and on the other hand the evaluation and development of techniques for the digitisation of and research into literary heritage material. The project is one among a handful of projects that have worked on emblem digitisation.

The sites of emblem digitisation projects that have been created over the last fifteen years can best be characterised as thematic research collections. Thematic research collections have been described by Palmer (2004) as thematic digital collections that aim to support humanities research, that are usually coherent, structured but open-ended, often created by scholars, and often supporting interdisciplinary work. The concept is more modest than the concepts of either the 'work-site' advanced by Eggert (2005), which is a site that houses transcriptions and tag sets that embody multiple interpretive renderings of a text based on non-intrusive technology, or the 'knowledge site' proposed by Shillingsburg (2006), a site that houses not just texts, facsimiles, new editions, adaptations and translations of a work, but also the functionality for users to add new mark-up, variant texts, emendations and annotations. Stronks (2007b) has rightly proposed the 'knowledge site' as a model for what the Emblem Project Utrecht aspires to be, but the thematic research collection seems a more adequate description of what it is right now.

Though many of the emblem sites correspond to the definition of thematic research collections, their contents, structure, functionality and background are very diverse. If we accept Sahle's definition (2003: 76) of what an 'edition' is ('die erschliessende Wiedergabe von historischen Dokumenten', a representation that opens up historical documents), these sites are all collections of editions of emblem books. The ways the books are represented and made accessible on the web, however, are very dissimilar. The purpose of this chapter is to sketch a model for digital editions, which will allow us to

---

[41] For an overview of the project, its aims and result, see Stronks & Boot (2007a). For an external review, see Daly (2007). For technical background, see Tilstra (2007). In this chapter, I rework material from two earlier papers about the project (2008c, 2008d) to develop a model of digital emblem editions (papers published in 2008, but written in 2004 and 2003 respectively).

discuss these different representations. A model, in this context, is a coherent vocabulary for discussion of the aims and set-up of the editions and a set of parameters describing the choices that edition designers have to make. I will discuss the Emblem Project Utrecht and the other emblem sites in terms of the model; thereby hopefully both clarifying the model and providing insight in the choices that have been made in the EPU.

The model that I propose consists of the following dimensions: (1) a basic orientation: does the edition aspire to be a facsimile, a scholarly edition or an index; (2) a set of generic functionalities: searching/selection, hyperlinking, comparison and annotation; (3) a set of parameters; and (4) the technical infrastructure of the site. A look at the origins of emblem digitisation, in the next section, will serve as an introduction to the model.

The references to emblem digitisation projects are followed by numbers between square brackets. These numbers refer to the list of projects in the appendix of this chapter.

4.1   The Origins of Emblem Digitisation

The first to write on emblem digitisation was David Graham, reporting on his Macintosh Emblem Project (1991b, 1991a). Graham used the Macintosh Hypercard software to create a database containing texts and pictures of four sixteenth century French emblem books. His articles describe how he was able to create a searchable and hyperlinked collection, where users could add notes, create links between emblems, describe motifs on pictures, and create private collections of emblems. Graham (2004) reviews this work, and stresses later developments that made further progress in digital emblem studies possible: open standards for text and pictures, cheaper hardware, and the arrival of the internet. Yet, Graham's early articles show an acute awareness of the digital emblem scholar's needs, and he lists a number of requirements that even now have been only partially fulfilled: (i) user added links and notes, (ii) link typing and description, (iii) link bidirectionality, (iv) complex searches and ad hoc concordances, (v) views of the distribution of search hits in a book (or corpus), (vi) site maps, (vii) diagrams that visualise the links that have been created between nodes, and (viii) user-created working sets or collections of emblems (1991b).

Many of the issues that were to arise in emblem digitisation projects[42] had already been discussed earlier, viz. in a 1978 symposium organised by Peter Daly in preparation of the *Index Emblematicus* (Daly 1980). In 1980, Daly already foresaw that the computer might remove the limits that print imposed: 'However, both computer technology and to a lesser extent publishing methods

---

[42] Such as the need to index secondary motifs, the permissibility of interpretive indexing and the need to index commentary texts.

are progressing and changing at such a rate that presently undreamed of modes of dissemination may make the [*Index Emblematicus*], as an information system, completely accessible in the future' (versus partial publication in print). The ambitious scope of the planned index was further enhanced by plans for an *Index of Emblem Art*, discussed a decade later (Daly 1990). The aim of the proposed *Index of Emblem Art* was to index not just emblem books, but all occurrences of emblematic motifs, in books of imprese but also in other works of literature, in manuscripts, and all objects of art, though it was well understood that some limitations would be necessary. The ideal of comprehensiveness ('it appeared that the only solution was to index absolutely everything') is a frequently returning dream in emblem digitisation.

The indexes discussed in these symposiums were conceived of as finding aids, not as substitutes for consulting the emblems books themselves. Even less, of course, they were designed as working and writing environments. The emblem sites that are their closest descendants are probably München and La Coruña. The insistence on the need for a single model and a set of principles to describe the entire corpus is important, even though this single model may be impossible to attain. It is a pity that in his later ideas on emblem digitisation Daly seems to have retreated into a position that stresses problems and limitations rather than the very real gains that have been made and the possibilities they open up (Daly 2002, 2004).

The 1980 volume goes some way towards developing a model for the index entries. It is in effect a data model for an emblem site. Other data models were developed by López Poza *et al*. (1996) and later by Stephen Rawles (2004). The data model determines what data can be stored about each emblem, and is an ingredient in the more encompassing site model.

4.2   Basic Orientation

There seem to be three basic paradigms in emblem digitisation, and, interestingly, each of these approaches has a very natural counterpart in the world of the book: the facsimile edition, the scholarly edition and the index. These paradigms determine the content of the sites, the sites' organisation, their usability, target audience and user interface. In fact, different types of editions tend to be hosted by different institutions.

The facsimile paradigm which sometimes prevails in libraries is the digital equivalent of the facsimile edition on paper.[43] Digital editions of the facsimile type provide digital images and perhaps a table of contents. There is no doubt that they are very useful. They are simple, effective, and to some extent the facsimile edition can function as a substitute for the physical book. It allows

---

[43] Such as the 'Emblematisches Cabinet' series of facsimile reprints published by Olms, Hildesheim.

for wide access to a book while at the same time saving fragile original material. *Pace* Buzzetti & McGann (2006: 54), the creation of a digital facsimile requires less scholarly effort than a fuller edition, that establishes a trustworthy text, that unravels the web of quotations that surrounds each emblem book, that provides annotation and a commentary that discusses the book's meaning and its place in the tradition.[44] The scholarly edition is useful to the scholar, because the editor's work contextualises the book and clarifies obscure places in the text, but of course it comes at a considerable cost. Finally, a site that primarily indexes emblems would be the digital counterpart to Peter Daly's *Index Emblematicus* volumes (1985-1998) or the Henkel and Schöne *Handbuch* (1976). Like the index volumes, such a site may to some extent provide text, it will provide pictures, but its true strength is in its use as an aid in finding emblems. Indexes may be provided on the text, lemmatised or not, on a pictorial description, on a condensed or more extended statement of the emblem's meaning, or on any number of other book or emblem properties. It is essential for emblem researchers to be able to compare emblems with identical or similar mottos, subjects or motifs and index-type sites will help locate these.

There are probably no sites that correspond precisely to one of these approaches. The best example of a facsimile-type site would be the English Emblem Project [9].[45] The best example of the index-type site is probably the German Emblem Book site [5]. The scholarly edition approach is probably best exemplified by the French Emblems at Glasgow site [3]. A prototypical combination would be the Spanish site at La Coruña [11]: multiple indices (among other things on names, motifs, sources, and key words), but also a full facsimile. One of the challenges that projects face is combining the best aspects of the three approaches: to provide a straightforward representation of the physical source, to provide a trustworthy and helpful edition of that source, and to provide search interfaces that give immediate access to related emblems in that volume and in other books.

4.3   Four Generic Functionalities

Four essential facilities for any edition, digital or paper-based, are searching, referring, comparing, and annotation, four of the scholarly primitives identified by John Unsworth (2000).[46] The contributions of digital technology to these facilities are ease-of-use, responsiveness to the user's needs, and sheer power.

---

[44] An example of a similar edition on paper would be Porteman's Hooft edition (1983) or Luijten's Cats edition (1996).

[45] The numbers between square brackets refer to the appendix at the end of this chapter.

[46] In his perceptive essay on the relevance of hypertext for critical editing, Lavagnino (1995) requests four types of functionality: selection, comparison, and construction of text versions, and integration of these functionalities with editorial commentary.

4.3.1 Searching and Selection

Serious study of any work requires being able to find relevant locations within the text. Given the semantic gap between human understanding of a literary work and its computational representation (Smeulders *et al.* 2000), no system will be able to present all locations relevant to a scholar's question. Facilities that a site offers may be searching (filtering based on presence of a text string), selection (filtering based on a system of categorisations), or a combination of these. For a bi-medial genre such as the emblem, it is clear that searching the texts is by itself never sufficient for localising relevant fragments, because the content of the images must also be taken into account. However, in a text-only content too, pure string searching will often be insufficient, as there may be a need to consider other properties of the text: the languages used in a text fragment, a text type (to limit search to poetry or prose or e.g. to quoted text) or even more complex properties (restrict a search e.g. to lines spoken by a certain character in a play). Selection may also be based on some stylistic or interpretive encoding added to the text. One may want to select poetry in a certain metre, or texts that have been marked as referring to certain Petrarchist conventions.

It is clear that the more complex selections will only be possible if information about the text characteristics (prose or poetry, the speaking character, the metre) is available. The text must have been adequately marked-up, or in some other way have been provided with annotations, in order for this filtering to be feasible.

In the Emblem Project Utrecht [1], the search and selection facilities do not support all of these functionalities. Text search can be limited to motto, all original text or editorial text. The picture index can be searched separately. Naturally, index-type projects, such as the Illinois [5] and München [6] projects, offer more extensive search facilities. The Glasgow projects [2, 3, 4] also offer a more comprehensive search mask that does include e.g. searching by metre or number of lines per stanza.

4.3.2 Hyperlinking

A large part of any edition's editorial material consists of some sort of references: to sources of quotations, to earlier or later versions of the same work, to its contemporary reception or to secondary literature. As Robinson has noted (2005: 12), the (printed) scholarly edition is essentially a hypertext waiting for the advent of the computer.[47] In the digital environment a reference can become a hyperlink – that is, if the referenced resource is digitally available.

---

[47] According to Sahle the nature of any edition is fundamentally at odds with a printed environment (2003: 79): 'Die Grundstruktur einer Edition mit ihren vielstufigen Repräsentationsformen und deren Vernetzung mit erschliessenden und kontextualisierenden Informationen widerspricht grundsätzlich der linearen Struktur typographischer Publikationsformen'.

It is therefore important not just to create hyperlinks where possible, but also to create adequate conditions for others to hyperlink into our own editions. This is where many projects fail to make use of the digital environment. Each resource should have its own URL (Uniform Resource Locator or 'web address'). 'Resource' here should be taken in a wide sense: digitised books, emblems, poems and even verse lines should have an individual URL in order to be able to point to (address) that specific resource.

An essential ingredient in a good addressing scheme is that one should be able to point not just to a location within a text ('this is where the Spanish epigram in emblem x begins'), but one should also be able to point to a chunk of text ('this and exactly this stretch of text is the Spanish epigram in emblem x'). Only once we are able to address chunks of texts rather than just locations, we can base advanced processing on these addresses ('fetch me the Spanish poem and its translation and show them to me in the context where I refer to it').

The Glasgow projects [2, 3, 4] and the Emblem Project Utrecht [1] are based on XML and are in principle equipped to offer this sort of addressing. A complication in this type of processing is that the XML source file contains the authoritative source of the text; for most visitors however, the web pages will be the natural point of access, and there is no intrinsic connection between these two sets of addresses. This problem also comes up when creating annotation facilities: the annotation should refer to the XML source, but should be displayed in the context of the web-based edition.[48]

4.3.3   *Comparison*

Comparison is a process of identifying similarities and dissimilarities between two or more objects. If items in some way or another can be thought of as alternative versions (perhaps a manuscript and a printed version, perhaps two prints from different presses, perhaps a text and its translation), an investigator may want to compare these versions. But the process of comparison is not necessarily limited to texts that may be considered alternative versions. In the case of the emblem, a listing of, say, all mottoes in a certain emblem book can also be studied from the perspective of 'what do they share, where do they differ'. Necessary for ease of comparison however is physical proximity. The eye should be able to jump from one text to the other and back, and the user interface of the edition should thus support joint display of comparable objects. A comparison facility might include functionality for semi-automatic collation.

A compare facility should be closely integrated into the search and select facility. Search and select operations result in objects that share some characteristic (e.g. having a parallel in Alciato). A process of comparison is what naturally follows (What else do they share? In what respects are they still very different emblems?). The result of comparison in its turn will prompt new

---

[48] See chapter 6.

queries (Does a certain shared characteristic apply across a wider selection of emblems?)

At present, as far as I know, only the Emblem Project Utrecht [1] offers a facility for emblem comparison. It can be invoked from the 'links to related emblems' section, with the two related emblems preset; users can also invoke the comparison facility manually and navigate to the emblems they want to see displayed in parallel.

4.3.4   Annotation

In a paper edition, readers can scribble in the margins of their books, but when each reader owns his individual copy, these marginalia are doomed to remain private. In the digital (online) environment, there is only a single copy, and when the environment allows readers to add their marginalia to that copy, subsequent displays of the edition can incorporate these marginalia. For this functionality, unlike the others, the edition must offer some sort of write access to some or all of its users. That is not to say readers can necessarily also modify the edited texts; but still, it is a fundamental step for the edition to take. What type of annotations to support then is a comparatively minor issue.

No current online emblem projects offer annotation facilities. As noted in chapter 3, Daly and Young's CD-ROM based edition (2002) [10] offers extensive annotation facilities, but from the viewpoint of the wider networked world, these annotations are like the scribblings in the margins of individually owned books: there is no easy way of sharing them with the world.

4.4   Parameters

In contrast with the basic orientation and generic functionalities discussed above, the parameters which I will discuss in this section concern the nitty-gritty details of site design. Many of these choices relate to the modelling decisions discussed in the next chapter.

4.4.1   Corpus

The most basic choice that any digitisation project has to make is about the corpus: which books will be digitised. Researchers may want to digitise emblems from the language and era of their specialisation. Libraries will want to digitise books from their own holdings. Other criteria may be the wish to create a coherent collection, or the wish to make available books that are not represented in the printed emblem indices. Trivially, the size of the corpus determines the amount of detailed study one will be able to expend on individual books. The La Coruña project [11] digitises Spanish emblems and the two German projects [6, 7] digitise German emblems, but nationality need not

determine corpus choices, as the French project [2] in Glasgow and the German project in Illinois show.

4.4.2   Full or Partial Digitisation

The choice between full and partial digitisation must be made at various levels. An edition is locally complete when the items it contains (i.e. emblems) are digitised in full. Index-type projects, that intend to help discovery of emblems rather than provide full editions, may limit the representation to motto, pictura and perhaps a main epigram, and ignore the other texts belonging to the emblem. For the edition type projects this is clearly not an option.

On the book level too projects have to decide whether to include or not non-emblem material (such as title pages, introductions, congratulatory poems, but also, e.g., the sermons and meditations which many later emblem books contain). Again it is clear that a project which perceives itself as providing an emblem database will make other choices than a project which aims at providing complete editions.

The Emblem Project Utrecht [1] chose to provide a complete representation both at emblem and at book level. The digital version should provide a substitute for consulting an original copy. At emblem level, there can be no full understanding if epigrams, constituents in various languages, prose commentaries, bible quotations, or other constituents are left out. At book level, title pages, dedications, prefaces and liminary poetry offer invaluable clues to the intentions of the books' makers and the reactions of their audiences. However, from Heinsius' collected Dutch poetry, the EPU reproduces only the emblems.

The same choice was made by a facsimile-type project such as the one based at Wolfenbüttel [7]. But the German Emblem Book project [5] in some cases only provides pages with emblem pictures and transcribed mottoes.

4.4.3   Transcription and/or Facsimile

Both text transcription and (facsimile) images have their irreplaceable strengths. A transcription is necessary for searchability, for ease of reading, and as a basis for further operations on the text. For an emblem edition, an image of the emblem *pictura* is obviously essential. But images of text portions of the book are also important. The transcription will probably ignore all those material aspects of the original that can easily be seen from an inspection of a digital image of the original (such as the distribution of texts over the page, use of typeface, dropped initials, etcetera). As Flanders argues (1997), facsimile images are also rhetorically important: they function as a claim to scholarly status.

The Emblem Project Utrecht [1] provides both, as do the Glasgow projects [2, 3, 4]. Facsimile- and index-type projects such as the German Emblem Book project [5] or the München project [6] do not give full transcriptions. An

older project, such as the Newfoundland Alciato site [8], gives full text but the only images are those of the *picturae*.

4.4.4   Text Enrichment

The transcribed text can be left as it is, or be 'improved' in various ways: normalisation of spelling, punctuation and capitalisation, correction of errors, expansion of abbreviations. In a digital environment, these choices need not be irreversible: it is always possible to show both the normalised and the original forms.

The Emblem Project Utrecht [1] has experimented with normalisation of spelling and punctuation. In the present interface these normalisations are not available (ideally, user should be able to switch normalisation display on or off). Nor does any of the other projects show multiple text versions. The French emblems at Glasgow site [2] does have normalised spellings (to present-day French orthography), but these are used for search purposes only. Presenting the normalised spelling to the reader would suggest a translation into modern French, which is not the case.

4.4.5   Editorial Enhancement

Emblem sites have many options with respect to the amount of editorial enrichment, commentary and indexing to be provided. In this respect many of the decisions will be based on the project's preference for an edition or an index-type site.

Picture description: Most projects give some kind of picture description, either in free format text or in keywords. The keywords may come from a controlled vocabulary, either a pre-existing one or one made up during the digitisation process. The most thorough form of picture description is no doubt provided by the Iconclass system (Brandhorst 2004). Iconclass is a hierarchical system of numeric codes with predefined notations for thousands of subjects from western art history. It facilitates description of pictures with great accuracy, without ambiguity and independent of natural language. Its adequate use however, necessitates the employment of a trained art historian.

References to emblems: References to emblems elsewhere, traditionally part of a free text commentary, are especially relevant in a digital environment, as they can take the form of hyperlinks to that other emblem – provided that one is digitally available too. The reference may be to a source emblem, to an emblem which is in some respect parallel to the current one, or to a later imitation.

References to sources: Emblem writers' often imprecise references to their textual sources can be checked and corrected where necessary. An adequate list of references to biblical sources e.g. will facilitate the creation of an additional index to the emblem book and to the whole site's collection.

Statement of meaning: Many sites will also provide a brief description of the meaning of each individual emblem. Sites differ in just how brief this description should be. The Illinois site [5], e.g., has an index on 'theme', like 'Christ as key to the next world'. In München [7], similar information is given in a 'Kommentar', which also describes just how the emblem expresses this meaning.

Other indices: Other indices may be created, for instance on the emblems' dedicatees, or on the days in the liturgical year to which the emblems were meant to be especially relevant. In Utrecht [1], we have experimented with an index on Petrarchist motifs.

Translation: The emblem texts may not be immediately comprehensible to modern readers, as they are either in Latin or in an early modern vernacular. French Emblems at Glasgow [2] has translated the Latin texts. The Emblem Project Utrecht [1] has translated Latin, and in very a limited number of cases the vernacular texts as well. The German Emblem Book [5] project translates the mottoes to modern German.

Emblem commentary: On the emblem level, there might finally be a full commentary. No site has attempted to create these.

Book introduction: Most sites, not all, do attempt to give some sort of introduction to the digitised volume. The Emblem Project Utrecht [1] and the Glasgow sites [2, 3, 4] introduce the books themselves and give some information on the digitisation procedures in free text. The German Emblem Book [5] project provides more structured metadata.

4.4.6   Access Paths

Sites provide various sorts of entries into the emblem collections. The simplest access path is a table of contents; another would be a thumbnail overview of the emblem *picturae*. This is very appropriate for edition-type sites. The indexes that have been created on the emblem can also be used for browsing. Sites that have been indexed using Iconclass can use the hierarchical Iconclass browser developed by Mnemosyne. Another option, chosen by the München project [6], is to begin by presenting the visitor with a search window which gives access to the emblems in all of its books. Though there is more to the site than this, it begins by showing to the world what you might call its 'database face'.

4.5   Technical Infrastructure

The last aspect to be discussed is the choice of a technical solution for storage and retrieval of the emblem book data. This should never be the primary consideration, yet it is a choice which very much determines what can be done later on, and it needs careful thought. I will leave aside the choices to be made with respect to the digital images; what I will discuss here is the technical infrastructure required for the emblem texts and indices.

Some of the older emblem sites, such as the Newfoundland Alciato site [8], use HTML (HyperText Markup Language) as a medium for storing the site's contents. HTML is the language which is used to define web pages. The advantage of this approach is that it is rather straightforward: the page which the visitor sees is the page which the editor makes. For serious work however, HTML as a storage medium is inadvisable. Of course, the visitor to the site will see HTML pages, but these HTML pages should be generated from the stored emblem information, which should not itself be in HTML. The reason for this is that HTML is meant for presentation, and each new way of presenting the information needs new HTML. If we want our site to offer multiple views of emblems, or flexible search options, or if we want our site to be easily adaptable for new display devices or other output media, we will need to continually reuse the same information in different ways. The only solution is to store the information in a richer format than HTML has to offer. For this richer format, some projects use a (relational) database, other projects use XML.

XML, or eXtensible Markup Language, is a language much like HTML, but much more flexible and powerful. Using XML implies marking up the transcribed text using 'tags'. The tags delimit structural, syntactic, semantic or stylistic units of the text. Grouping these units can define hierarchies that reflect the text's structure (see Renear 2004). The Text Encoding Initiative (TEI) is a consortium that has provided guidelines for the adequate use of XML and its predecessor, SGML (Sperberg-McQueen 1996). Utrecht [1] uses XML/TEI, as described in Boot (2004a) and Tilstra (2007). XML is the superior choice for projects that want complex text presentation and the possibility of complex text manipulation (Schreibman 2003). Projects that focus on indexing, rather than on full text features, will prefer a database, among other things for its simplicity in data entry.[49] Illinois [5] is a typical example of a site based on a database.

4.6   Edition Models *Ex Ante* and *Ex Post*

One of the items on McCarty's agenda for humanities computing is studying the primary genres of scholarly communication (2005: 207). All of the sites are to some extent targeted at a scholarly public, and it is important to assess their different approaches. The model I sketch here is an empirical model: it results from an inspection of existing emblem digitisation projects and their differences. I have avoided being judgmental in most of this chapter, but that does not imply I have no opinions on what is desirable. A truly useful emblem site must offer complete books that are transcribed fully. It should also contain a full facsimile of these books. It should use Iconclass for the indexing of the emblem pictures. It should be based on XML, as only XML offers guarantees for durability,

---

[49] More about XML in the following chapter.

application independence, and addressability. In this last section I will further examine the use of XML and discuss some other edition models.

As an empirical model for digital representations of one literary genre, it would be interesting to compare the model that I sketched here with other models for digital editions. Wiering (2009) outlines a conceptual model for digital critical music editions. It distinguishes between the content of the multidimensional archive underlying the edition (sources, including performance data, encodings, links and annotations), and the views on that content being shown to users (e.g. apparatus, stemma and edition). A sample of problems in music editing is discussed to show how the model might be applied. In general terms, Wiering's model is not specific to music editions. Its implementations would have to be, but the paper does not examine these. In contrast, many of the issues I discussed above are implementation-specific. The model I gave is an *ex post* model; Wiering's model is *ex ante*. A fuller model of the digital edition would have to take account of both. Another difference between the two models is that Wiering's model specifically targets critical editions while the model above takes into account any type of edition.

      The distinction between the underlying collection of material and the user views generated from that collection is similar to the distinction Vanhoutte (1999) draws between the archive and the museum functions of the digital edition. For Vanhoutte, the archive function comprises 'the preservation of the literary artefact in its historical form and the historical-critical research of a literary work', the museum function 'the presentation by an editor of the physical appearance and/or the contents of the literary artefact in a documentary, aesthetic, sociological, authorial or bibliographical contextualisation, intended for a specific public and published in a specific form and lay-out.' Unlike the user views in Wiering a museum-type edition is created by a human editor out of the mass of material stored in the archive.[50]

      Another *ex ante* model is given by Sperberg-McQueen (2009), who distinguishes between facts, selection and presentation. The facts correspond to the archive, an edition can be seen as a selection from these facts which can be presented in multiple ways. A key term for Sperberg-McQueen is mutability: mutability of facts, as new sources are discovered or seen in another light; of presentation, as output devices vary and develop; and of selection, as users can ask for many different subsets of the facts (*views*, in Wiering's terminology). One of the lessons that Sperberg-McQueen draws from this ubiquitous mutability is the importance of a software-independent storage of the underlying facts, and of as much as possible of the selection and presentation layers, which brings us back to the need for XML as a storage format for these editions.

---

[50] Vanhoutte's thinking on the issue may have evolved, as he has since argued for editions that are generated out of the underlying archive upon user request (2007).

I have no answer to the question why it should be that *ex ante* models are easier to find[51] than *ex post* models. Maybe a preference for the *ex ante* or *a priori* model corresponds to scholars' deep-seated desire to proceed from first principles, whereas the *ex post* approach confronts one with the unavoidable messiness of any practical application of first principles. Still, we can learn from that messiness.

Appendix

This appendix lists the emblem digitisation initiatives mentioned in this chapter. It does not pretend to provide a complete overview.

[1] Emblem Project Utrecht
URL: http://emblems.let.uu.nl/
Based at: Utrecht University
Description: Has digitised 25 books of Dutch love emblems.

[2] French Emblems at Glasgow
URL: http://www.emblems.arts.gla.ac.uk/french/
Based at: Glasgow University
Description: Digitises 'all the French Emblem Books of the sixteenth century, along with their Latin versions when appropriate'.

[3] The Study and Digitisation of Italian Emblems
URL: http://www.italianemblems.arts.gla.ac.uk/
Based at: Glasgow University
Description: 'an EU funded project for the study and digitisation of Italian emblem books'.

 [4] Alciato at Glasgow
URL: http://www.emblems.arts.gla.ac.uk/alciato/
Based at: Glasgow University
Description: Digitises 'twenty-two editions of the emblems of Andrea Alciato (1492-1550), the *pater et princeps* of the emblem genre'.

 [5] German Emblem Book. Also known as Digital emblematica.
URL: http://images.library.uiuc.edu/projects/emblems/
Based at: University of Illinois at Urbana-Champaign
Description: Initially, the German emblem books in the University of Illinois Library (67). At present (April 2009), 21 books are available.

---

[51] Two more *ex ante* models are given by Lucía Megías ([In print]) and Sahle (2003).

[6] Digitalisierung von ausgewählten Emblembüchern der frühen Neuzeit
URL: http://mdz1.bib-bvb.de/~emblem/index.html
Based at: Bavarian State Library & Ludwig-Maximilians-Universität, München
Description: 139 mainly German books are available in a prototype database
(some only partially).

[7] Emblematica online
URL: http://www.hab.de/forschung/projekte/emblematica-e.htm
Based at: Herzog August Bibliothek, Wolfenbüttel, Germany
Description: Mainly German emblem books from the collection of the Herzog
August Bibliothek.

[8] Alciato's Book of Emblems. The Memorial web edition
URL: http://www.mun.ca/alciato/
Based at: Memorial University, Newfoundland, Canada
Description: Alciato's *Emblematum liber* (in several editions), and (as secondary
material) Whitney's *Choice of Emblemes*.

[9] English Emblem Book Project
URL: http://emblem.libraries.psu.edu/
Based at: Pennsylvania State University
Description: Digitises 9 English language emblem books.

[10] The Emblems of George Wither and Gabriel Rollenhagen. A Collection of
Emblemes
URL: N/A. (CD-ROM)
Based at: N/A.
Description: The emblems of George Wither and his source Gabriel Rollenhagen.

[11] Proyecto de Investigación sobre Literatura Emblemática Hispánica
URL: http://rosalia.dc.fi.udc.es/emblematica/
Based at: La Coruña University, Spain
Description: Digitises the Spanish emblem corpus.

5   Digital Editing and Text Modelling: The Case of the Emblem Project Utrecht

In the previous chapter I sketched a model of digital emblem sites, based on our experiences at Utrecht and those of others. In this chapter I will stay closer to home and look at the role of text modelling in the creation of a site such as that of the Emblem Project Utrecht. In the annotation chapters of this thesis I will argue that modelling is an essential ingredient of annotation, if annotation is to become a navigable resource connecting the primary and secondary texts of scholarship. For this to become a reality, however, the texts that are to be annotated themselves need to be modelled. A analysis of their structural properties is necessary, among other things, to create a text structure that annotations can be attached to.

In this chapter I will describe some aspects of modelling which went into the design of the Emblem Project Utrecht: considerations that reflect the emblem genre (section 5.2), the goals of the project (section 5.3), the properties of XML and TEI Encoding (section 5.4) and some other considerations (section 5.5). In the last section I also draw some conclusions. Unavoidably, the discussions in this chapter will use some technical terminology, of which I will give the briefest of explanations in the first section.

5.1   XML, XML Schemas and the Text Encoding Initiative

The eXtensible Markup Language (XML) is a formalism that describes how tags can be added to text in order to describe properties of that text (Bray *et al.* 2004). This is a very simple example of what a poem may look like in XML (Boot 2004a):

```
<lg>
  <l>Ama à Dios de coraçon, </l>
  <l>Paraque se satisfaga</l>
  <l>Que amor con amor se paga. </l>
</lg>
```

The example shows several elements (*l*, verse lines, and *lg*, line group) in a hierarchic structure. Elements can have attributes which specify properties of

the elements, e.g. a poem type and a language: `<lg type="motto" lang="lat">`.

XML Schemas describe which elements and attributes can occur in a certain class of XML documents. They can also describe what combinations are allowed: in the example just given, a schema might prescribe that *l* elements must occur within *lg* elements, and not the other way around. There exist a number of XML Schema languages, of which the best known is the Document Type Description (DTD).

Markup schemas consist of rules that the computer can understand and check, but they should also explain in natural language what the elements are for and how they should be used. The Text Encoding Initiative (TEI) has developed a broad set of guidelines for encoding humanities texts, now in its fifth version (Burnard & Bauman 2007). The *Guidelines* are accompanied by a set of schemas that can help enforce the *Guidelines*' recommendations.

The most important reason for the introduction of XML is that it allows one to describe text structures without committing to a decision on how to use or represent the text. The reason why the EPU chose to use XML is precisely that while ideas about presentation and platforms for presentation change at alarming speed, the modelling of text structures should be more or less stable (Sperberg-McQueen 2009).

5.2   Modelling the Emblem Genre

A markup schema describes a class of documents. The phenomena that the markup highlights in the text to some extent correspond to the salient features of the text genre. The markup schema, Piez suggests, can be thought of as an expression of a theory of the text (2007). Piez does not distinguish between text genre and text; I would say that just as the markup schema can be a theory of the text genre (in our case, the emblem genre), the markup instance (the individual XML document) is a model of the individual text (in our case, emblem book). One of the strengths of XML is the expectation that a schema will be created that describes the shared, structural properties of the individual texts. That a schema creates the possibility of checking the individual document against the schema is a very big boon, but the deeper motivation for creating schemas is that schema creation is an attempt to account for the shared structural characteristics of a text genre.

Two fundamental modelling decisions taken by the EPU were:
1.   the decision to model an emblem book as a collection of individual texts (=emblems), rather than as a single text with subdivisions. We felt the unit of research in emblem scholars' work is often the individual emblem rather than the work. Often, it is also individual emblems and emblem motifs that are being imitated or borrowed elsewhere. To be

sure, there are coherent emblem books that treat a single subject, to be sure, but even in these cases there is not usually a very strong connection between one emblem and the next;

2. the decision to see the emblem motto as a constituent on the same level with the others. The tendency to consider the motto as a title, and therefore in some sense hierarchically above the rest of the emblem, is a very natural one. We felt we should not enshrine that preferred status in the markup, but chose rather to use neutral *div* (for division) elements for all emblem constituents.

I call these decisions fundamental because they are based on ideas about the nature of the emblem genre. On the other hand, they have no major technical consequences. Had we made the reverse decision, we could have built exactly the same site.

A model is always only a partial representation of the realities it models. In the EPU we decided not to represent many typographical features that seemed not especially meaningful: we ignored running titles, catchwords and predictable font switches. This example shows that text modelling is not a theoretical exercise, motivated only by a desire for faithful representation, but also a pragmatically motivated activity.

5.3   Emblem Modelling as a Pragmatic Activity

Markup does not just reflect properties of the text 'in itself', if such a thing would exist. The text model always is always also a model of what the editors want to do with the text and of what the editors expect users will want to do with it. Therefore, the text model is always, at least in part, motivated by pragmatic considerations. Text modelling is always modelling *for* as well as modelling *of* (McCarty 2005: 24).

Perhaps the most fundamental pragmatic decision that we took was to let emblem unity take precedence over physical contiguity. To given an example: in some books the emblem pictures had been bound separately from the other constituents of the emblems; in another book the left hand pages contained the text components of 15 emblems, the facing page the emblems' pictures. In cases like these, we collected the components that we felt belonged together into 'whole' emblems. The decision was motivated by the fact that emblems are the object of study that we wanted to present to the readers of our site. It is clear this decision would never have been taken by a library-based digitisation project, which would probably think in terms of physical objects rather than of collections of interpretively reconstituted units.

Pragmatic considerations also have effect at a more practical level. The EPU markup scheme contains elements that hold translations for text fragments, references to secondary literature and references to thumbnails of the emblem

pictures. These elements are not in any deeper sense part of the model of the emblem genre.

5.4   Emblem Modelling and TEI/XML

A model for texts belonging to certain genre will also be influenced by other attempts at modelling similar and other genres. It is obvious an emblem model will be influenced by emblem theories and especially emblem theories that have been implemented in emblem sites. A model for a genre such as the emblem will also be influenced by more generic text models. Specifically, modelling genre characteristics in XML is unavoidably influenced by the encoding practices of the TEI *Guidelines*, even where these have nothing specific to say about emblems.

The EPU was the first emblem project to use TEI for cover-to-cover representation of the emblem books.[52] An important factor in the decision to do so was the desire to adhere to a standardised, well thought out and inclusive way of digitising textual material. One of the strengths of the TEI is that while stressing the importance of standards, it fully accepts the need for customisation and extension when working with textual features that the *Guidelines* do not cater for. The TEI is not a straitjacket. Still, the decision to use TEI is likely to influence the modelling choices to be made later, perhaps in privileging content over physical properties or in favouring hierarchic structures. It is a decision that requires consideration.

The use of TEI/XML for serious digital editions is the current orthodoxy. If an edition is not based on TEI/XML, the *Guidelines for Editors of Scholarly Editions* require the editor to 'provide a rationale for departing from community practice' (Committee on Scholarly Editions 2007: 2.V.23.2). The dominance of TEI/XML is perhaps best expressed by the fact that the MLA volume on electronic textual editing contains a chapter by John Lavagnino titled 'When not to use TEI' – implying that the use of TEI, and therefore XML, is the default choice (Lavagnino 2006). Lavagnino's essay fully shares that assumption. Reasons not to use TEI could be either the availability of a more appropriate XML vocabulary (DocBook or EAD), or the undesirability or impossibility of transcribing the text. The rest of the essay suggests that even in a TEI project, out-of-the-box TEI may not be the solution for each and every textual feature.

That is not to say that TEI has never met with fundamental opposition. Lancashire has been an early and vocal critic of TEI. Some of his objections are that TEI devotes insufficient attention to material aspects of the encoded sources, that there is insufficient tool support, and that it cannot deal with multiple hierarchies (1995). In his view, TEI development was dominated by

---

[52] The Herzog August Bibliothek in Wolfenbüttel has experimented with using TEI for indexing rather than transcribing emblem books (Opitz 2004).

computer professionals rather than humanities scholars and the result is the undesirable imposition of a single format on the humanities.

Especially the issue of multiple overlapping hierarchies has been a problem for TEI. The problem is that XML requires a document to be encoded as a single hierarchy of elements, which does not fit very well with the realities of text. Van Zundert & Van Dalen-Oskam (2005) argue that therefore XML cannot account for non-linear phenomena that textual and literary scholars want to study. It is certainly true that most projects will have to deal with at least some amount of textual overlap: pages overlap with paragraphs, speeches in plays overlap with lines of poetry, quotations overlap with paragraphs, etc. The problem worsens with the encoding of interpretive phenomena,[53] especially in multi-user settings. The *Guidelines* provide a number of XML-based solutions, each with their own advantages and drawbacks in specific situations (Burnard & Bauman 2007: ch. 20). There are also other possibilities, some further removed from the concept of an XML document than others. TexMECS e.g. is based on the encoding developed for Wittgenstein's writings (Huitfeldt & Sperberg-McQueen 2003): TexMECS elements can overlap, and can do many other things that XML elements cannot (e.g. suspend and resume), but like an XML document, a TexMECS document still consists of text interspersed with tags. More radical approaches split text and markup. An example is the Just In Time Markup technology described by Eggert (2005).

Another issue that has been seen as a problematic aspect of XML has been termed 'the fallacy of prescience' by Liu & Smith (2008): it is the perceived need to know the structure of the text before one starts encoding. It is true that usually XML encoding is based on a Document Type Definition (DTD) or schema that restricts the allowed encodings. This is usually seen as helpful rather than problematic, because it helps prevent mistakes, but it can be seen as stifling and a brake on creativity. Another often-made criticism of XML that Liu and Smith raise is that an XML-encoded text can be hard to read and creates a cognitive distance between the scholar and the object of his/her investigation. As an alternative to XML, Liu & Smith propose a database model of text encoding that stores individual words in a database and creates elements by attaching what are called annotations to sequences of words.[54]

Although some of these problems are real enough, they are not necessarily related to XML's hierarchic object model, but rather to a processing model that tries to squeeze all information related to a text into a single document. It is usually unproblematic to encode basic document structure in a

---

[53] All markup is interpretive, but not all markup is equally so.

[54] A somewhat similar approach has been described by Burghart (2004). For Burghart, however, the database is only used as a step towards a TEI-encoded document, which is generated from the database. Tummarello *et al.* (2008) use essentially the same method in RDF: each word being a resource that can be referred to by other resources representing what used to be XML elements.

single XML document. Multiple layers of interpretive encoding representing non-linear, non-hierarchic phenomena can then be attached to this base document, if so desired, using stand-off techniques. In the latest TEI version, notes can point to locations in external documents using the W3C-standardised XPointer Framework (Grosso *et al.* 2003). This is much more robust than storing text and annotations in a database based on a proprietary schema, which makes continued accessibility of valuable scholarly work dependent on the functioning of database and application. That is not to say a well supported TexMECS-like system wouldn't sometimes be helpful in avoiding some of the more unwieldy XML constructs used to circumvent the single-hierarchy limitation.

As to the 'fallacy of prescience': during very exploratory work, use of a fixed DTD may indeed be problematic. However, no one forces the researcher to use DTD at that stage. For a project that aims at publication, at a certain stage the advantages of validation and increased productivity will outweigh those of experimentation. Software that exploits the encoding will have to know what to expect.

Most current debate about XML in the Digital Humanities community is not about either using XML or not, but is directed at finding out the respective suitabilities of XML and databases in digital projects. Bradley (2005) used database-like structures in XML documents, unlike the structural hierarchies that characterise traditional text encoding, to store information data from the *Durham Liber Vitae* project.[55] Bradley & Short (2005) use a relational database for a prosopography project, rather than opt for a text edition. Buzzetti & Rehbein (2008) argue for the need to combine databases and XML to represent both the text-like and data-like aspects of source texts. Ciula *et al.* (2008) combine XML and semantic web technology.

It would not have been impossible for the EPU to combine XML and database technologies in the project. The representation of the books, their structure, the basic annotation and translation could be represented in XML; the indexing, the bibliographies and the relations between emblems could be stored in a database.

An important consequence of choosing to use XML for modelling is that the process results in a text where the structural components of that text have been labelled, provided with identifiers and subsumed in a hierarchic structure. This creates the possibility to create more or less stable references to these components. This is an essential ingredient in an infrastructure that supports annotation.

---

[55] See chapter 10 for a similar attempt using TEI feature structures.

5.5 Discussion

The preceding sections argued that the modelling activity that results in a markup scheme is informed by considerations about the text genre, by pragmatic considerations with respect to the proposed site, and by the wider markup paradigm to which one adheres. There are yet other ways in which modelling enters into creating a digital edition:

1. Modelling scholarly categories: The site editors may want to index the site's emblems using multiple scholarly categories. They can be very diverse and will depend on the specific properties of the sub-corpus that is being digitised. As mentioned in the previous chapter, the EPU experimented with an index on Petrarchist motifs. The model that we used there was the model given by Forster in *Icy Fire* (1969). In Iconclass indexing of pictorial motifs, the model is given by the Iconclass conception of the world.

2. Choice of paradigmatic volumes: One important modelling decision is the choice of which volumes to consider paradigmatic for the genre. When Henkel and Schöne selected the emblem books they wanted to index for their handbook (1996, first impression 1967), they followed their view of what was emblematic. But in the act they also determined what books the next generation of scholars would consider quintessentially emblematic.

3. Decisions about appropriate functionality: Decisions about which functionality to provide on a site reflect, presumably, a vision about what is appropriate for the genre. If the EPU offers a comparison facility and includes hyperlinks to side-by-side views of related emblems, this reflects a view of emblem books as usually only loosely organised collections of individual emblems. Similarly, the Glasgow and La Coruña emblem projects include search options by metre and stanza type, presumably because metre and stanza type are essential elements in the emblem model being used. Again, like many other modelling decisions, choices such as these are self-perpetuating: the compare facility is created because it seems an appropriate activity, and it seems an appropriate activity because there is a compare facility.

It is however not just the text that needs to be modelled to create a site. The site itself is modelled, based on ideal, *ex ante*, models and based on *ex post* models that generalise experiences with other sites. I discussed these models in the preceding chapter. This in turn depends on a model of one's visitors and their interests.

Modelling is clearly an essential ingredient in the creation of a digital edition. In the context of this book, the main reason for it is that annotations can be attached to the resulting structure. The implications however are much

wider. Modelling is necessary to create editions that fit the material that is being edited. Modelling is also important because it prepares the text editions for inclusion in larger corpora of scholarly edited electronic texts, and because it helps increase and improve the collection of electronic models that the text encoding community has at its disposal. Modelling is therefore also a social act, which has benefits that transcend the edition at hand.

Part III. Digital Annotation Tools

6 Digital Edition Annotation using EDITOR

6.1 Introduction

The EDITOR program, developed at the Huygens Institute, is an annotation tool for scholarly users of electronic editions.[56] 'Annotation' should be understood in the wide sense discussed in chapter 3: it includes text commentary, categorisation of text fragments according to any typology, creation of references to other works, and creation of connections between text fragments. EDITOR allows researchers to describe, comment on, categorise and link text fragments of any sort. The annotations can be uploaded to the web and shared with other researchers.

The specific problem that EDITOR addresses is that, for many modern scholarly editions, there is a distinction between its source format (probably XML) and (multiple) presentation formats (probably HTML). Usually, annotation tools annotate web pages. These tools are unsuitable for scholarly use, as web pages are only transitory representations of the scholarly objects that need annotation. EDITOR is used to annotate the XML source of the edition. Annotations that unambiguously refer to locations in the XML source file of an edition are suitable for exchange with other researchers basing themselves on the same edition. Display of the annotations can also be integrated into the web edition itself, if the provider of the web edition feels they represent a worthwhile enhancement to that edition. A facility like SANE (discussed in the next chapter) could be used to display the annotations in the edition's presentation format.

This chapter discusses EDITOR's background, requirements and design. The thinking behind the program will be clarified by showing how it was applied

---

in an investigation into some aspects of theatricality in *Amoris Divini Emblemata* (Van Veen 1615).[57]

6.2   Context and Motivation

6.2.1   *Theoretical background*

Peter Robinson, in his much quoted article about the future of the digital edition (2003), states the scholarly digital edition is not just a collection of transcribed, edited, sifted, hyperlinked and commented texts. Beyond that, the digital edition is also a collection of tools that can be used on those texts, and it is a living collaboratory for people with a shared research interest in these texts. The edition is not, or should never be, finished. Readers should become contributors of corrections, additions, translations and annotations. In Robinson's view, this will remove a long-standing barrier between textual and literary scholarship.

The edition, then, should be home to a number of 'tools to augment scholarly activity', to quote John Bradley (2004). Bradley focuses on annotation, and envisages the use of scientific visualisation techniques to display the presence in the text of e.g. ideas, themes or motifs, in application of a 'strategy of using the machine to record and organise issues found in a text'. He warns against the temptation to limit ourselves to a linguistic/word-centred view of the importance of textual objects. As he put it elsewhere: '(…) our interest should be on building software that supports the presentation and manipulation *not* of the presentation of the text (as in a word processor) or the words in it (as in word-oriented search tools such as TACT) but of the analysts' conceptual model of their texts. By allowing the machine to store, present and manipulate the analyst's textual model, it can help the analyst to more fully develop the model itself, and support the better development of insights into the model and its components' (2003: 191).

The tools that we need should therefore be useful to the researcher at work on refining his/her model of the text, but their usefulness does not end there. At the stage of publication of the research findings, Bradley suggests the researcher may very well prefer to publish his/her model of the text itself, as a digital object, rather than a conventional paper. This brings us back to Robinson's argument for the edition as a collaborative enterprise, and the possibility of feeding the researcher's textual model back into the edition and having it displayed from inside the edition.

The EDITOR toolset discussed here was not designed to do everything that Bradley asks for. Especially in the area of high-level visual manipulation functionality would be missing. Researchers can however use EDITOR to create annotations (based on user-defined annotation types with multiple fields) to

---

[57] This chapter reworks Boot (2005) and (forthcoming 2009).

arbitrary (contiguous) elements or stretches of text in the edition. The use of a SANE-like facility would then allow the display of these annotations in conjunction with the edition text, either from within the edition or immediately from a scholarly publication. A finished version of EDITOR would go some way towards realising the 'Bradley-Robinson' vision of the edition. EDITOR also extends ideas initially developed at the Emblem Project Utrecht (EPU). To allow external researchers to create indices on the emblem material, work started on a 'Public Index Feature' (Boot 2004b). Annotation creation, display and manipulation should be accessible to all, not just to a privileged set of persons working at or associated with the editorial institute that creates the edition. It became clear soon, however, that the development of a truly flexible annotation toolset was beyond EPU's resources. Perhaps more importantly, annotation is one of Unsworth's 'scholarly primitives', and as he argues: 'software intended to enable these primitives should be developed and tested in the context of real scholarly use, but it should resist customisation, because purpose-built or project-centred software is unlikely to provide broad support for functional primitives' (2000). It is important to develop software that is potentially widely applicable in cooperation with potential users in many fields. EDITOR is not specifically targeted at emblem books, and its functionality was designed with colleagues from medieval studies and modern literature departments.

One of the factors that motivated EDITOR's development was the Huygens Institute's ambition to create editions that are useful research tools for scholars at universities and elsewhere. In a period in which the Institute is switching to the creation of digital editions, there is a need to experiment with and develop research functionality that will ensure usability in a digital context. EDITOR was meant to contribute to that usability, and thus to the feasibility of text-based digital scholarship. Despite suggestions to the contrary (Robinson 2005), EDITOR development was based on the assumption that new scholarly editions will be digital and based on XML. The true potential of the digital edition will only be realised once the edition can serve as a vehicle for digital scholarship. To accomplish that, we need an annotation program:

- that is suitable for the creation, manipulation and display
  - of diverse annotation types,
  - on XML based digital editions;
- that allows for displays integrated into the edition;
- where annotations will be visualised at multiple levels and in multiple modes.

Institutions that host digital editions, such as the Huygens Institute, could then facilitate the use of EDITOR (or similar programs) on these editions. Scholars would be able to annotate the editions using these programs, and if the scholar

agrees, the hosting institution could decide to offer access to his or her annotations from within the digital edition.

6.2.2   Related work

There are many programs that offer annotation facilities. I have mentioned some of these in section 3.4. Especially researchers in linguistics and the social sciences have developed annotation software. Social science programs for qualitative data analysis often come with facilities for annotation, linking, highlighting, and sometimes include facilities for automated text analysis or text mining software. EDITOR differs from these programs mainly in that it was conceived to annotate digital editions, i.e. publicly available XML files.

What sets EDITOR apart from much of the text analysis software developed for use in the humanities is, again, EDITOR's focus on sharing annotations in a web environment, but also the assumption that most of the annotation will be done by hand. It is based on the expectation that most of the interesting observations about edited texts will be made by humans, assisted by computer programs, and not the other way around.

A program that often comes up when discussing EDITOR is Annotea, developed at the W3C (Kahan *et al.* 2001). Annotea is server software for storing and retrieving annotations. It needs an annotation client for actually creating and displaying annotations. We felt existing Annotea clients were mainly targeted at simple web pages and did not provide an adequate platform for annotating large edition files. We might have used Annotea itself for storage of annotations, but felt there was little point in using it outside of its intended scope. Storage of annotation data in a server component would also raise the privacy and data ownership concerns I discuss below.

A very interesting annotation product is APE, the 'Assistant for Philological Explorations', developed by Dieter Köhler (Köhler 2006). APE specialises in annotation creation, not display. Kohler has spent much thought on the issue of canonical reference systems as used in referring to e.g. locations in the Bible or in the works of Aristotle. With APE, the researcher's annotations can use these canonical references rather than pointers to a single digital edition. This is clearly an important issue, which EDITOR does not address.

6.3   Requirements and Issues

6.3.1   EDITOR requirements

Development of EDITOR has been guided by an analysis of the requirements that an annotation tool should fulfil. The analysis was based on literature about existing annotation tools[58] and discussions of how annotation could help in

---

[58] Among other literature Gertz & Sattler (2003) and Bird & Liberman (2001).

researching the books that EDITOR's designers were working on at the time. General requirements are:

1. The software should be available as open source.
2. The output and storage formats should be based on public standards.
3. It should be possible to share annotations and annotation display definitions.
4. The software should work with any XML-based modern digital edition.
5. Results should be accessible from the digital edition, and offer entry into the digital edition.

More specifically, basic requirements for the creation of annotations are:

1. The annotations should be safe and durable. The creation of annotations represents a considerable investment of time; there should be no doubt about the continued availability of the annotations, whatever the fate of the EDITOR program.
2. The researcher should be in control. Annotations may be experimental or personal. The researcher should feel no concerns about the privacy of work that he/she has not explicitly decided to share with others.
3. The annotations should unambiguously identify locations (text fragments) in a publicly available edition. The annotations should not, therefore, refer to web pages (which will change as technology and interface design develop). Instead, they should refer to a presumably durable edition source format.
4. The edition itself should remain untouched. The edition source file may be the result of years of editorial work. Storing the annotations inside this file introduces the possibility of corrupting the editorial work. It would also create a bottleneck in the annotation process, as multiple scholars may be working on the same texts.
5. The researcher should not need special ICT expertise in order to work with the annotation toolset.
6. The annotation types should be user-defined. EDITOR should be useful to a wide range of humanities (sub-)disciplines, and no amount of predefined annotation types will be sufficient to satisfy scholars from all fields. (And of course, it is in the nature of scholarly work to want to introduce new distinctions, rather than to rehash the old ones).

If these criteria are met, the annotations will be in a suitable format for display in (some sort of) conjunction with the edition to which the annotations refer. The connection may be one-way only (hyperlinking from the annotation display to the edition), but in other cases will go both ways, making available the annotations from the existing edition.

Requirements for the annotation display are:

1. There should be a filter facility for the annotations (by annotation type, by annotation value, by annotated fragment location).
2. There should be a sorting facility for the annotations (again, by annotation type, by annotation value, by annotated fragment location, perhaps by annotated fragment type).
3. Annotation display should be optionally based on a tree view of the edition XML structure, making it immediately clear which text fragments carry which annotations.
4. Annotation display should be optionally aggregated at higher edition hierarchy levels (for instance counting all annotations of a certain type that occur on verse lines and showing their number on stanza level).
5. It should be possible to introduce a 'break' level or break locations; thereby facilitating side-by-side display of edition sections and their annotations. Side-by-side display of sections (for instance: different versions of poems; earlier and later sections in a manuscript) makes it possible for the researcher to e.g. compare properties (vocabulary, style, etc.) between sections.
6. In annotation display, it should be possible to switch the document hierarchy to an alternate hierarchy (for instance: from a hierarchy based on books, chapters and section to a hierarchy based on the physical document structure, in quires, pages and columns).
7. It should be possible to base annotation display on a joint display of multiple hierarchies.
8. It should be possible to visualise annotation counts (in any display unit) using diagrams. If for instance each dialect word has been annotated, it should be possible to display the number of dialect words per stanza in a diagram.
9. It should be possible to store annotation display definitions. Researchers should be able to store multiple display definitions for use by visitors. Visitors should themselves have the option of modifying these definitions and storing them for their own private use on later visits.
10. The main platform for annotation display should be the Web. The Web is presumably the place where digital editions themselves will be available, and increasingly, the platform of choice for the distribution of all scholarship.
11. Annotation display should offer a clear API to the world, in order to facilitate the display of annotations in external contexts.
12. Annotation display should be, wherever possible, hyperlinked to, and from, the web edition.

Some of the implications of these requirements will be discussed in the next section.

6.3.2   Some issues

Writing down requirements is the easy part of software development. Making the decisions about how to implement them and actually implementing them is the hard part. This section will discuss some of the decisions made about EDITOR implementation.

Local vs. central annotation entry

It is one of EDITOR's central requirements that annotations can be shared with the world. It was clear from the beginning that annotation display should be web-based – though not necessarily always on a publicly accessible web site. Though a strong case can be made for web-based annotation entry too, we decided otherwise.

The advantages of web-based annotation entry would be: sharing of software functionality between annotation entry and display components, accessibility from anywhere in the world, no requirements for local installation.

However, we felt the advantages of local processing outweighed these. Perhaps the most important of these advantages relate to EDITOR's trustworthiness for the researcher. The researcher's privacy should be guaranteed: his possibly experimental annotations should not be exposed to preliminary scrutiny by others. Perhaps more importantly, his intellectual property should remain his own. Storing annotations in a central server would expose them to risks that PC-based storage would avoid. A second aspect of trustworthiness is the annotations' guaranteed availability. Web server based storage is beyond the control of the individual scholar. The scholar can however save and backup files on his own computer.

A second important argument for local processing is the need for EDITOR to handle large edition files and large numbers of annotations. We anticipate edition files of tens of megabytes in size, carrying tens of thousands of annotations. It is true modern web programming techniques reduce the need for wholesale transfer of large files, and high speed internet access is nearly universal. Still, we felt the dedicated processing capability of a personal workstation was the best bet for the sub-second response times that we wanted to achieve.

Supported edition formats

EDITOR will presumably be used with digital editions that are based on TEI/XML. Beyond these edition's HTML presentation, there will always be a permanent and durable source format that is the mother format from which the HTML representation derives. This durable source format will use XML for its syntax and the TEI tagset for the definition of the edition contents.

However, the diverse nature of scholarly edition projects precludes prescribing the use of a single DTD or XML schema. An edition of a medieval manuscript tradition needs other elements than one that edits a volume of

modern verse. EDITOR cannot therefore assume the use of a single DTD. Even if the Huygens Institute were to decide upon a single DTD, we still would not want to limit the use of EDITOR to editions that use this single DTD. EDITOR was meant to be potentially useful for every scholarly edition. What EDITOR does need, however, is an XML file.[59]

Local edition display
The decision to annotate the XML source format of a digital edition implies that EDITOR needs to display the edition XML to the researcher. Again because of the differences between the edition XML files, it should be possible to use different display definitions for different editions. What is needed therefore is a flexible mechanism to apply style to XML files.

It was clear to EDITOR's designers that we did not want to develop our own standard for XML display, which left us with a choice between CSS and XSLT. XSLT, while very powerful, would not just style, but also transform the XML, and thus break the connection between the displayed, annotated text locations and the corresponding locations in the original XML file. That is why we chose to use CSS for defining the display of the edition within EDITOR.

For those editions that are based on TEI, the CSS style sheets to be used will likely be very similar. No two editions are the same, however, and especially identifiers like text numbers, line numbers and columns numbers – very essential to ease of navigation – depend very much on the texts being edited. Using EDITOR on a new edition will therefore often entail creating a custom CSS style sheet. These style sheets should probably be created by the institution hosting that digital edition, not by the individual researcher.[60]

6.4   Design

6.4.1   Development vision
In the original vision, the EDITOR annotation toolset would consist of three components: an annotation creation component, an annotation display component and an annotation server. Their configuration is shown in Figure 2. The left hand side of the figure shows the standard situation of an edition XML file, a generation process and resulting HTML pages on an editorial institute's web server. EDITOR would add the three components indicated in italics.

---

[59] The file does not need to be in TEI format. However, if EDITOR encounters n-attributes (used in TEI for text, page and line numbers) it may use these to facilitate navigation.
[60] Technically, the choice was made to use JRex, a Java wrapper around the (Mozilla) Gecko rendering engine, to render the edition XML. The style sheets therefore can use Mozilla-specific CSS-extensions. This proved essential for acceptable display of the edition. Whether the decision is entirely fortunate is debatable.
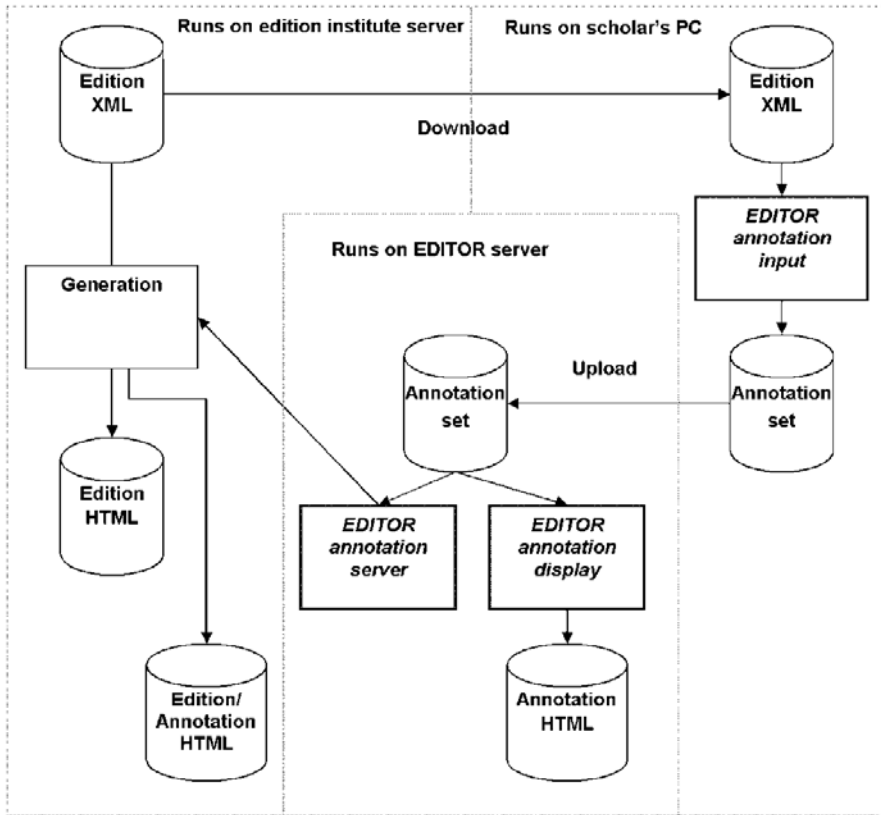
Figure 2 EDITOR components

The EDITOR annotation input component runs on the scholar's PC. It takes for its input a downloaded edition XML file and outputs the annotations created by the scholar. The scholar uploads the annotations to an EDITOR server, which may very well be hosted by the editorial institute (it may be the same server that houses the edition), but other configurations are possible. A research institute might for instance host a dedicated EDITOR server as a service to a future community of EDITOR users.

The EDITOR annotation display component runs on the EDITOR server. It displays the annotations on the web, possibly in conjunction with the edition XML file. It will also be able to show graphical visualisation of the annotations, in order to assist the scholar in the generation and testing of hypotheses concerning his/her material. The EDITOR annotation server was thought to facilitate a full integration of the annotations into the original digital edition. The process that generates the edition's HTML pages would be able to request annotation information from the EDITOR annotation server, which would

respond by sending the annotations over to the generation process. This would then be able to show the annotations in the context of the digital edition. Later it was decided that the annotation server functionality need not be specific to EDITOR. The SANE annotation exchange functionality discussed in the next chapter is meant to fulfil the role of an annotation server.

We decided to store the annotations in annotation sets, which are files on the computer disk. One can have any number of annotation sets. The annotations are not stored within the edition XML. This would duplicate the XML with each annotation set, and probably create the need for a complex process of merging annotation sets. Instead, our annotation files are really overlay files for the edition XML, similar to the approach taken by Eggert (2005).

6.4.2 Data model

The data model that the annotation sets use is given in Figure 3. An annotation set can contain annotations of multiple types. Each annotation can contain multiple fields, that are defined at the annotation type level. Annotation types also have a colour, that is used to indicate their presence in the text. Annotation fields have a name and data type: text, memo (multi-line text field), URL, category (a list of enumerated values), and related annotation.[61] A related annotation field refers to an existing annotation and hence allows annotation of relations between text fragments. For fields that are used to categorise text fragments, the system remembers the values used before in order to facilitate data entry. They are stored in the Default Value entity.

The annotations are typed by the Annotation Type entities that they refer to. They point to fragments in the edition by describing their begin and end points. The points are stored using an xpath-expression to identify a document node and an offset within that node. Input fields within the annotations get their name and type from the Annotation Type Field defined at the Annotation Type level.

EDITOR stores its data as RDF (Klyne & Carroll 2004).

---

[61] Other data types (e.g. numeric) were not considered necessary at the time but have obvious applications.
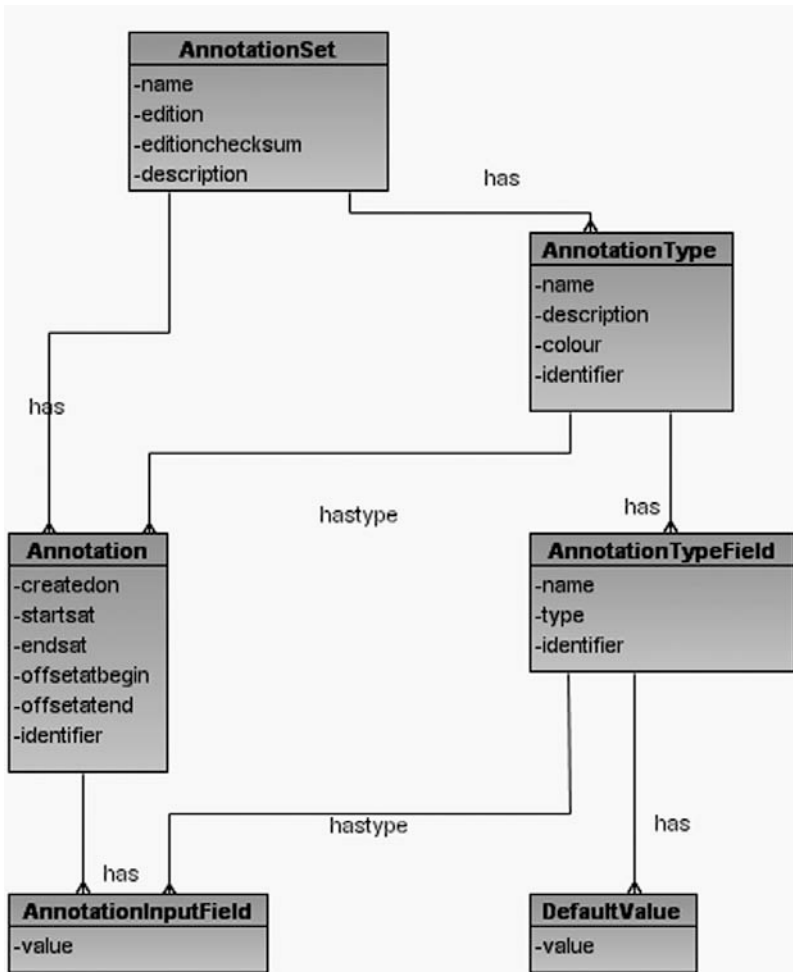
Figure 3 EDITOR data model. The set contains ('has') annotations, the annotation is typed by ('hastype') the annotation type.

6.5   The Program

Initial work on EDITOR concentrated on the annotation entry facility. Work on web display started later. The main window of the annotation entry program offers a view of the CSS-styled edition XML. A secondary window on the left displays a tree view of the XML document, which can be used to navigate through the document. Originally, the user could also access a raw XML view, but this view turned out to be hard to synchronize with the main display and was found to serve no real purpose. Accordingly, it was dropped.

Annotations can be created on text selected in the presentation window or on nodes selected in the tree view. The program will prompt for the

annotation type and then present a window where the annotation text and other data can be entered. The presence of an annotation is marked in the main view using brackets, in the tree view using bullet nodes. A separate sub-window shows all existing annotations (or a subset of them, based on type or value). Figure 4 shows the main elements of the user interface in overview. The next section will show these elements in more detail.

6.6   EDITOR in Use

6.6.1   Theatricality in Amoris Divini Emblemata

In the remainder of this chapter, I will highlight some EDITOR features based on an investigation into some aspects of Otto van Veen's *Amoris Divini Emblemata*. For a fuller report on this investigation see chapter 12.[62] One of the major innovations in this emblem book is the presence of Divine Love and the Soul in all emblem pictures but one. They are not just any two *Hinweiser*, figures that draw the viewer's attention to some aspect of the depicted objects. Divine Love uses these objects as a means of saying things to the Soul that are of crucial importance to her salvation, and jointly they are engaged in showing to the reader these objects and their importance. In many respects, Divine Love and the Soul are actors that stage condensed representations of highly symbolic scenes from the human Soul's road to salvation. In a sense, their presence turns the emblem pictures into something very much like theatre stills.

In investigation of this line of thought, using EDITOR I have been looking at, among other things, the following issues:

- How do these participants' actions (their gestures, the things they look at, the positions of their bodies) serve to bring home the emblem's message?
- In what sense does the presence of dramatic action influence the mode of reading applicable to the emblem? How, if at all, does the emblem address its reader or spectator?
- Is there a parallel to the theatrical aspects of the pictura in the emblem texts?

---

[62] I repeat here a number of passages from that chapter in order to keep the present chapter more or less self-contained.

Figure 4 EDITOR user interface. The annotation displayed shows how a single annotation can consist of multiple fields. The Annotation type Overview window displays the corresponding Annotation type definition.

6.6.2 Categories of gestures

An investigation of the participants' actions in the emblem pictures could well begin by categorising the pictures in a number of respects. In some pictures, the Soul expresses her willingness to be guided by traditional gestures of devotion, or prayer, or awe. Examples would be hands crossed before the chest[63] as a gesture of ardour, fervour, deepness,[64] folded hands,[65] hand(s) turned up in receiving or submissive mode.[66] Often they are accompanied by a kneeling attitude.[67]

EDITOR can help drawing up an inventory of occurrences of these and many other gestures, like the hand on the heart (as a gesture of assurance) or gestures of rejection. Similarly we can list the emblems on the basis of the type of relationship that obtains between the protagonists.
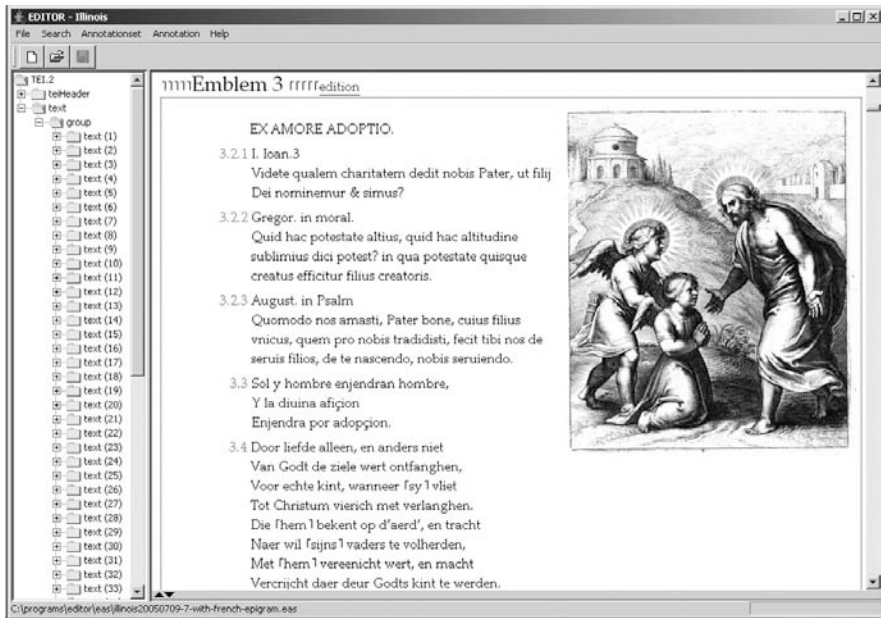
Figure 5 The edition opened for annotation

To annotate these phenomena, one would begin by creating a new annotation set in the EDITOR annotation entry tool, and selecting the edition XML file for annotation. The result is displayed in Figure 5. The right side of the screen displays the full book (the non-emblem material has been left out here). On the

---

[63] E.g. MICAT INTER OMNES ANIMA (41), see Figure 52.
[64] See Barash (1987) and Weise & Otto (1938).
[65] E.g. A MALO TUETUR (21), see Figure 56.
[66] E.g. IN UNITATE PERFECTIO (11), see Figure 64.
[67] E.g. AMORIS MERCES AMPLISSIMA (7), see Figure 57.

left side there is a tree view of the same material, which allows easy navigation through the book. There is a search function and there are hyperlinks to the emblem displays at the Utrecht site.

Now it is possible to start the process of annotation by creating an annotation type. One of the basic ideas in creating EDITOR was that it should be useful to scholars in many different fields. It is inconceivable that they should all use the same annotation types. The scholar therefore is expected to create the annotation types necessary to his or her field of study.[68] In our case we might create an annotation type 'Gestures by Anima' (see Figure 6). An annotation type can contain multiple fields. The figure shows a type with two fields: a memo field (which contains text), and a category field, meant to hold the information about Anima's folding hands or not. This field will have two possible values, yes or no.
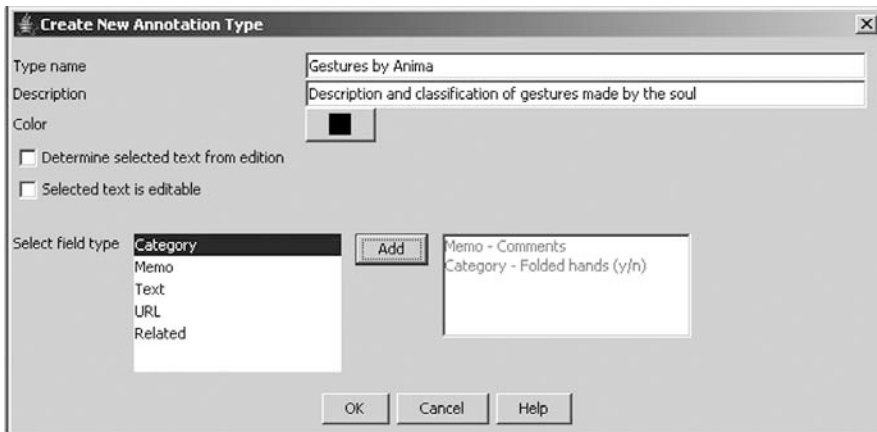
Figure 6 Creating a new annotation type

Now, why should we create a separate field for that, while we might just as well have written in the memo field: 'folded hands: yes'? The answer is more or less the same as the answer to the question why XML is used to create editions: one can do more with structured data. We will see an example below. Doubts about the suitability of an annotation and more complex interpretive issues can be addressed in the memo field.

Existing annotation types can be managed using the annotation type overview (see Figure 7). From the overview, fields can be added, edited and re-ordered.

---

[68] An import-export functionality would facilitate the sharing of complex annotation type definitions.
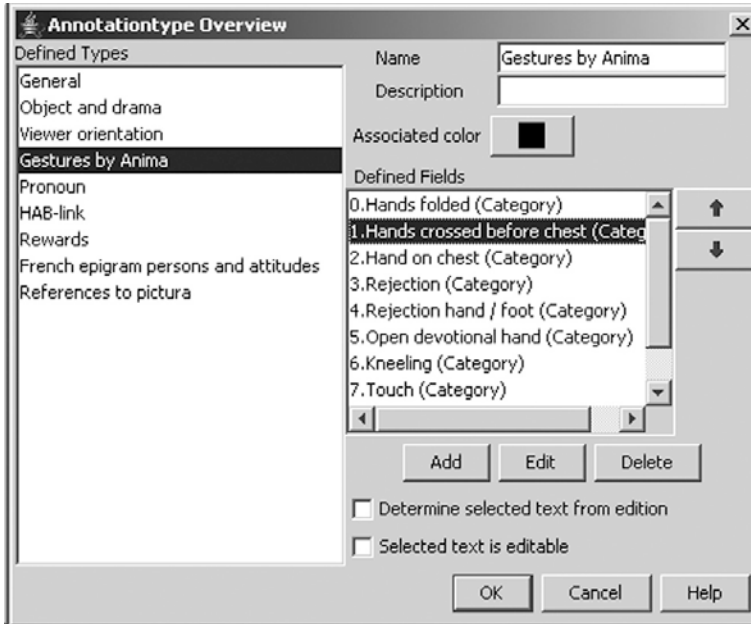
Figure 7 Managing annotation types

Annotations can be created by selecting either an element in the tree view or a stretch of text in the presentation view, and then selecting the correct annotation type in the Context Menu. The program will show a window with the fields defined for the annotation type (see Figure 8). The 'Gestures of Anima' annotation type I considered to be at emblem level, so for each emblem the appropriate information has been entered.

The annotation overview shows all annotations made in the active set of annotations. Using the annotation overview, the information can be used immediately in order to find one's way around the book. This brings out one of the advantages of using separate categories to encode different pieces of the information. The annotation overview can be filtered, by annotation type but also by category value (Figure 9). The result is a list of all annotations where, for instance, Anima is stated to have folded her hands (Figure 10). From the annotations, the related emblems are immediately accessible. This would be quite impossible if the annotation had been made using plain text in a memo field, or, heaven forbid, in a word processor.

Figure 8 Entering annotation data using the defined fields

Figure 9 Selection by category

Annotation contents	Type	Created	Selected text	Position
y, n, n, n, n, n, y, n,	Gestures by Anima	Sun Jul 10 17:28:49 CEST 2005	EX AMORE ADOPTIO. I. Ioan.3...	114
y, n, n, n, n, n, y, n,	Gestures by Anima	Sun Jul 10 17:35:44 CEST 2005	AMORIS MERCES AMPLISSIMA...	286
y, n, n, n, n, n, n, y,	Gestures by Anima	Sun Jul 10 17:42:18 CEST 2005	A MALO TVETVR. Psal.96. Qui	935
y, n, n, n, n, n, y, n,	Gestures by Anima	Sun Jul 10 17:52:51 CEST 2005	STERNIT ITER DEO. Chrysost. ...	1493

Figure 10 Result of selection by category value

6.6.3   Commenting on the position of the reader

One of the issues that come up in a theatrical analysis of the emblem is how the emblem addresses its audience, the reader-viewer of the emblem text and picture. The question is a complex one. Some of the emblem pictures address their viewers immediately, as for instance AMOR RECTUS (5, see Figure 48): Anima and Divine Love hold up the plumb line for us to see. The pictura requires our presence; the only way to make sense of the action in the picture is to assume that they are holding up the plummet for our inspection.

A theatrically important element in this picture is that there is a kind of stage area where the action takes place. There is a semicircular area bounded and delimited from the rest of the world by heavier shadows. This happens in many pictures, though by no means all (there are also many pictures where the action takes part on a path or by the roadside). The presence of a stage area suggests the depicted action is taking place not just by itself; it is meant to be seen.

Noticeable is also what in theatre studies has been called the 'technique of double address' (Barnett 1987), the fact that ostensibly the characters engage with each other, but ultimately they always engage with the public. In the case of AMOR RECTUS both Anima and Divine Love are turned towards the spectator, rather than towards each other. This is one extreme in a continuum, of which the other extreme would be Anima and Divine Love turned fully towards each other and excluding the audience. NEC VIDISSE SAT EST (47, see Figure 61) comes close to this other extreme, but there is still a slight opening up between the two lovers, which is a concession to the need to let the public in. The actors cannot perform their roles if they lose contact with the audience. The characteristic position of Anima and Divine Love is somewhat intermediate: body turned somewhat towards us, face turned more towards each other.

Returning to EDITOR: to store information about issues like these, we might create the category fields that we saw in the previous section, e.g. about there being a delimited stage area or not. But at least some of these issues perhaps do not lend themselves very well to yes or no decisions, and the appropriate way to say something about them may be a more extended annotation. What we could do therefore is to create an annotation type 'viewer orientation', have

some categories for clearly definable issues (like a look out of the picture), and one or more memo type fields for the less tangible issues.

The question may arise, again, why we should want to do that. Again the answer is that structured data offer possibilities for use well beyond those of unstructured data. Even when the annotation text has no further structure, the annotation is still of a named type, it has been placed in a named field, and it is unambiguously linked to the location that it refers to. Because of this we can sort these annotations by multiple criteria in the annotation overview; we can either group annotations that pertain to one emblem, or group by time of annotation, or group by annotation type, etc.

But having explicitly linked our annotation to a location in the edition, a multitude of other possibilities exist. E.g., if we feel confident enough about the value of our annotations we may put them on a website and present them with the edition text. That is one of the things that the EDITOR display component allows us to do. As a next step, the annotations may be made accessible from the 'official' edition site.

Figure 11 Display of the annotation set in the EDITOR display component

At present, a display of our annotation set in the EDITOR display component may look as in Figure 11. Again, on the left a tree view of the edition XML, in the center the CSS-styled edition XML itself, and on the right the annotations. The edition text links to the full edition at the Utrecht site. The edition text and

the annotations are mutually hyperlinked. The annotations can be filtered by type, and by category value. They can also be sorted on any field or on the annotation metadata (date, type, annotated text).

It is important to note that the researcher need not make this annotation display accessible to the whole world. Access rights can be set for each individual display element. It is perfectly possible to restrict access initially to a few named colleagues, and grant wider access only after deliberation, or a review. On the other hand, there is nothing to stop someone with access to an EDITOR display server from publishing his/her unreviewed annotations to the world at large. That is, however, no different from the usual situation on the web.

6.6.4   Annotating text

Going back to the emblem book once again, one may wonder whether the emblem texts contain something equivalent or parallel to the double address in the pictures. We can e.g. investigate the presence of second person pronouns, as an indication of someone being addressed, thus of the reader listening in on this conversation, and thus of the reader being the ultimate addressee of the conversation. Another interesting textual phenomenon are those places where the Soul is described not just as being guided, but as being willing and eager to be guided. It is clear that in an investigation of this type, it is important to be able to annotate arbitrary stretches of text, rather than structural elements.

To show the possibilities, we might create an annotation type 'Pronoun'. To describe the annotated pieces of text, we add some category fields: person, number, gender, and perhaps more. We might add a field that indicates the person or object that the pronoun refers to.

As before, this gives the possibility of selecting all first person pronouns, and again, the annotations can be uploaded for display on the web. Here we begin to miss the possibility of making selections by multiple criteria (as in: give me all first person pronouns that refer to the soul); and what we also begin to see is that we really need more advanced means of finding our way around in a mass of data.

Needed are visual displays of the information that we have encoded that will help us test our hunches about relations in the texts. It is all well and good to be able to sort and group annotations and to display them in the context of the edition, but it would also be helpful to have things like pie charts and histograms, and other tools that visualise the distribution of phenomena over the text.

To give a foretaste of what is needed, the annotation display component can display bar charts of numbers of annotations or numbers of occurrences of certain category values per textual element. Figure 12 displays a demo chart of first, second and third person Dutch pronouns per emblem.
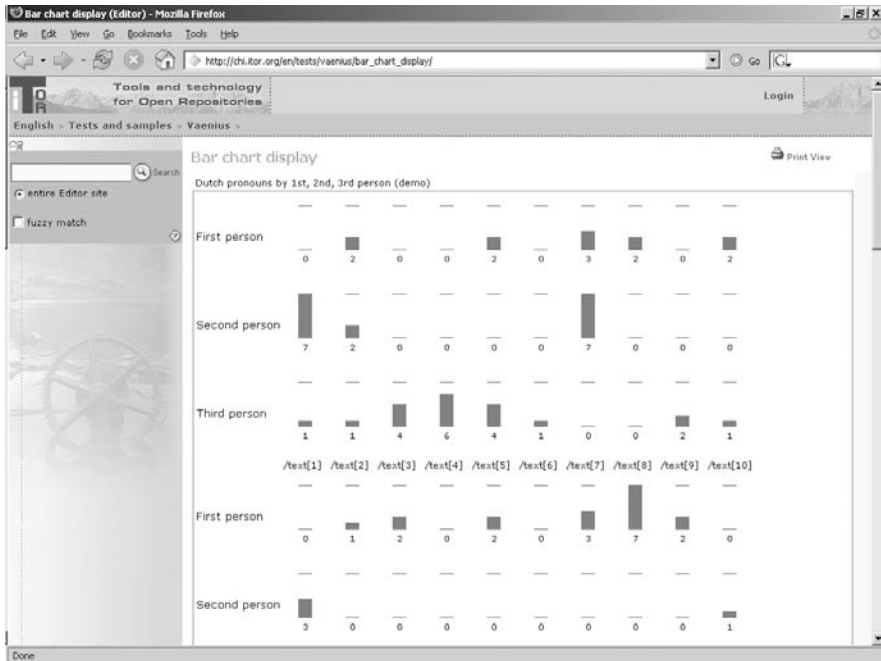
Figure 12 Partial bar chart of pronouns by person (demo)

What an annotation tool should do then is also to aggregate annotation data at a higher level, to bring together multiple collections of annotations and to help the investigator make sense of their relations.

6.7   Conclusion

At present (April 2009), the EDITOR annotation input component can be downloaded for testing.[69] There is a demo available of the annotation display component, in a very elementary state.[70] EDITOR development is frozen. Probably EDITOR functionality will be integrated in eLaborate (Van Dalen-Oskam & Van Dijk 2006).[71]

Even though EDITOR funding ran out before a product could be released to the outside world, work on its development has taught a number of valuable lessons. It has shown one way of approaching the issue of XML annotation. It has shown how a researcher can approach multi-faceted

---

[69] The program is available from SourceForge, at http://sourceforge.net/projects/editor-hi/.

[70] The display component can be accessed at http://chi.itor.org/nl/.

[71] For the latest developments see http://www.huygensinstituut.knaw.nl/editor.

questions by creating annotation types whose fields reflect these facets. It has also shown how a simple annotation display can help in navigating the texts being studied and how graphic visualization of annotation can help making sense of the phenomena under study.

What was not addressed in during EDITOR development is the integration between annotation display and the (existing) web display of the digital edition. How will the existing web edition of, say, *Amoris Divini Emblemata* 'know' that there are annotations to be displayed with certain text fragments? Moreover, if it 'knows', how is the process that generates the web edition from the XML source to take account of that information? This is one of the larger issues that any program like EDITOR will have to face. One possible answer is given in chapter 7, 'A SANE Approach to Annotation in the Digital Edition'.

EDITOR may help bring about a future where the digital edition will be, as Hrachovec and Köhler say, 'in the center of a net-based infrastructure of academic communication' (2002: 158). Hrachovec and Köhler are working on the analysis of Wittgenstein's *Nachlass*. To clarify the arguments of Wittgenstein's manuscripts they create a structure analysis and a running commentary using APE. With sophisticated technology their annotations refer back to the edition text. Their sample analysis of Wittgenstein's manuscript 115 is a good example of how scholarship can change when it begins to refer to the digital text (Hrachovec *et al.* 2004).

Scholarship will no longer necessarily take the form of articles published in journals, and, if we are lucky, as PDF files. Scholarly work of the future, created with the likes of APE, EDITOR, and their successors, will be fully integrated into online digital editions. The edition will then become a never-finished collaboratory – in its richness, variety, and growth not unlike the wider Web of which it is a part.

7 A SANE Approach to Annotation in the Digital Edition

7.1 Introduction[72]

Recent discussion about the scholarly digital edition has focused on ways to change the edition from a passive text, only there to be consulted, into a dynamic research environment. Siemens (2005) asks why as yet we have seen no convincing integration between text analysis tools and the hypertext edition. An essential step towards seamless integration of text analysis tools into the digital edition (TAML, the Text Analysis Mark-up Language) has been suggested by Sinclair (2005). The most visionary statement of the dynamic edition's potential is no doubt given by Robinson (2003). A dynamic edition, in his view, while offering text analysis and other tools that may shed light on some aspect or other of the edited texts, would also be open to the addition of new content, of corrections, and of many different types of annotations.

This chapter will focus on making available third-party annotations in the digital edition. A digital edition, in my perspective, consists of (1) an XML source file (or a collection of XML source files), probably encoded according to the *Guidelines* of the Text Encoding Initiative, (2) a set of HTML pages, image files and/or other output formats, used to display the edition contents to the online reader, and (3) a transformation process or other application that produces (2) from (1). The main challenge in annotation of a digital edition (and that which makes the task different from web annotation in general) is that the annotations should refer to the XML source file, but should be displayed to the reader in the context of the output files.

Research as described in the chapter continues work on the EDITOR annotation tool discussed in chapter 6. EDITOR facilitates creating annotations to XML files. The annotations can have user-defined types, they consist of multiple fields, and those fields can have their own data type. The annotations are created in sets that are expected to have some internal coherence (perhaps studying a single object, perhaps made to support the argument of a single article). They are not stored within the edition XML but in a separate file on disk. EDITOR also includes a facility (very limited) for display of these annotations

---

[72] This chapter is a lightly revised version of Boot (2006b).

over the web. In this chapter I describe work on a prototype for a component that facilitates the inclusion of EDITOR-created annotations in the digital edition. The chapter proposes exchange standards for requesting annotation information from annotation servers, and these standards will create the possibility for connecting other annotation creation tools and annotation display tools.

I will begin by discussing related work, in section 2. The remainder of the chapter introduces a markup language and a request protocol for the exchange of annotation information (in section 3); section 4 will discuss a prototype built using this protocol and markup language. In section 5 I will examine some issues arising from the approach. Conclusions and future work will be outlined in section 6.

7.2   Related Work

Annotation technology is a wide field, that is not easy to map. Most of the work in annotation studies, however, seems to centre on web annotation, annotation in digital libraries or annotation as a collaboration tool. While these terrains have much in common with annotating the digital edition, there are also important differences.

As said before, the essential difference between web annotation and annotation of the digital edition is that the object of annotation in this last case is not a (series of locations in a) web page. It is a location in an XML source file that is nontrivially transformed into one or more web pages. The annotated file itself is not visible to the reader. In the case of edition annotation, therefore, the edition server must insert the annotations into generated HTML pages based on knowledge of the transformations applied to the XML source file. This makes all but impossible what is quite common for web annotation, i.e. that the annotated pages are typically not 'aware' of having been annotated. That is why web annotation systems, such as W3C's Annotea project (Kahan *et al.* 2001; Koivunen 2005), can have annotation clients (such as Annozilla)[73] that insert the annotation information into the HTML pages that the user sees. While that approach makes for simplicity with regards to the annotated objects, it also necessarily limits the amount of integration that can be achieved between the annotation display and the annotated HTML pages. In our case, the edition server must be aware of the existence of annotations, and it must take responsibility for showing these to the user. While this may be an extra complication, it also makes possible an integrated display of content and annotations that is essential for facilitating in-depth study of either.[74]

---

[73] http://annozilla.mozdev.org/.
[74] The limitations of Annotea's annotation model are another reason why it is not really suitable for the task of scholarly annotation.

A number of studies have reported on digital library annotation systems. In the annotation typology developed by Marshall (1998a) these systems seem mostly to focus on (hyper-)extensive annotation. Scholars use the system described in Agosti *et al.* (2005) to annotate relations between images in digital manuscripts. Arko *et al*. (2006) describe a system where patrons can annotate resources in a digital library as to their usefulness. For me, one of the interesting things about their approach is that they describe a procedure for annotation exchange using the Open Archive Initiative's Protocol for Metadata Harvesting (OAI-PMH).[75] My proposal for an annotation exchange protocol was in some respects inspired by the OAI-PMH. The OAI-PMH protocol itself is not suitable for the kind of exchange envisioned here, for a number of reasons:

- the requested annotation information will not always be plain (meta-)data, but may also be counts of items (annotations) or annotation information already formatted for display;
- OAI-PMH requests are typically made to get all of the data in a repository and store them for later use; the requests I envisage are made to serve a specific user request and therefore must support fine-grained 'selective harvesting': getting annotations selected by e.g. annotation type and/or annotation value and/or annotated location ('give me all annotations of type pronoun, where number is singular, that apply to the last stanza of poem x').

Annotation as a tool for collaboration is in itself a very wide field. Its application ranges from the creation of business documents in a workgroup (Cadiz *et al.* 2000) to scholarly cooperation (Van Zundert & Van Dalen-Oskam 2005). Most initiatives seem to assume however that there is a single context for annotation creation and annotation display, that is shared between the collaborating persons. In many cases, this will be a reasonable assumption. In the case of digital edition annotation, however, the creation and display contexts are crucially different. In the present chapter we will also look at exchange of annotations between multiple tools or repositories.

7.3   Introducing SANE

7.3.1   Design Considerations
The protocols and software for annotation exchange that I discuss here will be collectively called SANE, short for Scholarly ANnotation Exchange. Designing SANE I was guided by the following considerations:

- While some annotations may be stored within edition source files, these are not the annotations that are our present concern. The editors that create an edition will want to include the annotations they regard

---

[75] http://www.openarchives.org/OAI/openarchivesprotocol.html.

as necessary. SANE focuses on providing annotations from third parties, and even if it would be feasible to include all third-party annotations in the edition source file, that file should not be exposed to the risk of corruption that comes with any kind of change. The annotations that we are concerned with are therefore stored outside the edition source files.

- Annotations refer to stable locations in edition source files. This assumes that if the edition should change, the older version, including its source, will remain available.[76]

- I propose a model where the annotation handling is relegated as much as possible to a specialised annotation server. This server may, but need not, run on the same physical machine as the edition application. There are multiple reasons why this model makes sense: (1) annotation processing may be a heavy task, requiring the presence of infrastructural software (e.g. databases, indexing software) that the edition machine not necessarily houses and the edition systems staff is not necessarily comfortable with; (2) annotations may have a complex internal structure that the edition should not have to be aware of (EDITOR-created annotations e.g. have that complexity); and (3) annotation display functionality may be complex.[77]

- While leaving the annotation handling to the annotation server wherever possible, in the context of the digital edition it is still the edition server that is ultimately responsible for the annotation display. A decision where and how to display the annotations requires knowledge of the structure and semantics of the edition source XML and knowledge of the logical structure of the edition's user interface.

- Apart from a protocol for annotation exchange, we also need agreement on a storage format. It should be possible to display a selection of annotations in a specific context (which is what the request protocol will allow us to do), but it should also be possible to do wholesale import or export into or from multiple annotation tools. Because of this, I will introduce both a markup vocabulary and an exchange protocol for annotation information. See below for a further discussion of the question why we need both of these.

The next two subsections will introduce the SANE Markup Language and Request Protocol that facilitate the communication between annotation

---

[76] With editions and annotations made for scholarly purposes, this should be self-evident. To make the annotations available in the newer edition version, the references in the annotations may have to be updated. That issue is outside the scope of this chapter.
[77] Display complexity can result from the complexity of the annotations themselves and/or from the complexity of the desired functionality. We will e.g. require overviews, categorisations, counts and diagrams.

creation tools, annotation servers and edition servers. In the last subsection of this section we will look at other possible usage scenarios.

7.3.2   SANE Markup Language

The SANE Markup Language, or SANE-ML for short, is an XML markup vocabulary that can be used to store or exchange annotation information. It is defined by a Relax NG schema. A SANE-ML document describes a single annotation set. The data model at present reflects the EDITOR data model (see chapter 6), but it is intended to be tool-independent. An initial, annotated version of the schema is available in appendix A.

A SANE-ML document contains multiple sections. Their contents are briefly described in Table 1. Please refer to the schema annotations for more detail.

Table 1 Sane-ML sections

setData	Contains information at set level: name, author, creation and modification dates, rights information, program and program version used to create the annotation set, contact information, (references to) extended prose description of the purpose and meaning of the annotations. This section, like the next one, holds information stored in the AnnotationSet entity in the data model shown in Figure 3 on page 85, as well as some information that EDITOR did not store.
editionData	Contains information about the edition(s) annotated in the current set: URL, checksum and a brief description.
typeData	Contains information about the scholar-defined annotation types: name, description, and the fields of which they consist. The fields in their turn have a name and description. Besides that, fields have a data type, and they may have a number of default values. This section holds the information stored in the entities AnnotationType, AnnotationTypeField and Defaultvalue in the data model.
annotationData	Contains information about the annotations themselves: the editions and the locations in the editions that they refer to, the annotation type that they instantiate, the text selected for the annotation and the annotation's values for each of the annotation type's fields. Holds information from the Annotation

	and AnnotationInputField entities in the data model.

Besides (meta-)data about the annotations, a SANE-ML document may also contain information about appropriate display characteristics for the annotations. At the set level, a default colour scheme can be defined. A property at the annotation type level specifies whether highlighting the selected text in the edition is appropriate for this annotation type (for annotation types that are typically applied at word or phrase level this property will probably be set to true; annotation types that apply to larger text units, such as chapters, will have the property set to false).

7.3.3   SANE Request Protocol

The SANE Request Protocol, or SANE-RP, defines a protocol for the exchange of annotation information between an annotation server and an annotation requestor. As in the Open Archive Initiative's OAI-PMH protocol, requests consist of a verb with a number of parameters in the URL querystring.

The SANE-RP `mode`-parameter is used to request either output as XML data, to be processed by the requesting application (`mode=data`), or as XHTML and/or SVG data, ready to be displayed to the user (`mode=display`). With `mode=display`, only minimal effort is required from the application for displaying the annotation information.

Presently defined SANE-RP verbs are:

Table 2 Sane-RP Verbs

`identify`	**Description**: requests identifying information from the annotation server, such as name, SANE-RP protocol level that it understands, default rights information, contact information, etc. **Required parameters**: mode.
`getSetData`	**Description**: requests set level information, including information about the annotated editions and some information about the available annotation types. **Required parameters**: mode, setName.
`getTypeData`	**Description**: requests annotation type information, including the description of the annotation fields that the type contains. **Required parameters**: mode, setName,

	typeName.
`getTypeCount`	**Description**: requests the number of annotations of a certain type, and distributions of values for its constituent fields (of type category). **Required parameters**: mode, setName, typeName.
`getTypeCountUnit`	**Description**: requests the number of annotations of a certain type, by textual unit. The textual units are defined by a parameter (nodeListURL) that refers to a file that lists the top nodes that define the units. **Required parameters**: mode, setName, typeName, nodeListURL. `mode=display` will return an SVG bar chart of the annotation count by textual unit.
`getAnnotation`	**Description**: requests data about a single annotation. **Required parameters**: mode, setName, annotationId.
`getAnnotationData`	**Description**: requests data about annotations to a single node and its descendants (editionNode parameter) or to a series of nodes (nodeListURL parameter). Returns node elements with the applicable annotations. Optionally filters the annotations by annotation type, field, and/or value. **Required parameters**: mode, setName, editionNode or nodeListURL. **Optional parameters**: typeName, typeFieldName, searchTerm.
`getAnnotationNodes`	**Description**: from a list of nodes (nodeListURL parameter), getAnnotationNodes returns the subset of those nodes of which a descendant has been annotated. Optionally filters the

	annotations by annotation type, field, and/or value. **Required parameters**: mode, setName, nodeListURL. **Optional parameters**: typeName, typeFieldName, searchTerm.
`getAnnotationFragments`	**Description**: requests XML edition fragments with included annotation markers, either for a single node (editionNode parameter) or for a series of nodes (nodeListURL parameter). The annotation markers signal the span of text that the annotation applies to. Attributes give annotation id and the annotation type. Optionally filters the annotations by annotation type, field, and/or value. **Required parameters**: mode, setName, editionNode or nodeListURL. **Optional parameters**: typeName, typeFieldName, searchTerm.
`getAnnotationCatCounts`	**Description**: For a category-type field in an annotation type, returns the values of that field and their number of occurrences. With `mode=display`, the result is formatted as a pie chart. **Required parameters**: mode, setName, typeName, typeFieldName.
`getDisplayData`	**Description**: request suggested display information for an annotation set. **Required parameters**: mode, setName.

7.3.4   *Usage Scenarios*

The principal usage scenario for which SANE is being designed is the display of EDITOR annotations in the digital edition. It is easy, however, to imagine other applications for SANE (see Figure 13):

1. Displaying annotations from other annotation tools in the digital edition: annotations created using tools such as Pliny (J. Bradley 2006),

Atlas.ti[78] or APE (Köhler 2006) could be made accessible from a digital edition if either (1) they can be exported into SANE-ML or (2) these tools provide SANE-RP-based access to their data.

2.   Displaying automatically generated information in the edition: the information displayed need not result from manual annotation. The information to be displayed might be the output from a concordance tool, it might be a page popularity count or a list of frequent referrers.

3.   Statistical analysis of annotations: the requesting application is not necessarily a digital edition. It might be a piece of software that can be used for statistical analysis of the annotations (perhaps for computing correlations or doing regression analysis).

4.   Text mining of annotations: the requesting application might also be a system like nora (Plaisant *et al.* 2006). Based on the categorisations provided by an annotator it might try to predict a categorisation for other texts in the corpus.

5.   Archiving an EDITOR annotation set: researchers may want to safeguard their annotations against eventual disappearance of the EDITOR annotation tool. One possible way to do that is to create a single XML file that contains both the edition source and the annotations that the researcher created. An archiving application could use the SANE-RP `getAnnotationFragments` verb to request the edition XML enhanced with annotation markers and the `getAnnotation` verb to request the annotations themselves.
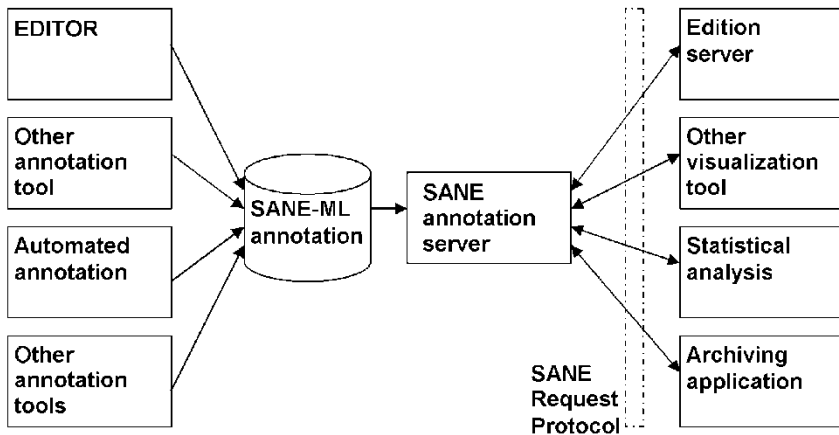
Figure 13 SANE-ML and SANE-RP

In most of these potential applications, there is a use for both SANE-RP and for SANE-ML. From the point of view of the requesting application, what matters is the protocol. The requesting application (the digital edition, the archival tool, the nora-like text mining application) is not concerned with the annotation storage format (SANE-ML). However, from the point of view of the application that creates the annotations or annotation-like information (e.g. page visit counts), the availability of SANE-ML is essential. If there were no SANE Markup Language, these tools would have to write an annotation server of their own. For annotations stored in a standard format, an annotation server that knows that standard will know how to handle requests for these annotations.

7.4   Prototype

In order to test the feasibility of the SANE approach, a prototype has been developed. It consists of an edition of Otto van Veen's emblem book *Amoris Divini Emblemata*, as digitised in the Emblem Project Utrecht, and an annotation server that implements SANE-RP. Both are built in Cocoon[79], the Apache Foundation's XML publication framework.

        The main annotation set used for testing was an annotation set developed to study some aspects of theatricality in the emblem book. Most annotations in this set are made at the level of the emblem, others at the level of individual words. Most of the annotation types contain a number of categorisation fields.

To prepare the files for showing them through the prototype, a number of actions were taken:
- The XML edition had been simplified for the purpose of testing EDITOR. Some of the simplifications were removed, and the EDITOR annotation set was changed correspondingly (to keep references correct). This should not ordinarily be necessary. Both actions were done using an XSLT stylesheet.
- The EDITOR annotation set was transformed into a file that conforms to the SANE-ML schema. This involved a change from RDF/XML (which EDITOR uses) into 'plain' XML. During the transformation some information was added that EDITOR does not store. The transformation was effected using an XSLT stylesheet. Future EDITOR-like programs could use SANE-ML as an export format, making this step redundant.
- To aid performance, a copy of the edition XML source was created where the annotated fragments are marked by `seg` elements. Overlap between annotations and between annotations and structural elements was handled by, where necessary, creating multiple `seg`

---

[79] http://cocoon.apache.org/2.1/.

elements for a single annotation. The file was created using two PHP scripts.

A future annotation server would probably include a front-end interface for entering new annotation sets into the annotation repository. One of the tasks to be triggered upon receiving a new annotation set would be the creation of this edition-plus-annotation XML. Alternatively, a production version of the annotation server could use a DOM-structured database for storing the edition XML and the annotations, as suggested by Bradley (2004).

Figure 14 Showing available annotation sets

The edition of the emblem book uses the left-hand side of the screen. The edition is very basic and consists of a table of contents and text pages. In a fuller edition, the right-hand side of the screen would be used for translations, commentary, introduction, etc. Here, however, we are interested in showing external annotations. In overview, they may be presented as in Figure 14. The options for limiting the display to reviewed annotation sets and for sorting the display are at present inactive. Clicking the server name will fire a SANE-RP `identify` request, the result of which will be shown below the display options. The identify request, as most of the other requests we will encounter,

uses the `mode=display` option. The results therefore can be included 'as is' in the HTML page that the edition creates.[80]

Selection of an annotation set for display sends a SANE-RP `getSetData` verb to the annotation server. The results include set level, edition level and annotation type level data, as shown in Figure 15. From here, we have a number of options. Pressing the copyright sign will show us the copyright information for both the edition and the annotation set. We can ask for the annotations to be displayed right away, using the 'add to layout' icon on the set or on the annotation type level. On the type level, we can ask for more information about the type (clicking the type name), ask for counts (using the calculator icon), or ask for a bar graph of the number of annotations by textual unit.
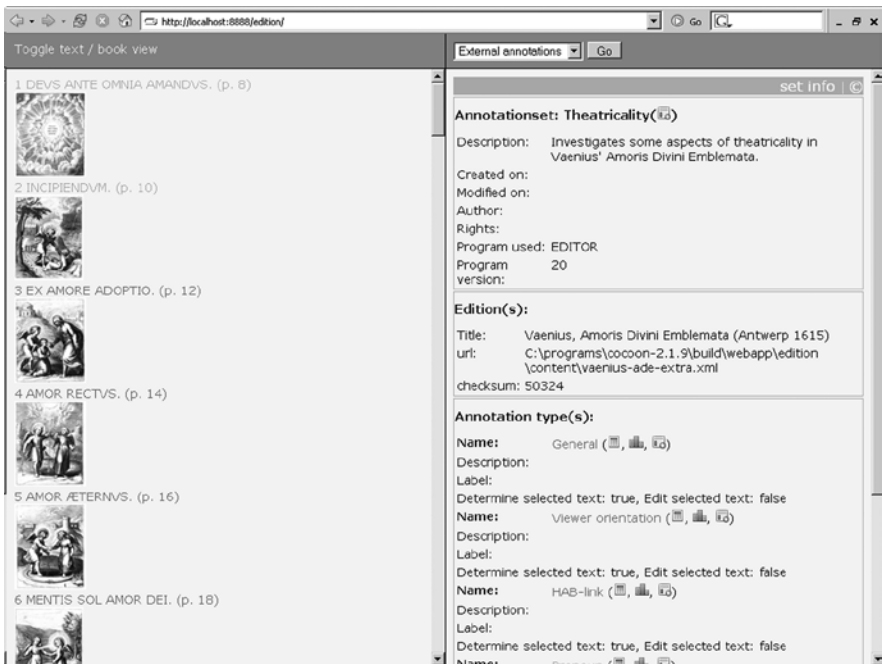
Figure 15 Showing an annotation set

Choosing the latter option on the Pronoun annotation type sends a `getTypeCountUnit` request to the server. If `mode=display`, the server will answer with an SVG chart of Pronoun annotations by textual unit, as seen in

---

Figure 16. The bar numbers in the chart are hyperlinked to displays of the corresponding emblems with the Pronoun annotations being shown.
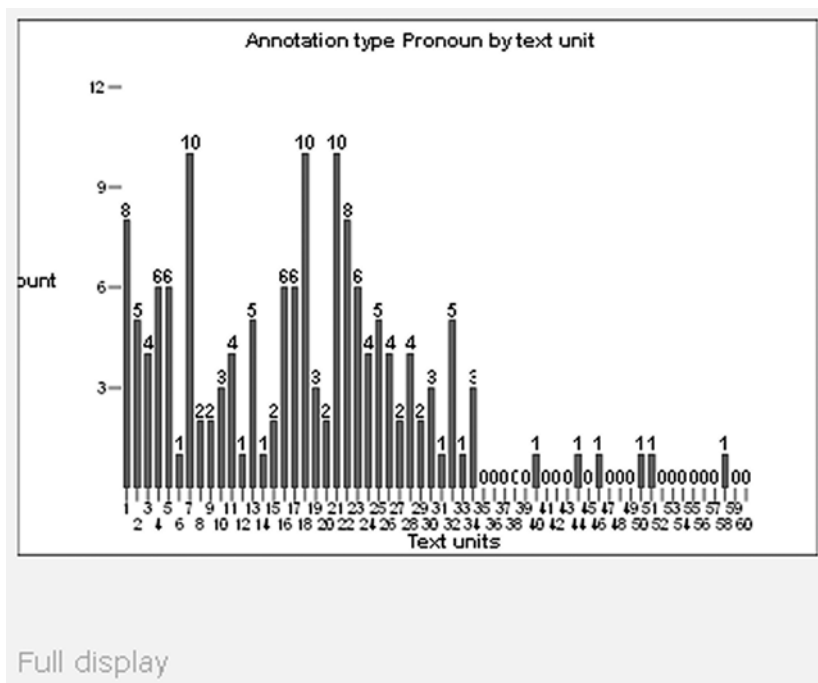
Figure 16 Showing the number of annotations of a certain type by text unit.

Rather than asking for the bar chart, we might also have asked for more information about the annotation type (which would have caused a `getTypeData` request to be sent to the server). That would result in a display including the annotation type fields. For fields that are categories we could have asked for distribution information (`getAnnotationCatCounts` verb), which would result in the display of a pie chart of the field's values. The pie slices again would be hyperlinked to a display of the book with the corresponding annotations being shown (Figure 17).

To display the book with (a selection of) the annotations, the edition sends a number of requests to the annotation server. A `getAnnotationNodes` request returns the nodes that have (descendants that have) annotations that conform to the annotation filter. The resulting list of nodes is used to create a filtered table of contents, which is shown on the left-hand side of the window. A `getAnnotationData` request returns the annotations themselves, which will be shown at the right hand side. Here the system also displays the active annotation filter. See Figure 18.
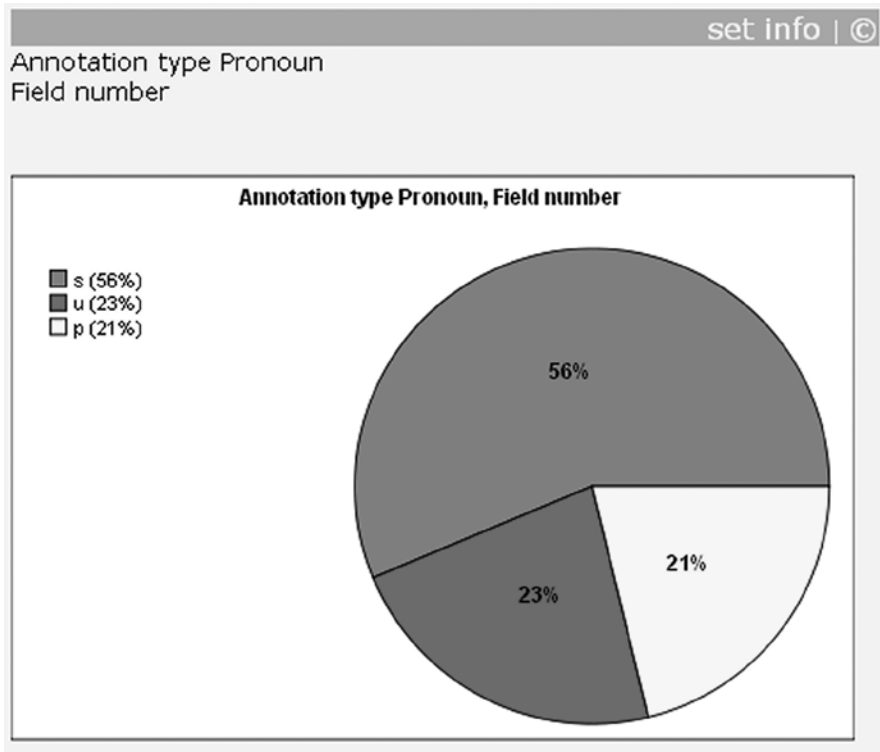
Figure 17 Showing the distribution of values of a category field

From here, we have multiple possibilities. Clicking an item in the table of contents, we can ask for an annotated display of an individual emblem. We can also remove one or more of the filter criteria, thus including a larger number of annotations in the display. Depending on the strictness of the filter applied, the annotation display itself might include hyperlinks that add extra selections to the filter.

In Figure 19, we see a display of one emblem from *Amoris Divini Emblemata* with all available annotations. The XML underlying this display was delivered by the annotation server after a `getAnnotationFragment` request. In the display, the start and end locations of annotations are marked using coloured triangles. One of the annotations is being shown in full, an action that results from placing the mouse over the opening triangle (which fires a `getAnnotation` request). Placing the mouse over the closing triangle of an annotation underlines the annotated fragment. This is useful in the case of overlapping, nested or coinciding annotations. Notice that some of the annotations have their annotated text highlighted, others do not. Notice too that each annotation type has its identifying colour. These aspects of the display result from data returned from a `getDisplayData` verb.
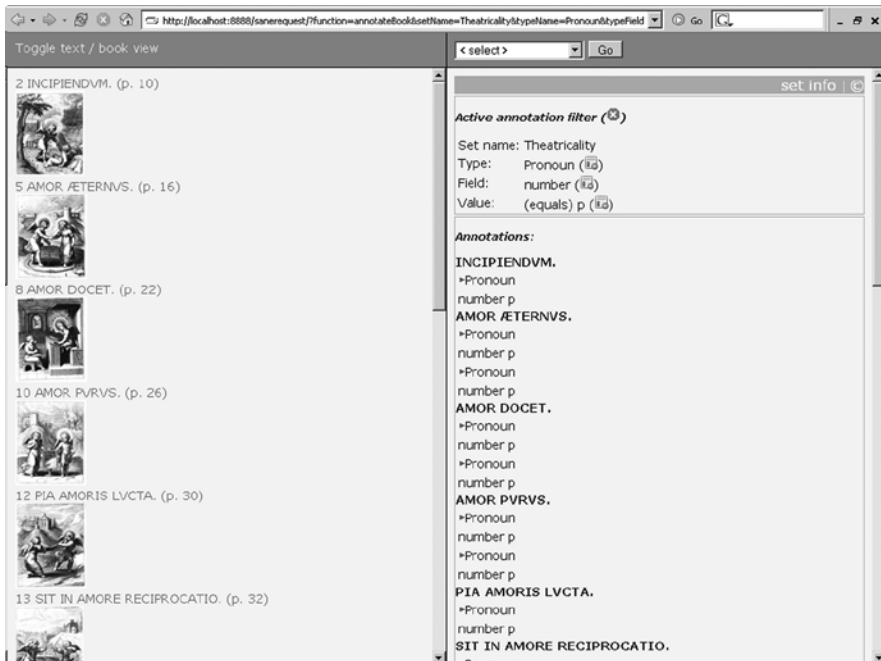
Figure 18 Showing a table of contents, the annotations and the annotation filter.

7.5   Discussion

What the prototype does show is that it is feasible to integrate the display of external annotations in the digital edition. It also suggests, I believe, that this integration will prove to be a useful thing. Being a prototype, it lacks a number of desirable features:

- it should be possible to change the order of displayed annotations;
- it should be possible to display annotations from multiple annotation sets;
- it should be possible to integrate (a simple visualisation of) annotation counts in the edition table of contents;
- it should be possible to search the edition and the annotations conjointly.
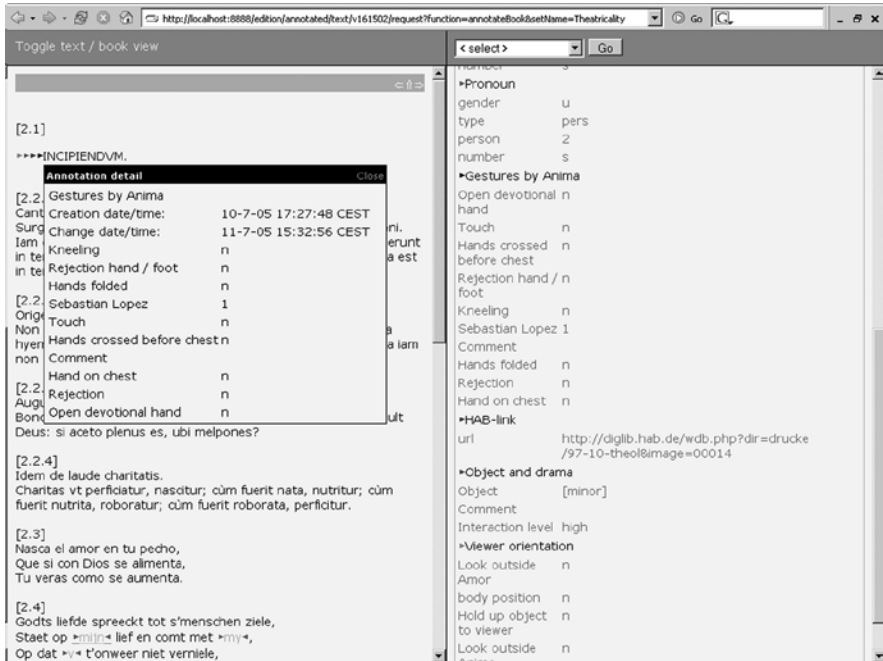
Figure 19 Showing all annotations for an emblem.

The main problem in the creation of the prototype has been the demarcation of functionality between annotation server and edition. A fundamental decision in SANE-RP design was to have the annotation server do as much of the annotation handling as possible. Unlike the OAI-PMH, SANE-RP can be used to create HTML fragments ready for display, using the `mode=display` parameter. The advantage of this approach is that the display functionality does not need to be built into the editions that request the annotations. However, as the display of the annotations is meant to be integrated with the edition display, responsibility for display can never lie with the annotation server alone. This creates a situation where the spheres of concern of the annotation server and the edition are not always clearly demarcated.[81] Especially tricky are the hyperlinks included in the annotation display, as they are created by the annotation server but need to be acted on by the edition.

A perennial issue in XML annotation is the fact that annotations do not necessarily fit in the document hierarchy. In our case, we assume the

---

[81] As stated above, the edition and the annotation server were built using Cocoon. I am sure that for some of the things I grappled with Cocoon has to offer more elegant solutions than the ones I came up with. The fact that the issues mentioned in this section arose at all, however, is I believe due to the problems of demarcation between edition and server. Perhaps mode=display is not a good idea.

annotations are stored externally, and refer to the annotated fragments using pointers to their begin and end points. They can therefore overlap, nest or coincide both among themselves and with(in) the elements of the XML hierarchy. For display purposes, an hierarchy that aggregates the document and its annotations is built behind the screens, but that is a temporary document, not meant to be seen by the human eye.

7.6   Conclusions and Future Work

I opened this chapter by mentioning Peter Robinson's vision for the future of the digital edition. The ability to include externally created annotations is only a small step in the realisation of that vision. What I propose here is in some ways even conservative. SANE is about annotation of an existing and stable edition. Within SANE, there is no way for a scholar to add new texts, to correct decisions made by the editor, or to annotate another person's annotation.

What can SANE mean then for the edition, 'our mutual enterprise', as Robinson terms it?

A facility like SANE (in conjunction with an annotation creation tool like EDITOR) will probably be used in one of two ways:

- to add a body of annotations to an edition as a service to the reader, in order to facilitate comprehension of the text and its contexts (annotations to personal names, to obscure places, references to parallel texts elsewhere, to illustrations, etc.);
- to add a body of annotations that support an interpretive or theoretical claim about the edited work. Here the annotations may serve as a linking pin between the online edition and a piece of scholarship, perhaps an article, that they provide the details for.

Both of these applications show the collaborative nature of scholarly work. In the first case, the edition originally created by an editor is enhanced by another scholar for the benefit of a third person, the reader. In the second case the scholar is no longer merely using the edition to create his or her own scholarship (the article), but in return makes the details of that scholarship accessible to the user of the digital edition, and thus to other scholars.

Whether facilities like these will prove to be a success is hard to predict. What seems clear is that a proposal along the lines of SANE creates an opportunity for new tools that create annotations and new tools that know how to do interesting things with them. The next steps to be taken in order to make annotation exchange along these lines a true possibility would be:

- create a formal description for SANE-RP: even though it may be impossible to create a definitive formal description at this stage, people that want to use the protocol need a document to base their implementation upon;

- experiment with a TEI version of SANE-ML;[82]
- create an experimental server that serves the SANE request protocol based on annotation sets that use the SANE Markup Language. This would entail both creating the software and run it as a service for those willing to experiment with SANE client software;
- change EDITOR to use SANE-ML as an export format;
- create a demo edition that shows the benefits of being able to create annotations to a scholarly text, and having these annotations show up in the context of the annotated text.

Taken together, these steps create the conditions for joint experiments in annotation exchange. The editions to be created will profit from being able to include annotations from many different sources.

---

[82]For that experiment, see chapter 10.

8   Decoding Emblem Semantics

In the two previous chapters, I described an annotation approach where each annotation is basically a distinct piece of information about some aspect of the annotated text(s). In this chapter and the next, I investigate an approach where research can result in a network of interrelated annotations. This network is structured by an ontology of the domain on interest.

8.1   Introduction[83]

Texts carry meanings. Authors encode meanings into texts, and readers decode meanings from texts. Literary scholars study the ways in which meaning has been encoded into texts, and their expertise can help them decode older or otherwise more complex texts.

　　This decoding process results, presumably, in a new or increased understanding of the text in question. It may result in a new text, which explains this new understanding of the original text. These interpretational texts may have a merit of their own, independent of their relation to the original text. Nevertheless, one of the criteria by which we rank interpretations is the light they shed on the texts they comment on.

　　Ultimately, there is no criterion for judging the validity of an interpretation. What we can do, to a certain extent, is to verify the interpretation's claims as to the original text. We may surmise that this verifiability depends at least partially on two things: the frequency and specificity of the interpretation's references to the original text, and the clarity and applicability of the theoretical notions which it applies to that text.

The texts this thesis is concerned with are emblems. An emblem may be loosely characterised as a combination of a motto, an image (*pictura*), and a subscription (one or more poems or prose texts accompanying the image). The emblem is an example of a fairly complex 'text' genre, and the relations among its parts have been subject of scrutiny by modern literary theorists.

　　Figure 20 shows an emblem from Otto van Veen's *Amoris Divini Emblemata* (Antwerp 1615). The text consists of a motto ('Superna respicit', 'She looks upward'), a number of quotations from the church fathers, and three poems (in Spanish, Dutch, and French). The image depicts a sunflower, turned

---

[83] This chapter is a lightly revised version of Boot (2006a)

towards the sun, and the Cupid-like figure of Divine Love is pointing this out to the Soul (the winged girl), who is told to do likewise. The poems help to understand the scene: the sun represents God, and the Soul learns from Divine Love that she should be like the sunflower: as the sunflower follows the sun, the Soul should follow God.
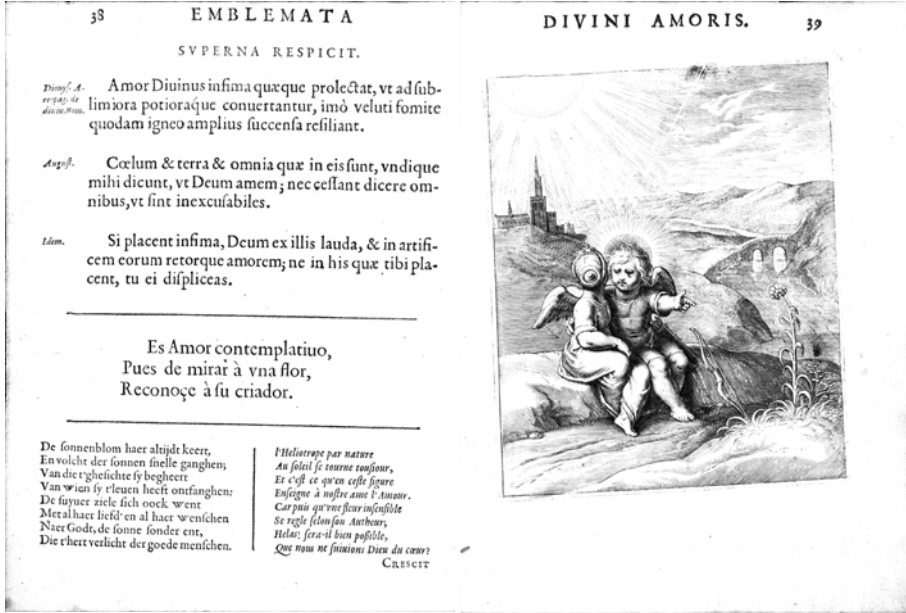
Figure 20 SUPERNA RESPICIT (16). From Otto van Veen, *Amoris Divini Emblemata* (Antwerp, 1615)

This loose description represents the beginning of an interpretation. It may be developed by looking at the second quotation, where Saint Augustine is quoted as saying 'heaven and earth, and everything which they contain, on every side they say that I should love God'.[84] Van Veen may have added the quotation to strengthen his argument: the sunflower is one of the things which, as Augustine states here, teach us to love God.

We see the interpretation resulting from the interplay between textual and visual elements in the emblem. And it is precisely the ways in which these textual and visual elements interact which have been the subject of constant speculation by twentieth-century emblem scholars.

What one would like to do, therefore, is to formally describe the way in which text and image constituents contribute to the overall meaning of the

---

[84] Coelum & terra & omnia quae in eis sunt, undique mihi dicunt, vt Deum amem. *Confessiones* 10.6.8 (O'Donnell 1992).

emblem. I will take a semiotic approach, and consider emblem texts and images as vehicles for signs. I will therefore need to model these signs, their bearers in text and image, and the semantic structures they carry.

This chapter focuses on some of the more technical aspects of this desideratum. I will discuss the kind of questions one would like to see answered by the model (section 8.2) and look at some other approaches (section 8.3). The chapter then goes on to describe a basic sign model (section 8.4) and its implementation (section 8.5). Also, I discuss how to employ the resulting encodings to answer the questions previously asked (section 8.6), as well as some future work in section 8.7. Finally, in section 8.8, I return to some issues discussed in the Introduction.

8.2 Use Cases: Emblematological Questions

Modelling sign structures takes significant effort, which requires some justification. Here I mention some of the questions which a sign model should help in answering.

Querying image content

Ways should be provided to query image content. For example: 'Does the sun appear in all of these emblems?', 'What objects with allegorical meaning are used?', 'What or who is the Soul looking at?'

Complex text queries

In the context of the discussed emblem book, one might want to search for emblems where Divine Love is drawing the Soul's attention to some object. Even if one were to restrict a query to the text of the emblems, it is clear that text search facilities alone will never be sufficient.

Querying semiotic relations

Scholars would similarly want to investigate the presence of semiotic relations, such as: 'In which emblems is a certain aspect of meaning present in both text and image?' Where do the quotations provide support for the line of reasoning established in the epigrams?' 'If there is a metaphor, where do the pictures contain elements from both tenor and vehicle?'

8.3 Related Work

In some sense, what this chapter does is to devise a way to store metadata about digital objects. This section will mention some similar projects, and discuss how they relate to what I hope to do.

Iconclass is a very large cataloguing system, meant to describe the contents of visual works of art (Van den Berg 1995).[85] Iconclass provides a

---

[85] See also http://www.iconclass.nl/.

hierarchy of categories. For describing the pictorial content of emblems, probably no classification system is better suited than Iconclass. What is missing in Iconclass, from my perspective, is a predefined vocabulary to describe relations between pictorial objects.

Wielemaker *et al*. (2003) have redefined Iconclass (and other thesauri) as a Resource Description Framework (RDF) Schema.[86] In the schema, relations between depicted objects can be expressed using a sentence template with four variables: agent, action, object, and setting. For our sample emblem, this would allow us to express, e.g., that the sunflower (agent) is turned towards (action) the sun (object). The approach is clearly an important step towards structured description of image content.[87]

Using *METS* (Metadata Encoding and Transmission Standard), descriptions can be added to regions within digital material of any kind (Library of Congress 2003).[88] The facility seems to be used mainly for high-level description, but it could also be used to provide detailed markup. As METS allows the inclusion of externally defined elements, it is certainly possible to integrate METS with other approaches.

Annotea is an extensible framework for sharing annotations on the World Wide Web (Kahan *et al.* 2001).[89] The basic annotation model is very simple. However, the model has been defined in RDF Schema, and is meant to be extended for more specific purposes. One could easily envisage the model presented in the following sections to be stored in an Annotea server.

On a concluding note, after this brief review, it appears that many possible ingredients for documenting meaning aspects of text and images already exist. Especially in the field of ontology definition (see section 5.1) vigorous activity is going on (McGuinness 2002).

8.4   An Informal Sign Model

Sign models may take many forms (Nöth 1995: 79-102). For our purposes, it is especially important that our model should be able to describe the interaction between signs.

In my analysis, a sign consists of a sign vehicle (signifier) and a meaning (signified). In the simple case, a sign vehicle will be just a fragment of the emblem image or text. In the more interesting case, the sign's vehicle will be another sign, or perhaps multiple signs, each having a specific role in the constitution of the new sign.

---

[86] See section 8.5.1.
[87] See also Schreiber 2001.
[88] See also http://www.loc.gov/standards/mets/.
[89] See also http://www.w3.org/2001/Annotea/.

To give an example: in the emblem image, one sees the sunflower and the sun. Both are simple signs, fragments of the image which carry meaning. One also sees that the sunflower is directed towards the sun. This is a complex sign, not reducible to either of the simpler signs. The complex sign has for its vehicles the image region 'sunflower + sun', plus the simple signs 'sunflower' and 'sun'.

In a more abstract example, the sunflower's behaviour (looking at the sun) may be considered a metaphor for the desirable behaviour of the Soul (to direct itself towards God). Here the metaphor sign has as its vehicles the metaphor tenor (the behaviour of the Soul) and vehicle (that of the sunflower).[90]

The meaning of a sign may be given, informally, as a simple phrase. To enable intelligent querying, however, it is vital that at least part of the meaning is captured formally. In my model, I will accomplish this by assigning each sign to a class of signs, and define a hierarchy of classes, which represents an ontology (see subsequently) of the domain. For example, in the emblem discussed above, the figures of Divine Love and the Soul both are members of a class of person-like beings. The class of person-like beings has subclasses for male and female person-like beings, and Divine Love belongs to the former, the Soul to the latter.

Lower-level sign may have specific 'roles' in higher-level signs. When e.g. Divine Love points out the sunflower to the Soul, this higher-level sign contains the three lower-level signs of Divine Love, the sunflower, and the Soul, in the respective roles of actor, object, and beneficiary. These roles will need formal representation in order to allow the model to answer questions like 'What objects are being pointed at for the Soul's benefit?'. I will define a hierarchy of properties to model these roles.

8.5   Representation of the Sign Model

8.5.1   RDF and RDF Schema

The emblem book from which the sample emblem was taken has been digitised in the *Emblem Project Utrecht*. The project applies the TEI (Text Encoding Initiative) P4 *Guidelines* (Burnard & Sperberg-McQueen 2001-2004). In the context of a TEI-encoded corpus, it may seem natural to use the TEI tagsets in encoding the signs. While this is certainly a possibility, it is not the procedure chosen in this chapter. Here I will discuss the RDF(S) representation as used in the remainder of this chapter.

RDF, the Resource Description Framework, is a basic component of the World Wide Web Consortium's vision of the Semantic Web (Klyne & Carroll

---

[90] Unfortunately, there are two separate senses of the word *vehicle* here. In metaphor theory (Richards 1936) the 'vehicle' is the term used to describe the metaphorical image, applied to the 'tenor'. This should not be confused with what I have been calling the sign's 'vehicle', Saussure's 'signifier' (Nöth 1995: 88).

2004). RDF defines a very generic architecture for the description of 'resources' – usually, but not necessarily, resources on the Web. RDF 'models' (collections of RDF statements) consist of triples, which connect a subject, a predicate and an object. Because of this very generic data model, RDF can be used to express almost anything. Applications of RDF use a specific data model within the framework of this generic architecture. RDF Schema is an RDF extension in which these RDF applications can be defined (Brickley & Guha 2004).[91]

Though other facilities are available, the basic functionality of RDF Schema is the definition of classes and their properties. Classes may have subclasses, a class may be a subclass of several superclasses. Properties define relations between classes: a property 'wrotebook', e.g., may define a relation between the class of persons (the domain of the property) and the class of books (its range). Properties may have multiple values. As classes have subclasses, properties have subproperties.

RDF Schema, therefore, allows the definition of what is termed an 'ontology': a description of the types of things that exist in a specific domain of knowledge, and the kinds of relations these things are involved in. RDF Schema is a lightweight ontology language; other languages build on the RDF Schema foundations to be able to express other aspects of a knowledge domain (M. K. Smith *et al.* 2004). An RDF(S) model of a domain can be expressed using RDF/XML syntax.

8.5.2   An RDF(S) Representation of the Sign Model

In order to formally represent the model discussed above, I need to describe three types of objects: (1) text and image fragments; (2) an ontology for sign classes; and (3) individual signs. I will use RDF Schema and RDF, respectively, to handle (2) and (3). The identification of the text and image fragments (1) uses TEI facilities. For ease of processing (especially validation), the three types of information will be stored in separate documents. While each emblem will obviously have its own text and image fragments and sign description (1 and 3), the ontology (2) is meant to be shared between emblems, as the shared ontology guarantees the commensurability of the individual emblem descriptions.

8.5.2.1   Identification of Text and Image Fragments

Identification of text fragments is fairly straightforward. Structural divisions in the text have been encoded as such in the earlier process of digitisation. Other relevant parts of the text are marked up as <seg> elements (segments). Text fragments are identified as <ptr> elements (pointers) to these structural divisions and/or segments. To be precise: the pointers in the TEI encoded

---

[91] I will refer to these technologies (RDF and RDF Schema) as RDF(S).

emblem file identify relevant text parts, while the RDF files describe them as representing text fragments.

The extra level of indirection (the use of <ptr> elements) is used because a single fragment may consist of multiple segments. This can happen in the case of intervening markup (a word split over metrical lines) or in the case of non-contiguous segments (a verb and a corresponding preposition).

The image fragments are represented using <xptr> elements (extended pointers). The xptr's 'doc' attribute names an image file. A region in the file is identified using the 'space' keyword ('location term') in the 'from' attribute (Burnard & Sperberg-McQueen 2001-2004: 14.2.2.17).

8.5.2.2   The Sign Ontology

The sign class ontology is described in an RDF Schema document. In the sign model sketched above it was stated that signs could always be sign vehicles for further signs. Accordingly, in the sign class ontology the class 'SignVehicle' is hierarchically highest. For its subclasses, it has those sign vehicles that can only be vehicles (i.e. text and image fragments), and signs proper. The upper levels of the sign class hierarchy are shown in Figure 21.[92] In this and the next few diagrams, ovals correspond to sign classes and arrows denote the subset relation.
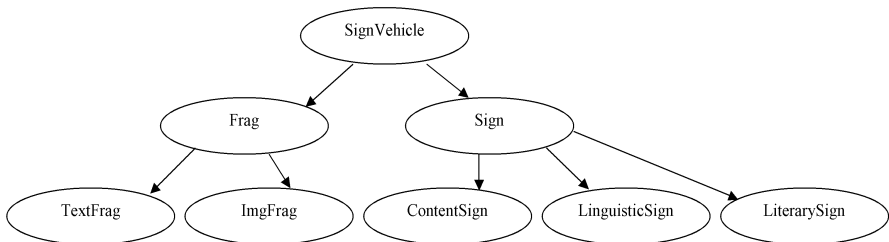
Figure 21 Top of the sign class hierarchy

The TextFrag and ImgFrag (text and image fragments) classes are subclasses of the Frag (fragment) class. [93] Up to now, three types of sign proper have been recognised: linguistic signs, literary signs, and a type dubbed ContentSign. The classes of linguistic and literary signs correspond to features of the text rather than to the possible reality it depicts or describes. An important subdivision of

---

[92] Some of the ideas for this ontology were borrowed from the openCyc ontology (http://www.opencyc.org/).
[93] This paper will use diagrams to represent the signs and sign classes under discussion. An interactive version of these diagrams and source files are accessible through http://peterboot.nl/thesis/. The RDF Schema ontology is also given in Appendix B.

the literary sign classes is formed by the structural divisions employed in encoding the text. The ContentSign class contains those signs that refer to aspects of a (possible) reality. In the present model it consists of the subclasses SpatialThing, Action, and MetaphysicalBeing (Figure 22); the Action class groups signs representing events and signs representing states of affairs.
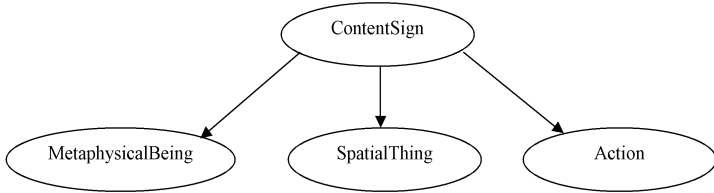
Figure 22 Subtypes of ContentSign

It is important to understand that the ontology is an ontology of interpretations. It is neither a formal semantic model nor a model of the world. It does not mean to explain how language can be used to transmit meaning, and it does not say anything about the world. What it does do, is to categorise the signs which result from an interpretative analysis of the emblem.

To clarify, let us look at a fragment of the subtree of the model containing the classes of signs of person-like beings, itself contained within the subset SpatialThing (Figure 23).
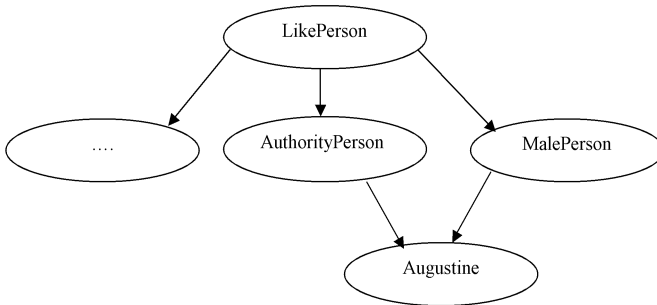
Figure 23 Fragment of the sign class hierarchy. Signs representing St. Augustine are a subclass of signs representing male persons and signs representing authoritative persons

Among the subclasses of person-like beings are male persons, and persons that are deemed authoritative. Augustine is male and authoritative, and the class of

signs of Augustine is therefore a subclass of both. But nothing in the model corresponds to Augustine himself. The model only describes signs of Augustine, and states that if a sign of Augustine appears in a text, this is also a sign of a male authoritative person.

Further aspects of the sign hierarchy are described using RDF properties. At the Sign class level, two properties apply. The SignMeaningProp property, a specialisation of the RDFS comment property, describes the meaning of the sign in English. The SignVehicleProp property is used to relate the sign to its vehicles. All other properties of signs are specialisations of the SignVehicleProp property. Figure 24 shows the SignVehicleProp property and some of its specialisations.
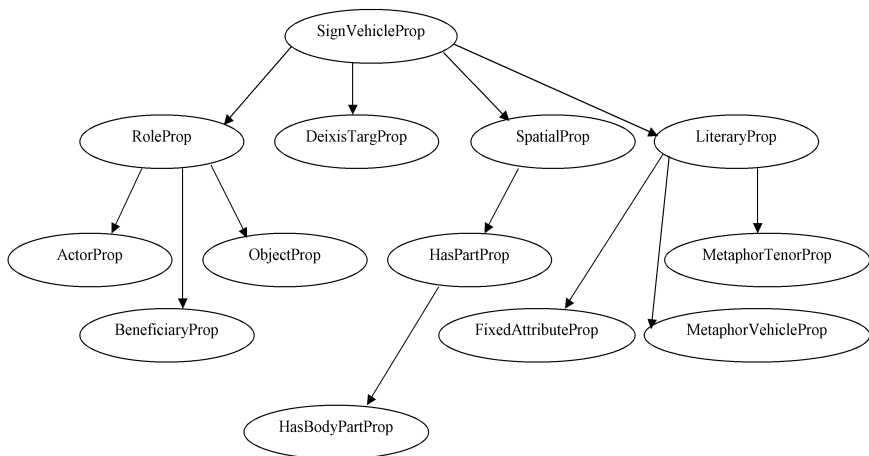
Figure 24 The SignVehicleProp property and some of its subproperties. Ovals correspond to sign class properties, arrows point from a property to its subproperties

Where applicable, ranges and domains for these properties have been defined at the appropriate level. For instance, the ActorProp (actor) property relates Actions to Signs. It makes sense to ascribe it to a looking-at event, but it would not make sense to ascribe it to a metaphor. The above diagram does not show domains and ranges of properties.

8.5.2.3   Signs and Sign Vehicles

To see how signs and sign vehicles are represented in an RDF document, let us look first at the text fragment 'Coelum & terra & omnia quæ in eis sunt' ('heaven and earth, and everything which they contain'), taken from the Augustine quotation. The emblem document contains a pointer to the segment containing

this text. In the RDF document, I identify the segment as representing a text fragment and describe the first sign resulting from this fragment. This sign belongs to the class SpatialThing (Figure 25). As in all diagrams in this section, ovals represent sign instances (identified by their 'id' attribute and their type) and arrows represent instance properties. The objects of literal properties (here: SignMeaningProp) are represented by rectangles.
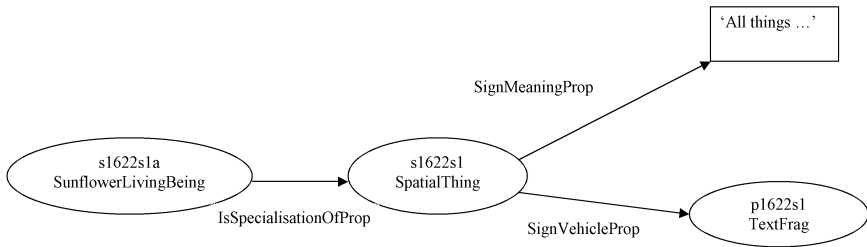
Figure 25 Text fragment from St. Augustine quote, analysed as referring indirectly to the sunflower

The figure also shows the inference that 'heaven and earth, and everything which they contain', in the context of this emblem, probably refers specifically to the sunflower. I represented that inference using the 'IsSpecialisationOfProp' property (itself a specialisation of the 'InferenceBasisProp' property). This sunflower is also the inferred actor of an AssertionAction: it (the sunflower) says (the assertion) that one should love God (the object of the assertion). See Figure 26.

It is clear that the number of possible inferences has no limit, and that it is therefore impossible to state all of them. When the Augustine quote has 'mihi' ('me'),[94] this may be understood as 'us', or even 'you'. One may choose to state this inference, recording both meanings, one may limit oneself to the literal meaning, or one may immediately proceed to the wider meaning. The choice between the possibilities is a practical one and should be decided in accordance with the necessities of research. The network of signs is not meant to give an account of the psychological process of understanding. It only records possible steps in an analysis of the work of art.

---

[94] See note 84.
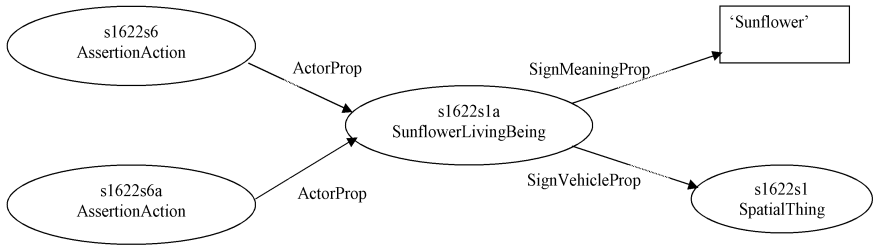
Figure 26 The inferred sunflower sign as actor of assertions

The signs primarily based on image fragments are recorded similarly. Figure 27 shows the sunflower, which has for its vehicle an image fragment, and functions within a PointingAction (Divine Love pointing it out to the Soul) and two LookAtActions (it being turned towards the sun, and therefore, implicitly, to God) as object and actor respectively.
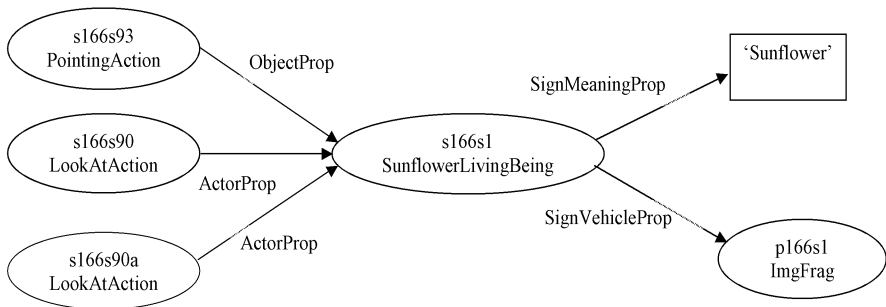
Figure 27 The sunflower in the pictura, actor in two actions and object in another

In the examples so far, the sign structure has been encoded in considerable detail. In some cases this will be useful, in other cases it may be overkill, or just impossible. One may choose to represent some aspects of meaning at a higher level. For instance, one may want to record that the Augustine quotation endorses the Dutch epigram simply by relating the corresponding structural divisions (Figure 28).
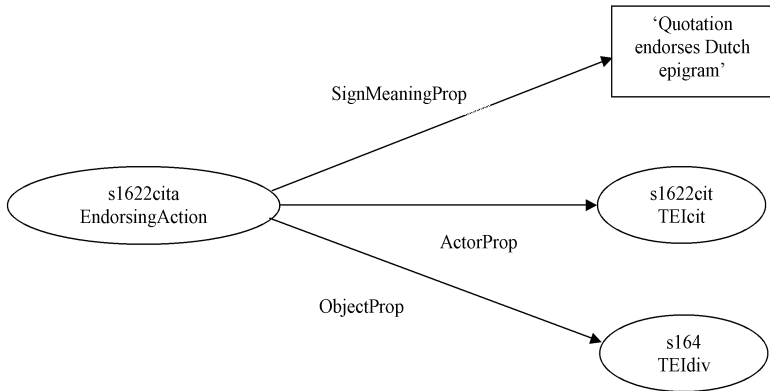
Figure 28 St. Augustine quote endorsing content of Dutch epigram

It should be noted that many of the signs are not purely 'text' or 'image' signs. In fact, to an emblem scholar, the more interesting signs are those that depend on both text and image. In our initial discussion of the sample emblem, we noted some of the relations between text and image. Another example is provided by the French epigram, which explicitly points to the image ('en ceste figure': 'in this figure').

Finally, after defining a sign model and a corresponding class model, one should check the models for validity, on several levels. As XML files, they should be well-formed XML; as RDF files, they should be syntactically correct RDF; but they should also be semantically correct. The object of a 'HasBodyPart' property should, for instance, be a BodyPart. One of the advantages of using RDF(S) is that many tools exist which can perform this type of validation.[95]

8.6   Answering the Questions

Once the sign and sign class models have been encoded, the question arises how to make the results accessible for emblem research. I will discuss three approaches: (1) navigating the emblem and the corresponding model; (2) hyperlinks from scholarly discussion; and (3) ontology-based querying. The last approach, potentially the most exciting one, will be discussed in the next chapter. I believe these approaches show the possibilities of what i have done so far.

---

[95] In preparing this article, I have used VRP (Validating RDF Parser), available from ICS-Forth (http://139.91.183.30:9090/RDF/VRP/).

8.6.1   Navigating the Model

The elements indicating the fragments in the emblem file, the sign model, and the class model which it instantiates, all refer to each other. Therefore, it is fairly straightforward to show a view of the emblem from which information about the recorded signs is directly accessible. A prototype[96] has been developed in which

- the presence of sign vehicles in text and image is indicated, using square brackets in the text and an outline of the image regions in the picture;
- elementary information about the fragments and the signs in which they participate is available in pop-up or 'alt' texts;
- there are hyperlinks from the fragments to diagrams detailing the sign structure;
- selection lists allow selection of a sign class or sign for further manipulation.

Using the prototype, one can explore sign model and emblem in tandem. See the next chapter for experiments in making the information encoded about the sign structures intelligible to the viewer.

8.6.2   Hyperlinks from Scholarly Discussion

On linking to the prototype pages, the prototype interprets the linked-to URI's query component.[97] From scholarly discussion (if conducted in HTML), a scholar can reference the prototype and point at, e.g., the sign that describes how the sunflower looks at the sun, which the prototype will show as in Figure 29.

Similarly, one can point to text fragments and have them highlighted for the benefit of the user. The facility is not limited to showing concrete objects like sunflowers. It may be used to highlight those text fragments that support a proposed interpretation. A link to the diagram pages may show, e.g., the constituents of a specific metaphor. If the scholar wants to talk about metaphor in general, he can have the prototype display his metaphor model, and the reader can immediately verify whether the concrete metaphors found in the corpus fit in this model.

---

[96] The prototype is available through http://peterboot.nl/thesis/. It needs a DOM2 capable browser and an SVG viewer. It has been tested on Windows XP, using Internet Explorer 6.0 and the Adobe SVG Viewer. It is available from the mentioned web page (see note 93). Stylesheets were used to generate HTML pages from the model. Diagrams were generated by ATT's DOT program (part of the GRAPHVIZ suite, available from http://www.research.att.com/sw/tools/graphviz/). SVG (Scalable Vector Graphics, http://www.w3.org/Graphics/SVG/) files are used to highlight the image regions.

[9797] URI: Uniform Resource Identifier, see http://www.ietf.org/rfc/rfc2396.txt. A URI may contain a query component in the form of text following a question mark.

Figure 29 In response to a click on a hyperlink, the prototype highlights relevant regions in the pictura

8.6.3   Ontology Based Querying

The kind of questions asked in section 2 (how to capture image content, textual content, and the semiotic relationships among them that make up the emblem) is easily put, but not that easy to answer. I have developed a data model that can, in principle, accommodate all bits of information which are needed to answer these questions, but the model's flexibility does also make it rather complex. A single taxonomy of concepts can be easily queried. A model like this one, that can consist of any number of interrelated hierarchies, is rather more difficult to handle. What is needed is a user interface that can help the user formulate a query at a conceptual level, hiding the complications of the RDF query composed in the background. See chapter 9 for a prototype of an interface that can do this.

8.7   Future Work

In this chapter, I developed a basic sign model, and used it to express some insights in a sample emblem. More work, however, is necessary before non-experimental use of the modelling techniques proposed in this chapter is appropriate. Interesting issues to pursue would be:

- Without doubt one of the main issues is the one of software support. While it is certainly possible to enter the relevant information using any XML editor, it is a time-consuming and error-prone process. Even specialised RDF editors seem too generic in nature, and need

- enhancements for the specific tasks they may be called upon to do here.
- Another important issue is the choice of data model. In this chapter, RDF(S) has been used to model signs, but the ontology might be reformulated in a more powerful ontology language. The next chapter will use OWL rather than RDF Schema. It should be investigated whether perhaps Topic Maps are better suited to this task. Or would it be better to stick to TEI, perhaps using TEI graphs?
- Finally, after a high-level model has been chosen, the ontology should be expanded. A careful modelling of literary categories such as the metaphor is of particular importance for emblem studies. Furthermore, the ontology will have to be integrated with existing ontologies (Iconclass, OpenCyc).

8.8   Discussion

I opened this chapter with a discussion about interpretation. Interpretation is a concept with a long history behind it. For present purposes I will assume that an interpretation of a work of art is any statement, or set of statements, about, or purporting to be about, (a) possible meaning(s) of that work of art. In this sense, the sign model which I constructed for the sample emblem may be legitimately termed an interpretation. I am now in a position to be more precise about the introduction's claim that the verifiability of an interpretation's claims about the text depends at least partially on the frequency and specificity of the interpretation's references to the original text, and the clarity and applicability of the theoretical notions which it applies to that text.

An interpretation of the kind I have arrived at, has two distinguishing characteristics: first, all statements are, directly or indirectly, linked to the text locations they discuss; and secondly, all categories employed in the interpretation have been explicitly defined, and the interpretation is formulated in terms of these definitions. To give an example: the ontology defines a metaphor to consist of a vehicle and a tenor (clarity of notion); in positing the existence of a specific metaphor, the sign model identifies which signs fulfil the vehicle and tenor roles (applicability of notion); all of these signs are identified in direct reference to their constituent signs (frequency of reference) and finally to the text and image fragments which are their ultimate vehicles; and these fragments can be identified with any desired degree of precision (specificity of reference).

The point of this is not that my interpretation of the metaphor is necessarily a good one. On the contrary; the point is that its possible weaknesses are being exposed for my own and my fellow scholars' scrutiny, and that it was the very act of formulating the interpretation which so exposed them. Any interpretation which uses mechanisms like the ones proposed in this

chapter (if not the concrete syntax), will be thus exposed, and will benefit from that exposition. It is likely that this exposition will increase our interpretations' robustness.

9   Creating a Metaphor Index

9.1   Introduction

This chapter describes an index on metaphor in the emblem book *Amoris Divini Emblemata*.[98] The index should be interesting both for its contents (that is, for the information about the use of metaphor in the book) and as an example of modelling a complex literary phenomenon. Modelling a complex phenomenon creates the possibility to formulate complex queries on the descriptions that are based on the model. The chapter describes an application that uses this possibility. The application user can interrogate the metaphor data in multiple ways, ranging from canned queries to complex selections built in the application's guided query interface.

      Unlike other emblem indices, such as Peter Daly's *Index Emblematicus* or Henkel and Schöne's *Handbuch*, the metaphor index is not meant to be a tool for resource discovery, i.e. a tool that helps emblem scholars find emblems relevant to their own research. It presents research output rather than input. The modelling techniques that it exemplifies should help a researcher formulate detailed observations or findings about his research subject – in this case, metaphor – and make these findings amenable to further processing. The result is an index, embedded in an overview or explanation of the data for the reader. The researcher should integrate the presentation of data in a narrative or argument. Here I describe one possible way of effecting this integration.

      This chapter builds on the techniques developed in the previous chapter. It uses RDF to formulate the observations (see section 9.3.2 for further explanation). For the model, it uses the web ontology language OWL. The previous chapter used OWL's predecessor, RDF Schema. New here, as mentioned in the previous chapter, are the complex queries that the application facilitates, a modelling of the text structure that facilitates filtering by type of text unit, the integration between secondary text and research data also suggested there, and the pre-built queries that are one ingredient in effecting that integration.

      The metaphor index was instrumental in doing the research for chapter 11. For background on metaphor I refer to that chapter.

---

[98] Parts of this chapter were presented at a number of conferences, most completely at Digital Humanities 2008 in Oulu (Boot 2008b).

The metaphor index is not complete. The metaphor model that it employs is not subtle enough to describe the full complexity of metaphor and not all metaphors have been (fully) described. Still, it is useful to describe the index and the modelling procedures behind it, as in the humanities the explicit modelling of complex phenomena and their relations is still unusual. Use of the OWL language in these models is even rarer. Unprecedented in the humanities is, as far as I know, the interactive exploration of the modelled phenomena that the metaphor index offers. Also new is the close integration between the index, the primary text that is being indexed, and an essay that describes the indexing. A prototype of the index that I describe here is available on the web.[99] It is described in section 4.

As to the readership of this chapter, I do hope it will include literary scholars as well as information scientists. Modelling literary phenomena should be done by literary scholars, and only they can evaluate the developing techniques for doing it. More about modelling literature in the next section.

9.2   Literary Modelling

The centrality of modelling to the practice of humanities computing is one of the essential tenets of McCarty's *Humanities Computing* (2005). The terms 'model' and 'modelling' may sound foreign to some humanities scholars. In this section I will argue that many of the theories and conceptions created by humanities scholars can be considered as models of a subdomain of literary studies. After giving a definition of a model I will give a number of examples of models employed in literary studies. A number of these models have been implemented in computer programs, notably in the analysis and creation of narrative. I will mention a number of these, with special attention for computer-based studies of metaphor.

For McCarty, the process of modelling is what drives humanities computing; the models themselves are but 'temporary states in a process of coming to know' (2005: 27). In the noncomputational models that I will mention, however, the process is less important than the distinctions a proven model allows us to make. McCarty defines modelling as a 'continual process of coming to know by manipulating things' (2005: 28), as 'the heuristic process of constructing and manipulating models', and a model as 'a representation of something for purposes of study' (2005: 24). He also describes a model as 'an explicit, delimited conception of the world' (2005: 21).

I will define a model as a representation of a domain, that distinguishes, names and classifies the objects and their constituents and that describes their interaction and contribution to the larger whole. A model may focus on some aspects of reality at the expense of others. In this sense, quite a number of

---

[99] Accessible from http://peterboot.nl/thesis/.

literary theories would qualify as models of their domains. There is nothing in this definition that requires the use of a computer. Yet, the mere fact of classification creates an opportunity for using one.

Perhaps the most detailed example of a model in humanities scholarship is classical rhetoric. Its elaborate terminology of the parts of the oration, the branches of oratory and the tasks of the orator provides a complex model for discussing the merits of speeches and (by extension) of literary works. In its discussion of persuasive appeal (logos, pathos, ethos) it studies how these components affect the audience and thus contribute to a speech's purpose. Rhetoric is also a radical simplification, in that an individual oration has to take into account not just the precepts of rhetoric but also the individual characteristics of the case to be argued, the audience to be addressed, etc.

Most models in the humanities are less complex than that of classical rhetoric. Some models are no more than a pair of opposed concepts, such as Saussure's powerful distinction between *langue* and *parole*, the language as a system vs. the individual utterances. The complexity of other models is somewhere between these two extremes. An example from emblem studies might be Albrecht Schöne's emblem concept, that describes the emblem's components, their mutual interaction, the relation to the depicted reality, and the way this is perceived by the emblem reader. For Schöne, his model is 'die Beschreibung einer idealtypischen emblematischen Grundform, von der aus die Fülle der Sonderformen erfassbar und systemfähig wird' (Schöne 1993: 30). Finley considered Weber's ideal types as the type of model best suited to history (1985).

As McCarty states, a model in the humanities (as elsewhere) can be either a model *of* something, or a model *for* something. I would suggest historical periodisation is an obvious example of a model-of; an example of a model-for might be the abstract that is written for an upcoming conference paper. An interesting case would be the model that informs a bibliography: it is both a model of the reality that the bibliography captures, and a model for the items in the bibliography. The same holds for things like repertories and (biographical) dictionaries: an idea about the salient characteristics of the described objects is what informs the items that describe the objects.

Other humanities fields where models have been applied include edition studies (where the planning of a scholarly edition demands careful modelling of the components of that edition and their relations), communication studies (with its well-known model of sender, message and receiver, possibly complemented by the context or code) and stemmatology.

A purely literary field that has often been the domain of modelling is narrative. One of the subjects of modelling here is e.g. the treatment of time, with the familiar distinction between *fabula* and *sujet*. Cavazza & Pizzi (2006) in this context mention Propp's theories about narrative functions in the folktale, Greimas' description of character roles, and Barthes' description of narrative

structures. The case of narrative is especially interesting, because what used to be models *of* existing narrative are now used as models *for* the construction of new narratives in story generation systems. These models have been transformed into computer representations and are used for generating interactive narrative. Bal (for the distinction between fabula, story and narrative text) and Campbell (for the analysis of the structure of heroic myths) are other theorists whose work is used in what is currently a very lively area of study (Gervás *et al.* 2006; Swartjes & Theune 2006; Tuffield *et al.* 2006). A formal model for drama is given by Damiano *et al.* (2005)

What these examples show is that modelling, in the sense of distinguishing and naming parts and describing their relations, is not something foreign to the study of literature or the wider humanities. For a model to be implemented on a computer, it may need a more explicit formulation, and of course the computer enforces consistency in its application (or at least exposes inconsistencies), but what the computer gives back is the possibility of manipulation. A number of the mentioned models are, like the narrative model, being used in programs for e.g. stemmatology and bibliography.

Other areas where computer-based modelling is used for the study of literary subjects include the traditionally elusive concept of style (in a host of studies, see e.g. Craig 2004; Luyckx *et al.* 2006), the study of intertextuality (Schellner *et al.* 2003), of literary characters (Zöllner-Weber 2005), of the process of understanding while reading (Moorman & Ram 1996), and modelling the claims in (scholarly) argument (Uren *et al.* 2003). The subject too wide to be discussed here.

Even the number of studies that use the computer for modelling and investigating the subject of metaphor is very large. As does the subject of metaphor itself, computer-based analysis of metaphors spans the fields of literature, linguistics, philosophy, psychology and cognitive science. Again, the computer treatment of metaphor is based on or inspired by pre-computational models: the vehicle-tenor model formulated by the literary critic Richards (1936), the interaction view formulated by the philosopher Max Black (1962, 1979), and the cognitive science view formulated by Lakoff and Johnson (1980). I can only mention a few approaches. Cognitive scientists model the process of understanding metaphor (Indurkhya 1987; Kintsch 2000; Martin 1990). Philosophers model metaphor semantics (Steinhart 2001). Linguists devise programs that locate metaphors in language corpora (Ferrari 1996; Mason 2004). Others try to model metaphoric word meanings in lexicographic resources like Wordnet (Lönneker & Eilts 2004). Handling metaphoric language is obviously important for information retrieval software. Czejdo *et al.* (2007) analyse metaphors to hunt down potential terrorists on the Web. I have not found many predecessors in the computational modelling of metaphor in a specifically literary corpus. Rommel (1996), rather than modelling the metaphors themselves, uses lexical and syntactical clues to find metaphors in Smith's

Wealth of Nations. Pasanek has collected a database of metaphors of mind in eighteenth century British literature and categorised them by field (2006; Pasanek & Sculley 2006). Noë (2001) lists the metaphors in fourteenth century mystic Ruusbroec's *Dat rijcke der ghelieven*. As far as I am aware, no one has tried to list and model the metaphors in one literary work and offer an interactive interface for their study.

9.3   Modelling

9.3.1   Metaphors and Their Context

There is a number of aspects of metaphor and the texts where metaphors occur that I want to represent in the index.[100] A metaphor has a vehicle and a tenor, in the terminology of Richards (1936). When love, for its strength and endurance in adversity, is compared to a tree, the tree is the vehicle, love is the tenor. I want to be able to create hierarchies, both for the comparands (that is, vehicles and tenors) and for the metaphors. It should be possible to state that 'love as a tree' belongs to a wider class of 'love as a plant' metaphors. It should also be possible, as argued in Lakoff & Johnson (1980), to state that a tree is a plant, and that it, with roots, fruit, leaves and seeds belongs to the vegetal kingdom. It often happens that an emblem contains references to an object invested with metaphorical meaning elsewhere in the book. We should be able to state its presence without necessarily indicating something we suppose it stands for.

In the index, I also want to represent the locations in the emblem (text and image fragments) that refer to the vehicles and tenors. The text fragments are stretches of emblem text, the image fragments are, for now, rectangular regions in the emblem pictures. It is necessary to represent the text structure, in order to relate occurrences of the comparands to locations in the text. Knowing that the emblem book under scrutiny contains a Spanish, a Dutch and a French epigram for each emblem, we should be able to ask for references e.g. to fruit in Spanish epigrams. On the metaphor level too we should be able to state the text and image components the metaphors are based on.

9.3.2   Modelling Languages: RDF, OWL and TEI

The modelling language used here to express the metaphor model is RDF, the Resource Description Framework,[101] designed by the World Wide Web Consortium for use in the 'semantic web' (Berners-Lee *et al.* 2001).[102] I will model the metaphors in *Amoris Divini Emblemata* using RDF triples.

---

[100] For more about metaphor see chapter 11 in the present book.

[101] More about RDF and ontologies in the previous chapter.

[102] For more about RDF, see section 8.5.1.

What may be termed the 'meta-model' of this collection of RDF triples is provided by an ontology. This ontology, unlike the philosophical concept, should have no metaphysical implications. An ontology is just a formal model of a domain under study; it models the kind of objects that exist in that domain, their relationships and their properties; it provides a shared understanding of a domain (Antoniou & Van Harmelen 2004: 10-12). Ontologies are an essential ingredient in the semantic web; they can be expressed using the Web Ontology Language OWL (McGuinness & Van Harmelen 2004). It is the ontology that defines the vocabulary that can be used in the RDF statements in our model. The ontology thus limits the things that we can say; it provides, in McCarty's words, the 'explicit, delimited conception of the world' (McCarty 2005: 21) that makes meaningful manipulation possible. The ontology is also what 'drives' the application built for consultation of the metaphor index (see section 9.4).

OWL provides three different sublanguages for ontology definition: OWL Lite, OWL DL and OWL Full, in order of increasing expressiveness. The distinction between the languages is important for applications that want to use the ontologies for reasoning. For my purposes this is largely irrelevant. All of the features that the metaphor index uses are from OWL DL. It is important to note that the use of semantic web technology does not imply the expectation that the metaphor ontology will one day be integrated into a larger 'federated knowledgebase' on metaphor (Marshall & Shipman III 2003). I do not expect semantic web technology will provide the magical glue to make this index interoperable with other metaphor collections. Criticism that considers this aspect of the semantic web vision to be unrealistic therefore does not apply to the metaphor index. I will return to other issues that Marshall and Shipman's paper raises in the Discussion section of this chapter.

Applications of OWL in the humanities include the description of tangible objects (in museums and in archaeology), lexicography, and the development of portal functionality. Ciula *et al.* (2008) describe using OWL to describe the network of persons and associations underlying complex historical documents. Pasin & Motta (2006) use an OWL ontology of the domain of philosophy in an educational setting. Use of OWL for modelling and recording interpretation is still very much an experiment. Zöllner-Weber (2008) used OWL for modelling literary characters. The question of whether interpretive phenomena lend themselves to being modelled in OWL is still open.

The initial modelling of the text is given by its TEI XML encoding. In order to facilitate queries that address both metaphorical aspects and text structure, the text structure will be translated into RDF triples. In order to answer a query for, e.g., metaphors for love based in both picture and text, we need to collect information about metaphor and text structure in a single place.

9.3.3   *A Metaphor Ontology*

The metaphor model will be embedded in a larger model of signs in texts, as described in the previous chapter. I will not go into the details of this, but will show a part of the ontology as it concerns us here in Figure 30. In this figure, arrows point to subclasses. Everything that is noticeable about a work of literature is considered a sign. All signs are (potential) signifiers, signs in their turn are either based on content or on text structure, and the 'content' signs in this model are either metaphors or comparands. The emblem structure consists of textual units (among which are epigrams) and pictures.
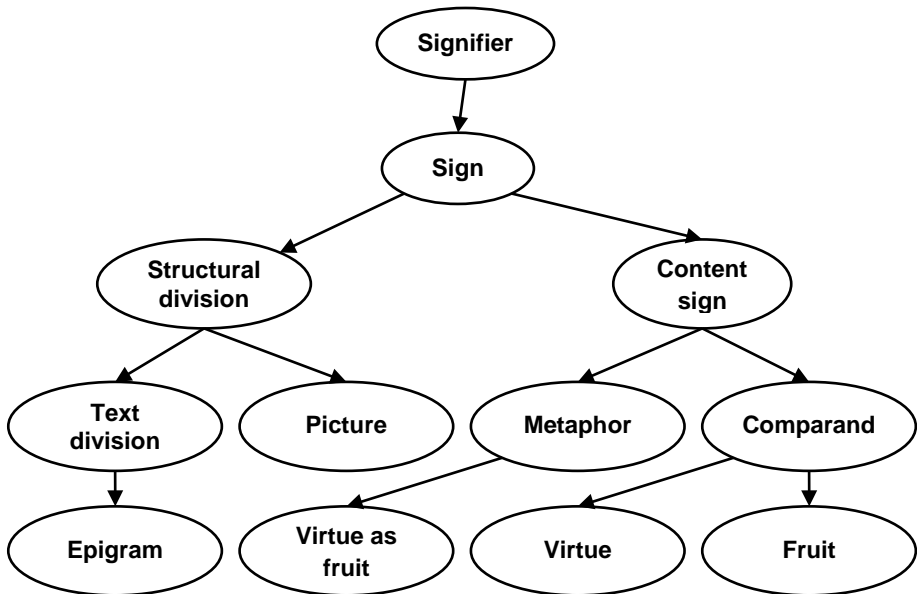
Figure 30 Part of the classes that make up the metaphor ontology. Arrows point to subclasses. The classes at the bottom level are just examples; many other ones could have been shown if more space were available. For simplicity, this diagram ignores class properties (see Figure 31 and Table 3).

Besides these subclass-superclass relations, the model defines, at class level, a number of relations between objects and properties of objects.[103] Metaphors and comparands are connected on the base of tenor and vehicle properties (Figure 31).

All of the properties defined in the ontology are listed in Table 3. The table gives the name of the property and a description. It also indicates to what

---

[103] In OWL terminology, what one would usually call a property is called a 'datatype property', a relation is called an 'object property'.

type of objects the property applies (its domain, in the 'from' column) and the type of objects that it can take for its value (its range, in the 'to' column). The properties enabe querying objects by their relations to other objects: to ask for all metaphors based in an emblem picture, to ask for all metaphors for love, or to combine these criteria (Figure 32).
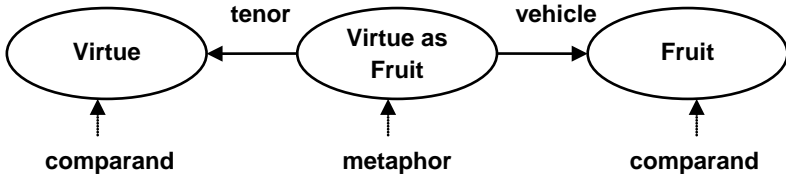
Figure 31 A metaphor and the properties relating it to the comparands

Name	From	To	Description
Sign vehicle	Sign	Signifier	Used to relate a sign to its signifiers
Sign meaning	Sign	String	Indicates the meaning of the sign
Based-In	Sign	Structural division	Links sign and the structural units of the emblem that it is based in
Contains	Structural division	Signifier	Links structural divisions and the image and text fragments that they contain
Tenor	Metaphor	Comparand	Links a metaphor and its tenor
Vehicle	Metaphor	Comparand	Links a metaphor and its vehicle
Tenor-In	Comparand	Metaphor	Links a tenor to the metaphor(s) it is the tenor of
Vehicle-In	Comparand	Metaphor	Links a vehicle to the metaphor(s) it is the vehicle of

Table 3 Properties described in the metaphor ontology

Figure 32 Objects can be queried by their relations

The RDF model that this ontology describes is based on objects in the TEI file, identified by URL. These objects are the text/image fragments that the RDF

model defines as belonging to the appropriate OWL class. The comparands are defined as the corresponding signs (fruit, tree, virtue, love). The metaphors in turn are defined with reference to their comparands (fruit + virtue, tree + love).[104]

There are many cases where the simple ontology used here is inadequate for capturing the structure of the metaphors in the text. Even if the model is adequate, there often is a need for further explanation. For this purpose, the meaning property, available for any sign, allows us to describe the meaning of a sign in natural language. Any other comment that is deemed useful can be attached at any desired level of the index using the RDF Schema comment property.

9.3.4   OWL Special Constructs

Those interested in OWL may want to know which OWL features were used beyond the (sub)class, (sub)property, range and domain that were already available in RDF Schema. They are listed in Table 4.

owl:inverseOf	The *Tenor* and *Tenor-In* properties are inverse properties (if *a* and *b* are related by the Tenor property, *b* and *a* are related by the Tenor-In property). The same holds for the *Vehicle* and *Vehicle-In* properties.
owl:Thing	All things in the domain of discourse are potential signs, and thus sign vehicles. The class of sign vehicles is equivalent to the class of all things, owl:Thing.
owl:equivalentClass	See owl:Thing, above, and owl:unionOf, below.
owl:unionOf	The class of fragments is equivalent to the union of the classes of text and of image fragments.
owl:disjointWith	The classes of text and of image fragments are mutually disjoint (they have no members in common).
owl:TransitiveProperty	The *Sign vehicle* property relates signs to their vehicles (signifiers). If a sign is a vehicle for a second sign, the first sign's signifiers are also signifiers of the second sign.

Table 4 OWL features used in the metaphor ontology

The application described in the next section does not know how to exploit these features, except for the owl:inverseOf property. The owl:inverseOf property is used to infer the presence of the inverse triples. This facilitates

---

[104] The top level of the ontology is given in appendix C. All definitions are available from the prototype.

queries that rather than ask for 'metaphor that use a tree as a vehicle' ask for 'trees that are used as metaphors'.

9.4   An Application Prototype

In order to present the metaphor index to a reader, I have developed an web application that allows readers to consult and explore the index. The application is an example of an ontology-driven application as discussed in (Guarino 1998): the data model, the application logic and the user interface are all based on the metaphor ontology. This section describes the application.

9.4.1   Technical Environment

The point of departure for the metaphor index was a TEI-encoded digital text of *Amoris Divini Emblemata* from the Emblem Project Utrecht, stripped down considerably to simplify encoding of text and image fragments. XML files that contained the RDF and OWL statements were created manually, with no other support than an occasional script or editor macro.[105]

The application that makes the data accessible to others was created using PHP and a MySQL database backend. To handle RDF I used RAP, the RDF API for PHP. An upload process was written that reads the RDF and OWL files and stores their contents in a single RDF model in the database. The process uses a RAP inference model to create inferred triples, whose existence is deduced from other triples (if an object is a metaphor, it is also a sign, by virtue of the fact the ontology defines metaphors as an (indirect) subclass of signs).

The upload process also reads the TEI XML file that contains the digital text of the emblem book and creates RDF triples that represent part of the structure of the emblem book texts. These triples state e.g. the fact that emblem x contains Spanish language epigram y. Another inferred relation is the 'based-in' relation, that binds together signs and the textual divisions that contain the text or image fragments that are a sign's vehicles (or the vehicles of its vehicles, etc.). Finally, the process creates a number of auxiliary triples for the application, unrelated to the metaphor model.

9.4.2   Information Perspectives

The application has to provide insight into three basic layers of information: the primary text (the emblems), the database-like collection of metaphor data, and a secondary text that should make these three layers into a coherent whole. The

---

[105] Programs exist that help in RDF editing or ontology creation. I needed to create RDF statements that related to fragments of the TEI files, and have found no tools that could help me with that. Tool support for the creation of models like these and their connection to TEI fragments is essential to successful development of useful indexes.

application organises this in three perspectives: an overview perspective, an emblem perspective and an ontology perspective. Each of these perspectives offers one or more views on the data. The views have been realised as HTML frames.

The overview perspective displays, at an equal level of prominence, two ways to access the data (Figure 33). The left frame shows a basic selection interface into the metaphor index. Users can select metaphors by metaphor class and/or by tenor, vehicle and text constituent. Using the buttons at the top will replace the selection panel with a table of contents of the book, a list of emblem thumbnails, a list of all indexed metaphors or a guided interface for creating expert queries. Executing a query will replace the contents of the panel by a list of hits. Using the left hand frame then, the user can explore the metaphor data or the book.
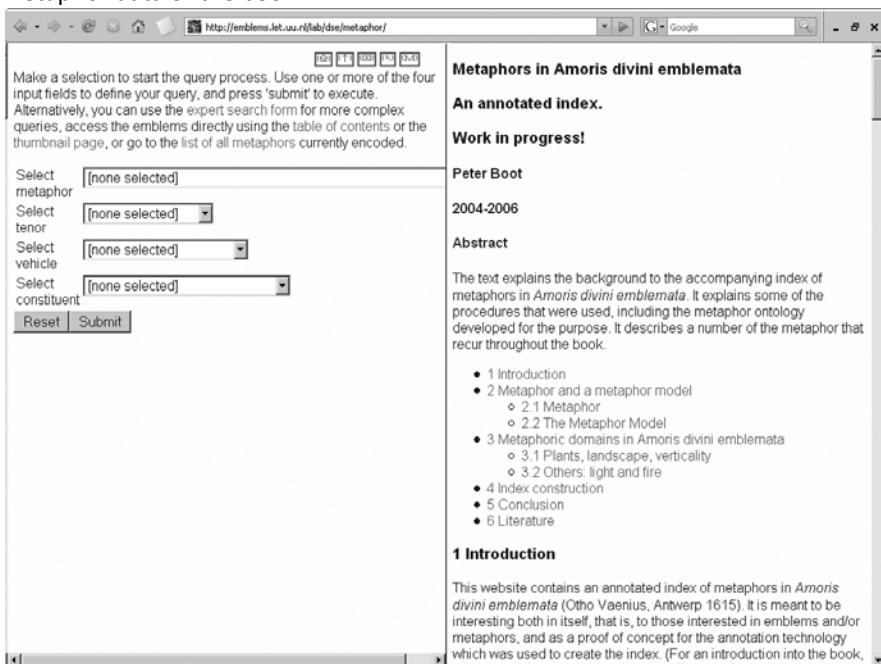
Figure 33 Overview perspective

The right hand frame of the overview perspective displays an essay that explains the metaphor index and its workings, but that also treats the use and meaning of metaphor in this emblem collection. This secondary text provides the context that I argue in chapter 13 an index like this cannot do without. The essay will resemble a scholarly article and should obey the conventions for scholarly writing. It should mention its sources, point to literature, and come with an abstract. In one respect it should be different from today's scholarly articles, viz. in that it should be deeply hyperlinked into the surrounding index

and emblem book. The essay is an exposition of what is there, a guide into the findings on the use of metaphor in the emblem book. It hyperlinks to emblems, naturally, but also to the description of specific metaphor usage and to elements of the ontology that the index is based on. Activating other hyperlinks will execute a query in the metaphor index and list the result in the left hand panel (Figure 34).
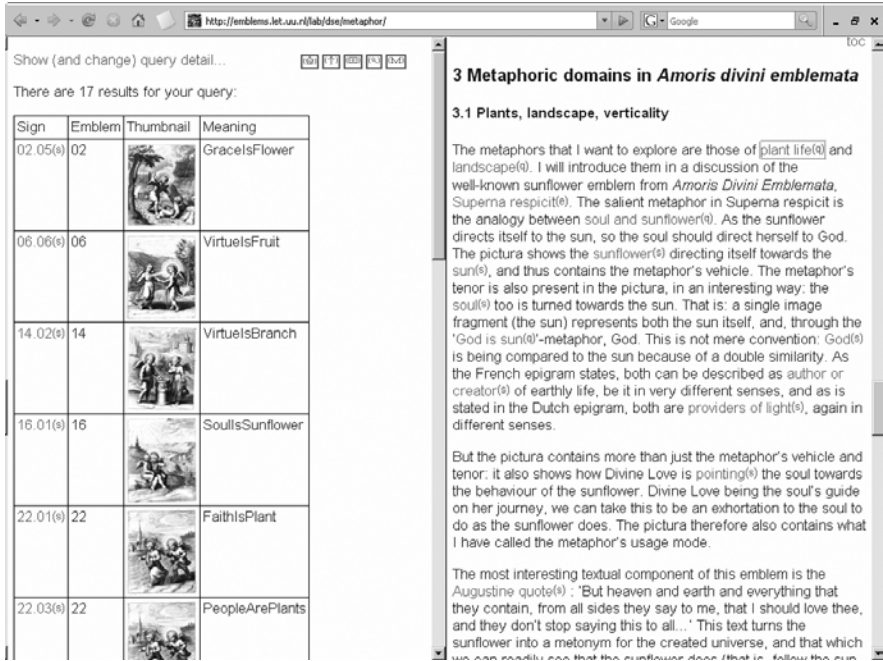
Figure 34 Clicking the hyperlink 'plant life' (top right) executes a query with hits shown in the left panel

Like the overview perspective, the emblem perspective consists of two frames (Figure 35). The right hand frame displays a single emblem from the book, hyperlinked to a full display on the Emblem Project Utrecht site. The left hand view displays information about metaphor use in the emblem. The metaphors are all hyperlinked to, first, an index query that will bring up all their occurrences in the book, and, second, to the corresponding concept in the ontology (see below).

When activating the emblem perspective, it is possible to request display of a specific metaphor or, more generally, a sign. In the right hand view, this will cause the corresponding signifiers (text and image fragments) to be highlighted in red. In the left hand view, this will display information about the definition and properties of that specific sign.

Figure 35 Emblem view with one metaphor highlighted in picture and text

The ontology perspective, finally, displays the ontology defined for the metaphor index. It is hyperlinked both from the emblem perspective and from the essay. The ontology perspective in its present state is created using a ready-made tool for the display of OWL ontologies, OWLDoc (Figure 36). As one can see, on the right it displays a class definition with commentary, if available, and below that a hierarchical display of superclasses. Under the 'Usage' heading subclasses and applicable properties are shown. The frames on the left display an overview of all definitions available in the ontology.

In the current application, the ontology perspective is not hyperlinked back into the index. The OWLDoc software clearly does not 'know' about the index and how to address it. It would be desirable to link the classes and properties that the ontology perspective displays to overviews of all occurrences of these classes and properties in the index.
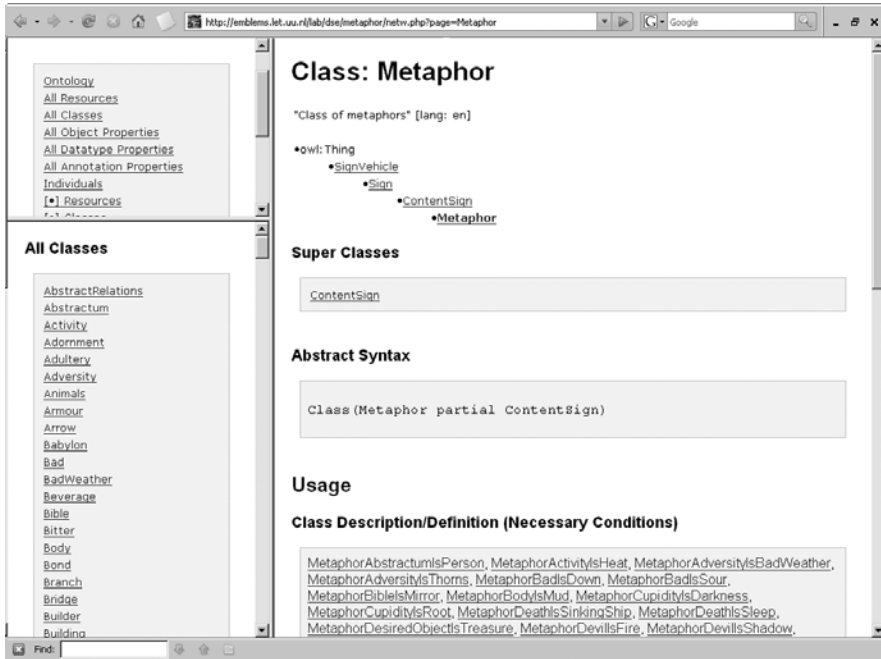
Figure 36 Ontology view, with display of class metaphor

9.4.3   Exploring the Index

The metaphor index, as said before, is meant to be explored. The application provides five ways to explore the data in the index. Exploration may proceed:

1.   by emblem: for each emblem, the application displays the relevant metaphors about which the user can ask for more information;
2.   from a list of all metaphors;
3.   through a selection made from the overview perspective;
4.   through 'canned' queries activated from the essay;
5.   through the guided interface for creating expert queries.

The first three of these are unproblematic. The last two however merit further discussion.

As to (4), the essay contains a number of 'canned' queries, stored as a query string ready to be executed. The queries use the RDQL query syntax for RDF (Seaborne 2003).[106] The application reads the query from the TEI document that contains the essay and executes the query using the RAP RDQL query engine.

---

[106] The RDQL query language has no official status. It is a 'member submission' to the World Wide Web Consortium and was input for the formulation of the SPARQL query language (Prud'Hommeaux & Seaborne 2008). The RDQL queries used are very simple, and their reformulation into SPARQL will be a minor effort.

Figure 37 Expert search. Start building query	**Figure 38 Expert search. Click '+' to create more criteria**
Figure 39 Expert search. Select the desired criterion	**Figure 40 Expert search. Final state of query**

As to (5), the expert query interface guides the user through the steps needed to create a complex query, such as one that asks for metaphors for love that have a pictorial component ('based in picture'), where the tenor (love) occurs in the Dutch epigram. The user effectively creates an RDQL query with possibly very complex WHERE-clauses without having to know the RDQL syntax. The expert query interface opens by showing the user a single field where he or she can select a sign type, metaphor or otherwise, for querying (Figure 37). Pressing 'submit' will show a panel where the query can either be executed or modified (Figure 38). 'Submit' will now execute the query, but the '+'-button will open a panel where an extra criterion can be added (Figure 39). The finished query, consisting of multiple search criteria at multiple levels, is shown in Figure 40.[107]

In the process of refining the query, the application builds a representation of the query in the database. The representation closely matches the structure of RDQL queries. When the query is executed, the application translates the query representation into RDQL before executing it.

---

[107] The query display uses internal names (such as 'MetaphorTenorProp') for the OWL classes and properties. It is clearly desirable to replace these names with labels more suited to human readers.

In the hits panel, the user can ask for a display of the search criteria and, if desired, the generated RDQL (Figure 41). From the display, clicking the '+'- or '-'-buttons, he or she can start modifying the query right away.

Figure 41 Expert search. Display of executed query and generated RDQL in results panel

9.5   Discussion

As said in the introduction to this chapter, the metaphor index is not complete. That is not necessarily a very bad thing. The index is an experiment, it is an attempt to find out something about the feasibility and usefulness of encoding complex literary phenomena.

The feasibility obviously depends to some extent on the availability of tool support. Entry of the metaphor data was done largely by hand. The tool that displays the metaphor data had to be designed and programmed, and was revised a number of times. Yet the availability of tools is not the deciding factor. For one thing, there is quite a large number of metaphors in the book, and the data are complex. As it is still the human eye that recognises metaphorical structures in this type of material – as I suppose it will remain for some time – entry of the data, even if supported by very efficient tools, will always be time-consuming. Furthermore, while it is obvious that an existing application framework would have saved a not insignificant amount of time, it is also true that information display always has to be tailored to the nature of the information. The opening window shows a search for metaphor by type, tenor, vehicle and location, and this display is clearly specific to the subject of the index. I am very sceptical about 'no-programming', wizard-built displays of complex information.

There is also a certain tension between the experimental nature of this index and the need to collect a body of material and create a display application.

One can question the need for a large amount of data. If the aim is not to support resource discovery, but solely to provide insight, do we then need this large amount of data? Is all software meant to be discarded, as McCarty quotes Perlis, 'the whole point is to always see it as soap bubble?' (2005: 21). Is an 'experimental index' not a contradiction in terms? It is certainly true that, while working on this, I often experienced the conflict between the desire to introduce another aspect of metaphor into the model, and a need to create a body of material that it would be worthwhile to explore. It is also true, however, that insight does not come from subtlety alone. There is no insight without numbers.

McCarty writes about the computer as 'a rigorously disciplined means of implementing trial-and-error (…) to help the scholar refine an inevitable mismatch between a representation and reality (as he or she conceives it) to the point at which the epistemological yield of the representation has been realised' (2005: 5). It is true that the computer helps us be rigorous and disciplined, but perhaps for that very reason the representations that the computer helps us build may become a burden. Computing can slow us down. It might have been wise to exercise some restraint and e.g. describe only the metaphors in the first five emblems, or not describe the locations in text or image (or just point to a line rather than to precise words). After all, the index should clarify the conceptual structure of metaphor as it is used in the book; it need not be a work of reference.

I would like to address this issue by way of the notion of 'premature formalisation', which is one of the motivating factors for Neil Audenaert and others's *Critspace* tool (2008). Formality comes with cognitive overhead, it does not deal well with tacit or evolving knowledge, it cannot represent the situated nature of knowledge (Marshall & Shipman III 2003). Shipman and Marshall even wrote a paper 'Formality Considered Harmful' (1999). Still, obviously formality is in some way essential to what we as computing humanists do: the explicitness and consistency and manipulability of our models depend on it. But it is true it comes with an unavoidable lack of flexibility, unless we are willing to start from scratch repeatedly, and thus indeed discard what we have done so far.[108]

Approached from another angle, the semantic web technology that this chapter uses can also be employed at a very large scale, as in the Finnish *CultureSampo* project (Hyvonen *et al.* 2006). CultureSampo employs an ontology of the entire cultural domain in order to make cultural heritage accessible. This is an example of formalisation at a very grand level. It may be criticised for setting in stone today's ideological limitations,[109] but I doubt this criticism is fully justified. After all, the work of cataloguing and classifying and building thesauri did not start with the advent of the computer; there has always been an aspect of formalisation and rigour in attempts to come to terms with the plenitude of

---

[108] We met a similar criticism of XML and its fixed schemas in chapter 5.
[109] As it was during the discussions following Eero Hyvonen's keynote presentation of CultureSampo at Digital Humanities 2008.

cultural artefacts, and there have always been things that were made invisible because of this. To give an example: Iconclass may be biased towards western, Christian, medieval and early modern art, but fortunately that has not kept people from working on it. So there is a stage where formalisation is necessary and useful. The CultureSampo 'prototypes' are obviously prototypes in a very different sense and on a very different scale than the prototype that I have presented in this chapter. From a research perspective, one would say, there is no need ever to go beyond prototypes of varying degrees of refinement. Maybe we should distinguish different types of research, some more experimental, oriented towards suggesting ways of approaching issues, some more of an applied type and oriented towards engineering solutions. Both approaches have their value in humanities computing as they have in any other discipline. There is no sharp boundary between people doing fundamental physics and people that design and build aircraft, and neither should there be one in the field of digital humanities.

What we have seen in this chapter then is an example of the first type of research, trying out an approach toward modelling literary phenomena and making the results accessible to exploration, querying and embedding in research reports. This indeed is software meant to be discarded, but not therefore without a purpose: the purpose of suggesting there may be some merit in connecting our primary sources and our scholarship, by way of modelling the phenomena that we study. I have called this concept mesotext, and will discuss it in chapter 13.

10 Towards a TEI-Based Encoding Scheme for the Annotation of Parallel Texts

10.1 Introduction

The translation and adaptation of emblems was a very frequent phenomenon in early modern emblem book culture.[110] Adaptation or translation could be applied to a single emblem, but also to a series of emblems or to a whole book. A related phenomenon is the presence of epigrams in multiple languages in the emblems of a single book. These processes create bodies of parallel texts that merit study. In the Emblem Project Utrecht (EPU) corpus,[111] we encounter this situation within the confines of a single book (e.g. the vernacular epigrams in multiple languages in Otto van Veen's *Amoris Divini Emblemata*), between books in different languages (e.g. the Dutch translation of Hugo's *Pia Desideria*), and between books in a single language (e.g. Jan Suderman's reworking of Otto van Veen's emblems in *De godlievende ziel*).

I conceive of parallel text analysis as a task of annotation creation and visualisation. It is clear that other perspectives are possible and desirable. Automated analysis of vocabulary and stylistics would be very interesting. I see these processes as potential contributors of machine-created annotations. The questions of how to store parallel-text annotations and how to display them in a manner conducive to insight are independent of the way the annotations are created.

There are a number of technical challenges to overcome in annotating parallel texts. (i) Before the parallel texts can be annotated, they have to be aligned. In order for an annotation to be able to comment on two parallel fragments, it must either itself refer to both fragments or refer to some kind of link between the fragments. It is clear that creating a set of links between parallel texts (an alignment) facilitates annotation and indeed is a condition for serious study of the parallelism. The alignment may be quite complex: in the

---

[110] This chapter is a lightly revised version of Boot (2009).
[111] Herman Hugo's *Pia Desideria* (Antwerp 1624) was a book that gave rise to a series of imitations and emulations. A research project led by by Els Stronks and Marc Van Vaeck investigates the shape of that tradition in the Northern and Southern Netherlands. The present chapter is part of the Huygens Institute's contribution to that research effort.

process of adapting the text, its units may have been reordered, parts of it may have been removed, and units may have been split or merged. (ii) An application that facilitates annotating parallel texts will have to provide for an aligned display of these texts (with all of the complexities that the alignment may have), where both the individual texts and the aligned fragments can be subject to annotation. (iii) As useful as a body of annotations may be, it will only provide insight if it is possible to visualise the annotations based on the text alignment – only then can we can handle the higher-level differences in style, content and imagery that the parallel texts may reveal.

In this chapter I will focus on the first issue, while also continuing work on the SANE (Scholarly ANnotation Exchange) Markup Language, discussed in chapter 7. The chapter will formulate a TEI-based annotation format, and will replace the proprietary way of addressing the annotated resources used in the earlier chapter by the xpointer-based standard references of TEI P5, the most recent version of TEI. The format will use the mechanism of feature structures, recently standardised by ISO. I will begin by discussing some related work (in section 10.2), and then discuss the requirements for the annotation format in section 10.3. Section 10.4 proposes a TEI encoding for alignment and annotations. Section 10.5 discusses a prototype application that uses this format for storing annotations. The chapter assumes an overall familiarity with the TEI *Guidelines* (Burnard & Bauman 2007).

10.2 Related Work

Annotation is a subject that has drawn much interest in recent years. Researchers have worked on annotation frameworks from a number of perspectives. Linguistically oriented frameworks are Bird and Liberman's work on annotation graphs (2001), Carletta *et al*.'s work on the Nite XML toolkit (2005), and the Linguistic Annotation Framework under development by Ide and Romary (2003). Ide and Romary use a model based on feature structures. An annotation model targeted towards humanities collaboratories is developed by Agosti *et al*. (2004). Microsoft researchers developed a Common Annotation Framework directed towards collaborative functionality in office applications Bargeron *et al.* (2001).

The focus of this chapter is more limited. The chapter does not attempt to design a generic annotation model, but proposes a TEI implementation of such a model. One respect in which this implementation differs from other annotation models is that it assumes the annotated documents are accessible over the web. The annotation document should be similarly accessible. While in general it may be good system design to hide implementation details behind an application programming interface, there are also important reasons for exposing stored data to the world. One of these reasons is digital durability (applications will inevitably stop working, while data may survive), another is the

scholarly requirement of accountability, yet another the fundamental unpredictability of scholarly research needs (see also section 10.4.1 below).

In a series of papers, John Bradley has been exploring appropriate data structures for annotation in the humanities (J. Bradley 2004, 2005; J. Bradley & Short 2005). A recurring topic in these papers is the suitability of XML and relational databases for storing annotation-like information. In its exploration of the use of 'data'-like feature structures for annotation, the present chapter can be thought of as a contribution to that debate.[112]

10.3   Requirements

This section explores a number of requirements which an annotation format, more specifically a format that deals with text parallelism, will have to fulfil.

10.3.1   Parallel Texts

Text parallelisms are not given in the text, but are created by the researcher. In creating a parallelism between texts one stresses some aspects of the text at the expense of others. In the simplest case, one can construct a 1-1 correspondence between the constituents of two or more parallel text structures, but usually the situation is more complicated. For one thing, one may want to create a parallelism not between two but between many texts. We may, for example, consider the parallelism between Van Veen and Suderman, but we might as well include one or more of the other adaptations of Van Veen's book.

The text alignment that I develop here will only deal with structured parallelism, of the sort discussed in the introduction. It should be possible to discuss individual parallels or agreements between text fragments, but one does not need a larger alignment of parallel fragments to do so.

There are many other complications that attempts to create an alignment between parallel texts may have to deal with:

- new units may have been introduced, old units removed, as in Harvey's reworking (*School of the heart*) of Van Haeften's *Schola cordis*. (See Bath 2005);
- there may be different paratexts: Suderman's reworking of Van Veen's book is introduced by three Dutch prayers to the three persons of the Trinity; Van Veen introduced his book by (a.o. things) a Latin-language 'Carmen de amore'. This represents a dramatic change of context which influences the reading of the parallel texts;
- there may be changes in ordering of aligned units, as in the re-ordering of Alciato's emblems into a thematic collection, bringing it in line with the tradition of the commonplace book. (See Bath 1994: 31-35). What is of interest

---

[112] We already encountered this issue in chapter 5.

here is not so much changes in the texts themselves as the way they are grouped into meaningful units;

- part of a unit may have no counterpart in parallel text;
- units may have been merged or split;
- units may have been subsumed under a new hierarchy.

An alignment between text fragments has to make clear what is aligned, but also what cannot be aligned: the items that have no counterpart in the parallel text.

10.3.2   Annotanda

It follows from the above that annotation of parallel texts should not just be annotation of the parallel text fragments. It should also be possible to annotate those texts that have no equivalent in the other text. It should also be possible to address them as a group. Therefore, things that it should be possible to annotate (annotanda) are:

- texts alignments as a whole: it should be possible to discuss the merits and demerits of a proposed parallelism. If, for example, we have created a parallelism at the level of individual poems in two books, it should be possible to argue that the parallelism had better be defined at the higher (emblem) level, or vice versa;
- the individual entries in the text alignment, i.e. two or more aligned text units (it should be possible to describe e.g. their relation in terms of similarity, the way they compare in a certain respect, etc.);
- structural units in the text (poems, lines, pictures, emblems, pages);
- fragments of text (a phrase, a sentence) or fragments of pictures;
- groups or classes of text units ('the Spanish poems in *Amoris Divini Emblemata*', 'the preliminary material in the book', 'the poems without a counterpart in book x'). It should be possible to create these groups either by enumeration or by referring to some property these texts have (e.g. the xml:lang attribute in an XML file).

Finally, it should also be possible to create second-order annotations: to annotate other annotations and annotation types. It should be possible to comment on one's own annotations and on annotations made by others, as well as on e.g. the usefulness of a certain annotation type. As the annotation types embody a researcher's theoretical position, it is probable that serious discussion of another scholar's annotations will need to target the annotation types as well as the individual annotations.

10.3.3   Annotation Types

In the context of the EDITOR annotation tool (chapter 6) I have argued that usually researchers will want to define their own annotation types. The phenomena that a researcher studies are defined by his or her special interest and the texts being studied. The researcher will probably be looking for

something new: therefore pre-existing annotation types will usually be of limited interest.

However, I want to add that work on the EDITOR annotation tool bore out that an annotation system may want to make available a small number of default annotation types (such as 'comment', 'question', 'todo') in order to get annotators started. Following the ubiquitous web 2.0 practice, users will also expect a 'tag' annotation type. It should also be possible to re-use annotation types defined elsewhere, either by copying the definitions or by referring to them.[113]

Other requirements for the annotation types are largely similar to the requirements used for EDITOR:

- annotation types should be named;
- they should consist of one or more fields;
- these fields should have their own data type: in EDITOR the allowed data types were boolean, string, hyperlink and symbol (an enumeration of valid values), but other desirable data types include number, date, and rich text. Rich text is needed to do all those things that one usually does in prose: create lists, italicise, create hyperlinks, include figures, and create tables;[114]
- it should be possible to group fields together;
- it should be possible to have repeating (groups of) fields (see examples in section 10.4.2).

In addition to the fields that the researcher defines, an annotation system should create fields to hold information about the date the annotation was created, the date of last modification, and the users involved in these acts. These should be system-maintained fields available with each annotation type.

Using these components, scholars can create annotation types relevant to their research. An investigation into a reworking of an emblem book might define an annotation type that contains a number of rich text fields to describe possible changes in imagery, in the way of addressing an audience, and in vocabulary; a more quantitative approach might compare the rhetoric appeal of texts using a number field to hold the number of second person personal pronouns; another approach might characterise the adaptation process by choosing one or more of a limited set of available characteristics (e.g. 'literal translation', 'free translation', 'inspiration', 'modernisation', 're-telling').

---

[113] Compare Ide and Romary's proposal (2004) for a Data Category Registry.

[114] Another data type that might be thought desirable is a reference to an external taxonomy, such as Iconclass. I am assuming, however, that most of these taxonomies will shortly be reformulated as in terms of e.g. the W3C's upcoming SKOS recommendation (Isaac & Summers 2009). It will then be possible to refer to an entry in such a taxonomy using a URI. Thus a hyperlink data type should be sufficient.

10.3.4 Annotations

An annotation connects an annotandum and an annotation type, and assigns values to the annotation type's fields. Annotations should be subject to further annotation, and should therefore be addressable.

Annotations will be created, stored and displayed in sets. An annotation set will usually be stored in a single document. Besides the annotations, the document will contain information at the annotation set level (ownership, description, motivation, rights information), information about the documents that have been annotated, and about the annotation types.

10.4 TEI Representation of Annotations

10.4.1 Annotation Sets

I assume a situation where the texts that will be annotated are available in XML form. The parallel texts may coexist in a single document, they may each be stored in their own document, or the individual text units may be stored in separate documents.[115] For the purpose of annotating the parallel texts, this is irrelevant. If one of the texts is not available as an XML document, we may want to create a TEI proxy document for that text (see section 10.4.6).

For a number of reasons, the annotations will not be stored within the XML documents that contain the annotated texts:

- the choice would be either to create a new copy of the original XML file(s) or to modify the original copy. The first option would multiply the number of file copies and potentially create difficulties for annotation exchange. The second option may not be open to me (I will probably not have the right to change the original file) and would open up the files to corruption;
- as the number of annotations increases the size of the files might become unmanageable;
- as multiple people may want to annotate the same file, a complex locking scheme would be needed.

I will therefore assume that the annotations are stored externally in a single TEI document for each annotation set. For performance reasons an actual annotation system may need a database backend, but conceptually I envisage a set of annotations as a document, available at a URL. One reason for this is that, unlike a database, a document is something that one can store and come back to a few years later with a reasonable chance of the document still being readable. Another reason is that from elsewhere in the world one can point to a location in a document, not to a record in a database.

---

[115] In the Glasgow Emblem Project, each emblem is a separate XML file. Studying the parallelism between several Alciato versions there would entail creating an alignment between as many files as there are emblems.

The information at the annotation set level is best stored in the TEI header. A number of applicable elements is given in Table 5.

Element	Information category
title	Annotation set name
author	Annotation set author
availability	Rights information
aboutDesc	To identify and describe the files that are being annotated, I suggest the creation of a new element, aboutDesc, that contains either a series of paragraphs or a list of bibliographical entries (listBibl). This new element would form part of the file description. (The source description (sourceDesc) is unsuitable as these files are not sources).
projectDesc	Describes the purpose of the annotations and the thinking behind it
fsdDecl	Declares available feature structures.
creation	Annotation set creation date
revisionDesc	Revision log for the annotation set
appInfo	Records information about an application which has modified the annotation set

Table 5 (Proposed) TEI elements to store annotation set level information

I propose to use a new element dataSection to hold data about alignment and annotations. The dataSection is an element at the level of the text element in annotation documents. It would not be impossible to store annotation data in a text element, but it is inappropriate. In 'regular' TEI documents, elements like this are usually relegated to the text's back element, but there too, it might be cleaner to introduce a dedicated element.

10.4.2   Annotation Types

I will use the TEI mechanism of feature structures to model and store annotations. Feature structures are a mechanism originally developed for linguistic annotation (Langendoen & Simons 1995), but, as noted long ago by Susan Hockey, 'there is no reason they should not be used for literary and historical interpretation' (1996: 8). The TEI chapter on feature structures (Burnard & Bauman 2007: ch. 18) has recently been reformulated as an ISO standard (ISO TC 37/SC 4 2006). There is a separate mechanism (the feature system declaration) that is used to define acceptable feature structures and their allowable values. The feature system declaration is in the process of being standardised by ISO. The relevant elements are listed in Table 6.

Element	Description

f	feature
fDecl	feature declaration (field)
fDescr	feature description
fs	feature structure
fsDecl	feature structure declaration (annotation type)
fsDescr	feature structure description
fsDescr	feature structure description
fsdLink	link to feature structure declaration elsewhere
vColl	collection of feature values
vRange	feature value

Table 6 TEI feature structure elements used to store annotation and annotation type information

The feature system declaration is an element containing feature structure declarations, and each annotation type will correspond to a feature structure declaration. The feature structure declaration (an `<fsDecl>` element) can be defined in the annotation document, or it can be defined elsewhere and referred to by an `<fsdLink>` element. The `<fsDecl>` elements contain `<fDecl>` elements for the features (fields), and feature structure (`<fs>`) elements for the groups of features. The feature data types ('atomic values') allowed in the TEI *Guidelines* are numeric, binary, string and symbol. I will extend this by also allowing `<date>`, `<note>`, and `<ptr>`. The availability of the `<note>` element makes it possible to create annotations that use rich text. The availability of `<ptr>` makes it possible to refer to a web site (e.g. a Wikipedia page) as well as to point to any other resource accessible by URI: perhaps an image file, perhaps an entry in an external taxonomy, perhaps even another annotation. For the details of declaring feature structures I refer to the *Guidelines*, section 18.11.

An annotation application would most likely include a number of predefined annotation types (see section 10.3.3), i.e. a number of feature structure declarations. Others will be created by the annotator. Such an application should provide a guided interface for the creation of these annotation types.

A number of examples follow.

A 'tag' annotation type would consist of a single string field/feature, as follows:

```
<fsDecl type="tag" xml:id="...">
  <fsDescr>Tag</fsDescr>
  <fDecl name="tag">
   <vRange><string/></vRange>
  </fDecl>
```

```
</fsDecl>
```

The `<vRange>` element defines the permitted content of the feature.

A researcher that wants to characterise a text reworking in terms of use of imagery, rhetoric and vocabulary, could define the following feature structure declaration. It would allow him or her to create a note about each of these subjects for each alignment entry or for the alignment as a whole.

```
<fsDecl type="reworking" xml:id="...">
 <fsDescr>Describes reworking in terms of three
     characteristics</fsDescr>
 <fDecl name="charimag">
   <fDescr>Describes changes in imagery</fDescr>
   <vRange><note/></vRange>
 </fDecl>
 <fDecl name="charrhet">
   <fDescr>Describes changes in rhetoric</fDescr>
   <vRange><note/></vRange>
 </fDecl>
 <fDecl name="charvocab">
   <fDescr>Describes changes in vocabulary</fDescr>
   <vRange><note/></vRange>
 </fDecl>
</fsDecl>
```

Apart from the context of annotation of parallel text, a researcher interested in the representation of conversations in novels, e.g., might define an elaborate annotation type that, besides a general description, contains information about the participants (name and age), about the approximate duration of the conversation (number of minutes), a reference to a discourse typology, and a series of references to earlier conversations that this conversation refers to. Such a feature structure might look like this:

```
<fsDecl type="conversation" xml:id="...">
   <fsDescr>Describes conversations and their connections
</fsDescr>
   <fDecl name="general_description">
     <fDescr>Describes general aspect of
       conversations</fDescr>
     <vRange><note/></vRange>
   </fDecl>
   <fDecl name="duration">
     <fDescr>Estimated duration in minutes</fDescr>
     <vRange><numeric value="0"/></vRange>
   </fDecl>
   <fDecl name="conversation_type">
     <fDescr>Reference to xyz's conversation
       typology</fDescr>
     <vRange><ptr/></vRange>
```

```
      </fDecl>
      <fDecl name="participants">
         <fDescr>A list of participants</fDescr>
         <vRange>
            <vColl org="list">
               <fs type="participant"/>
            </vColl>
         </vRange>
      </fDecl>
      <fDecl name="conversations_referred_to">
         <fDescr>References to discussed conversations (refers
            to xml:id of corresponding annotations)</fDescr>
         <vRange>
            <vColl org="list">
               <ptr/>
            </vColl>
         </vRange>
      </fDecl>
   </fsDecl>
<fsDecl type="participant" xml:id="...">
   <fsDescr>Describes participant</fsDescr>
   <fDecl name="name">
      <vRange><string/></vRange>
   </fDecl>
   <fDecl name="age">
      <vRange><numeric value="0"/></vRange>
   </fDecl>
</fsDecl>
```

'Participant' is a separate feature structure here, used in the larger
'conversation' feature structure. Notice the repeating field
'conversations_referred_to' and the repeating group 'participants' (see section
10.3.3). The `<vColl>` element indicates the feature's content is a collection of
items.

It is clear that feature structures, especially as extended here, offer a very
powerful means of attaching many different types of annotation to texts in
general and to parallel texts in particular. A question that may arise is how this
general purpose mechanism relates to the more specific data structures that TEI
already offers. Suppose that in annotating a text we need to describe a number
of persons in terms of date of birth, death, affiliation, etc. Obviously, we would
want to use the prosopography elements made available with TEI P5. On the
other hand, there is something to be said for having a single uniform mechanism
for all annotations. As for at least some annotation types there will not exist
ready-made TEI elements, using feature structures for all annotations remains
an attractive solution – even more so if we should have an application that
creates entry templates for feature structures based on arbitrary feature
structure declarations.

One possible solution to this dilemma would be to reformulate the abstract TEI model for these elements, e.g. `<person>`, now formulated as an XML structure, in terms of feature structures. The attributes and sub-elements of the `<person>` element would become features within a corresponding person feature structure. One could even imagine a feature system declaration that functions as a library of ready-made feature structure definitions corresponding to all or most data-like TEI elements (`<person>`, `<event>`, `<biblStruct>`, etc). A user who needs to annotate a text using a bibliographical item could just point to the biblStruct feature structure and the application would present him/her with a template ready to fill out. It should not be too hard to convert this feature structure biblStruct into a canonical `<biblStruct>` when needed.

10.4.3   *Alignment*

Alignment will be represented by a `<linkGrp>` containing `<link>` elements. The `<linkGrp>` element will have type="alignment". The targFunc attribute will specify the functions of the link elements' targets, the first value corresponding with the first target, etc. All `<link>` and `<linkGrp>` elements carry xml:id-attributes, as they are all potential annotation targets (either within the current annotation set or from outside, perhaps in an annotation set created by a later scholar). Alternative alignments will be represented by a second `<linkGrp>` (or a third, etc). In another extension to the *Guidelines* we use the `<head>` element to label the alignment.

An example:

```
<linkGrp
   targFunc="English_epigram Dutch_epigram"
   type="alignment" xml:id="lg-1">
   <head>Alignment Van Veen Amorum Emblemata dut-eng</head>
   <link n="emblem 1"
      targets="http.../vae.xml#v012 http.../vae.xml#v014"
      xml:id="lg-1-001"/>
   <link n="emblem 2"
      targets="http.../vae.xml#v022 http.../vae.xml#v024"
      xml:id="lg-1-002"/>
   <!-- more link-elements -->
</linkGrp>
```

This alignment aligns poems in Dutch and English within a single book. The targFunc attribute describes which is which. We might have chosen to align poems in three languages, or poems from multiple books.

One possible complication in an alignment of more than two text series is that not all elements of the alignment are necessarily complete. In that case, the correspondence between the series of labels in the targFunc attribute and

the series of pointers in the links' targets attribute will break down. The easiest solution would be to place the individual targets, with their own labels, in separate elements. This would result in the following:

```
<link n="emblem 1" xml:id="lg-1-001">
    <target target="http.../vae.xml#v012"
        targFunc="English_epigram"/>
    <target target=" http.../vae.xml#v014"
        targFunc="Dutch_epigram "/>
</link>
```

It may sometimes be desirable to make clear which text units cannot be aligned in a given alignment. In the `<linkGrp>` that defines the alignment, I propose to include pointer (`<ptr>`) elements with type="unaligned" to point to the unaligned text units.

10.4.4   Annotanda and Annotations

An annotation links something that is being annotated and the annotation's content. In our case, the annotation's content is a feature structure (an `<fs>` element). I propose to create annotations by creating a pointer (`<ptr>`) element that points at the annotandum using the target attribute, and at the annotation content (the `<fs>` element) using the ana attribute.[116] The `<ptr>` element will be assigned an xml:id attribute, so that the annotation can be a target for further annotation. The feature structure's type attribute refers to the type attribute of the corresponding feature structure declaration. Its feature (`<f>` element) children give the values for the fields defined in the feature structure declaration.

An alternative for using the `<ptr>` element to connect annotandum and feature structure would be to use `<link>` elements. However, the advantage of using a `<ptr>` element to point at the annotandum is that its target attribute can use multiple whitespace separated URI's to refer to a number of locations that are to be considered a single annotation target. This can be done while still unambiguously identifying the annotation content (the feature structure) in the ana attribute. When using the link element, both annotandum and feature structure are pointed to from the link element's targets attribute, causing room for confusion.

According to the *Guidelines* (section 16.2), pointers can point at nodes, node sets, points (a location of length zero) and ranges (a range is a stretch of text between two points). TEI P5 uses the W3C's xpointer framework, for which it has registered a number of extension schemes. This seems sufficient for

---

[116] The ana attribute, which takes a URI for its value, 'indicates one or more elements containing interpretations of the element on which the ana attribute appears' (Burnard & Bauman 2007: 17.2).

pointing at fragments of the annotated texts, as they can be identified by an xml:id attribute, by an xpath expression, or even by a range of text within some element. Where necessary it is possible to point at a number of nodes, e.g. by selecting them in an xpath expression that refers to some value for the xml:lang attribute.

In the case of parallel texts, the annotation target will typically be an entry in an alignment, that is, a link element part of the same document that houses the annotation. In that case the pointer will not point not at the external document, but at the link element that represents the alignment entry in the annotation document. The same holds true for any situation where what one wants to annotate is not a resource or group of resources but a relation between resources.

A number of examples of annotations encoded as feature structures follows. First an example of the 'reworking' annotation type defined in section 10.4.2. The <ptr>'s ana attribute connects <ptr> and <fs>.

```
<ptr target="#lg-1-001" ana="#fs-001" xml:id="a-1"/>
<fs type="reworking" xml:id="fs-001">
   <f name="charimage">
      <note>about the <emph>imagery</emph></note>
   </f>
   <f name="charrhet">
      <note>about the <emph>rhetorics</emph></note>
   </f>
   <f name="charvocab">
      <note>about the <emph>vocabulary</emph></note>
   </f>
</fs>
```

Now an example of a more elaborate feature structure, involving the 'conversation' annotation type, also defined earlier:

```
<fs type="conversation" xml:id="fs-002">
   <f name="general_description">
      <note>Conversation taking place ...</note>
   </f>
   <f name="duration"><numeric value="15"/></f>
   <f name="conversation_type">
      <ptr target="http://example.org/typology/a34"/>
   </f>
   <f name="participants">
      <vColl>
         <fs type="participant">
            <f name="name"><string>John</string></f>
            <f name="age"><numeric value="19"/></f>
         </fs>
         <fs type="participant">
```

```
            <f name="name"><string>Alice</string></f>
            <f name="age"><numeric value="34"/></f>
        </fs>
    </vColl>
  </f>
</fs>
```

Finally a number of examples of valid TEI P5 pointer targets. They point at, respectively, an element by its id attribute (lg-1-1), an element that is the second child of the fourth child of an element with id 'emb01', a range of text between elements, a range of text given by offset and length within an element, a point within a text, and a series of `<lg>` elements having 'fr' as the value of its xml:lang attribute. It is clear these are very powerful facilities.

```
http://.../vae.xml#lg-1-1
vae.xml#element(emb01/4/2)
vae.xml#range(element(emb01/4/2),element(emb01/4/5))
vae.xml#string-range(element(emb01/4/2),5,3)
vae.xml#left(string-range(element(emb01/4/2),5))
vae.xml#xpath1(//lg[@xml:lang='fr'])
```

10.4.5 TEI Conformance

The proposed encoding scheme extends the TEI in a number of respects:

- it allows a number of existing TEI elements (`<note>`, `<ptr>`, `<date>`) to appear as atomic features, where the *Guidelines* do not permit them to occur;
- it allows `<head>` within `<linkGrp>`;
- it proposes a new element (`<dataSection>`) to store 'data'-like information;
- it proposes a new element (`<aboutDesc>`) to store data about the files that this document is about;
- it proposes a new element `<vDescr>` to describe the individual symbol values.

These modifications to the *Guidelines* have been documented in an ODD document (*Guidelines* section 23.4).[117] In terms of the section on TEI conformance (23.3), this makes the proposed encoding scheme a TEI extension. Documents conforming to this scheme can be transformed into TEI conformant documents, replacing the non-canonical `<note>`, `<ptr>` and `<date>` elements by string, replacing `<dataSection>` by `<text>`, and ignoring the `<aboutDesc>`, `<vDescr>` and the offending `<head>`.[118]

---

[117] The document is available in appendix D.

[118] Documents conforming to this scheme are not 'TEI conformable' (Burnard & Bauman 2007: 23.3.6), as information is lost in this transformation.

10.4.6   TEI Proxy Documents

If one of the texts is not available as an XML document,[119] we may want to create a *TEI proxy document* for that text. A TEI proxy document I define to be a TEI document that contains enough of the structural aspects of the texts to be able to serve as a basis for defining an alignment between the texts involved and for annotations to be attached to. In the case of an emblem book a TEI proxy would contain structural divisions for the emblems, the mottoes, the epigrams, the pictures, the quotations, but probably not for each line in a poem, depending on the level of detail needed for the annotation. The TEI proxy would not contain a transcription of texts. It might contain pointers to a digital facsimile instead, for instance one digitised in Google Books.

A reason for creating a proxy document rather than a full transcription is that it, presumably, saves work. Having a full transcription is preferable, but in a world of limited resources a proxy document may be an acceptable compromise between having all and having nothing. Once we can attach annotations to portions of a proxy document we can see which annotations refer to the same document fragment. We can also create links between document fragments. If we have related the text structure to facsimile images, we can also fetch the page that contains the text that is being annotated.

Apart from the usual TEI element and attributes, a TEI proxy document will have one new element (`<proxy>`), and one new attribute (proxyFor). The `<proxy>` element is used as a placeholder for those elements not transcribed in the proxy document. The result attribute[120] may be used to give the name of the element which the `<proxy>` element replaces. The proxyFor attribute contains a URI and can be used on the `<sourceDesc>` element. It can be used to point to a document that this document proxies for – e.g. by pointing to an entry in a library catalogue or to a book in a digital library somewhere.

To create links between text structure and a facsimile elsewhere, I use the TEI P5 `<facsimile>` element. The `<facsimile>` element contains `<surface>` elements, corresponding to pages or other objects that text is written or printed on. `<Zone>` elements can be used to identify regions in surfaces. The `<graphic>` element relates surfaces and zones to image files. The images can reside on a local server but might as well form part of a digital library elsewhere. From the text structure, the facs attribute points to the corresponding `<surface>`s and `<zone>`s in the facsimile.

Example
Google Books contains a copy of Edmund Arwaker's translation of Herman Hugo's *Pia Desideria*. A fragment of a TEI proxy document for this book might look like:

---

[119] One reason for this might be that there *is* an XML source for a digital edition, but it is inaccessible because of intellectual property concerns.
[120] Borrowed from the <join> element.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
   <teiHeader>
      <fileDesc> …
         <sourceDesc proxyFor="http://books.google.com/
            books?id=uyc3AAAAMAAJ">
         </sourceDesc>

         …
      </fileDesc>
   </teiHeader>
   <facsimile>
      ...
      <surface xml:id="s150" ulx="0" uly="0"
         lry="925" lrx="575">
         <graphic url="http://books.google.com/
            books?id=uyc3AAAAMAAJ&pg=PA150"/>
      </surface>
      ...
   </facsimile>
   <text>
      ...
      <div type="emblem" n="15" xml:id="e15">
         <pb facs="#s150" n="150"/>
         <proxy result="figure" xml:id="e15.1"/>
         <div type="quote" xml:id="e15.2">
            <proxy result="quote">
               <desc> Bible quotation</desc>
            </proxy>
         </div>
         <pb facs="#s151" n="151"/>
         <div type="motto" xml:id="e15.3">
            <head><num>XV</num></head>
            <cit>
               <quote>How shall we sing the Lord's Song
                  in a strange land?</quote>
               <proxy result="bibl"/>
            </cit>
         </div>
         <div type="subscriptio" xml:id="e15.4">
            <proxy result="lg">
               <desc>Poem consisting of multiple
                  stanzas</desc>
            </proxy>
         </div>
      </div>
      ...
   </text>
</TEI>
```

External annotations can unambiguously identify the structural components of
this text, e.g. by using xml:id attributes.

Other functions of such a proxy document might be to serve as a table of contents in digital libraries that do not provide one. In Google Books, for example, the generated table of contents is often nearly useless. A proxy document could also provide the basis for a 'jump to next section'-facility.

10.5   Conclusion

I developed a prototype application called PAT (Parallel Annotation of Texts) in order to experiment with the proposed annotation scheme.[121] Using PAT, a researcher can create an alignment between two series of texts, display the aligned texts, create and modify annotation types and annotations, and export the annotations either in the annotation encoding proposed in this chapter[122] or as a TEI conformant document.

PAT's main window by default shows a text alignment, the available annotation types and the created annotations. The user can change window contents and layout. Using the File menu, the user can create, modify and export annotation sets. Once an annotation set is open, the user can (1) select an alignment or create a new one; or (2) create or modify annotation types; or (3) create or modify annotations. Things that can be annotated are alignment entries, the aligned text units, annotation types and annotations. The application uses the annotation types' definitions to create the fields where the user can enter data. For 'symbol' fields, the user can select one of the valid values.

A number of lessons were learned from building this prototype.
1.   The most important lesson is that even though the feature structure formalism may seem intimidating to an encoder without adequate tool support, with the aid of a suitable application feature structures and feature structure declarations are very expressive and easy to create. The user need not even be aware of the fact that feature structures are being used to store his or her annotations. Feature structures provide a solid basis for a generically applicable TEI annotation format.
2.   There is no easy solution to the problem of modifying an annotation type that has already been used, that is, to change a feature structure declaration that some existing feature structures conform to. This is a

---

[121] The application was programmed in OpenLaszlo (client component) and Cocoon (server component).
[122] Actually it produces documents conforming to an earlier version of this schema. That version does not allow notes in feature structures, and it does not use the dataSection element. It stores annotations as link elements that link annotandum and feature structure. It implements only a subset of the desirable functionality as sketched in section 4. It does not, for example, facilitate the annotation of text or image fragments. There is as yet no possibility to use externally defined annotation types. The application does not handle nested or repeating feature structures.

3. problem that all systems for structured annotation share, but the problem is aggravated by the fact that the feature structure's type attribute, and the feature's name attribute, are repeated in the declaration and in each occurrence. The same problem would also occur, however, in more conventional TEI-XML encoding.

3. Even using the best of tools, the creation of text alignments is never going to be intuitive. At present the user can create parameterised templates for the URIs of the text units that should be aligned; the application will then increase the parameters' values a specified number of times in order to create the desired number of alignment entries. An alternative procedure would be to manually relate and label a large number of text units. Creating the alignment seems an activity best left to the technically trained user, and an XML editor might be the most convenient tool for creating it.

4. It is not self-evident how the aligned text units should be presented to the user. If the units are short poems, it is fairly simple to fetch these and display them next to each other. But it is hard, or impossible, to build a general-purpose display tool that will adequately handle any TEI text fragment (possibly including notes, images, references to other texts, perhaps modern translations, etc.) and create a suitable parallel display for an arbitrary number of these aligned texts.

Taken together these lessons suggest that, even though a general purpose tool for annotating text parallelism may not yet be feasible, the proposed annotation encoding format is sound. Feature structures, even though they are expressed in XML, present a type of encoding that is very different from what is usually associated with XML. They differ from conventional XML encoding in that (i) textual properties that are expressed using feature structures are not encoded as sub-elements within the text, but are associated with the text through a pointing mechanism; and (ii) a single set of abstract elements (feature, feature value, etc.) is used to encode all sorts of properties, unlike the more conventional practice of assigning each property its own element or attribute. This extra level of indirection and abstraction implies a need for application support, and feature structure declarations can provide a basis for the dynamic creation of windows for entering the corresponding feature structure data.

As to the alignment of parallel text units, the important point of the approach described here is that annotations are explicitly attached to a link between text units. This makes possible later filtering of annotations based on annotation target. For texts that do not draw much attention this filtering may not seem very important. For texts that we can expect to draw a large amount of scholarly attention in the foreseeable future, it will become important to distinguish annotations on the texts themselves from annotations that address the relation between the texts and its adaptations. Future generations of

scholars will continue to study many of the texts that we study today. Because we live in an electronic age, the output of their studies will remain accessible to later scholars. If we do not want later scholars to be overwhelmed by the amounts of scholarly publications of the generations before them, we must assure that those results identify as precisely as possible the objects that they are about. When they say something about the relation between two text units, they should make that explicit. This chapter is a contribution to an infrastructure where they can do so.

Part IV. Emblem Studies

11 A Mirror to the Eyes of the Mind. Metaphor in Otto van Veen's *Amoris Divini Emblemata* (Antwerp 1615)

11.1 Introduction[123]

In *Die Liebesemblematik des Otto van Veen* Anne Buschhoff states 'daß sich Van Veens Liebesemblemkunst durch eine besonders konkrete Anschaulichkeit und neuartige Verständlichkeit auszeichnet, da sie vielfach metaforisch verfährt' (2004: 292). This chapter will give a partial account of this metaphorical process in *Amoris Divini Emblemata* (Antwerp 1615). Buschhoff thoroughly explains many of the metaphors that van Veen employs, and the traditions from which they derive. She largely ignores, however, the mutual connections between many of these metaphors. I will argue that to some extent at least the book's coherence can be attributed to the coherence of the metaphors it employs. To arrive at a better understanding of this coherence we should, among other things, pay more attention to the *Hintergrundbildlichkeit*, which Buschhoff discusses only briefly (287). The metaphors that provide the visual, but also literary, context to each emblem's central metaphor do largely define the web of interrelated meanings that create the book's overall significance.

      The book itself will not need a long introduction. After his earlier book of secular love emblems, Otto van Veen published his religious emblem book in Antwerp in 1615. The book contains sixty emblems, with figures on the right page, texts on the left; for each emblem, there is a group of quotations in Latin, from the Bible and the church fathers, and three epigrams, in Spanish, Dutch and French respectively. Each of the emblems, except the first one, shows the protagonists Amor Divinus and Anima, Divine Love and the Soul, involved in a symbolic action that represents a stage on the Soul's journey towards heaven or God.

      A look at the emblem MENTIS SOL AMOR DEI[124] will show how the role of metaphor in *Amoris Divini Emblemata* goes far beyond the presence of a

---

[123] This chapter was published before as Boot (2007). Part of it was researched with the help of the metaphor index that I report on in chapter 9.

[124] 'The love of God is the sun for the mind' p. 18.

metaphorical object or activity in the heart of the pictura. After this introduction, I will discuss my understanding of what a metaphor is, and mention a number of the recurring metaphors that are introduced in the first few emblems of the book. Then, in a discussion of a sample of emblems, I will examine some of the issues surrounding the use of metaphor: spatial, landscape and plant metaphors, the relation to allegory, and the emblem as a representation of its own communicative situation.[125]

11.2 Mentis sol amor dei

Figure 42 MENTIS SOL AMOR DEI (6)

Many of the metaphors that create the intellectual coherence of the book are clearly present in MENTIS SOL AMOR DEI (6, Figure 42). First of all there is the light. The fullest statement of the metaphor structuring the emblem is given in the first quotation, attributed to Augustine: 'What the sun is to the senses, love is to the mind, and as the sun illuminates earthly things so that they may be seen, so God's love illuminates spiritual things'.[126] The pictura shows the sun, and a beam of light emanating from Divine Love's nimbus and shining on the Soul's heart. Metaphors of light are ubiquitous in *Amoris Divini Emblemata*. The Dutch and French epigrams mention the fruits born by nature under the influence of sunlight and compare these to the fruits of Divine Love. Imagery of fruits, of

---

[125] I will not discuss each significant detail in the emblems that I mention. My intention is rather to draw attention to some of the lesser-noticed features of the book as a whole .
[126] 'Quod sol est sensibus, hoc Amor est mentibus: & sicut sol terrena omnia illustrat ut videri possint: ita intelligibilia illustrat Amor Dei'.

plant life, of trees, and therefore of agriculture, recurs throughout the book. Mario Praz severely criticised Van Veen for the bleakness and barrenness of his landscapes in *Amoris Divini Emblemata*, but in fact most pictures show vegetation – as this one amply does (Praz 1975: 134-38). And in the context of this book, each tree and each plant is proof of God's presence in the world.

In the pictura, there is a noticeable number of vertical lines: both the tree and the chapel's steeple pointing upwards, the steeple firmly towards the sun. The metaphors of up and down, high and low, heaven and earth, good and bad, again are omnipresent in picture and text. The church, furthermore, is not just there to point upwards. Its presence is also a reminder of the overriding importance of religion in man's life. Just as the presence of Cupid may turn a picture of a candle into a sign of the dangers of love, the presence of a church turns an indifferent landscape into a religious allegory. Another aspect of this allegory is water: in the context of *Amoris Divini Emblemata,* water may be compared to virtue, or to love; moreover, in the pictures it seems very often to be related to the river of death –water that is to be crossed in order to reach a final destination. Hence also the importance of bridges.

Another aspect of metaphorical representation in *Amoris Divini Emblemata* prominently present in MENTIS SOL AMOR DEI is its allegorical embedding. The picture shows not just the metaphor, not just the illumination of the mind; it also shows Amor Divinus and the Soul in action. They are holding hands, which shows their intimacy; her body language seems to express her willingness to receive – the scene may even remind us of an annunciation, the Soul like Mary gratefully accepting. In fact, the vernacular epigrams explain that the light of the sun helps the earth bear fruit. The spiritual fruits which Divine Love's light helps grow are 'La vertu, le repos, la paix'. Even details like the Soul's wings are significant, as one of the quotations interprets them as the love and faith that help the Soul reach God.

Finally, I want to draw attention to a very pervasive parallelism in the emblems of *Amoris Divini Emblemata*: Divine Love illuminates spiritual things and in doing so illuminates the Soul; and in a certain sense this relation between Divine Love and the Soul mirrors the relation between the implied author of the emblem and its implied reader. The author illuminates the reader, just as Divine Love illuminates Anima. This will turn out to be a recurring feature in many emblems in the book and I will return to it at the end of the chapter.

11.3 Metaphor

So what is a metaphor? In the field of the bimedial genre of the emblem, we are fortunately immune to the lures of theorists that consider metaphor to be a linguistic or literary phenomenon only. For the purposes of this investigation, I will consider metaphor to be a conceptual entity, consisting in the conventional association between two concepts, the tenor concept and the vehicle concept.

The tenor is conceived in terms of the vehicle. The tenor and vehicle terminology stems from from the theories of I.A. Richards (1936), with a cognitive twist by Lakoff and Johnson (1980). Usually a metaphor is based on an agreement or structural analogy between the concepts. The vehicle concept, with its surrounding conceptual structure, provides a schema which helps organise and form the tenor concept (Goodman 1976).

The most important reason to view metaphors as primarily conventional associations between concepts is that metaphors typically occur in clusters: 'human being conceived of as a plant' may define such a cluster, leading to metaphors where stages of man's life may be compared to a plant's life cycle, or parts of his body to roots, branches, leafs. These metaphorical associations are indispensable for speaking and thinking about abstract matters.

The metaphors that appear in texts or images are examples of these conceptual metaphors. Parts of the vehicle's schema, as transferred to the tenor, will occur in the text and refer to the schema's original domain (love 'illuminates' the Soul). Other schema constituents may tacitly contribute to the significance of the metaphorical statement.

Metaphors are thus primarily a tool for understanding the world. But the promotion of a certain understanding of the world may also be employed for other reasons. A metaphor may be used to convince, but also to make a promise, as in 'He will cover you with his feathers', or in order to exhort, as in 'be like a mirror without blemish'. I will refer to this aspect as the metaphor's usage mode.

In her book, Buschhoff mentions some of the metaphors that the opening emblems of *Amoris Divini Emblemata* create or activate. Some of those we already met in the description of MENTIS SOL AMOR DEI. The heavenly light of the opening emblem will return on every picture in the guise of the nimbus of Amor Divinus, and often also as the sun. Another powerful metaphor which the first emblems evoke is the concept of life as a journey. It is a metaphor with a particularly rich source schema, and it provides a particularly forceful example of metaphors coming in groups: a journey may have a purpose, there may be ambushes, one can get lost, one needs a vehicle, a road to travel on, one may meet enemies or highwaymen, etc. Each of these aspects appears in *Amoris Divini Emblemata*. And each road or path in picture or text, even if not the subject of explicit comment, will still contribute to the overall significance of the emblem and the book as a whole.

11.4 Superna respicit

The first metaphors that I want to explore are those of plant life and landscape.[127] I will introduce them in a discussion of the well-known sunflower emblem from *Amoris Divini Emblemata*, SUPERNA RESPICIT (16).[128] The salient metaphor in SUPERNA RESPICIT (Figure 20) is the analogy between Soul and sunflower. As the sunflower directs itself to the sun, so the Soul should direct herself to God. The pictura shows the sunflower directing itself towards the sun, and thus contains the metaphor's vehicle. The metaphor's tenor is also present in the pictura, in an interesting way: the Soul too is turned towards the sun. That is: a single image fragment (the sun) represents both the sun itself, and, through the 'God is sun'-metaphor, God. This is not mere convention: God is being compared to the sun because of a double similarity. As the French epigram states, both can be described as author or creator of earthly life, be it in very different senses, and as is stated in the Dutch epigram, both are providers of light, again in different senses.

But the pictura contains more than just the metaphor's vehicle and tenor: it also shows how Divine Love is pointing the Soul towards the behaviour of the sunflower. Divine Love being the Soul's guide on her journey, we can take this to be an exhortation to the Soul to do as the sunflower does. The pictura therefore also contains what I have called the metaphor's usage mode.

The most interesting textual component of this emblem is the Augustine quotation:[129] 'But heaven and earth and everything that they contain, from all sides they say to me, that I should love thee, and they do not stop saying this to all…'[130] This text turns the sunflower into a metonym for the created universe, and that which we can readily see that the sunflower does (that is, follow the sun, orient itself towards its creator) is asserted to be true for all created things. Conversely, this also allows us to look at the landscape that surrounds the main metaphorical objects as a bearer of significance. The recurring features of this metaphorical landscape are: vegetation, light, churches, water and bridges, and verticality.

Many of these metaphorical fields, such as those of vegetable life, provide fine examples of pre-existing analogy that *Amoris Divini Emblemata* could just take up. In fact, the second emblem in the book, INCIPIENDUM (2, Figure 43),[131] already activates the plant metaphor: the texts quote the *Song of Songs* and state that winter is over, that a time of sweetness has come, and that

---

[127] This section takes up material that was discussed from a technical point of view in chapter 10. Here, I focus on content.
[128] '[He] looks upwards', p. 38.
[129] As discussed earlier in chapter 8.
[130] 'Coelum & terra & omnia quae in eis sunt, undique mihi dicunt, ut Deum amem; nec cessant dicere omnibus'.
[131] 'To start with', p. 10.

flowers appear on the earth. The pictura shows these flowers, which the vernacular epigrams describe as the flowers of grace that should adorn our hearts.

Figure 43 INCIPIENDUM (2)

Figure 44 AB UNO AMORE MULTA BONA (27)

There is one emblem where plant imagery appears as the central metaphor, and that is the emblem AB UNO AMORE MULTA BONA (27, Figure 44).[132] In this emblem, the texts present Love as the root of all good things: 'as many tree branches proceed from a single root, so too many virtues proceed from a single love. And the branch of good work has no greenness, if it does not stay in the root of love'. The pictura shows the virtues (strength, faith, hope, etc.) being offered for display by the Soul and Love conjointly. They are framed by trees with prominent roots.

In other emblems, as in MENTIS SOL AMOR DEI, the virtues are not compared to branches, but to fruit. Elsewhere again, the virtues are the plants themselves, and love is no longer compared to the plant's root, but to the water that feeds the plant, as in IN SPIRITU SEMINAT (22, Figure 45).[133] The French epigram says: 'Ainsi qu'en arrosant la plante / Elle s'accroist de iour en iour: / De mesme nostre foy s'augmente / Estant arrosée d'Amour'. The pictura shows Divine Love watering the plants that Anima has just sown.

The root can be metaphorically replaced by water, as both can be conceived of as an origin, which also connects to metaphors of sources and fountains. Water of course is necessary not just to plants, but to all living beings.

---

[132] 'Many good things from one love', p. 60.
[133] '[He] sows in the spirit', p. 50.

The life-giving forces of water are a central subject in three other emblems, among which is Sɪᴛɪᴍ ᴇxᴛɪɴɢᴜɪᴛ (39, Figure 46).[134] The water mentioned here is the water which, in a reference to John 4:14 quenches thirst forever.

Figure 45 Iɴ sᴘɪʀɪᴛᴜ sᴇᴍɪɴᴀᴛ (22)

Figure 46 Sɪᴛɪᴍ ᴇxᴛɪɴɢᴜɪᴛ (39)

The metaphors of plants and water are interesting, because they lend themselves very well for background scenery. The same holds true for the high-low opposition. Many pictures contain hills or mountains, churches on hilltops, narrow roads leading uphill. One of the emblems most explicit in creating and exploiting an equivalence between high, up, heaven, the good on the one hand, and low, down, hell and evil on the other hand, is Cᴏɴsᴄɪᴇɴᴛɪᴀ ᴛᴇsᴛɪs (52, Figure 47).[135] The Soul is toying with an arrow while she has to choose between the heavenly Jerusalem, on a mountain, a steep path leading uphill, and Babylon, on the plain; Amor Divinus is standing up proudly, wings turned up, head erect, a young man raising the cross, while earthly love is smaller and bent, a baby boy with wings dejectedly downwards.

In almost all emblems of *Amoris Divini Emblemata,* one ore more of these fundamental metaphors help set the scene for the emblem's central metaphor. And even though only the vehicle of these surrounding metaphors is present in the images, one can safely say that their presence is no coincidence. A good final example is Aᴍᴏʀ ʀᴇᴄᴛᴜs (4, Figure 48),[136] where the picture shows light, water, trees, a hill (meaningful verticality) and a church. There is of course no simple one-to-one translation, like 'water always means a or b or c', but what

---

[134] '[He] quenches thirst', p. 84.
[135] 'Conscience is witness', p. 110.
[136] 'Righteous love', p. 14.

the background elements do achieve is to present the landscape as meaningful and charged with moral tensions.

Figure 47 Conscientia testis (52)

Figure 48 Amor rectus (4)

11.5   Gravata respuit – Allegory

The second aspect of metaphor in *Amoris Divini Emblemata* that I want to discuss is the question of its relation to allegory. What I will try to show is how deeply the allegorical and the metaphorical are intertwined in this book. As we did for metaphor, we should realise that allegory is not just a trope. Allegory is not a painter's trick, invented to circumvent the difficulties of representation of abstract entities. Abstract entities are personified because, as Lakoff and Johnson write '[Personification] allows us to comprehend a wide variety of experience with nonhuman entities in terms of human motivations, characteristics and activities' (1980: 33). This will become clear once we look at the personifications in *Amoris Divini Emblemata*: in the first place of course Divine Love and the Soul themselves, but also Nature, Hope, and the virtues.

        A good example of the allegorical embedding of metaphor is Gravata respuit (23).[137] Gravata respuit (Figure 49) contains two main metaphors: the heavy geese that, as the text explains, are unable to fly, and the metaphor of 'up = good' which Amor Divinus draws attention to. The geese are weighed down by their useless fat. And as a quotation from Augustine makes clear: 'A soul in the

---

[137] '[He] rejects heavy things', p. 52.

bonds of earthly love has as it were meat on her wings and cannot fly, but once she is cleansed of that impure worldly desire, she spreads her wings and flies'.

Figure 49 GRAVATA RESPUIT (23)

Figure 50 NATURAM VINCIT (20)

As already seen in the context of MENTIS SOL AMOR DEI and SUPERNA RESPICIT, the presence of the allegorical figures of Amor Divinus and the Soul makes it possible for the picture to state the preferred attitude towards the metaphorical object. In the sunflower emblem, Amor Divinus indicated the sunflower to the Soul. In MENTIS SOL AMOR DEI, we saw the Soul accepting and taking to heart the illumination by Amor Divinus. In GRAVATA RESPUIT there is a complex interaction. Amor Divinus points upwards, for the Soul's benefit, as appears from his looking at her. She is looking at him, showing with her right hand that she is taking his message to heart, with her left hand rejecting the example given by the geese.

　　　The allegorical figures therefore indicate what I have called the usage mode of the metaphor. Anima makes clear the geese are to be rejected, Amor Divinus shows heaven is to be preferred. Of course the reader who witnesses this complex communication between the protagonists is also the ultimate destination of this communication. Very often the pictures employ the fixed and so-called allegorical attributes of Divine Love (the nimbus, the wings, the quiver, bow and arrows), and to a lesser extent those of the Soul, in order to highlight or add a shade of significance to the concepts of Divine Love or Soul. In GRAVATA RESPUIT the wings of the Soul are essential to contrast her ability to rise above the

earth with the geese's inability to do so. In the case of Divine Love, the pictures often use his bow with humorous effect.[138]

The other allegorical figures in the book are similarly employed within a metaphorical context. The personification of Nature in the emblem NATURAM VINCIT (20, Figure 50)[139] is clearly an allegory, in the sense of a personification of an abstract entity by a woman in a state of partial undress. Nature is being personified[140] in order to make her into an antagonist, someone of doubtful morals that may and should be pushed away. Personification is a metaphor because it allows to think about the abstraction in terms of a person, and thus e.g. as someone that may need to be vanquished. Similar remarks might be made about Hope, in ANIMÆ SPES OPTIMA NUTRIX (30, Figure 51).[141] Hope with her anchor is not a purely conventional sign: she provides the Soul with essential (metaphoric) nourishment, and their intimacy and the proximity of Divine Love add a meaning far beyond the merely conventional. Even where allegory is at its most conventional, as in the group of virtues in MICAT INTER OMNES AMOR VIRTUTES (41, Figure 52),[142] these are not just representations of the virtues. Their portrayal intimates, e.g., that they are subservient to Love, and that they belong together. The directions of their eyes may indicate that they rely on heaven for their strength. Being *bona fide* personifications, they are fully dressed. They are illuminated by the light shining from heaven.

To sum up: an investigation into the use of metaphor in *Amoris Divini Emblemata* shows (1) that because of the shared metaphors the book is very much a coherent whole, perhaps more so than has been previously thought; (2) that the embedding of a metaphor in an allegorical setting creates the possibility for communicating complex attitudes towards the metaphorical objects; and (3) that because of (1) and (2) the density of meaning in the pictures may be higher than hitherto recognised.

11.6 Author and Reader

To conclude, I want to return to the issue of the relationship between Amor Divinus and Anima, which I will approach by discussing the only picture from which they are absent, the pictura of the first emblem in the book, DEUS ANTE OMNIA AMANDUS (Figure 53).[143] We see a glory of clouds, with the words 'Oculus

---

[138] For instance in ODIT MORAS (31, '[Love] hates delay', p. 68) or ODIT TIMOREM (50, '[Love] hates fear', p. 106).

[139] '[He] conquers Nature', p. 46.

[140] As the many-breasted Diana of Ephesus.

[141] 'Hope is the best foster mother for the soul', p. 66.

[142] 'Love shines among the other virtues', p. 88.

[143] 'God is to be loved above all', p. 8.

non vidit, nec auris audivit.'[144] This Bible quotation is taken up in the epigrams to convince the reader that he should love God, and that God will give something in return that the eye has not seen and the ear has not heard.

Figure 51 Animæ spes optima nutrix (30)

Figure 52 Micat inter omnes amor virtutes (41)

Margit Thøfner, in her article on Theresian mysticism in *Amoris Divini Emblemata* (2002), considers these words to be only apparently anti-sensory, but what she does not take into account is the fact that it is in the picture of the very first emblem, in the most prominent place imaginable, that Van Veen reverts to using text rather than image. Even quite apart from the contents of the text, which seems an unequivocal repudiation of the senses, to me this is a very clear statement that the senses by themselves are untrustworthy and need guidance. What this emblem book sets out to do, then, is not just to show and explain metaphorically significant objects or actions, but also to thematise the process of guidance; which, I suggest, is why Amor Divinus and Anima are there in the first place, and why there is a natural similarity between Amor Divinus and the implied author on the one hand, and Anima and the implied reader on the other hand.

In some of the emblems that we have seen, this process is very clear: Divine Love indicates the sunflower to the Soul, Divine Love points her towards heaven, Divine Love points at the water she is drawing from the well,[145] and on each of these occasions, he draws her attention to the symbolic meaning of the objects – just as, by this very token, the author does to the reader. In other cases

---

[144] 'The eye has not seen, nor the ear heard', 1 Corinthians 2:9.
[145] in Sitim extinguit.

there is no symbolic object that Amor Divinus draws attention to, but still he is teaching, illuminating, enriching - all things that, metaphorically, the emblem itself accomplishes.

Figure 53 DEUS ANTE OMNIA AMANDUS (1)

Figure 54 VIRTUS CHARACTER AMORIS (14)

A very interesting emblem in this context is VIRTUS CHARACTER AMORIS (14, Figure 54), [146] as it shows the whole process (fig. 5): Divine Love points out a passage in a book, the Soul takes that to heart, and as a result, she is '[D]e bonnes œuures reuestuz', as the French epigram says, clothed in the virtues (bridle, scales, mirror) that the inscription states are the proof of Love. Now the book, presumably, is primarily the Bible, but I think we are justified in thinking of the emblem book as well, which would equate Divine Love with the emblem's author, Anima with the emblem's reader.

Similarly, in SOLLICITUS EST (54, Figure 55) [147] Divine Love holds up the Gospel to the attentive Soul, while in the background a woman uses a mirror to inspect her hairdo. The background scene glosses the Gospel as providing a mirror to the Soul, and the emblem texts concur. Van Veen quotes Gregory: 'Holy Scripture is placed as a mirror before the eyes of the mind so that there we see our inner face'. Again, to provide a mirror to the Soul is also what this very emblem does, and in that sense, the emblem pictures itself, and the communicative situation in which it functions. [148] Thus, the metaphors in *Amoris*

---

[146] 'Virtue is the mark of love', p. 34.

[147] '[He] is wakeful', p. 114.

[148] Gelderblom encounters a number of emblems in Luyken's *Het leerzaam huisraad* that similarly represent the emblem book itself and the process of guidance it should fulfil (2003).

Divini Emblemata do not just represent the properties of divine love, their avowed tenors. Ultimately, they may also mirror the book's aspiration of providing guidance to the Soul.

Figure 55 Sollicitus est (54)

To us, however, these mirrors cannot pretend to be without distortion. As many have said, the Soul in *Amoris Divini Emblemata* represents the book's reader.[149] However, to some extent this is no more than a rhetorical manoeuvre. The Soul as being depicted in the pictures is not just any reader. In her willingness to accept the teachings proffered in the book, in her ostentatious display of agreement with Divine Love, her representation is every inch as much calculated to convince the actual reader as is that of Divine Love. In some respects, rather than represent the reader, Anima is *playing* the reader, she is pretending to be the reader, in order to tell an actual reader what to do. This is a subject that we will discuss in the next chapter.

---

[149] See e.g. Thøfner (2002: 98)

12   Playing and Displaying Love. Theatricality in Otto van Veen's *Amoris Divini Emblemata* (Antwerp 1615)[150]

12.1   Introduction

In Emblem 23, GRAVATA RESPUIT (Figure 49),[151] in Otto van Veen's *Amoris Divini Emblemata*, Divine Love directs the Soul towards heaven. He points her away from the attachment to the world which, as her left hand shows, she too is rejecting. As she looks into his eyes, she places her right hand on her heart, in order to assure him that she understands his commands, perhaps even to request his approval.

One of the major innovations in Van Veen's *Amoris Divini Emblemata* is the presence of Divine Love and the Soul in all emblem pictures but one. Divine Love and the Soul are not just any two *Hinweiser*, however. The Soul is portrayed as in need of salvation and the help of *Amor Divinus*, or Divine Love. In many respects, the Soul and Divine Love exist because they are being looked at; they are actors staging condensed representations of highly symbolic scenes from the narrative of human salvation. In a sense, when Van Veen added the Soul and Divine Love to the pictures he set himself a task not unlike that of a playwright or even a play director: he had to make characters act convincingly for an audience. Their presence turns the emblem pictures into something very much like theatre stills.

Agnès Guiderdoni-Bruslé recently noticed the dramatic features of *Amoris Divini Emblemata*, stating: '[Anima and Amor Divinus] allowed for the dramatisation of the spiritual life and the love relationship between the Soul and God. [...] Each emblem, as a dramatic nucleus, unveiled a new step of the

---

[151] 'The Soul rejects being burdened'.

narrative to the eyes of the meditating reader' (2004b: 302). As early as the seventeenth century, playwrights such as Scherer and Lang recognised the theatrical potential of the spiritual love emblem and wrote plays that staged the vicissitudes of the Soul and Divine Love (or Grace) in silent scenes (Bauer 1982).

In the past, relations between emblem and theatre have been studied mainly from the point of view of the emblem's influence on the theatre. Schöne's *Emblematik und Drama* (1993) studied dramatic texts, language, and structure, even stage construction, from the perspective of the emblem, while countless studies attest to the influence of the emblem on Elizabethan drama (Szőnyi 2003). And while Schöne is aware of theatrical influence in the emblem, he unambiguously states that the theatre was the main beneficiary in the exchange: 'Vom Bild zum Spiel führt die Entwicklung. Sie trägt dem Spiele seinen Bildcharakter zu' (1993: 225). Van Gemert (1991) compares the functions of structural components of emblem and drama. When scholarly attention does focus on the theatre in the emblem, the subject is usually the theatre as a motif—the theatre as a stock metaphor for the deceptiveness and changeability of the world (Schilling 1979). Yet another strand of writing that relates theatre and emblem is concerned primarily with the theatre of memory and the role of the emblem in the art of memory (Bolzoni 2004).

This chapter examines *Amoris Divini Emblemata* from a theatrical perspective. The parallelism between the theatre and this emblem book is based on the shared characteristic of participants being assigned parts to perform in front of an audience. It traces some consequences of this basic parallelism for Van Veen's emblems rather than focusing on contemporary dramatic practice and its possible influences on the book. As an introduction, I discuss some general parallels between emblem books such as Van Veen's and the theatre, more specifically the theatre still. The body of the chapter is devoted to exploring concrete examples.

12.2   Frozen Theatricality

Emblem scholars tend to feel that, among the literary arts, the emblem is very special in being both a visual and a textual art at the same time. Both media, however, are also present in the theatre. Just like the emblem, the theatre requires the spectator simultaneously to attend to a text and to visual events, events that have a spatial dimension. As George Kernodle writes: 'It is time to recognise that the theatre is one of the visual arts. It is an offspring of literature on one side of the family tree, it is no less a descendant of painting and sculpture on the other side' (1944: 2). Some of the parallels between the theatre and the emblem, once pointed out, may seem self-evident, but it is still useful to mention them as they widen the perspective on the emblem.

Adding to the representation of an emblematic object, the portrayal of the interested parties of the Soul and Divine Love as actors turns the object into

an event, transforming the object's setting into a stage where this event can take place. The larger narrative of the Soul's salvation turns these events into scenes in an encompassing drama. The essential parallels between the *Amoris Divini Emblemata* emblems and the theatre derive from the presence of an ensemble of participants who play roles in front of an audience to stage a larger story. In both emblems and theatre, the human, or humanoid, figures are the principal bearers of meaning. Both show meaningful interaction among the participants. It may not be self-evident to consider the human figures to be the main vehicles for significance in Van Veen's emblem pictures. The pictures also contain emblematic objects such as the sunflower and the peacock. The preferred attitude towards the objects, however, is defined by the behavior and gestures of the human participants.

The basis for the analogy therefore is the personification of the abstract concepts of the Soul (*Anima*) and of Divine Love (*Amor Divinus*). As personifications they can act, and hence also play-act. It is important to notice that the participants play roles. Both representations (emblem and theatre) have been set up; they are not the representation of a naturally occurring event. They do not depict something pre-existing, but rather have been constructed, thought out, and designed in order to evoke a specific response from their readers or viewers. They only pretend to represent.

There is another parallel between the theatre and the emblems from *Amoris Divini Emblemata*, i.e. in the relation between participants and spectators. In the theatre actors can address the audience as well as other characters on stage. However, when ostensibly addressing the characters on stage, they implicitly also address the audience. Similarly, the participants in the emblem pictures can address the viewer and each other, but when addressing each other they are still also addressing the audience. Because of this, as in the theatre, in the emblem one can either play to the audience or pretend to ignore it.

In the wake of Schöne, Kirchner has argued for the existence of a much wider correspondence between emblem and drama, which is supposed to hold true for the whole emblem genre, not just for emblems showing human beings in interaction. This parallel is based on the emblem's and the drama's shared characteristic of being able to both represent and explain: 'Beide Kunstformen sind auf Grund ihrer deutungshaltigen Geschlossenheit in der Lage, hinter dem Schein der dargestellten Objekte und Ereignisse deren wirkliche Beschaffenheit Aufzudecken. [...] Solche Verweisungskraft begründet letztlich das wechselseitige Durchdringen von Emblem und Trauerspiel des Barock' (Kirchner 1970: 73). The present chapter does not treat this possible wider parallelism.

In some respects the emblem pictures correspond to the theatre painting or still, rather than the theatre itself, in that both emblem and still must represent a single instant from a larger scene. A frozen, unchanging image has to represent a scene that involves movement, change of position, and action. The

condensation of meaning from the larger scene into a single instant may naturally lead to gestures that are larger than life. Because of this, the emblem setting may foster playing to the audience rather than to the other participants, thus reinforcing the emblem's inherent tendency towards didacticism.

If one compares emblem pictures to theatre stills, one can view the emblem pictures as the record of a performance. What one engages in then is primarily the theatre in the sense of 'what takes place between and among performers and spectators', rather than in the drama, a 'mode of fiction designed for stage representation' (Elam 1980: 2). The emblems' didacticism, however, also suggests viewing the emblem pictures as a sort of screenplay, as a set of instructions for the reader to perform. The approaches are complementary and fit well into the notion of the Soul as an idealised reader.

The emblem texts, too, can sometimes be considered from both perspectives. The emblem text, viewed as a record of what is spoken in a production of a play, can alternatively be viewed as a dramatic text to be performed by the characters in the picture (and/or by the reader). In the case of *Amoris Divini Emblemata,* however, much of the epigrammatic text should be considered explanatory text, spoken by a commentator. This issue will be discussed in greater detail below.

The remainder of this chapter examines the emblems of *Amoris Divini Emblemata* from the point of view of this 'frozen' and possibly magnified, theatricality, posing the following questions:

(1)   How do the participants' actions (their gestures, the things they look at, the positions of their bodies) drive home the emblem's message?

(2)   How, if at all, does the emblem address its reader or spectator? How is the reader drawn into the dramatic situation?

(3)   Where does this leave the emblematic object? Readers tend to expect an emblem to present an object, or activity, or even a concept, which is explained in order to teach them a lesson. How does the presence of the dramatic action influence the presentation of this emblematic object?

The final section briefly extends this line of investigation to the emblem texts.

In addition to the artistic correspondences between the procedures of the emblem and those of the theatre, there are also common elements between these arts that derive from general social norms of decorum. Herman Roodenburg (2004) has pointed out how certain codes of civility permeated both social life and the arts (painting, theatre, even preaching). In Van Veen's emblems, figures do not simply place their feet next to each other. The norms of *contrapposto* (one leg forward, a swelling of the hip, one shoulder lower than the other) are usually observed. For the gestures of rejection, discussed in the next section, the Soul uses her left hand, turning her head away, as theorists

about eloquence and the theatre agree that she should.[152] Interesting as these parallels may be, at present they do not concern us.

12.3   Gestures

Figure 56 A MALO TUETUR (21)

Figure 57 AMORIS MERCES AMPLISSIMA (7)

The central element in the pictorial language of the Baroque was, to quote Andreas Henning, 'die affektgeladene Inszenierung der Gestik' (1999). In a number of emblems, the Soul uses gestures with fixed, symbolic significance to express her willingness to be guided. Examples of these traditional gestures of devotion, prayer, and awe are the hands crossed on her chest in Emblem 41, MICAT INTER OMNES AMOR VIRTUTES (Figure 52);[153] the folded hands in Emblem 21, A MALO TUETUR (Figure 56);[154] and the open submissive hand of Emblem 6, MENTIS SOL AMOR DEI (Figure 42).[155] The gestures are often accompanied by a kneeling

---

[152] See references to Quintilian, Caussin 1643, and Mallet 1753 in Barnett (1987: 60-61).
[153] 'Love shines among the virtues'. This is an "Inbrunstgestus", a gesture of fervor and deepness, according to Weise & Otto (1938: 28). The gesture also has a liturgical meaning, according to Barasch (1987). See also emblems NULLUS LIBER ERIT, SI QUIS AMARE VOLET (Emblem 40, 'No one will be free, if he wishes to love') and AMORIS FLAGELLUM DULCE (Emblem 28, 'Love's birch is sweet').
[154] 'Love protects us from evil'.
[155] 'The love of God is the sun of the mind'. This is an "Ergebengheitsgestus", or gesture of submission and resignation, according to Weise & Otto (1938: 5).

attitude, e.g., in Emblem 7, AMORIS MERCES AMPLISSIMA (Figure 57)[156], and Emblem 30, ANIMAE SPES OPTIMA NUTRIX (Figure 51).[157]

Figure 58 STERNIT ITER DEO (33)

Figure 59 AMORIS UMBRA INVIDIA (25)

In other cases, *Anima's* hand is on her chest, to confirm she is taking the lesson to heart, as in Emblem 23, GRAVATA RESPUIT (Figure 49).[158] Other examples are Emblem 3, EX AMORE ADOPTIO (Figure 62);[159] Emblem 7, AMORIS MERCES AMPLISSIMA (Figure 57); and Emblem 33, STERNIT ITER DEO (Figure 58).[160] Usually her gaze is directed towards Divine Love, to the object to which he draws attention, or to where he is leading her as, e.g., in Emblem 25, AMORIS UMBRA INVIDIA (Figure 59).[161] In some cases where she is looking at Divine Love, her look appears to be a silent request for approval, as in Emblem 46, OMNIA SPERNIT (Figure 60).[162] The latter also shows a characteristic instance of *Anima* rejecting what she is supposed to reject. It happens six times in all, in five of which *Anima* uses the left hand or foot, as mentioned above, as a sign of rejection. Another very characteristic event in OMNIA SPERNIT is Divine Love taking the Soul by the hand, leading her forward, while she does her best to prove her correct

---

[156] 'The reward of love is very rich'.
[157] 'Hope nourishes the soul best'.
[158] It is a gesture of asseveration and assurance, according to Weise & Otto (1938). It is done using the right hand (Barnett 1987). It also occurs in e.g. AMOR PURUS (10) and VIRTUS CHARACTER AMORIS (14).
[159] 'Adoption out of love'.
[160] 'Love paves the path for God'.
[161] 'Jealousy is the shadow of love'.
[162] 'Love spurns all'.

understanding of, and compliance with, his wishes. Nearly always, she is being led, being guided, being taught, being shown things, and being protected. When their bodies touch, if he is not seeking to guide her, she is looking for help and seeking support.

Figure 60 OMNIA SPERNIT (46)

Figure 61 NEC VIDISSE SAT EST (47)

Anima is not just doing what a Catholic soul should do in order to be saved, rather, she is overdoing it, and she is exaggerating for the audience, in order to make the desirable behavior absolutely clear. This process is not, I believe, related to the ecstatic rapture that some may find overdone in Baroque art (Treffers 1995). Ecstasy or rapture would not fit at all in Van Veen's book. What one sees is the magnifying process of the emblem at work, enlarging the representation of gestures which it could not actually perform.

12.4   The Emblem and the Audience

How does the presence of the participants influence the relation between emblem picture and viewer? As I argued in the previous chapter, the behavior of the Soul and of Divine Love gives an intrapictorial explanation of how the, usually metaphorical, emblematic object is to be understood. In doing so, the pictorial composition also echoes the relationship between the emblem book author and its reader.

To look specifically at how the viewer interacts with the dramatic aspects of the emblem pictures, let us first examine a case where the drama is

only subtly present, as in Van Veen's Emblem 4, AMOR RECTUS (Figure 48),[163] in which *Anima* and Divine Love hold up a plumb line for us to see. The first thing to note is that the picture requires the presence of a spectator; the only way to make sense of the action in the picture is to assume that *Anima* and Divine Love display the plummet for our inspection.

However static the composition may be, there is nevertheless a number of elements that make the picture more than the sum of its parts. The very fact of *Anima* and Divine Love being there together, standing next to each other, cooperating, and holding the plumb line together, is highly significant. That his hand is above hers is also meaningful. The wind that stirs *Anima*'s and Divine Love's robes—representing perhaps God's life-giving breath—contributes to making this a living event rather than a timeless moment. *Anima* looks up and extends her hand, presumably directing herself to God.

In this picture both *Anima* and *Amor Divinus* are turned towards the spectator, rather than towards each other. This is an extreme position in a continuum, of which another extreme would be *Anima* and *Amor Divinus* turned fully towards each other and excluding the audience. Emblem 47, NEC VIDISSE SAT EST (Figure 61),[164] comes close to this other extreme, but actually there is a slight opening up between the two lovers, which is, of course, a way of letting the spectator see. It is a concession to the need for theatre. In fact, the characteristic position of the Soul and Divine Love is intermediate, with their bodies turned somewhat towards an audience, while their faces are turned more towards each other. Barnett discusses this as the 'technique of double address', which is necessitated by the fact that ostensibly the characters engage with each other, but ultimately, of course, they always engage with the viewer (1987: 435-40).

Pictures where *Anima* and *Amor Divinus* are not turned towards each other at all are usually those where *Anima* is turned towards someone else; it would have been rude for a theatre actor to play with his side or back to the audience. In Emblem 3, EX AMORE ADOPTIO (Figure 62), *Anima* is turned towards Jesus, while *Amor Divinus*, usually the leader, is relegated to a supporting role. But again, though we have no part in this, all characters, Christ included, are turned slightly towards us, because it is for us that they stage the performance. It will come as no surprise that even though all action is ultimately for our benefit, the degree to which the protagonists are turned towards each other is also a measure of the degree to which we are, at least initially, excluded. One obvious consequence of the protagonists facing each other is that it becomes more difficult for them to look out of the frame and at us, as in Emblem 42,

---

[163] 'Right love'.
[164] 'Nor is it enough to have seen'.

OMNIA VINCIT AMOR (Figure 63)[165] and Emblem 11, IN UNITATE PERFECTIO (Figure 64).[166]

Figure 62 EX AMORE ADOPTIO **(3)**

Figure 63 OMNIA VINCIT AMOR **(42)**

Figure 64 IN UNITATE PERFECTIO **(11)**

Figure 65 AMOR PURUS **(10)**

---

[165] 'Love conquers all'.
[166] 'Perfection lies in unity'.

12.5   Drama and the Emblematic Object

The introduction of *Anima* and Divine Love in the emblem pictures brings with it the reference to a larger narrative. This narrative can be construed as the Soul's journey towards mystical union or, alternatively, as the Soul's journey towards heaven.

There are a number of possibilities for the pictorial relationship between the emblematic object and the larger narrative in the emblem.[167] Four broad groupings can be delineated: (1) the emblems where the main concern is the display of an emblematic object (or concept, or activity) without a significant amount of interaction between the protagonists; (2) emblems that still show an emblematic object, but bring out its significance in the interaction between the pictures' protagonists; (3) emblems that, in displaying the protagonists' actions towards an emblematic object, also display an event in the spiritual life of the Soul; and finally (4), emblems where one could say that 'drama takes over'. Here the significance of the picture is determined by the larger story in which the participants are involved, and any emblematic objects here will be of only secondary importance.

Figure 66 Sᴏᴛ ɪɴ ᴀᴍᴏʀᴇ ʀᴇᴄɪᴘʀᴏᴄᴀᴛɪᴏ (13)

Figure 67 Pɪᴀ ᴀᴍᴏʀɪs ʟᴜᴄᴛᴀ (12)

In the first group one finds cases where *Amor Divinus* and *Anima* together show an object to the viewer. An example is provided by Emblem 4, Aᴍᴏʀ ʀᴇᴄᴛᴜs (Figure 48), and similar cases can be found in Emblem 11, Iɴ ᴜɴɪᴛᴀᴛᴇ ᴘᴇʀꜰᴇᴄᴛɪᴏ;

---

[167] For another analysis of the way in which the addition of protagonists and setting may change the functioning of an emblematic object, see Guiderdoni-Bruslé (2004a).

Emblem 10 AMOR PURUS (Figure 65);[168] and Emblem 27 AB UNO AMORE MULTA BONA (Figure 44).[169] The picture is inconceivable without the supposition of an audience to which the emblematic object is shown. Similar are the cases where Divine Love and the Soul together symbolise a single concept, as, e.g., the concept of mutuality in Emblem 13, SIT IN AMORE RECIPROCATIO (Figure 66).[170] Similar configurations can be demonstrated for Emblem 12, PIA AMORIS LUCTA (Figure 67),[171] and Emblem 22, IN SPIRITU SEMINAT (Figure 45).[172] Again, the pictures where Divine Love and the Soul act jointly against a vice, as in Emblem 31, ODIT MORAS,[173] share the characteristic of the participants showing a meaningful object or activity to the viewer. A comparable situation is also portrayed in Emblem 53, SUPERBIAM ODIT (Figure 68).[174] Interaction between the participants here is typically limited, as the main thrust of communication is directed towards the viewer.

Figure 68 SUPERBIAM ODIT (53)

Figure 69 IUCUNDUM SPIRAT ODOREM (37)

In the second group, the interaction between Divine Love and the Soul highlights the importance of the emblematic object. That is what happens when Divine Love gives the Soul something for her consideration, such as the vessel of

---

[168] 'Pure love'.
[169] 'From one love many good things come forth'.
[170] 'May love come from both sides'.
[171] 'The pious struggle for love'.
[172] 'Love sows in the spirit'.
[173] 'Love hates delays'.
[174] 'Love hates pride'.

perfume in Emblem 37, IUCUNDUM SPIRAT ODOREM (Figure 69),[175] or is pointing at something for her benefit, as in Emblem 14, VIRTUS CHARACTER AMORIS (Figure 54).[176] In a variation on this procedure, Divine Love may help the Soul in an activity in which she is engaged, thus expressing his support and the importance of that activity, as in building the Christian life in Emblem 36, AMOR AEDIFICAT (Figure 70).[177]

Figure 70 AMOR AEDIFICAT (36)

Figure 71 PIETATE IN PARENTES POTIOR (18)

　　In a somewhat more complex form, a number of pictures represent a choice of some kind between different objects. A good example is Emblem 18, PIETATE IN PARENTES POTIOR (Figure 71),[178] where the choice is between the storks, representing parental love, and the crucifix. In this case, the choice is between something good and something better; in the example of Emblem 23, GRAVATA RESPUIT (Figure 49), the choice is between good and evil, heaven and earth. Other emblems representing similar choices include Emblem 46, OMNIA SPERNIT (Figure 60), and Emblem 51, AMORIS FELICITAS (Figure 72).[179]

　　In the third group of emblems the focus has moved away from the emblematic object towards an event on the spiritual journey of the Soul. Emblem 7, AMORIS MERCES AMPLISSIMA (Figure 57), for instance, depicts a scene where the Soul is being rewarded; the emblematic object (the laurel crown) is

---

[175] 'Love exudes a pleasant aroma'.
[176] 'Virtue is the hallmark of love'. See also Emblem 16, SUPERNA RESPICIT ('Love looks up at what is on high').
[177] 'Love builds'.
[178] 'Superior to the piety towards one's parents'.
[179] 'The happiness of love'.

just an element in this scene and not in itself an object for contemplation. The picture shows a scene from the drama rather than an object or activity to contemplate. Similarly, in Emblem 33, STERNIT ITER DEO (Figure 58),[180] though the road between God and humankind, mentioned in the motto and identified in the epigrams with Divine Love, is visible in the picture, its pictorial significance is less important owing to the encounter of Jesus and the Soul for which the road provides the setting.

Figure 72 AMORIS FELICITAS (51)

Figure 73 SINE AMORE MORS (55)

Lastly, the fourth group contains emblems where the *pictura* focuses on an allegorical depiction of crucial events in the spiritual life of the Soul. There are still metaphorically significant objects, but they play only a minor role. In Emblem 2, INCIPIENDUM (Figure 43),[181] Divine Love and the Soul set out on their spiritual journey. In a sense, the scene is repeated in Emblem 55, SINE AMORE MORS (Figure 73),[182] where Divine Love approaches the Soul, who is spiritually dead because she lives without Divine Love. Similarly, Emblem 3, EX AMORE ADOPTIO (Figure 62), shows an allegorical representation of God's adoption of the Soul in which metaphoric objects only appear on a second plane. A subgroup of these emblems is alike in that they depict the dangers that the Christian Soul may encounter and the protection offered by Divine Love. In Emblem 21, A MALO TUETUR (Figure 56), Divine Love has slain the enemies of the Soul and further protects her with his shield; both Emblem 25, AMORIS UMBRA INVIDIA (Figure 59),

---

[180] 'Love paves the path for God'.
[181] 'To start with'.
[182] 'Without love there is death'.

197

and Emblem 38, AMORIS SECURITAS (Figure 74),[183] portray similar situations. The significance of the *pictura* is in the portrayal of the protection Divine Love affords her and of the Soul's willingness to be protected.

Figure 74 AMORIS SECURITAS (38)

Figure 75 AMOR DOCET (8)

12.6   Emblem Texts

In his article on speech bubbles in emblems, Laurence Grove has pointed out that 'emblematic works that have speech bubbles are generally those that have something to sell, and this most commonly is a religious belief' (2005: 102). Grove argues that the conjunction of the intimacy of dialogue and the immediacy of the image in speech bubble emblems is a way of drawing the viewer into the conversational exchange.

Since speech bubbles express direct speech and possibly dialogue, in the emblem one would look for these primarily in the *subscriptio*. In the *subscriptiones* of *Amoris Divini Emblemata*, however, the amount of direct speech, let alone dialogue, is minimal. Most of the epigrams contain commentary rather than text spoken by the participants. Nevertheless, the theatrical analogy could be maintained, as many theatre genres, like the early cinema, employ commentators or *bonimenteurs* (Hummelen 1992; Lacasse 2004). Interestingly, however, there is an eighteenth-century adaptation of *Amoris Divini Emblemata* by Jan Suderman (Amsterdam 1724), who wrote new epigrams in Dutch to accompany the mottoes and pictures of Van Veen's

---

[183] 'Love's security'.

book.[184] In his preface, Suderman declares that he has tried to give speech to the 'silent poetry' of the pictures, and that he has avoided reading the original texts. While in Van Veen's vernacular epigrams usually the speaker cannot be identified with either the Soul or Divine Love, in Suderman's texts about half of the epigrams are spoken by the Soul, usually addressing Jesus.[185] Four *subscriptiones* are explicitly marked as dialogues between Jesus and the Soul.

If indeed there is a parallel between Van Veen's emblems and theatre scenes, what Suderman does, for those four pictures, is to provide a possible stage text for the protagonists. In, e.g., Emblem 8, AMOR DOCET (Figure 75),[186] he has Divine Love explain to the Soul the importance of the narrow gate depicted in the background of the picture. The Soul answers that she 'hear[s] and write[s] down your dear lessons, my bridegroom' and exclaims that she will enter that gate at whatever cost. Divine Love then says the gate is wide enough for those who have laid down their earthly burdens. The very simplicity and plainness of the exchange are the best proof that the remarks are meant for the audience rather than for the primary recipients.

In the dialogue to Emblem 4, AMOR RECTUS (fig. 8), the Soul exclaims to Jesus: 'My Beloved, that this painting may instruct me.' Within the emblem the Soul exclaims the sentiment that the devout reader should exclaim outside it. In the dialogue to Emblem 18, PIETATE IN PARENTES POTIOR (fig. 17), referring to Jesus' finger pointing at the crucifix, the Soul says: 'So it is, my Love. Your finger shows me the pains you have suffered'. Again, the additional text spells out what is implicit in the picture, i.e. the Soul's awareness of Jesus' suffering. The Soul therefore, in Suderman's interpretations of these pictures, plays the part that, I believe, Van Veen had already assigned to her, and in the process also demonstrates to readers the script they have to follow.

12.7　Conclusion

Van Veen's religious emblem book is unique in its thematic unity and pictorial consistency. These qualities are, to some extent, what makes an analysis along the preceding lines possible. Another element that may make such an analysis especially relevant in the case of Van Veen is his training as a painter and experience in representing complex compositions and human emotions. In later books of religious love emblems, such as Hugo's *Pia Desideria*, *Amoris Divini et Humani Antipathia*, and Luyken's *Jesus en de ziel*—all heavily influenced by Van Veen as they were—the gestures, looks, and behaviour of Divine Love and the

---

[184] The book also contains adaptations of Hugo's *Pia Desideria*.
[185] Suderman does not seem to distinguish between Divine Love and Christ.
[186] 'Love teaches'.

Soul are sometimes harder to recognise than in the case of Van Veen.[187] Nevertheless, a theatrical analysis of these and other emblem books would be worthwhile. Human or human-like figures in emblem pictures, as soon as they begin to interact among themselves or with the spectators, become vehicles for shades of meaning that merit analysis. They create the potential for the internalisation of the emblem's lesson, through identification with the protagonists and immersion in the dramatic situation.

Jan Luyken's pictures would provide an interesting subject for further study. The emblems of Pieter Huygen would be of special interest (Amsterdam 1689), as Luyken provided the pictures to which Huygen wrote the texts. As Vinken has shown (1960), Huygen did not always fully understand Luyken's intentions, and these misunderstandings concern, among other things, the part played by the ubiquitous spectators Luyken included in the pictures. Vinken makes clear that the spectators have the formal role of guiding the reader's eye into the pictures, while they also simultaneously represent us, the book's pious readers as the spectators outside the frame. For some reason, though, Vinken considers this to be a non-literary, purely stylistic matter. That Luyken should have wanted devout and pious characters for his spectators is surely not just a question of style, however. Here, as in the theatre, it is true that 'all that is on the stage is a sign'.[188] The spectators stand, and look, and even move in ways that are meaningful—they make gestures and they interact. Even if Huygen has read these signs incorrectly, he is still their earliest interpreter.

The abstract form of dialogue is also encountered in Luyken's own emblems from *Jesus en de ziel* [*Jesus and the Soul*] (Amsterdam 1678).[189] In this book the epigrams are usually spoken by the Soul, and are followed by a scriptural quotation which Luyken headed 'divine response.' Without this 'divine response' we might not even have conceived of these emblem texts as dialogues. In most cases, the Bible quotation does indeed seem to answer a concern raised by the Soul in the epigram. To be able to see the emblem text as a dialogue, as a text therefore that is performed in front of an audience, brings out the double addressee of the text. It makes apparent that no text is ever as thoroughly lyrical as it appears at first glance, and that the phenomenon of theatricality may be more widespread than previously recognised.

---

[187] Among other things because the figures here are sometimes turned with their backs to the audience.

[188] Veltruský, quoted in Elam, 7.

[189] Besides the epigram and divine response, the emblem text also contains another Bible quotation (placed below the picture), a prose commentary, and sometimes other texts. A divine (or 'sacred') response also occurs in the emblems in Luyken's *Vonken der liefde Jezus*.

Part V. Reflections

13   Mesotext. Framing and Exploring Annotations

13.1   Introduction

The ultimate justification for digitisation efforts is not mere electronic availability of texts, however important that is. The wider issue is to make the content of works available as potential nodes in a larger digital network that will include not just the sources but also the tools, the output and the intermediate products of scholarship.[190] This chapter explores what this wider digital network of scholarly material may look like, and it will argue that annotation will fulfil a central role in this network.

Annotation is, as we have seen in chapter 3, an important item on the wish list for digital scholarly tools. It is one of John Unsworth's primitives of scholarship (2000). It was identified as a potential 'killer application' for digital humanities (Juola 2008). In edition studies, Peter Robinson (2003) has called for it to be included in future digital editions. At Brown University's Virtual Humanities Lab, work on annotation facilities for its Boccaccio editions is in progress (Zafrin & Armstrong 2005). Wittgenstein students are working on the integration of annotation into a digital edition (Hrachovec & Köhler 2002). A number of annotation systems has been discussed in the context of this book.

The word I will use for a body of annotations is *mesotext*. 'Mesotext' because it is text that can be located somewhere in between the primary texts of scholarship (the sources that scholarship is based on), and its secondary texts (the books and articles that result from it). In the next section (13.2) I examine the concept of mesotext. In section 13.3 I describe a number of instances of the concept, some referring to earlier chapters of this book, some pointing elsewhere. The conclusion lists a number of conditions for successful mesotext.

---

[190] When we set out on what was to become the Emblem Project Utrecht, Els Stronks and I wrote a paper on the kinds of analysis we wanted to facilitate in this digital network (Boot & Stronks 2002). Subsequently, we have researched Petrarchist imagery in Heinsius' emblems (in collaboration with Jan de Boer) and rhetorical elements in Jacob Cats' emblems (Boot & Stronks 2003). In addition, Els Stronks has analysed the presence of churches in our material (2007a).

13.2   Mesotext

Mesotext is text that is about primary texts that are the object of scholarship. It provides the argument for secondary texts that are the output of scholarship. The basic ideas of mesotext may be formulated as follows: (a) The business of (our area of) scholarship is the creation of theories about or on the basis of texts; (b) to a large extent, scholarship consists of taking notes; (c) these notes record observations that refer to texts or fragments of texts; (d) the notes consist of statements that apply theoretical notions to these texts or text fragments; (e) these notes can provide micro-arguments that validate the theoretical notions and their application; (f) in the digital era, if properly realised, the notes define a hypertext structure that seamlessly connects the primary texts, the notes themselves, the concepts that they instantiate, and the secondary texts that explain these concepts to the public. Figure 76 shows the resulting network, and shows how mesotext is positioned between primary and secondary text. Following e.g. Leggett and Shipman (2004), I will refer to the secondary, scholarly text, as opposed to the underlying data, as the 'narrative'. The narrative explains and contextualises the data. Mesotext consists of the theoretical notions (the model) and the annotations that implement the notions, which I call *mesodata*.
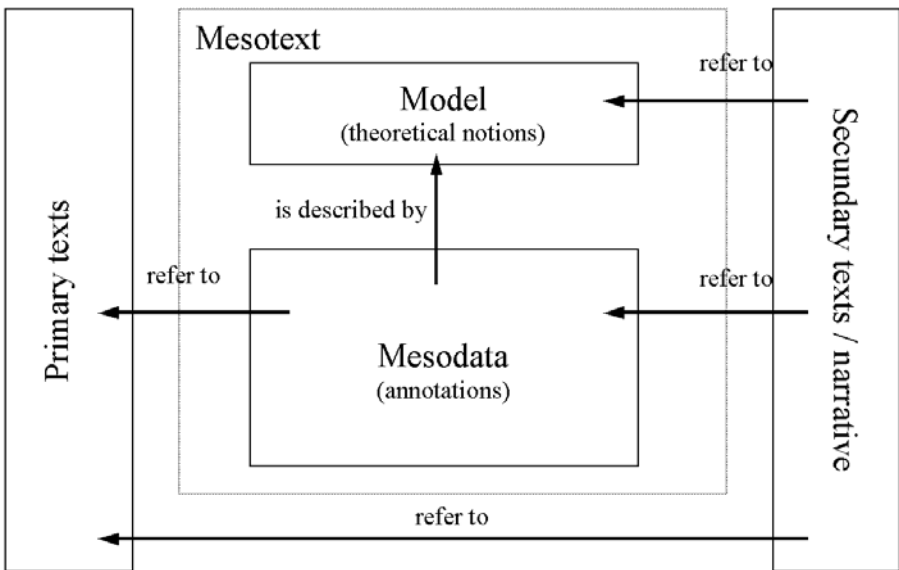

Figure 76. Mesotext (model and data) between primary text and narrative

One can conceive of mesotext in a number of different ways. One can view mesotext as (1) primarily a reaction to a primary work. John Bradley's Pliny annotation tool sets out to do just that: 'the addition of annotations or the creation of notes to record one's initial reactions to what one has read' – though it does considerably more (2006). One can also move away from the primary text and view mesotext as (2) a body of supporting evidence for a scholarly argument. And lastly, one can try to bracket out both primary text and secondary text and look at mesotext as (3) merely a collection of data, or observations.

From the first perspective, mesotext consists of notes, scribblings, lists and tables that are created in coming to grips with the sources that are the subject of scholarship. When the writing is done, the mesotext tends to be discarded. If considered at all, it is seen as preliminary, unripe, chaotic material. What I am arguing, however, is that in a digital context, mesotext can enrich scholarship. Mesotext contains the statements that support a scholarly article's argument – it is the material that underpins secondary texts. The underpinning consists of, to some extent at least, observations on the primary text(s). In a digital environment, the mesotext can be made accessible from and provide an entrance into the primary text that it is about, as well as into the secondary text it supports.

Mesotext is not limited to running text. As mentioned before, it can contain lists and tables. It may also contain categorisations and hyperlinks. Moreover, it need not even be just text (depending on how wide one's concept of text is). It may include statistical computations, diagrams, graphs, even pictures. It may also consist of word processor documents, spreadsheets, databases, or output of annotation programs.

Mesotext is nothing new. The creation of notes to text has been a scholarly practice since time immemorial. There have been periods when nearly all scholarship consisted in note taking and making, in the creation of brief glosses or longer commentaries on ancient or sacred texts. In modern editorial practice, annotation and commentary are still valuable end products of scholarly activity. In other branches of scholarship, however, annotation has been relegated to a subordinate, preliminary status. Scholars write notes in order to write articles; the notes are thrown away when the article has been completed.

In the second perspective, the view from the scholarly text, the notes are not necessarily thrown away, at least not all of them. Rather, they should be systematised and made to conform to the theoretic notions that have grown out of reading and studying the primary text, and that inform the secondary text. Inevitably, this will prompt further revisions of these notions, as any new contact with a primary text is bound to do. In addition, revised theoretical notions may prompt new rounds of systematisation of the notes. If we are fortunate, and if we have limited our subject wisely, these repeated revisions will end in a body of

notes that provides the supporting argument for the scholarly text and details the application of the theoretical notions to the primary text.

By 'theoretical notions' I do not necessarily mean anything very grand or abstract, such as Marxism or deconstruction. However, a note about a fragment of a primary work is never just a statement of fact. It is a statement that applies an idea, an aspect of a model, to that work. The note validates the model, extends its scope, or perhaps refutes it.

A successful model in humanities computing should, according to McCarty, be computationally tractable, i.e. be explicit and consistent, and it should be manipulable (2005: 25). A model that is explicit in the aspects, fields, variables, or categories that it describes, structures the observations that result from its application. That structure defines filtering and navigation facilities on these observations. For example, if a model for narrative characters includes gender and age and explicitly describes these traits, one can list the characters or their actions by gender and age and maybe even cross-tabulate behavioural aspects against these parameters.

From the third perspective, we can look at mesotext as being just data, not in the sense of being merely an unorganised collection of crude facts, but rather as a collection of observations, that may be more or less interpretive, but is not embedded within a narrative that organises the observations. Outside the emblem or humanities realm, an atlas may be a good example of what I mean: it is purely factual; nevertheless, its maps are organised according to a model. Its facts are not there, however, to support a specific argument. In this context, I will also speak of mesodata. Mesotext is mesodata, seen through a model. Mesodata is made accessible by the existence of the model.

When thinking of a collection of observations within the framework of emblem studies, the emblem index comes to mind, and in this repect one may want to think about the Henkel & Schöne *Handbuch zur Sinnbildkunst* (1976). Emblem scholars know that this book indexes emblems by their main iconographical motif and that it provides, for each of the indexed emblems, a very condensed description of the meaning. It contains a number of supplementary indices, by motto and by meaning. If only in dimension, it is a monumental feat of scholarship in itself. But it is also a collection of data that provides underlying argumentation for many actual and potential emblem studies.

Let us imagine for a moment that a far-sighted publisher would create a digitised, searchable version of the *Handbuch*. Furthermore, imagine that a publisher or learned society, or anyone else, should decide to create an online emblem journal. Then, try to envisage that with each mentioning of an iconographical motif in this journal, there would be a popup menu, and one item on this menu would be 'look up in digitised Henkel and Schöne'. And let us imagine one more thing: that the emblems in the digitised *Handbuch* should be hyperlinked to their display on one or more of our digital emblem sites. Lastly,

suppose that the digitised *Handbuch* creates a reverse link to the places where it is linked from, and thus probably to where it is referred to or quoted. There is nothing very revolutionary, or even difficult or sophisticated about these imaginings.

What we would have then would be a perfect, if straightforward, example of mesotext. A large amount of raw (meso-)data, helpful in exploring the primary works that have been indexed; accessible from the scholarship that is partially based upon the data; helpful in investigating the claims that scholarship makes; referring back to the scholarship that quotes it. The index of the *Handbuch* would in this case provide the model of the visual world that the observations are based on. The reason why I call this hypothetical example of mesotext straightforward, is that this index is not a very sophisticated model.

In fact, this is not that far from reality. Hans Brandhorst and his Mnemosyne group are working on a cumulative Iconclass index on the emblems in the Utrecht, Wolfenbüttel and Glasgow sites (Brandhorst 2004). Researchers will be able to link to that index and find emblems by iconographical motif, and the index will refer visitors to the site that displays them. In this case, the Iconclass system provides the model for the annotations.

I will not suggest the Henkel and Schöne *Handbuch* or the Mnemosyne Image Index are the kinds of works that come into existence as the by-product of an article, saving scrap paper from the waste bucket. However, the *Handbuch* and the scholars' notes share some important characteristics: with some degree of precision they refer to one or more primary texts, to some extent they conform to a model and they can support an argument.

McCarty notes that annotation and commentary have much in common (2005: 93). Annotations are the 'morsels' a commentary may bring together, commentary consists of morsels that might live as individual annotations. Annotation, however, is a much wider phenomenon than that which would fit in a commentary. Categorisation, illustration, hyperlinking, modelling: anything that can be said with regard to a text is annotation. Mesotext is an organised collection of such annotations. Mesotext is metatext. As Gérard Genette would put it, it is text that talks about another text (1982). But unlike the ordinary secondary scholarly text, mesotext in a sense is data. It has no explicit point of view. There is no thesis that it explicitly argues for – though it may be used to argue for one, clearly. As the word mesotext indicates, it is text that is framed by other texts - the texts it is about, the texts it supports, and, as we will see, the models that inform it. The concept of mesotext may help allay fears that the fragmented nature of the web will strip scholarship (and perhaps life) of some of its coherence and thus of meaning.

13.3   Approximations to Mesotext

I will discuss a number of examples that illustrate what mesotext may look like. The first one examines the reading notes of reading writers, and considers their intermediate position between the work they respond to and the work they may be preparing. The next two examples come from the chapters on SANE and the metaphor index in this book. The fourth one discusses some tendencies in present-day science publishing. The fifth example is taken from Vika Zafrin's work on a hypertextual rendering of the tradition of the *Chanson de Roland*.

13.3.1   Reading Notes and Commonplace Books
Readers' notes, as discussed by Dirk Van Hulle & Wim Van Mierlo in their introduction to the special issue on the subject of *Variants* (2004), reflect the reception of a book by the reading public, and are therefore of interest to book historians and historians of reading. When the reader is also a writer, besides a record of his reading, these notes are often also the first stage of his writing. They trace 'the interiorisation of the reading, the transformations and mutations of the intertext, the recreations of the source' (3). Van Hulle & Van Mierlo see reading notes as 'documents with a specific value in their own right, which deserve an appropriate critical and editorial treatment' (2).

One of the examples examined in the volume is that of John Dawson, an eighteenth-century British excise officer (Colclough 2004). Dawson was not an intellectual, but owned a sizable collection of books that he bequeathed to his parish, and the collection as it has come down to us is nearly complete. Dawson was an avid reader of contemporary history. While reading, he built manuscript tables of contents and indexes to the works he read. The sheets are now often bound with the relevant volumes. Sometimes the indexes turn into annals of their own, as in the lists of the parliaments associated with each English monarch. For each parliament Dawson lists a number of properties, such as the dates sessions began and ended and the date parliament was dissolved. This begins to look like an annotation type consisting of multiple fields, such as we met in previous chapters of this book. Dawson also worked on a volume called *My Own Life*, and some of his annotations are directed towards this writing, as he collected references to national events and events in his own life in his history books.

Ferrer (2004) looks at these Janus-like notes (looking backwards to the text they refer to, looking forward to the text they are working towards) from the point of view of genetic criticism. An important point he makes is the multiplicity of contexts that annotations are part of. An extreme case Ferrer mentions is the note that Coleridge made in a Shakespeare volume, a note that answered an earlier note by Wordsworth and that addressed Coleridge's infant son. The interesting consequence that Ferrer draws is that only hypertext can do justice to the multiple contexts annotations have. This is especially true in the

case of a writer like Stendhal, who dated his marginal notes and often revisited them, making further notes. A paper-based edition has to choose between date and textual context as an ordering principle. Hypertext can even show how the margin gradually fills up.

Reading notes are not scholarly annotations, but they occupy a similar position. Being located between a primary and a secondary text, they can illuminate both. They define pathways from source text to result and back. That can still be true when the writer who uses the notes is not the same as the one that wrote them down, as in the case of the printed commonplace books of the Renaissance. The classroom practice of culling quotations out of classical authors and transcribing them into a private commonplace book gave rise to large printed collections of commonplaces (Moss 1996). To facilitate retrieval by authors in need of appropriate quotations, these collections were structured by elaborate systems of subject headers. Emblem writers used the commonplace books as sources (Bath 1994: 34) and emblem books could function as a commonplace book (López Poza 2000). Luijten (1996) and Bloemendal (2002) have shown how many of the quotations in Dutch emblem books were taken from Langius' *Nova Polyanthea* (1604). It is important for these intermediate works to be accessible on the web. The group around López Poza has created a site that contains a number of works in their entirety (López Poza 2004-2009).[191] This opens the interesting possibility of hyperlinking the quotations as they occur in online emblem books to their potential sources in the digitised commonplace books.

13.3.2 *Showing and Aggregating Notes in the Scholarly Digital Edition*

The next example comes from the work on the EDITOR annotation toolset that I reported on in chapter 6 and the follow-up efforts towards Scholarly Annotation Exchange (SANE) discussed in chapter 7. EDITOR is a tool for the creation of annotations on XML-encoded digital editions. SANE is a protocol and a prototype implementation for exchanging annotation information between annotation tools and annotation clients. SANE might facilitate the actual presentation of the EDITOR-created annotations in the context of a web-based scholarly edition.

EDITOR facilitates the creation of annotations that are organised in annotation sets. Annotations correspond to annotation types that can consist of multiple fields, and these fields can have multiple types: a memo, a hyperlink, a categorisation, etc. Using the SANE prototype, a digital edition can request EDITOR-created annotation information from an annotation server. The edition might display available annotation sets; the user may select an annotation set and, e.g., ask for a pie chart with distribution information on a specific field.

---

[191] http://rosalia.dc.fi.udc.es/Poliantea/index.jsp.

Selecting a specific value could lead to a display of the annotated fragments. The user could then ask for the full annotation information.

However, this should be only the first half of the story. An annotation might say something about the word 'hem' being a third-person personal pronoun. It is the sort of annotation that by itself is really not that interesting. It is only interesting as part of a larger group of annotations; e.g., when being used to make a point about the nature of the speaking persona in a book. The second half of the story would be a way to point to groups of annotations from the secondary literature that discusses them. In their articles, scholars should be able to include something like the pie chart shown in Chapter 7 (Figure 17), and that pie chart should be 'live'. Just as in the edition, its slices should be hyperlinked to a display of the annotations represented by that slice.

The three main conditions for this type of annotation display to be possible are: (1) the annotations must refer to precise locations in the primary text in order for this to work properly; (2) the secondary text must refer to the annotations and to specific annotation displays; and (3) the annotations must be based on annotation types for the high-level displays (the diagrams and filtered views) to be possible. The annotation types, with their constituent fields and values, provide the structure that makes the annotations amenable to filtering and more generally manipulable. The annotation types, as suggested earlier, form a model of the phenomenon being studied, and the annotations provide an application of that model to the work being investigated.

13.3.3   Hypertext Generated from a Semantic Model

The notion that the annotation types provide a model that structures annotations is attractive, because of McCarty's claim that modelling is really the proper business of humanities computing. Annotations then are not something marginal; they have their place at the centre of what computers can do for the humanities.

Though model-making is implicit in much of humanities scholarship – indeed, as McCarty notes (2004: 255), any well-articulated idea would qualify as a model of its subject – what humanities computing adds is, as we have seen, the demand for computational tractability and manipulability. That implies the need for formal definition of the models. The quality, subtlety and richness of the model are, of course, largely functions of the insight that the researcher has into the phenomena and works that he or she studies. But the potential expressiveness is determined by the metamodel: the abstract types of data and relations that our modelling language allows us to express. The next example of 'annotations in action' is based on the richer metamodel of the index of metaphor in *Amoris Divini Emblemata* discussed in chapter 9. In this example the boundaries between annotations, primary and secondary texts will be harder to draw than in the previous example.

As we saw in Figure 33, the metaphor index opens with a view of an article-like text, a narrative (including things like abstract and bibliography), and a search interface into the annotations about metaphor. The user can choose to open a table of contents. The article text contains hyperlinks, and these links are either towards search queries on the annotations, towards emblems, or to elements of the ontology that underlies the metaphor descriptions detailed in the annotations. Clicking a search for plant life will bring up a list of hits (Figure 34) , and clicking on a hit will display the relevant emblem with the particular sample of plant life highlighted (Figure 35). The system will mention the other metaphors and metaphor-like structures described for this emblem, again hyperlinked to queries for that metaphor, but also provided with hyperlinks to a display of information about the relevant class in the ontology (Figure 36).

What is noticeable about the metaphor index is that at surface level it really is a single hypertext - from a user's perspective it is hard to see where the narrative stops and the mesotext begins, or where the mesotext ends and the primary text begins. The surface hypertext however is a derivative product. It is not really authored.[192] It is generated behind the scenes from a technical representation of the narrative, the mesotext and the primary texts.[193] In the metaphor index the primary text is no longer the point of reference. In the annotated edition discussed before, the edition was the point of reference for the annotations: even though the annotations could be used to select which fragments of the edition to view, it was still the edition that was being studied with the help of the annotations. In the metaphor index, the narrative is the natural point of departure for the reader.

13.3.4   *Trends in Science Publishing*

Related ideas have motivated a number of initiatives in science publishing. Examples are the 'scientific publication packages' that have been discussed in the field of digital curation, 'datuments' in molecular science and 'open notebook science' in chemistry.

Digital curation has been defined by Van Horik as 'the maintenance of research materials, such as research data over their entire life circle and over time for current and future generations of users' (2007: 132). One of the central concerns of digital curation is the accessibility and preservation of the scientific data that underlie scientific publications. Hunter (2006) has proposed the creation of 'scientific publication packages' that should include raw data, visualisations, tables, spreadsheets and article texts. This approach is similar to the mesotext approach in the sense that research data are bundled with the

---

[192] This indirect authoring has been described by Obendorf (2006) as necessary for the creation of successful hypertexts for the web.
[193] In the terminology of Prätor (2007), the topology of the underlying texts makes possible the navigational facilities at the surface.

scholarly articles or narrative, which does facilitate inspection of the underlying data by other researchers. It differs from the mesotext approach in that there is no attempt to actually hyperlink the research data from the text. An important element in the scientific publication packages are workflow descriptions that describe research steps taken during the investigation, which may be needed to replicate the results. The packages should also document provenance and lineage of its various components. This will help in assigning academic credit to the multiple contributors to a package.[194]

A similar approach is that of the 'datuments' introduced by Murray-Rust & Rzepa (2004). Datuments contain machine-readable research data as part of the text. The motivation for the datument is that today's publishing industry's *lingua franca*, the PDF document, is acceptable for human readership but in fact it is very much a dumbed-down version of the research material underlying the text. Given the quantity of research articles in the field that the authors discuss, it is imperative that software can be used to inspect and archive the molecular formulae of the compounds discussed in the articles. Including these data as snippets of Chemical Markup Language (CML) in the article files would make it possible for software robots to collect the data. More relevant from a mesotext perspective, a human reader could load the CML into a 3D molecule viewer that would facilitate viewing the molecule from various angles, highlighting the different elements, zooming in on parts of the structure, etc. The mesotext and datument concepts diverge when Murray-Rust & Rzepa insist that all of the underlying components are copied and aggregated at the time of publication. The datument is self-contained and not part of a larger hypertext. Their argument is that one cannot rely on individual components elsewhere in the network to remain accessible and the only safe approach is therefore to create complete snapshots.

Also in chemistry, Jean-Claude Bradley is advocating 'open notebook science' (2008a). The fundamental idea is that science is a collaborative process that benefits from being conducted in the open. Researchers who join the open notebook initiative maintain their laboratory notebooks on publicly accessible wiki pages and make available links to their datasets and results, which can be inspected by means of open-source viewers. In a further step, Bradley *et al.* (2008) hope to create tools that help make these notebooks and data machine-readable.

What these initiatives share with the concept of mesotext is that research data underlying scholarly publications are made accessible to readers. Inspection of research procedures by colleagues becomes much easier. Colleagues can do secondary analyses on the research data and reach different conclusions. One respect in which these initiatives differ from the mesotext

---

[194] The Object Reuse and Exchange specification (Lagoze & Van de Sompel 2008) can provide part of the infrastructure for scientific publication packages, as shown by Cheung *et al.* (2007).

concept is that in the sciences there is not something like the primary text of the humanities. All data result from human measurements, and they are not 'about' a given text that is being studied. But even in this respect, there may be an analogy. Yet another mesotext-like aspect of science publishing are the curated databases of annotations to e.g. the human genome, that Stein (2001) compares with annotation to the Torah. Stein argues annotations do not count as publications, yet they are a major part of scholarly output. To efficiently fill the annotation databases, scholarly papers would need to be provided with some sort of a structured abstract (something like a datument, probably) that unambiguously identifies the relevant data. In the case of genome annotation that would produce a full analogy to the mesotext concept: the genome as the primary text, research publications as secondary text, the annotation databases as mesodata, and the ontologies that structure these annotations as the model the mesodata conform to.

13.3.5   *Roland and the Need for Closure*

In the past, the interlinear glosses in a manuscript that translated Latin words into a vernacular equivalent developed into separate works, dictionaries, thereby creating a distance between the word and its gloss (Hüllen 1989). Today, we can restore the proximity between word and gloss by hyperlinking the dictionary and the primary text. What mesotext can do is similar: move the annotation out of the notes and appendices into a separate and structured body of observations, yet now, thanks to the possibilities of the electronic age, keeping it still readily available from both work and article.

Figure 77 provides another interesting example of such of a structured body of annotations, which I will use to address the problem of closure. It is the work of Vika Zafrin. Her subject is the tradition of the *Chanson de Roland*, from its earliest sources through its later incarnations, such as *Orlando Furioso* (Zafrin 2007).

On the left there is a list of scenes from the Roland stories. Selecting one of the scenes takes one to a text fragment (on the right) narrating that scene from one of the text traditions. In this fragment, based on the annotations, a number of objects can be highlighted: references to characters, to imagery and to themes. The available ones are listed in the centre. A mouse-over on one of these highlights the corresponding occurrences in the text fragment. Clicking in the list brings up (on the left of the screen) the list of scenes where that person appears or that theme occurs. Naturally, one can click on a reference in that list and have that scene appear, and then click on a theme that appears in that scene – and so on, ad infinitum.
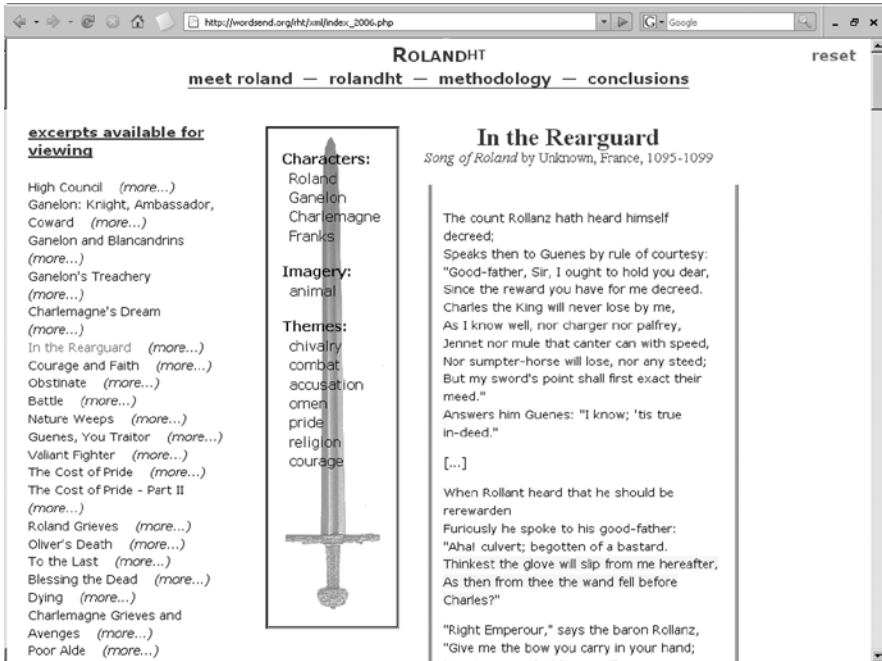
Figure 77 RolandHT, by Vika Zafrin: scenes, annotations, highlighted text fragments

It is this 'ad infinitum' that may be disconcerting. Are we never done with this? Merely amassing relevant material in a hypertext has been seen as an abdication of authorial responsibility (Atkinson *et al.* 2001: 463), that willy-nilly turns every reader into a reviewer (Smulyan 1999). Open-ended hypertext corpora, however, resist closure (Landow 2006: 302). Zafrin argues that the desire for closure is unproductive 'if the goal of reading is to acquire a broad contextual sense. (…) One can merely stop at a certain point, knowing there's always more to explore' (2006: 162). For Zafrin, the desire for closure is merely what she calls a 'function of our reading habits' (161). Against that view, I would argue that while reading, the way we process information depends among other things on our expectations of what we are going to find. The traditions of academic writing help us orient ourselves during the reading process and in fact enable us to reach some kind of conclusion about the worth of what we are reading. Scholarship knows no closure, but individual pieces of scholarship do and should offer at least provisional closure. 'Writers of hypertext cannot abdicate responsibility to lead' (Machan 1994: 303). We need to move on, after all. We do not read our indexes; we search them once we have a reason to do so. This argues for the presence of a narrative that can be seen not as closing, but as framing the mesotext; as providing context, orientation and motivation for exploring the mesodata (and Zafrin does in fact incorporate a number of critical essays in *Roland[HT]*). Unsworth writes 'instead of opening for the reader a single

path through a thicket of text, the critic can provide her with a map and a machete' (1997b). The narrative fulfills the functions of map and machete.

13.4   Conclusion

From the above approximations, we can identify a number of conditions for successful mesotext. Grouped by the components given in Figure 76 they are a follows.

With respect to the primary text:
1.   the primary text must be structured, so as to facilitate precise addressing of text fragments being annotated and to facilitate grouping and aggregation of annotations based on text structure (that is, there should be a text model and not just an annotation model).

With respect to the annotations:
2.   the annotations must refer to precise locations in the primary texts;
3.   the annotations should conform to the formal model.

With respect to the model:
4.   there should be a formalised model of the domain;
5.   the model should be a model of a single, coherent domain;
6.   the formal model should be a (partial) expression of the theory that the narrative espouses.

With respect to the narrative:
7.   the narrative should contain active links into the body of annotations: to individual annotations, groups of annotations, statistical computations done on annotations, charts of annotations, etc.

With respect to the system as a whole:
8.   systematic exploration should be possible from all sides – from the primary text, from the narrative, from the model and from the mesodata. One should be able to enter the mesotext either on the basis of the text unit of the primary text, or on the basis of the issues dealt with in the narrative, or on the basis of the types of annotations that have been defined, or finally by unguided browsing and searching;
9.   useful mesotext should be open: freely accessible and conforming to a public standard.

This last condition raises difficult issues about the infrastructure of scholarly publication. Obviously, the present infrastructure, where a PDF document is still something to be grateful for, is woefully inadequate for the exchange of

annotations. However, the cross-fertilizing effect that the conjunction of multiple bodies of annotations can have, will not occur unless we make these annotations, the texts, the articles and the models available in an open and publicly available format. SANE attempts to create conditions for open mesotext. The metaphor index however is a closed system (if it can be called a system), and so apparently is *Roland*^{HT}. More experimentation is clearly necessary in order to prepare the ground for annotation exchange. One of the biggest challenges will be the creation of models that are both sufficiently powerful to express significant insight, and still sufficiently flexible to adapt to changing circumstances and views. Models should be light-weight and open to extension, modification and reorganisation. They should be expressed using formalisms that are widely understood.

Now that the initial goals of the Emblem Project Utrecht have been met, it appears that creating useful annotation tools and devising ways to share the results is considerably harder to do than digitisation itself. However, the rise of Open Access and the concept of the 'Open source critical edition' (Bodard & Garcés 2009) bode well for sharing scholarly material and the related annotations.[195] The concept of mesotext, with McCarty's insistence on the importance of modelling, may help to create the conditions for successful exchange of annotations in the scholarly domain.

---

[195] The concept of the 'Information Commons' (Kranich 2004) and more specifically, the increasing popularity of path creation tools such as H2O playlists (Zittrain *et al.* 2006) and the Collex exhibition builder (Nowviskie 2005; Nowviskie & McGann 2005) are also good signs for scholarship based on digital primary material.

14   Conclusion

Many an emblematist has warned against trying to accomplish two things at once. Van Veen wrote 'Hee that two hares doth chase gets sildome anie one' (2006: 13). Disregarding this advice, in this thesis I have pursued multiple goals. My two main objectives have been to develop annotation tools for digital editions and to write about emblems. In this last chapter I will look at what has been accomplished. I will look at the results in emblem digitisation, in annotation and in emblem studies, and then suggest some future work.

14.1   Results in Emblem Digitisation

The Emblem Project Utrecht site is a well-visited site that is favourably commented upon on many web pages. It is used as supporting material in educational sites and encyclopaedia pages. Newer emblem sites have been to some extent modelled upon the Utrecht example. The Utrecht project actively participates in the international emblem digitisation community, and its emblem data have been harvested by the Open Emblem Portal. A number of articles have been published about the project, as was a book that has been favourably reviewed.[196]

In this thesis, chapter 4 has sketched a model of emblem digitisation sites. It characterised emblem sites in terms of a basic orientation, four generic functionalities, a number of parameters and technical infrastructure. Chapter 5 investigated how modelling plays a role in the design of a site such as that of the EPU.

As Peter Daly noted, prompted by Google's announcements that they would be scanning entire libraries, 'it is more important than ever to think about the needs of scholarship in this brave new world of virtual libraries' (2007: 198). Perhaps the most important aspect of the project is that it has helped us think about what emblem scholars need and what they can reasonably expect from digital editions. Google is not a charity. Even though many emblem books will become available over the coming years[197], Google will not take particular care

---

[196] Dietz (2007), Leerintveld (2008), Pierazzo (2008)

[197] Sagrario López Poza's catalog now lists 32 emblem books in Google books (2009). Many other books are available at various levels of sophistication in other facsimile-oriented digitisation initiatives.

of the specific needs of a small band of emblem scholars. It is up to the scholars themselves to do that, and the only way it can be done is through developing sites and seeing what works.[198]

14.2   Results in Annotation Technology

14.2.1   *Annotation's Promises*

In his interesting book about the markings sixteenth-century English readers made in the books they read, William Sherman devotes the afterword to a somewhat downbeat reflection on the future of reader's notes in a digital age (2008). Book digitisation succeeds in reproducing the book as an object, but as yet has not been very successful in facilitating the (often social) things that people want to do with books. In digital books we cannot leave marks of our reading and they do not inform us about marks left by others. This confirms the findings of the inspection of digital libraries and scholarly edition projects in chapter 3: they offer skimpy annotation facilities at best. Sherman ends his afterword with a call to scholars and digital designers to 'produce the kinds of tools that serve the readers of the past as well as the future'.

The main part of this book, part III, can be seen as a response to that request. It was devoted to developing annotation tools, and investigating their potential for digital emblem studies and the wider digital humanities. In chapter 3, I have given a wide definition of an annotation ('a resource brought to bear on another resource') and argued that our tools should facilitate the creation of sets of typed and structured annotations. For inspiration, I briefly looked at a number of annotation tools or approaches from outside the domain of literary studies. From these sources of inspiration, one can distil a number of promises (I summarise those mentioned in chapter 3, leaving aside those that I have not had occasion to work on):

1. computers can process annotations, and the effectiveness of this processing is proportional to the amount of structure that the annotations have;
2. one type of automated processing is selection, and selection facilitates quick and specific look-up of annotated resources;
3. corpora that are systematically annotated with large numbers of annotations, as in linguistics, are only feasible in a context of shared data and bode well for scholarship as a shared enterprise;
4. individual annotations can express this interpretation not just by content but also by conforming to a formal definition that is part of a text model;

---

[198] Facilities that seem unlikely to come about as a result of mass digitization initiatives: recognition of structural units in books as emblems, recognition of relations between emblems, recognition of linguistic parallelism between emblem fragments, recognition of objects in emblem pictures.

The discussion will refer back to these promises by number.

14.2.2   Discussion of Individual Approaches

The thesis contains work on three different approaches to structured annotation: the EDITOR/SANE approach, the semantic web approach and the Text Encoding Initiative's (TEI) feature structures approach. This subsection discusses each of these approaches individually and examines how they might contribute to fulfilling the promises mentioned before. The next subsection will compare and evaluate the approaches.

Not all aspects of the various approaches have received equal attention. For EDITOR, there is an annotation entry facility that is beyond the proof-of-concept stage. For the semantic web approach, attention has focused on display. For the feature structures approach, attention was devoted mainly to modelling issues. Not all of the promises mentioned above can therefore be usefully discussed for each of the approaches. All approaches however, should be judged as a way of tackling annotation's challenges rather than on their status as finished software (which none of them can lay claim to).

EDITOR and SANE

EDITOR facilitates the annotation of XML-based digital editions. Annotations are based on annotation types that consist of multiple fields. The edition XML is downloaded to a workstation and is annotated there. The process of annotation results in an annotation set that can be uploaded to a web server for an elementary web display. SANE is a proposal for an annotation exchange protocol and markup language. It should be possible to display annotations created by means of EDITOR-like tools in a digital edition using software that implements the SANE proposals. One of the fundamental ideas behind SANE is that annotation display should not be limited to textual display in the context of an annotated text fragment, but should include high-level annotation overviews such as charts.

What we have shown in EDITOR is how annotators can create annotation types containing multiple fields, and how these definitions can be used to create the forms that the user fills in to create annotations (4). The annotation overview window facilitates look-up of annotated passages (2). SANE is the component that should facilitate sharing the annotations with others (3). SANE was developed mainly for display of annotations in an online edition, but the verbs and parameters that it defines should also facilitate other uses (1).

Among the other strong points of EDITOR are the fundamental decision to keep the annotated text and the annotations separate; the possibility to annotate any fragment, from a single letter to an entire text; and the possibility to have overlapping annotations. SANE adds to that a flexible request protocol that includes the capacity to filter annotations by type, value or annotated location. This creates a possibility to display e.g. an edition table of contents that

includes information about the numbers of annotations of a certain type that the edition sections contain.

Probably the biggest obstacle to actual use of an EDITOR-like application is the assumption that there is an XML file which will be the annotation target, and the corresponding need to somehow display the XML to the annotator. As XML files, by design, contain no information about how they should be displayed, this requires the creation of a CSS stylesheet for each new (class of) XML edition. Another potential problem with locations in an XML file as annotation targets is that, if this edition is not the definitive edition of the edited work – and we know no edition is definitive – at some stage in the future the annotations will need to be transferred to the new edition. This is the problem that Dieter Köhler addresses in his APE toolkit (2006), by introducing the digital equivalent of the canonical citation schemes that have been in use for ages in citing e.g. classical authors or the Bible. The drawback of that approach is that it requires previous identification of potential annotation targets, and is therefore less suitable to annotating arbitrary stretches of text.

Semantic web approach
The most important aspect of the semantic web approach is that annotations are based on an ontology. The ontology describes a network of concepts with their relations and properties. Annotations are accessible as nodes in this network. The ontology can be used as a basis for integrated queries using complex conditions formulated in terms of the ontology's classes and properties. Semantic web technology can be used to handle the annotation data.

Here, individual annotations are expressed as instances of the classes defined at ontology level (4). That the properties and relations described at instance level conform to those defined at ontology level can be checked using OWL validating software. One of the properties defined in the ontology used in chapters 8 and 9 relates the annotation (here characterised as 'sign') to its sign vehicle, the text or image fragment that it is based in. This provides the basis for look-up functionality (2). There is in this approach no special facility for exchange of annotations with the wider world, but as the annotations are stored in RDF it should be easy enough for other applications to process the annotations for unanticipated purposes (1).

Some interesting points of the approach as sketched in this book are that queries in standard RDF query languages can be stored elsewhere and 'fired' from other texts, facilitating query execution in response to a hyperlink being clicked; the assisted creation of queries formulated in terms of the ontology; and the hierarchy of very general to more specific annotations, where display of the generic annotations will automatically include the more specific

ones. Of the approaches tried out in this book, the semantic web approach is the only one where I have tried to annotate not just text but also image regions.[199]

The most fundamental question to be asked about this approach of making annotations is probably whether it is worth the trouble. To give a formalised account of each interpretive step may be more tiresome than illuminating, if the aim is just that. Projects that use similar technology[200] are usually targeted at facilitating retrieval applications, where interpretive detail is not required. The technology could perhaps be applied in an educational setting, where a step-by-step methodology is useful. Maybe we need an 'interpretation browser', a general purpose application that will let users navigate any interpretation, from the minute detail of shades of meaning of individual words up to the critic's theoretical background informing the interpretation.

At a more practical level, a limitation of the approach as implemented here is that the annotated regions are indicated as such in the TEI XML file that stores the annotated text. As argued in chapter 5 and elsewhere, this will not work well when multiple people work on the same text or when the annotator does not have write access to the file. Future experiments will have to use a stand-off approach.

TEI feature structures approach
The feature structures approach is an attempt to model annotations using TEI feature structures. Feature structures originate in linguistics. They provide an abstract data structure that can be used to describe arbitrary features of text fragments. I have used feature structures in chapter 10 to be able to give a standard-based formulation for the annotations and annotation types created with EDITOR. This was done in the context of annotating the sort of parallel text structures that one often meets in the emblem corpus: series of translated or reworked emblems within or across books. A specific situation that this chapter tried to deal with (proposing a TEI proxy document as a possible answer) is the unavailability of an XML edition source file.

Experiments with this approach have not reached the stage of fulfilling the promises mentioned above, except for 4: the feature structure declaration defines the possible feature structure values. However, the use of something like SANE would make it possible to display these annotations in a digital edition, and to do many other things. An interesting question is exactly how and where in a digital edition we would display annotations that do not refer to a single text fragment, but to a fragment's relation with one or more other fragment(s).

One of the important aspects of this annotation solution would be the use of a public standard such as TEI. In TEI, feature structures may be the least frequently used facility and they are next to unknown outside linguistics. Feature

---

[199] There is no fundamental reason why this should be impossible in the other approaches.
[200] Such as those described by Ciula *et al.* (2008) and Hyvonen *et al.* (2007).

structures, unlike other TEI elements,[201] require application support at the encoding stage (Hockey 1996: 9). Feature structures introduce an extra layer of abstraction, and require the encoder to redo the modelling work already done by the TEI designers for conventional elements. As noted, it would be an interesting exercise to express the abstract model of conventional TEI elements such as <person> in terms of feature structures. As a minimum, to be useful in literary annotation, the data types allowed for features should be expanded to include date, URL and rich text. Feature structures do however provide a powerful mechanism to define and validate annotation structures.

14.2.3   Comparison of the Approaches

In all three methods of scholarly annotation of literary material it is possible to create a model of textual phenomena and to describe the phenomena in terms of the model. The approaches that I have essayed differ in a number of respects, but not all of these differences follow from essential characteristics of the approaches. I have, e.g., not created an application that supports data entry for the metaphor index discussed in chapter 9, but there is nothing that would make such an application impossible.

The two fundamental differences between on the one hand the EDITOR, SANE and the feature structures and on the other hand the semantic web approach are that in the latter case there is indeed a 'web' of interrelated annotations, and that this web can be queried with RDF query languages. The characteristics follow from the existence of an ontology that the annotations are based on. As the annotations acquire their meaning from being part of a network, they lose their meaning when displayed out of context. For the semantic web approach an exchange facility such as SANE makes little sense. What does make sense is to walk through the network using an appropriate query language.

When it comes to deciding when to use which type of tool, the decision should probably depend not on the research subject, but rather on envisaged modes of access to the created annotations. In this book I explored the possibilities of the semantic web approach in my work on a metaphor index, and the possibilities of EDITOR in work on theatrical aspects of the emblem. I do not think however that either subject could not have been researched using the other tool. As we also saw in chapter 13, the natural way to access output of ontology-based annotation is from a secondary scholarly text that motivates the ontology and its application; the annotations created with EDITOR are most naturally accessed from the edition that they annotate. More general considerations may be that EDITOR-type tools will probably be somewhat easier to use and will not require getting familiar with ontologies. In the absence of

---

[201] Depending on the technical know-how of the encoders.

finished and ready to use tools, however, these considerations can only be preliminary.

14.2.4   Assessment

Much of the inspiration for the work on this book came from John Bradley's paper 'Finding a Middle Ground between 'Determinism' and 'Aesthetic Indeterminacy': a Model for Text Analysis Tools' (2003). Bradley's central concern is to have the computer assist in the creation of the critic's conceptual model of the text. Since writing that paper, Bradley went on to develop the annotation tool Pliny, which to a large extent realises the program that he outlined in his paper (2008c, 2008b). The tools that I have been working on can be seen as facilitating some of the same aims. It may be opportune, before ending this book, to ask what the use of tools like these can mean for literary studies. Do they require a more data-oriented, more 'scientific' approach in the study of literature? Do they favour a specific method or orientation? If so, is this desirable? These questions came up in the discussion of the qualitative data analysis tools that social scientists use in chapter 3.

Susan Hockey has formulated what is perhaps the classic position on what computers can do in literary studies: '[Computer-based tools] can provide concrete evidence to support or refute hypotheses or interpretation which have in the past been based on human reading and the somewhat serendipitous noting of interesting features' (2000: 66). The tools that Hockey writes about, however, are tools that basically find and count features or patterns in text. They instantiate the 'algorithmic-transformational' approach for which Bradley wants to find an alternative. Hockey's formulations certainly sounds like a change of method compared to traditional literary studies: 'concrete evidence', 'refute hypotheses', and the unfortunate suggestion that 'human reading' may be something belonging to the past. Still, scholarly judgment is needed 'to interpret the result within the context of other research'. Sculley & Pasanek (2008) stress the point that any outcome of applying a program to text is a new text, which in its turn calls for interpretation.

In a series of influential articles Ramsay (2003, 2005, 2007) has called for an algorithmic literary criticism where the computer generates 'readings' of texts, just like literary critics do: 'paraphrased, elaborated, selected, truncated and transduced' (2007: 490). The computer finds pattern, 'and from pattern, the critic may move to the grander rhetorical formations that constitute critical reading.' In an article hat reports on his experiences developing the StageGraph program (a program that creates graphs of scene locations in Shakespeare) Ramsay responds to the Hockey's statements quoted above: his graphs do not provide 'concrete evidence', the word 'refute' is inappropriate in a literary context, and Ramsay is doing the reading with the help of the computer. 'To prove something (...) wasn't (and isn't) the goal. The goal is to say something new, provocative, noteworthy, challenging, inspiring—to put these texts back

into play as artefacts reconstituted by re-reading' (2005: 189). This vision of the computer's potential is related to the interpretive power of textual deformance as argued by McGann & Samuels (2004).

Even though computational literary criticism naturally comes with its own methods, these opinions do not suggest a revolutionary program, out to replace traditional methods of literary studies. *A fortiori* this will hold for Bradley's interpretive program. Bradley's simple model of the phases of scholarship looks like Figure 78 (2008b: 266). Pliny was developed to support work in these phases, not to propose a different model. The program supports all three stages: it helps in note-taking, in organising the notes into something like the mental model of the text mentioned earlier, and the developed model can be exported as a topic map to support presentation.
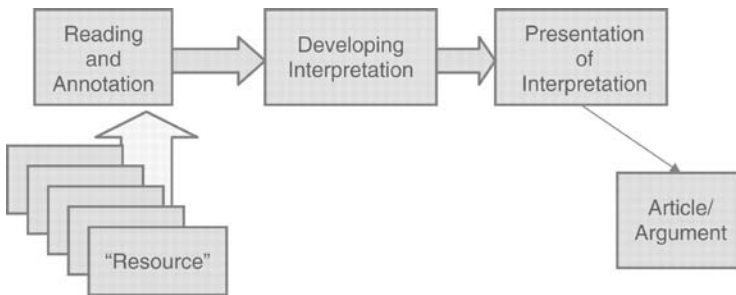
Figure 78 Phases of literary scholarship according to Bradley

The tools that I have been working on fit very well into this model. They differ from Pliny in a number of ways: (i) Pliny is explicitly pre-ontological. In Bradley's view, structure emerges and should not be imposed from the beginning. The tools that I have worked on use an explicit ontology (or annotation types) from the start; (ii) the tools that I have worked on are perhaps more than Pliny oriented towards publication: things like the metaphor index make no sense when not published, and EDITOR was conceived as a tool to annotate a publicly available digital edition.

I would say the tools I have been working on do not in any way make literary scholarship 'scientific', because scientific notions of knowledge are not applicable. What these tools bring about is that they aid precision, because interpretive statements can be linked to the locations that they discuss, and because an explicit definition of the categories employed in the interpretation helps to enforce consistency. This was discussed in the last section of chapter 8 and again in chapter 13. What I would also hope is that tools like these can contribute to changes in scholarly publication. Annotations illustrate the primary texts that they are about and the secondary texts that they support and they can fulfil a function in scholarly discourse. The concept of mesotext, introduced in the previous chapter, is meant to highlight that function.

14.3   Results in Emblem Studies

Much of the work on the thesis was motivated by the interest in emblem studies, and the final result contains two chapters of traditional emblem studies, both investigating aspects of *Amoris Divini Emblemata*. In chapter 11 I investigated the use of metaphor in the book. The chapter stated that shared background metaphors of landscape, water, plant life, light and verticality (to name some of the most salient) provided the setting for much of the foreground action in the book and helped make the book into a coherent whole. By embedding the metaphors in an allegorical setting, the book communicates attitudes towards the metaphorical objects. One of the metaphors that the book uses is that of the teacher-pupil relationship between Divine Love and the Soul, and in a sense the book 'acts out' that metaphor: the book is to the reader what Divine Love is to the Soul.

In chapter 12 I looked at the emblem pictures in *Amoris Divini Emblemata* as theatre stills. In some respects Divine love and the Soul are like actors that stage a scene in front of an audience. Their actions and gestures, ostensibly often directed at each other, have to drive home the emblem's message and to draw the reader/spectator into the dramatic situation. The interaction between the protagonists may support but also overrule the meaning of the emblematic object. In other emblem books, a similar effect can occur in texts that use a dialogue form, as the speaking characters always address the reader while speaking to each other.

I believe that these chapters contribute to a deeper understanding of *Amoris Divini Emblemata*. They address a number of shared characteristics of the book's emblems that hitherto have received little attention. It is perhaps no coincidence that the subjects of these chapters primarily require attention to be devoted to the book itself,[202] and not to other works, as this is what the annotation facility in development supported. The facility would not provide much help for investigations that need access to many different sources. The availability of an annotation facility could in this way influence the type of research being undertaken. This is probably unavoidable and no cause for regret, but also a stimulus to develop annotation tools that can target any book or text that we need.

---

[202] As do those of alternative subjects that I considered at one stage or another: aspects of practical virtue vs. mysticism, ways of addressing the reader, systematic differences in meaning between the various components of the emblem.

14.4 Future Work

In the annotation experiments that I have done there have not been external participants. It is clear that to design a truly useful annotation facility more user input is necessary. The best environment to test this is a large and well-used digital library that has a significant portion of scholarly users. In the Netherlands, the digital library that best fits that description is DBNL, the Digital Library of Dutch Literature.[203] A large experiment with a number of user groups in the DBNL context would be very useful. Perhaps, unlike the several approaches tried out in this book, an experiment like this should start with simple annotation facilities, and should try to get a number of user groups to cooperate. What may help in jump-starting such an experiment is a working demo, such as the demo exhibit that Jerome McGann built to showcase the capabilities of the Collex digital exhibit builder (McGann s.a.).

The annotation facility to be used in such an experiment should have a number of characteristics. People should be able either to share annotations, or to keep their annotations to themselves. It should be possible to group annotations in sets. Other users should have the option of turning annotation display on or off. One of the most important requirements is that annotations should be accessible in a number of ways: not just as a footnote to the text they annotate, but perhaps as a mini-essay in its own right, or as an anonymous contribution in a bar chart, or as one entry in a series of search hits. A SANE-like facility should make it possible to refer to collections of annotations from outside of the DBNL. The use case that I have in mind for that is to make annotations accessible from on-line scholarly articles.

In this thesis, I have argued a number of assertions about annotation: (i) that a note-taking facility available in the immediate context of a digital text will be welcome to scholars, that they will indeed use it to take notes and that these notes will lead to scholarly output; (ii) that at least some of the notes taken in that facility will be useful to others than the note-taker: either in themselves, as commentary on the text they respond to, or as support of the text they have may have led to; (iii) that annotation structure, as it creates the possibility to model the phenomena one is interested in, and thus to create theories about these phenomena, enhances the scholarly value of these annotations. The proposed experiment should test the truth of these assertions.

Annotation is a practice with a long history behind it. It is almost as old as writing itself. It is only since the advent of hypertext, however, that we can create live connections between annotations and the texts they annotate or prepare. I have called this network of connections mesotext, and in this thesis I have sketched its significance.

---

[203] About DBNL, see Van Stipriaan (2001, 2003)

Appendices

Some documents that reflect essential design decisions are given here as appendices. They have been edited for readability. An electronic version of these and other documents is available at http://peterboot.nl/thesis/.

Appendix A. SANE RNG Schema

What follows is a tentative Relax NG schema for the SANE-ML markup language. See chapter 7.

```
<?xml version="1.0" encoding="UTF-8"?>
<grammar
  xmlns="http://relaxng.org/ns/structure/1.0"
  xmlns:a="http://relaxng.org/ns/compatibility/annotations/1.0"
  datatypeLibrary="http://www.w3.org/2001/XMLSchema-datatypes">
  <a:documentation>
    This schema proposes a description format for the output of text
    annotation tools. It is based on the format used in the EDITOR
    tools being developed at the Huygens Instituut but meant to be
    tool independent.
  </a:documentation>
    <start>
      <element name="annotationSet">
        <attribute name="schemaVersion">
          <value>0.1</value>
        </attribute>
        <element name="setData">
          <a:documentation>Contains information about the annotation
            set.</a:documentation>
          <ref name="setDataContent"/>
        </element>
        <element name="editionData">
          <a:documentation>Contains information about the editions
            being annotated by this set.</a:documentation>
          <ref name="editionDataContent"/>
        </element>
        <element name="typeData">
        <a:documentation>Contains information about the annotation
          types in the annotation set.</a:documentation>
          <ref name="typeDataContent"/>
        </element>
        <element name="annotationData">
        <a:documentation>Contains information about the
          annotations.</a:documentation>
          <ref name="annotationDataContent"/>
        </element>
      </element>
    </start>
```

```
<define name="annotationDataContent">
  <zeroOrMore>
    <element name="annotation">
      <a:documentation>Each annotation corresponds to a single
       annotation element.</a:documentation>
      <attribute name="xml:id">
        <a:documentation>Annotations need an identifying attribute
          as it should be possible to point to
          them from other annotations.</a:documentation>
        <data type="ID"/></attribute>
      <attribute name="editionRef">
        <a:documentation>Refers to the edition using its url
          attribute.</a:documentation>
        <data type="anyURI"></data>
      </attribute>
      <attribute name="typeRef">
        <a:documentation>Refers to the annotation type that types
          this annotation.</a:documentation>
        <data type="anyURI"></data>
      </attribute>
      <ref name="accounting"/>
      <attribute name="startsAtPath">
        <a:documentation>At present, a simple xpath tree
          expression, eg.
          /TEI.2[1]/text[1]/group[1]/text[6].</a:documentation>
        <text/>
      </attribute>
      <attribute name="startsAtOffset">
        <a:documentation>Starting point of the annotation. Starts
          from zero. Zero if annotation is on full element. The
          offset is an offset within the text nodes of the element
          identified by the startsAtPath attribute, and does not
          consider text nodes at lower levels.
        </a:documentation>
        <data type="integer"/></attribute>
      <attribute name="endsAtPath">
        <a:documentation>At present, a simple xpath tree
          expression, eg.
          /TEI.2[1]/text[1]/group[1]/text[6].</a:documentation>
        <text/></attribute>
      <attribute name="endsAtOffset">
        <a:documentation>End point of the annotation. Starts from
          zero. -1 if annotation is on full
          element.</a:documentation>
        <data type="integer"/></attribute>
      <optional>
        <element name="selectedText">
          <a:documentation>For debugging, optimisation or other
            purposes, the annotation creation tool may provide (a
            fragment of) the annotated
            text.</a:documentation>
          <text/></element>
      </optional>
      <oneOrMore>
        <element name="annotationField">
          <a:documentation>Annotations can consist of multiple
            fields.</a:documentation>
          <attribute name="typeFieldRef">
            <a:documentation>Refers to the field that types this
              annotation field, defined at the annotation type
              level.</a:documentation>
            <data type="anyURI"></data>
          </attribute>
          <element name="value">
```

```
              <a:documentation>Value of the annotation
                field.</a:documentation>
              <text/></element>
          </element>
        </oneOrMore>
      </element>
    </zeroOrMore>
  </define>
  <define name="editionDataContent">
    <oneOrMore>
      <element name="edition">
        <attribute name="url">
          <a:documentation>Annotations use this url to refer to the
            edition.</a:documentation>
          <data type="anyURI"></data>
        </attribute>
        <attribute name="checksum">
          <a:documentation>May be used to check the integrity of the
            edition.
            [Needs extra info about algorithm sed?]</a:documentation>
        </attribute>
        <ref name="desc">
          <a:documentation>Used to give a brief description of the
            edition, friendly to the human eye. Suggested content:
            author, short title, year.</a:documentation>
        </ref>
      </element>
    </oneOrMore>
  </define>
  <define name="setDataContent">
    <attribute name="name">
      <text/>
    </attribute>
    <ref name="desc"/>
    <ref name="accounting"/>
    <optional>
      <element name="contactInfo">
        <a:documentation>Probably an email address of the set
          author.</a:documentation>
        <text/>
      </element>
    </optional>
    <element name="program">
      <a:documentation>Program used to create the
        annotations.</a:documentation>
      <text/></element>
    <element name="programVersion">
      <a:documentation>Program version used to create the
        annotations.</a:documentation>
      <text/></element>
    <element name="rights">
      <a:documentation>License that applies to this annotation
        set.</a:documentation>
      <text/></element>
    <zeroOrMore>
      <element name="setInfo">
        <a:documentation>Can be used either to describe the
          background and thinking behing the creation of the
          annotation set or to refer to an explanation elsewhere,
          possible in a scholarly article in a repository.
        </a:documentation>
        <oneOrMore>
          <element name="p">
            <a:documentation>Paragraphs of plain
```

```
                    text</a:documentation>
              <text/>
            </element>
            <element name="uri"><data type="anyURI"/></element>
          </oneOrMore>
        </element>
      </zeroOrMore>
      <optional>
        <element name="colorScheme">
          <a:documentation>Suggests color series to be used in
            displaying this annotation set.</a:documentation>
          <oneOrMore>
            <ref name="color"/>
          </oneOrMore>
        </element>
      </optional>
    </define>
    <define name="typeDataContent">
      <oneOrMore>
        <element name="annotationType">
          <attribute name="xml:id">
            <a:documentation>Use by annotations to refer to annotation
              type.</a:documentation>
            <data type="ID"/></attribute>
          <attribute name="name">
            <a:documentation>Should be unique within annotation
              set.</a:documentation>
          </attribute>
          <optional>
            <attribute name="selectedTextEditable">
              <a:documentation>Boolean value that determines whether
                the 'selected text' value of the annotation (see
                selectedTextDetermined attribute)
                can be modified by the researcher.</a:documentation>
              <data type="boolean"/></attribute>
          </optional>
          <optional>
            <attribute name="selectedTextDetermined">
              <a:documentation>Boolean value that determines whether
                the annotation creation application will store (a
                fragment of) the selected
                text.</a:documentation>
              <data type="boolean"/></attribute>
          </optional>
          <ref name="label"/>
          <ref name="desc"/>
          <optional>
            <ref name="color">
              <a:documentation>Annotation types may be assigned their
                own identifying color for consistent
                display.</a:documentation>
            </ref>
          </optional>
          <optional>
            <attribute name="highlighting">
              <a:documentation>Boolean value that suggests whether or
                not to highlight the annotated text fragment on
                display of the annotations. Typically false for an
                annotation type that is applied to larger
                textual units.</a:documentation>
              <data type="boolean"/>
            </attribute>
          </optional>
          <optional>
```

```
            <attribute name="highlightFieldRef">
              <a:documentation>When highlighting the annotated
                fragment is done by the value of one of its category
                fields, the highlightFieldRef attribute determines the
                default field to be used for that
                purpose.</a:documentation>
              <data type="anyURI"/></attribute>
        </optional>
        <oneOrMore>
          <element name="annotationTypeField">
            <a:documentation>Defines a field within an annotation
            type.</a:documentation>
            <attribute name="xml:id">
              <a:documentation>Used to refer to the type field from
                the annotation field that it
                describes.</a:documentation>
              <data type="ID"/></attribute>
            <attribute name="name">
              <a:documentation>Must be unique within the annotation
                type. Need not be unique within the annotation
                set.</a:documentation>
            </attribute>
            <attribute name="fieldType">
              <a:documentation>Data type of the field. Values:
                category (set of values), text (string), memo
                (multi-line text), URL and annotation
                (reference to another annotation by its
                id).</a:documentation>
              <choice>
                <value>Category</value> <!--   -->
                <value>Text</value> <!--  normalisedstring? -->
                <value>URL</value> <!-- anyURI -->
                <value>Memo</value> <!--  string? -->
                <value>Annotation</value> <!-- anyURI -->
              </choice>
            </attribute>
            <ref name="label"/>
            <ref name="desc"/>
            <zeroOrMore>
              <element name="fieldDefaultValue">
                <a:documentation>The values used for fields of
                  type category are stored at the type level. They
                  may be assigned a label, a description and an
                  identifying color.</a:documentation>
                <attribute name="value"/>
                <ref name="label"/>
                <ref name="desc"/>
                <optional>
                  <ref name="color"/>
                </optional>
              </element>
            </zeroOrMore>
          </element>
        </oneOrMore>
    </element>
  </oneOrMore>
</define>
<define name="desc">
  <optional>
    <element name="desc">
      <a:documentation>Description of an annotation set, type,
      field or value.</a:documentation>
      <text/>
    </element>
```

```
        </optional>
      </define>
      <define name="label">
        <optional>
          <attribute name="label">
            <a:documentation>Used as a brief label in presentations of
              annotation types, fields or default
              values.</a:documentation>
            <text/>
          </attribute>
        </optional>
      </define>
      <define name="accounting">
        <optional>
          <attribute name="createdOn">
            <data type="string"/>
          </attribute>
        </optional>
        <optional>
          <attribute name="changedOn">
            <data type="string"/>
          </attribute>
        </optional>
        <optional>
          <element name="author">
            <text/>
          </element>
        </optional>
      </define>
      <define name="color">
        <element name="color">
          <attribute name="value"><text/></attribute>
        </element>
      </define>
</grammar>
```

Appendix B. Decoding Emblem Semantics Ontology

What follows is the RDF-Schema ontology used in chapter 8.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
]>
<!-- rdf:resource expanded to generate correct triples -->
<rdf:RDF xml:base="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf"
xmlns:p01s="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdfs:Class rdf:ID="SignVehicle">
    <rdfs:comment xml:lang="en">Class of Sign Vehicles</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Frag">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicle"/>
    <rdfs:comment xml:lang="en">Class of Emblem Fragments</rdfs:comment>
  </rdfs:Class>
  <rdf:Property rdf:ID="AnnotProp">
    <rdfs:subPropertyOf rdf:resource="&rdfs;comment"/>
```

```
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Frag"/>
  </rdf:Property>
  <rdfs:Class rdf:ID="TextFrag">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Frag"/>
    <rdfs:comment xml:lang="en">Class of Text Fragments</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="ImgFrag">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Frag"/>
    <rdfs:comment xml:lang="en">Class of Image Fragments</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Sign">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicle"/>
    <rdfs:comment xml:lang="en">Class of Signs</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="ContentSign">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:comment xml:lang="en">Class of Literary signs</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="LinguisticSign">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:comment xml:lang="en">Class of Signs</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="LiterarySign">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:comment xml:lang="en">Class of Literary signs</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="StructDiv">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LiterarySign"/>
    <rdfs:comment xml:lang="en">Class of structural divisions of encoded
text</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="TEIbibl">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
    <rdfs:comment xml:lang="en">Class of TEI bibl elements (bibliographical
references)</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="TEIcit">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
    <rdfs:comment xml:lang="en">Class of TEI cit elements
(citations)</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="TEIdiv">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
    <rdfs:comment xml:lang="en">Class of TEI div elements (general-purpose
division)</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="TEIl">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
    <rdfs:comment xml:lang="en">Class of TEI l elements (metrical
lines)</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="TEIlg">
```

```
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
      <rdfs:comment xml:lang="en">Class of TEI lg elements (poems or stanzas in
poems)</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="TEIp">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
      <rdfs:comment xml:lang="en">Class of TEI p elements
(paragraphs)</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="TEIquote">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
      <rdfs:comment xml:lang="en">Class of TEI quote elements
(quotations)</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="TEItext">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#StructDiv"/>
      <rdfs:comment xml:lang="en">Class of TEI text elements
(texts)</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="MetaphysicalBeing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#ContentSign"/>
      <rdfs:comment xml:lang="en">Class of signs of metaphysical
beings</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="God">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#MetaphysicalBeing"/
>
      <rdfs:comment xml:lang="en">Class of signs of God</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="Deixis">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LinguisticSign"/>
      <rdfs:comment xml:lang="en">Class of deictic signs</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="Negation">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LinguisticSign"/>
      <rdfs:comment xml:lang="en">Class of signs of negation</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="SpatialThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#ContentSign"/>
      <rdfs:comment xml:lang="en">Class of Signs of Corporeal
Objects</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="SourceOfLightThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      <rdfs:comment xml:lang="en">Class of signs of light
sources</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="LivingBeing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      <rdfs:comment xml:lang="en">Class of Signs of Living
Beings</rdfs:comment>
    </rdfs:Class>
    <rdfs:Class rdf:ID="SunflowerLivingBeing">
```

```
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LivingBeing"/>
      <rdfs: comment xml: lang="en">Class of Signs of Sunflowers</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="BodyPart">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#SpatialThing"/>
      <rdfs: comment xml: lang="en">Class of Signs of Bodyparts</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="WingBodyPart">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#BodyPart"/>
      <rdfs: comment xml: lang="en">Class of Signs of Wings</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="HandBodyPart">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#BodyPart"/>
      <rdfs: comment xml: lang="en">Class of signs of hands</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="LikePerson">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LivingBeing"/>
      <rdfs: comment xml: lang="en">Class of Signs of Person-like
beings</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="YoungPerson">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LikePerson"/>
      <rdfs: comment xml: lang="en">Class of signs of young
persons</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="AuthorityPerson">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LikePerson"/>
      <rdfs: comment xml: lang="en">Class of signs of persons with
authority</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="Augustine">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#AuthorityPerson"/>
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#MalePerson"/>
      <rdfs: comment xml: lang="en">Class of signs of Augustine</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="MalePerson">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LikePerson"/>
      <rdfs: comment xml: lang="en">Class of signs of male persons</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="FemalePerson">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LikePerson"/>
      <rdfs: comment xml: lang="en">Class of signs of female
persons</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="AllegoryPerson">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#LikePerson"/>
      <rdfs: comment xml: lang="en">Class of signs of allegorical
persons</rdfs: comment>
   </rdfs: Class>
  <rdfs: Class rdf: ID="Girl">
      <rdfs: subClassOf
rdf: resource="http://www. xs4all. nl/~pboot/p01/p01rdfs. rdf#FemalePerson"/>
```

```
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#YoungPerson"/>
      <rdfs:comment xml:lang="en">Class of Signs of Girls</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="WingedGirl">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Girl"/>
      <rdfs:comment xml:lang="en">Class of Signs of Winged Girls</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="Soul">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#FemalePerson"/>
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#AllegoryPerson"/>
      <rdfs:comment xml:lang="en">Class of Signs of Soul</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="Cupid">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#MalePerson"/>
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#AllegoryPerson"/>
      <rdfs:comment xml:lang="en">Class of Signs of Cupid</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="DivineLove">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#MalePerson"/>
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#AllegoryPerson"/>
      <rdfs:comment xml:lang="en">Class of Signs of Divine Love</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="Reader">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LikePerson"/>
      <rdfs:comment xml:lang="en">Class of Signs of the reader</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="NonLivingThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      <rdfs:comment xml:lang="en">Class of Signs of Non-living
Things</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="ArchThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
      <rdfs:comment xml:lang="en">Class of signs of arches</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="BowThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
      <rdfs:comment xml:lang="en">Class of signs of bows</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="BridgeThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
      <rdfs:comment xml:lang="en">Class of signs of bridges</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="ChurchThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
      <rdfs:comment xml:lang="en">class of signs of haloes</rdfs:comment>
   </rdfs:Class>
   <rdfs:Class rdf:ID="ChurchSpireThing">
      <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
```

```
    <rdfs:comment xml:lang="en">Class of Signs of Halo</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="DressThing">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
    <rdfs:comment xml:lang="en">Class of signs of clothes</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="HaloThing">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SourceOfLightThing"
/>
    <rdfs:comment xml:lang="en">Class of Signs of Halo</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="QuiverThing">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
    <rdfs:comment xml:lang="en">Class of signs of quivers</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="RiverThing">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#NonLivingThing"/>
    <rdfs:comment xml:lang="en">Class of signs of rivers</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="SunThing">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SourceOfLightThing"
/>
    <rdfs:comment xml:lang="en">Class of Signs of Sun</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Action">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#ContentSign"/>
    <rdfs:comment xml:lang="en">Class of Actions (In this model, events and
states of affairs are actions).</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="AssertionAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of Actions consisting in actor
asserting object</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="BeDirectedAtAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of Actions consisting in being directed
at somebody/someone</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="ConformToAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of Actions consisting person (actor)
conforming to behaviour of other (object)</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="CreationAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of Actions consisting person (actor)
creating object (object)</rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="DrawAttentionToAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
```

```
      <rdfs:comment xml:lang="en">Class of events consisting in one person
(actor) drwaing attention to an object (object) for another person
(beneficiary)</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="EndorsingAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of events consisting in actor endorsing
object</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="GuidingAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of events consisting in one person
guiding another</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="LookAtAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#BeDirectedAtAction"
/>
    <rdfs:comment xml:lang="en">Class of Looking-At-Actions</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="LoveAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of Love Actions</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="PointingAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#DrawAttentionToActi
on"/>
    <rdfs:comment xml:lang="en">Class of events consisting in one person
(actor) pointing out an object (object) to another person
(beneficiary)</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="ProtectingAction">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:comment xml:lang="en">Class of events consisting in one person
protecting another</rdfs:comment>
    </rdfs:Class>
  <rdfs:Class rdf:ID="Metaphor">
    <rdfs:subClassOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LiterarySign"/>
    <rdfs:comment xml:lang="en">Class of Metaphors</rdfs:comment>
    </rdf:Class>
  <rdf:Property rdf:ID="SignVehicleProp">
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicle"/>
    </rdf:Property>
  <rdf:Property rdf:ID="SignMeaningProp">
    <rdfs:subPropertyOf rdf:resource="&rdfs;comment"/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    </rdf:Property>
  <rdf:Property rdf:ID="RoleProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicleProp"/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
```

```
      </rdf:Property>
    <rdf:Property rdf:ID="ActorProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#RoleProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
      </rdf:Property>
    <rdf:Property rdf:ID="BeneficiaryProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#RoleProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
      </rdf:Property>
    <rdf:Property rdf:ID="DeixisTargProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicleProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Deixis"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicle"/>
      </rdf:Property>
    <rdf:Property rdf:ID="ObjectProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#RoleProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Action"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
      </rdf:Property>
    <rdf:Property rdf:ID="SpatialProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicleProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      </rdf:Property>
    <rdf:Property rdf:ID="HasPartProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
      </rdf:Property>
    <rdf:Property rdf:ID="HasBodyPartProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#HasPartProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LikePerson"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#BodyPart"/>
      </rdf:Property>
    <rdf:Property rdf:ID="FixedAttributeProp">
      <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LiteraryProp"/>
      <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#AllegoryPerson"/>
      <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SpatialThing"/>
```

```
    </rdf:Property>
  <rdf:Property rdf:ID="IsGeneralisationOfProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#InferenceBasisProp"
/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
  </rdf:Property>
  <rdf:Property rdf:ID="LiteraryProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicleProp"/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
  </rdf:Property>
  <rdf:Property rdf:ID="MetaphorVehicleProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LiteraryProp"/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Metaphor"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
  </rdf:Property>
  <rdf:Property rdf:ID="MetaphorTenorProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#LiteraryProp"/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Metaphor"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
  </rdf:Property>
  <rdf:Property rdf:ID="InferenceBasisProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#SignVehicleProp"/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
  </rdf:Property>
  <rdf:Property rdf:ID="IsSpecialisationOfProp">
    <rdfs:subPropertyOf
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#InferenceBasisProp"
/>
    <rdfs:domain
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
    <rdfs:range
rdf:resource="http://www.xs4all.nl/~pboot/p01/p01rdfs.rdf#Sign"/>
  </rdf:Property>
</rdf:RDF>
```

Appendix C. Metaphor Index Ontology

The metaphor index discussed in chapter 9 is based on an OWL ontology.
Technically, the ontology consists of a general ontology which defines a general
framework for discussing signs in emblem books, and an ontology specially

targeted at the domain of metaphor. I include here the general ontology and a
fragment of the metaphor ontology.

The sign ontology:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY owl  "http://www.w3.org/2002/07/owl#" >
  <!ENTITY xsd  "http://www.w3.org/2001/XMLSchema#" >
]>
<rdf:RDF
  xml:base="http://emblems.let.uu.nl/lab/dse/rdf/p06owltop.rdf"
  xmlns:p06owltop="http://emblems.let.uu.nl/lab/dse/rdf/p06owltop.rdf#"
  xmlns:p06owl="http://emblems.let.uu.nl/lab/dse/rdf/p06owl.rdf#"
  xmlns:owl = "http://www.w3.org/2002/07/owl#"
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs= "http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd = "http://www.w3.org/2001/XMLSchema#">


  <owl:Ontology rdf:about="">
    <rdfs:comment>OWL basic ontology for Decoding Significance in
      Emblems</rdfs:comment>
    <rdfs:comment>Emblem Sign Ontology</rdfs:comment>
    <owl:versionInfo>2007-04-04</owl:versionInfo>
    </owl:Ontology>
  <owl:Class rdf:ID="SignVehicle">
    <rdfs:label xml:lang="en">Class of Sign Vehicles</rdfs:label>
    <rdfs:comment xml:lang="en">Everything can be provided with
      interpretation. Therefore, everything can become a sign vehicle.
      Therefore, the sign vehicle is at the top of the ontology.
      In our study of emblems sign vehicles proper are the fragments (Frag)
      of text or image that we attach meaning to. In a wider sense, signs too
      are sign vehicles. Signs are interpreted sign vehicles, and may in
      their turn become a vehicle for further interpretations, and thus for
      further signs. Because everything can be provided with interpretation,
      in our domain the class of sign vehicles is equivalent to
      owl:Thing. </rdfs:comment>
    <owl:equivalentClass rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
    <owl:equivalentClass>
      <owl:Class>
        <owl:unionOf rdf:parseType="Collection">
         <owl:Class rdf:about="#Sign" />
         <owl:Class rdf:about="#Frag" />
        </owl:unionOf>
      </owl:Class>
    </owl:equivalentClass>
  </owl:Class>
  <owl:Class rdf:ID="Frag">
    <rdfs:subClassOf rdf:resource="#SignVehicle"/>
    <rdfs:label xml:lang="en">Class of Emblem Fragments</rdfs:label>
    <rdfs:comment xml:lang="en">Fragments are the basic sign vehicles in the
     interpretational ontology. They can be either text or image fragments
     (TextFrag or ImgFrag).</rdfs:comment>
    <owl:equivalentClass>
      <owl:Class>
       <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#TextFrag" />
        <owl:Class rdf:about="#ImgFrag" />
       </owl:unionOf>
      </owl:Class>
```

```
      </owl:equivalentClass>
      <owl:disjointWith rdf:resource="#Sign" />
   </owl:Class>
   <owl:Class rdf:ID="TextFrag">
      <rdfs:subClassOf rdf:resource="#Frag"/>
      <rdfs:label xml:lang="en">Class of Text Fragments</rdfs:label>
      <rdfs:comment xml:lang="en">Text fragments are fragments. The class of
         text fragments is disjoint with the class of image
         fragments.</rdfs:comment>
      <owl:disjointWith rdf:resource="#ImgFrag" />
   </owl:Class>
   <owl:Class rdf:ID="ImgFrag">
      <rdfs:subClassOf rdf:resource="#Frag"/>
      <rdfs:label xml:lang="en">Class of Image Fragments</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Sign">
      <rdfs:subClassOf rdf:resource="#SignVehicle"/>
      <rdfs:label xml:lang="en">Class of Signs</rdfs:label>
      <rdfs:comment xml:lang="en">Signs are sign vehicles, provided with an
         interpretation. They can in their turn be interpreted. The
         SignVehicleProp connects signs and their vehicles.</rdfs:comment>
      <owl:disjointWith rdf:resource="#Frag" />
   </owl:Class>
   <owl:Class rdf:ID="StructDiv">
      <rdfs:subClassOf rdf:resource="#Sign"/>
      <rdfs:label xml:lang="en">Class of structural divisions of encoded
         text</rdfs:label>
      <rdfs:comment xml:lang="en">Structural divisions of text are not just
         fragments. Structural divisions are fragments that are recognised as
         having a structural function. Structural divisions are
         signs.</rdfs:comment>
   </owl:Class>
   <owl:Class rdf:ID="Emblem">
      <rdfs:subClassOf rdf:resource="#StructDiv"/>
      <rdfs:label xml:lang="en">Emblems</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Emblem_Text">
      <rdfs:subClassOf rdf:resource="#StructDiv"/>
      <rdfs:label xml:lang="en">Class of motto emblem constituents</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Motto">
      <rdfs:subClassOf rdf:resource="#Emblem_Text"/>
      <rdfs:label xml:lang="en">Class of motto emblem constituents</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Subscription">
      <rdfs:subClassOf rdf:resource="#Emblem_Text"/>
      <rdfs:label xml:lang="en">Class of quote emblem constituents</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Subscription_Quote">
      <rdfs:subClassOf rdf:resource="#Subscription"/>
      <rdfs:label xml:lang="en">Class of quote emblem constituents</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Subscription_Epigram">
      <rdfs:subClassOf rdf:resource="#Subscription"/>
      <rdfs:label xml:lang="en">Class of epigram emblem
      constituents</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Subscription_EpigramSpanish">
      <rdfs:subClassOf rdf:resource="#Subscription_Epigram"/>
      <rdfs:label xml:lang="en">Class of epigram emblem
         Constituents</rdfs:label>
   </owl:Class>
   <owl:Class rdf:ID="Subscription_EpigramDutch">
      <rdfs:subClassOf rdf:resource="#Subscription_Epigram"/>
```

```
    <rdfs:label xml:lang="en">Class of epigram emblem
      Constituents</rdfs:label>
  </owl:Class>
  <owl:Class rdf:ID="Subscription_EpigramFrench">
    <rdfs:subClassOf rdf:resource="#Subscription_Epigram"/>
    <rdfs:label xml:lang="en">Class of epigram emblem
      constituents</rdfs:label>
  </owl:Class>
  <owl:Class rdf:ID="Pictura">
    <rdfs:subClassOf rdf:resource="#StructDiv"/>
    <rdfs:label xml:lang="en">Class of pictura emblem
      constituents</rdfs:label>
  </owl:Class>
  <owl:Class rdf:ID="ContentSign">
    <rdfs:subClassOf rdf:resource="#Sign"/>
    <rdfs:label xml:lang="en">Groups the signs that relate to content rather
      than to emblem structure</rdfs:label>
  </owl:Class>
  <owl:ObjectProperty rdf:ID="SignVehicleProp">
    <rdfs:label>Sign vehicle</rdfs:label>
    <rdfs:domain  rdf:resource="#Sign"/>
    <rdfs:range rdf:resource="#SignVehicle"/>
    <rdf:type rdf:resource="&owl;TransitiveProperty" />
    <rdfs:comment>The SignVehicleProp connects signs to their vehicles, that
      is, their signifiers. SignVehicleProp is, in ontology terms, a
      transitive property: is a sign is based on another sign, the basic
      sign's signifiers are also signifiers of the sign based on
      it.</rdfs:comment>
  </owl:ObjectProperty>
  <owl:DatatypeProperty rdf:ID="SignMeaningProp">
    <rdfs:label>Meaning</rdfs:label>
    <rdfs:domain  rdf:resource="#Sign"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
    <rdfs:comment>SignMeaningProp briefly describes the meaning of a sign.
      The description is not supposed to be exhaustive.</rdfs:comment>
  </owl:DatatypeProperty>
  <owl:ObjectProperty rdf:ID="BasedInProp">
    <rdfs:label>Based in</rdfs:label>
    <rdfs:domain rdf:resource="#Sign"/>
    <rdfs:range  rdf:resource="#StructDiv"/>
    <owl:inverseOf rdf:resource="#IsBaseForProp"/>
  </owl:ObjectProperty>
    <owl:ObjectProperty rdf:ID="IsBaseForProp">
      <owl:inverseOf rdf:resource="#BasedInProp"/>
      <rdfs:label>Is base for</rdfs:label>
      <rdfs:range  rdf:resource="#Sign"/>
      <rdfs:domain  rdf:resource="#StructDiv"/>
    </owl:ObjectProperty>
    <owl:ObjectProperty rdf:ID="ContainsProp">
    <rdfs:label>Contains</rdfs:label>
    <rdfs:domain  rdf:resource="#StructDiv"/>
      <rdfs:range  rdf:resource="#SignVehicle"/>
      <owl:inverseOf rdf:resource="#ContainedInProp"/>
    </owl:ObjectProperty>
    <owl:ObjectProperty rdf:ID="ContainedInProp">
      <rdfs:label>Contains</rdfs:label>
      <rdfs:domain  rdf:resource="#SignVehicle"/>
      <rdfs:range  rdf:resource="#StructDiv"/>
      <owl:inverseOf rdf:resource="#ContainsProp"/>
    </owl:ObjectProperty>
    <owl:DatatypeProperty rdf:ID="ThumbProp">
    <rdfs:label>Thumbnail</rdfs:label>
    <rdfs:domain  rdf:resource="#Emblem"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
```

```
      <rdf:type rdf:resource="&owl;FunctionalProperty" />
    </owl:DatatypeProperty>
</rdf:RDF>
```

A fragment of the metaphor ontology:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY p06owltop "http://emblems.let.uu.nl/lab/dse/rdf/p06owltop.rdf#" >
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY owl  "http://www.w3.org/2002/07/owl#" >
  <!ENTITY xsd  "http://www.w3.org/2001/XMLSchema#" >
]>
<rdf:RDF
  xml:base="http://emblems.let.uu.nl/lab/dse/rdf/p06owl.rdf"
  xmlns:p06owltop="http://emblems.let.uu.nl/lab/dse/rdf/p06owltop.rdf#"
  xmlns:p06owl="http://emblems.let.uu.nl/lab/dse/rdf/p06owl.rdf#"
  xmlns:owl  = "http://www.w3.org/2002/07/owl#"
  xmlns:rdf  = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd  = "http://www.w3.org/2001/XMLSchema#">
  <owl:Ontology rdf:about="">
    <rdfs:comment>OWL ontology for metaphors in emblems</rdfs:comment>
    <owl:imports
rdf:resource="http://emblems.let.uu.nl/lab/dse/rdf/p06owltop.rdf"/>
    <rdfs:comment>Emblem Metaphor Ontology</rdfs:comment>
    <owl:versionInfo>2007-04-04</owl:versionInfo>
  </owl:Ontology>
  <owl:Class rdf:ID="Comparand">
    <rdfs:subClassOf rdf:resource="&p06owltop;ContentSign"/>
    <rdfs:label xml:lang="en">Class of compared entities</rdfs:label>
  </owl:Class>
  <owl:Class rdf:ID="Metaphor">
    <rdfs:subClassOf rdf:resource="&p06owltop;ContentSign"/>
    <rdfs:label xml:lang="en">Class of metaphors</rdfs:label>
  </owl:Class>
  <owl:ObjectProperty rdf:ID="MetaphorTenorProp">
    <rdfs:label>Tenor</rdfs:label>
    <rdfs:subPropertyOf rdf:resource="&p06owltop;SignVehicleProp"/>
    <rdfs:domain  rdf:resource="#Metaphor"/>
    <rdfs:range rdf:resource="#Comparand"/>
    <rdfs:comment>Connects a metaphor and it's tenor. The property is a
     subproperty of SignVehicleProp, because the metaphor can only be there
     if the tenor is also there. The presence of the reverse property,
     TenorOfMetaphorProp, is inferred.</rdfs:comment>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="MetaphorVehicleProp">
    <rdfs:label>Vehicle</rdfs:label>
    <rdfs:subPropertyOf rdf:resource="&p06owltop;SignVehicleProp"/>
    <rdfs:domain  rdf:resource="#Metaphor"/>
    <rdfs:range rdf:resource="#Comparand"/>
    <rdfs:comment>Connects a metaphor and it's vehicle. The property is a
     subproperty of SignVehicleProp, because the metaphor can only be there
     if the vehicle is also there. The presence of the reverse property,
     VehicleOfMetaphorProp, is inferred.</rdfs:comment>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="TenorInMetaphorProp">
    <rdfs:label>Tenor in Metaphor</rdfs:label>
    <owl:inverseOf rdf:resource="#MetaphorTenorProp"/>
    <rdfs:seeAlso rdf:resource="#MetaphorTenorProp"/>
    <rdfs:comment>Connects a tenor to the metaphor it is the tenor of.
     Inverse property of MetaphorTenorProp. Presence of this property is
```

```
        inferred by the system.</rdfs:comment>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="VehicleInMetaphorProp">
    <rdfs:label>Vehicle in Metaphor</rdfs:label>
    <owl:inverseOf rdf:resource="#MetaphorVehicleProp"/>
    <rdfs:seeAlso rdf:resource="#MetaphorVehicleProp"/>
    <rdfs:comment>Connects a vehicle to the metaphor it is the vehicle of.
     Inverse property of MetaphorVehicleProp. Presence of this property is
     inferred by the system.</rdfs:comment>
  </owl:ObjectProperty>
  <owl:Class rdf:ID="MetaphorUse">
    <rdfs:subClassOf rdf:resource="&p06owltop;ContentSign"/>
    <rdfs:label xml:lang="en">Class of metaphor uses</rdfs:label>
  </owl:Class>
  <owl:ObjectProperty rdf:ID="MetaphorUsedProp">
    <rdfs:label>Metaphor</rdfs:label>
    <rdfs:subPropertyOf rdf:resource="&p06owltop;SignVehicleProp"/>
    <rdfs:domain  rdf:resource="#MetaphorUse"/>
    <rdfs:range rdf:resource="#Metaphor"/>
  </owl:ObjectProperty>

(What follows are the definitions of the individual metaphors).

</rdf:RDF>
```

Appendix D. TEI ODD File for Annotation with Feature Structures

The following ODD file defines the TEI extension used for annotation with TEI feature structures, as described in chapter 10.

```
<?xml version="1.0"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="en">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Annotation of parallel texts</title>
        <author>generated by Roma 2.13; modified by hand</author>
      </titleStmt>
      <publicationStmt>
        <p>for use by whoever wants it</p>
      </publicationStmt>
      <sourceDesc>
        <p>created on Monday 13th August 2007 10:05:22 AM by the form at
          http://www.tei-c.org.uk/Roma/ and modified many times later</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <front>
      <divGen type="toc"/>
    </front>
    <body>
      <p>Used for annotation of parallel texts</p>
      <schemaSpec ident="partext" docLang="en" xml:lang="en" prefix="pat">
        <moduleRef key="core"/>
        <moduleRef key="tei"/>
        <moduleRef key="textstructure"/>
        <moduleRef key="analysis"/>
        <moduleRef key="linking"/>
```

```
        <moduleRef key="iso-fs"/>
        <moduleRef key="declarefs"/>
        <elementSpec ident="note" module="core" mode="change">
          <classes mode="change">
            <memberOf key="model.featureVal.single" mode="add"/>
          </classes>
        </elementSpec>
        <elementSpec ident="date" module="core" mode="change">
          <classes mode="change">
            <memberOf key="model.featureVal.single" mode="add"/>
          </classes>
        </elementSpec>
        <elementSpec ident="ptr" module="core" mode="change">
          <classes mode="change">
            <memberOf key="model.featureVal.single" mode="add"/>
          </classes>
        </elementSpec>
        <elementSpec ident="numeric" module="iso-fs" mode="change">
          <attList>
            <attDef ident="value" mode="change" usage="opt"/>
          </attList>
        </elementSpec>
        <moduleRef key="header"/>
        <elementSpec ident="vDescr" mode="add"
          ns="http://www.example.org/ns/nonTEI">
          <desc>Describes in prose the meaning of a feature's value</desc>
          <classes/>
          <content>
            <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
              name="macro.limitedContent"/>
          </content>
        </elementSpec>
        <elementSpec ident="symbol" module="iso-fs" mode="change">
          <content xmlns:rng="http://relaxng.org/ns/structure/1.0">
            <rng:optional>
              <rng:ref name="vDescr"/>
            </rng:optional>
          </content>
        </elementSpec>
        <elementSpec ident="dataSection"
ns="http://www.example.org/ns/nonTEI"
          mode="add">
          <desc>Will hold data-like elements (linkGrp, feature structures,
            interp's) if they have nowhere else to go.</desc>
          <classes>
            <memberOf key="model.resourceLike"/>
          </classes>
          <content>
            <rng:oneOrMore xmlns:rng="http://relaxng.org/ns/structure/1.0">
            <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
name="div"/>
            </rng:oneOrMore>
          </content>
        </elementSpec>
        <elementSpec ident="aboutDesc" ns="http://www.example.org/ns/nonTEI"
          mode="add">
          <desc>Used to describes the files that another file is about</desc>
          <content>
            <rng:choice xmlns:rng="http://relaxng.org/ns/structure/1.0">
              <rng:oneOrMore xmlns:rng="http://relaxng.org/ns/structure/1.0">
                <rng:ref
                  xmlns:rng="http://relaxng.org/ns/structure/1.0" name="p"/>
              </rng:oneOrMore>
              <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
```

```
                        name="listBibl"/>
                    </rng:choice>
                  </content>
                </elementSpec>
                <elementSpec ident="linkGrp" module="linking" mode="change">
                  <content xmlns:rng="http://relaxng.org/ns/structure/1.0">
                    <rng:optional xmlns:rng="http://relaxng.org/ns/structure/1.0">
                      <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                      name="head"/>
                    </rng:optional>
                    <rng:oneOrMore xmlns:rng="http://relaxng.org/ns/structure/1.0">
                      <rng:choice xmlns:rng="http://relaxng.org/ns/structure/1.0">
                        <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                          name="link"/>
                        <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                          name="ptr"/>
                      </rng:choice>
                    </rng:oneOrMore>
                  </content>
                </elementSpec>
                <elementSpec ident="fileDesc" module="header" mode="change">
                  <content xmlns:rng="http://relaxng.org/ns/structure/1.0">
                    <rng:group xmlns:rng="http://relaxng.org/ns/structure/1.0">
                      <rng:group xmlns:rng="http://relaxng.org/ns/structure/1.0">
                        <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                          name="titleStmt"/>
                        <rng:optional
                          xmlns:rng="http://relaxng.org/ns/structure/1.0">
                          <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                            name="editionStmt"/>
                        </rng:optional>
                        <rng:optional
                          xmlns:rng="http://relaxng.org/ns/structure/1.0">
                          <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                            name="extent"/>
                        </rng:optional>
                        <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                          name="publicationStmt"/>
                        <rng:optional
                          xmlns:rng="http://relaxng.org/ns/structure/1.0">
                          <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                            name="seriesStmt"/>
                        </rng:optional>
                        <rng:optional
                          xmlns:rng="http://relaxng.org/ns/structure/1.0">
                          <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                            name="notesStmt"/>
                        </rng:optional>
                      </rng:group>
                      <rng:oneOrMore xmlns:rng="http://relaxng.org/ns/structure/1.0">
                        <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                          name="sourceDesc"/>
                      </rng:oneOrMore>
                      <rng:optional xmlns:rng="http://relaxng.org/ns/structure/1.0">
                        <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
                          name="aboutDesc"/>
                      </rng:optional>
                    </rng:group>
                  </content>
                </elementSpec>
              </schemaSpec>
            </body>
          </text>
        </TEI>
```

Samenvatting

Geesteswetenschappers bestuderen primaire literatuur en schrijven secundaire literatuur. De centrale stelling van dit proefschrift is dat digitale annotaties het mogelijk maken om interactieve connecties te leggen tussen de primaire en secundaire literatuur. Het laat zien hoe onderzoekers met elektronische hulpmiddelen annotaties kunnen maken bij de bronnen die ze bestuderen en hoe ze de daarmee de onderwerpen die ze interesseren kunnen modelleren. De annotatietypen die daarbij worden gedefinieerd geven aan hoe de annotaties moeten worden geïnterpreteerd. De zo geïnterpreteerde annotaties bieden een mogelijkheid om de primaire tekst vanuit de annotaties te verkennen; ze bieden ook de ondersteunende argumentatie voor de secundaire teksten. Ik noem deze annotaties mesotekst.

Mesotekst kan worden gezien als het geesteswetenschappelijk equivalent van de experimentele data van de natuurwetenschappen. En zoals die natuurwetenschappelijke data niet voortkomen uit toevallige metingen, maar voortkomen uit hypothesen die weer gebaseerd zijn op theorieën, of modellen, zo zijn ook geesteswetenschappelijke data expliciet of impliciet gebaseerd op een model van de bestudeerde verschijnselen. Dankzij zo'n model is een collectie annotaties meer dan een verzameling toevallige waarnemingen. De centrale plaats die ik in dit proefschrift toeken aan de notie van het modelleren van annotaties is mede te danken aan de visie op het gebruik van de computer in de geesteswetenschappen die Willard McCarty heeft geformuleerd in *Humanities Computing* (2005). De gedisciplineerde toepassing van een model, en dat is waar de computer ons bij helpt, maakt de beperkingen van het model duidelijk en toont de kloof tussen het model en de verschijnselen die het probeert te beschrijven of verklaren. Dat proces helpt de onderzoeker het model te verfijnen, of eventueel te verwerpen, en zo komt wetenschappelijk inzicht tot stand.

Dit proefschrift begeeft zich op meerdere wetenschappelijke terreinen. De achtergrond is de studie van de emblematiek. De emblematiek was een genre in de Europese literatuur van de zestiende tot en met de achttiende eeuw, waarin een opschrift, een afbeelding en een onderschrift werden gecombineerd tot, meestal, een wijze les. Vaak werd het geheel verrijkt met geleerde aanhalingen of een commentaar in proza. De makers van embleemboeken waren zware gebruikers van wat we nu 'knippen en plakken' zouden noemen, en ze leenden zonder schroom uit de beeldende kunsten en uit de klassieke en

contemporaine literatuur. Omgekeerd werden de motieven uit de emblematiek weer vaak toegepast in de andere kunsten en in bijvoorbeeld de decoratie van gebruiksvoorwerpen. In emblemen konden heel verschillende onderwerpen aan de orde worden gesteld, zoals liefde, godsdienst of politiek. Het gebruik van verschillende media, de variëteit in onderwerpen en de vele intertekstuele relaties maken de emblematiek tot een geschikt terrein voor experimenten op het terrein van de alfa-informatica.[204]

Als ik een enkel veld zou moeten aangeven waartoe dit proefschrift behoort, is het alfa-informatica. De praktijk van het modelleren die McCarty aanwijst als essentieel voor alfa-informatica is wat wordt gefaciliteerd door de annotatietools die ik zal beschrijven. Alfa-informatica is van nature een interdisciplinair terrein. McCarty karakteriseert het als een 'methodological commons', waar begrippen en methoden uit andere vakgebieden worden geleend en elders toegepast. Een deel van mijn inspiratie ontleen ik aan het gebruik van annotatie, of annotatie-achtige verschijnselen, in andere wetenschapsgebieden. Ik wil er met dit boek aan bijdragen dat het begrip annotatie gaat behoren tot het methodologisch gemeengoed van de alfa-informatica, en ik ontwikkel methoden en hulpmiddelen voor het gebruik ervan.

Behalve op de gebieden van de alfa-informatica en de emblematiek begeeft dit proefschrift zich op nog enige vakgebieden. Een noodzakelijke voorwaarde voor digitale annotatie is de beschikbaarheid van digitale teksten. Twee hoofdstukken zijn daarom gewijd aan de embleemdigitalisering. Een ander terrein waaraan het proefschrift raakt is dat van de wetenschappelijke editie. De editie is een belangrijk onderwerp, omdat annotatie traditioneel een belangrijke rol speelt in de wetenschappelijke editie, maar ook omdat het wenselijk is de annotatie bij een digitale tekst vanuit edities van die tekst te ontsluiten. Tenslotte heeft het proefschrift ook betrekking op het terrein van het wetenschappelijk publiceren. Het idee van mesotekst is dat annotaties digitaal verkend moeten kunnen worden en moeten aansluiten op de teksten waarop ze betrekking hebben (de primaire teksten) en de teksten die ze ondersteunen (de secundaire teksten). Primaire en secundaire teksten moeten dus digitaal beschikbaar en toegankelijk zijn. In tegenstelling tot de PDF-documenten waar we het nu mee moeten doen, moeten die teksten adresseerbaar, interactief en uitbreidbaar zijn. Wat we nodig hebben zijn directe verbindingen tussen de wetenschappelijke output en de ondersteunende gegevens, en ons huidig publicatiesysteem is daarop niet ingericht.

Dit proefschrift bestaat uit vijf delen. Het eerste deel is inleidend en geeft de context voor wat in de volgende delen aan de orde komt. Het tweede deel bevat twee beschouwingen over embleemdigitalisering. In het derde doe ik een aantal voorstellen voor elektronische annotatiehulpmiddelen. Het vierde deel bevat

---

[204] Ik gebruik 'alfa-informatica' als Nederlands equivalent van 'humanities computing'.

twee hoofdstukken over emblematiek. Ze kwamen tot stand mede met behulp van de annotatiehulpmiddelen die worden besproken in het derde deel. In deel vijf beschrijf ik het concept mesotext en verken ik mogelijke toepassingen ervan in de context van het geesteswetenschappelijk onderzoek. Daarna volgt de conclusie.

Deel I

Hoofdstuk 1 beschrijft kort het terrein waarop mijn onderzoek zich begeeft en het beoogde publiek van dit proefschrift. Hoofdstuk 2 geeft een inleiding in de embleemstudie: een korte (voor)geschiedenis van het embleem en een bespreking van de verschillende vormen die het embleem in verschillende landen aannam. Vervolgens bespreek ik verschillende wetenschappelijke benaderingen van het embleem: van Praz, Heckscher en Wirth, Schöne, Daly, Scholz en anderen. Tenslotte introduceer ik 'digital emblem studies', digitale emblematiek, embleemstudie die digitale hulpmiddelen ontwikkelt en gebruikt. Als emblematologen voordeel wil hebben bij de ontwikkeling van digitale technieken, dan moeten ze zelf ook actief betrokken willen zijn bij de ontwikkeling daarvan.

In hoofdstuk 3 bespreek ik de rol van annotatie in de digitale editie. Ik begin met een heel algemene definitie van annotatie: een object dat wordt betrokken op een ander object. Een annotatie kan een verklarende opmerking bij verouderd woordgebruik zijn, maar ook een plaatje van een genoemd schilderij, een verwijzing naar Wikipedia, of een filmclip van een scène uit een toneelstuk. Annotaties kunnen gemaakt worden voor een breed publiek of voor privé doeleinden. Voor mijn doel gaat het vooral om verzamelingen van getypeerde annotaties. Een getypeerde annotatie is een annotatie waarvan een deel van de betekenis expliciet wordt gemaakt door een annotatietype dat de structuur van de annotatie definieert. Zulke annotatietypen zijn zinvol in een verzameling annotaties omdat het op basis daarvan mogelijk wordt de annotaties bijvoorbeeld te filteren of te visualiseren.

Vervolgens onderzoek ik de beschikbaarheid van annotatiehulpmiddelen in digitale bibliotheken en digitale edities. Het blijkt dat daarvan nog nauwelijks sprake is. Ik bekijk daarna hoe annotaties gebruikt worden in een aantal aangrenzende vakgebieden. De belangrijkste karakteristieken van die benaderingen zijn dat annoteren een sociaal proces kan zijn, dat structuur de mogelijkheden voor benutting van de annotaties vergroot, dat het zinvol kan zijn om annotaties op verschillende manieren te visualiseren en dat annotaties de basis kunnen zijn voor geavanceerde zoek- en selectiemogelijkheden.

Deel II

Deel II bevat twee hoofdstukken over embleemdigitalisering. In de aanloop van het werk aan dit proefschrift was ik betrokken bij de opzet van het Emblem Project Utrecht (EPU), en zijdelings ben ik daarbij betrokken gebleven. De teksten die in het project zijn gedigitaliseerd, met name Otto van Veen's *Amoris Divini Emblemata* (Antwerpen, 1615), waren de grondstof voor het werk in de volgende hoofdstukken.

        In hoofdstuk 4 schets ik een model van de digitale embleemsite. Globaal kunnen de sites gekarakteriseerd worden als 'thematic research collections' (Palmer 2004), maar het is mogelijk om de sites meer specifiek te beschrijven. Ik karakteriseer digitale embleemsites in termen van een basisoriëntatie (facsimile, wetenschappelijke editie of index), een aantal generieke functionaliteiten, een aantal specifieke parameters en vervolgens de technische keuzen. In de discussie contrasteer ik het gegeven model (een empirisch, *ex post* model) met een aantal *ex ante* modellen van de wetenschappelijke editie.

        In hoofdstuk 5 draait het opnieuw om modelleren en digitaliseren, maar in een ander verband: ik bespreek de rol van het modelleren in het ontwerp van een digitale editie. Daarbij gaat het vooral om het modelleren van de tekst, waarbij ook pragmatische aspecten een rol spelen. In het bijzonder komt hierbij het gebruik van XML voor de digitale editie aan de orde. Het Text Encoding Initiative (TEI) heeft voor de codering van digitale teksten richtlijnen ontwikkeld die het uitgangspunt waren voor het werk van het EPU. Eén van de redenen om te kiezen voor XML is juist de expliciete modellering waarom XML vraagt. Daarbij ontstaat een elektronisch hanteerbaar model van het uit te geven genre, waar ook de wijdere gemeenschap die zich bezig houdt met tekstcodering baat bij heeft.

Deel III

In deel III bespreek ik drie in dit onderzoek ontwikkelde annotatiehulpmiddelen, die verschillende benaderingen van het annoteren tonen.

        In hoofdstuk 6 bespreek ik een eerste benadering aan de hand van het annotatietool EDITOR. EDITOR veronderstelt dat een digitale editie bestaat uit een onderliggend XML-bestand, en HTML-bestanden (webpagina's) voor de weergave aan de lezer. Met behulp van EDITOR kunnen onderzoekers annotatietypen definiëren die bestaan uit meerdere velden met verschillende datatypen (bijvoorbeeld tekst of hyperlink). Onder verwijzing naar zo'n annotatietype kunnen ze vervolgens de editie-XML voorzien van annotaties. Het resultaat is een set van annotaties, die ook getoond kan worden op het web. Ik laat zien hoe EDITOR kan worden ingezet in een onderzoek naar theatrale aspecten van *Amoris Divini Emblemata*.

In hoofdstuk 7 toon ik hoe dergelijke annotaties kunnen worden uitgewisseld tussen programma's die annotaties maken en mogelijke gebruikers van annotaties. Eén zo'n mogelijke gebruiker is de digitale editie: het is wenselijk dat annotaties die bij een tekst zijn gemaakt kunnen worden getoond vanuit de digitale editie van die tekst. Ten behoeve van dergelijke uitwisseling schets ik een protocol en een markup vocabulaire. Het markup vocabulaire dient voor de opslag van annotaties, het protocol kan gebruikt worden om annotaties op te vragen. Ik heb een prototype van een programma ontwikkeld dat verzoeken op basis van dat protocol afhandelt. Ik laat zien hoe op basis daarvan annotaties zichtbaar kunnen worden gemaakt vanuit een digitale editie. Ook kunnen zo gefilterde overzichten en grafische weergaven van de annotaties worden getoond.

Een andere benadering wordt besproken in hoofdstuk 8 en hoofdstuk 9. Waar de annotaties van EDITOR in principe los van elkaar staan, wordt in deze hoofdstukken een netwerk van naar elkaar verwijzende annotaties gebouwd, vergelijkbaar met het semantisch web waar het W3C naar streeft. De structuur van dat netwerk wordt gedefinieerd door een ontologie, een model van de soorten objecten, met hun eigenschappen en relaties, zoals die in een bepaald domein voorkomen. Ik gebruik een ontologie die gebaseerd is op een algemeen semiotisch model: een teken berust op een tekendrager en kan zelf weer tekendrager zijn voor volgende tekens. Ik kies voor een opslag van het model in Resource Description Framework (RDF) formaat. In hoofdstuk 8 leg ik de basis van de benadering uit en argumenteer ik waarom de benadering zinvol kan zijn. In het kort luidt die argumentatie dat het hiermee mogelijk wordt een interpretatie formeel vast te leggen: de interpretatie maakt gebruik van begrippen die zijn gedefinieerd in de ontologie, en de annotaties verwijzen eenduidig naar de plaatsen in de tekst waarop ze betrekking hebben. Dat maakt interpretaties niet noodzakelijk juister, maar wel preciezer en eenduidiger.

In hoofdstuk 9 maak ik van de in 8 gelegde basis gebruik voor het ontwerp van een index op metaforen in *Amoris Divini Emblemata*. Ik ontleed metaforen in vehicle en tenor op basis van het model van Richards (1936). Daarvoor is een ontologie gemaakt in de ontologie-taal OWL. In een aantal emblemen heb ik op die manier een groot deel van de er in voorkomende metaforiek geanalyseerd. Ik heb een toepassing ontwikkeld waarin per embleem de onderkende metaforen worden getoond, en waarin van elke metafoor de plaatsen in tekst en beeld worden getoond waar van die metafoor gebruik wordt gemaakt. Het achterliggende RDF-model kan op verschillende manieren worden bevraagd, onder andere via een toepassing waarbij de gebruiker aan de hand van de ontologie een zoekopdracht kan samenstellen. Het geheel wordt gecompleteerd met een betoog dat kan worden gebruikt om de bevindingen over de metaforiek samen te vatten. Vanuit het betoog kan met hyperlinks worden gewezen naar afzonderlijke metaforen of andere tekens, naar emblemen, naar onderdelen van de ontologie, en naar kant-en-klare

zoekopdrachten op het RDF-model. Hier wordt gerealiseerd wat ik hierboven beschreef: de verbinding die annotaties (hier de observaties vastgelegd in RDF) leggen tussen primaire tekst (hier de embleembundel) en secundaire tekst (hier het betoog).

Hoofdstuk 10 kiest een derde benadering. Hier wordt gebruik gemaakt van annotaties met een vergelijkbare structuur als in het geval van EDITOR, maar nu worden ze vastgelegd met behulp van TEI feature structures. Feature structures zijn een formalisme, oorspronkelijk ontwikkeld voor taalkundige annotaties, waarvan ik aannemelijk maak dat het ook voor andersoortige annotaties, bijvoorbeeld in de letterkunde, goed bruikbaar kan zijn. Zgn. feature structure declarations spelen de rol van de annotatietypen; op basis van de declaraties kunnen schermen worden getoond waar gebruikers de toegestane velden (features) kunnen invullen. Ik maak hiervan gebruik in de opzet van een gegevensstructuur voor het annoteren van verschijnselen in parallelle tekststructuren, zoals ze in de emblematiek veel voorkomen. Daarnaast introduceer ik de notie van een TEI proxy document: een TEI document dat geen volledige transcriptie van een tekst bevat maar wel voldoende van de structuur om te kunnen dienen als aanhechtpunt voor digitale annotaties. Hiervan kan gebruik worden gemaakt als geen volledig XML document beschikbaar is, bijvoorbeeld in het geval van een werk dat gedigitaliseerd is in Google's library project.

Deel IV

Deel IV bevat twee emblematologische hoofdstukken. De verbinding met de rest van het proefschrift is dat ze tot stand zijn gekomen in samenhang met de ontwikkeling van de annotatiehulpmiddelen die besproken worden in deel III.

Hoofdstuk 11 bevat een beschouwing over metaforen in *Amoris Divini Emblemata*. Eén van de bevindingen is dat terugkerende achtergrondmetaforiek van landschap, water, plantaardig leven, licht en verticaliteit de setting vormen voor de handelingen op de voorgrond en verantwoordelijk zijn voor een belangrijk deel van de samenhang van het boek. De inbedding van metaforen in een allegorische context maakt het verder mogelijk om een gewenste houding ten opzichte van metaforische objecten over te brengen. Eén van de metaforen die het boek gebruikt is die van de leraar-leerling verhouding tussen de Goddelijke Liefde en de Ziel, en het boek belichaamt die metafoor: het boek wil voor de lezer zijn wat de Goddelijke Liefde is voor de Ziel.

In hoofdstuk 12 bekijk ik de afbeeldingen van *Amoris Divini Emblemata* als waren het afbeeldingen van theaterscènes. In sommige opzichten zijn de Goddelijke Liefde en de Ziel vergelijkbaar met acteurs die voor een publiek een scène opvoeren. Hun handelingen en gebaren, ogenschijnlijk gericht op de ander, zijn altijd ook bedoeld voor het publiek, dat wil zeggen voor de lezer van het boek. In andere embleemboeken kan een dergelijk effect optreden in

teksten met een dialoogvorm, omdat de sprekers terwijl ze met elkaar spreken altijd ook de lezer mee aanspreken.

Deel V

In hoofdstuk 13 introduceer ik het concept van mesotekst. Mesotekst kan worden gepositioneerd tussen de primaire en secundaire teksten van de geesteswetenschapper. Digitale annotatietechnologie maakt het mogelijk dat aantekeningen die beginnen als notities in de marge van de primaire tekst worden hergebruikt in de argumentatie voor een these in de secundaire tekst. Annotatietypen geven een model van de annotaties en dus van bepaalde aspecten van de tekst die wordt geannoteerd. Dankzij de door de annotatietypen aangebrachte structuur kan de verzameling annotaties vanuit verschillende richtingen (primaire tekst, secundaire tekst, model) worden verkend.

Ik vervolg de bespreking van mesotekst met een korte beschouwing over een aantal verwante verschijnselen. Eén van die verschijnselen zijn de aantekeningen van schrijvers bij hun lectuur, die vaak een eerste stadium op weg naar scheppend werk zijn. In natuurwetenschappelijke publicaties zien we een ander vergelijkbaar verschijnsel: de tendens om de data die hebben geleid tot een wetenschappelijke publicatie op één of andere wijze bij te sluiten bij de publicatie, zodat er 'scientific publication packages' ontstaan. Een andere tendens is om naast publicatie in PDF-vorm een gestructureerde weergave van een aantal sleutelgegevens bij te sluiten, zodat artikelen op zinvolle wijze automatisch doorzoekbaar worden.

In hoofdstuk 14 tenslotte trek ik enkele verdere conclusies. Ik bespreek de verschillende geteste benaderingen voor de ontwikkeling van annotatiehulpmiddelen. Het belangrijkste verschil tussen aan de ene kant de benadering van EDITOR en de TEI feature structures en aan de andere kant de semantisch-web-benadering is dat in het laatste geval inderdaad een web van annotaties ontstaat. De annotaties ontlenen daar een deel van hun betekenis aan het netwerk waar ze deel van uitmaken. Een faciliteit voor de uitwisseling van annotaties heeft in dit geval weinig zin. Een keuze tussen deze technologieën zou waarschijnlijk niet moeten afhangen van het object van onderzoek, maar eerder van de gewenste manier van benadering van de annotaties. De semantisch-web-benadering is waarschijnlijk wat lastiger voor wie niet vertrouwd is met een ontologische benadering.

Een belangrijke vraag is in hoeverre hulpmiddelen zoals degene die ik in dit proefschrift bespreek kunnen of zullen leiden tot een fundamenteel andere werkwijze in de geesteswetenschappen. Ik bespreek kort de ideeën van Ramsay (o.a. 2007) over een 'algorithmic criticism', en sluit me aan bij zijn stelling dat het in de studie van literatuur niet in de eerste plaats gaat om iets te bewijzen, maar om iets nieuws, opmerkelijks of inspirerends te laten zien. Ik zie de

hulpmiddelen die ik heb besproken eerder als een ondersteuning van de interpretatieve activiteit van de onderzoeker, vergelijkbaar met het programma van Bradley (2003), waarvan zijn annotatieprogramma Pliny de uitkomst is.

Tenslotte bepleit ik de organisatie van een groot proefproject rond annotatie van geesteswetenschappelijk interessante teksten. In de Nederlandse context zou de DBNL daar een goed platform voor bieden, gezien het volume aan beschikbare tekst en de aantallen bezoekers. In een dergelijk onderzoek moet worden onderzocht of wat ik in dit proefschrift aannemelijk heb willen maken in de praktijk ook werkt en bruikbaar is.

Glossary – List of Abbreviations

API: Application Programming Interface. A system provides an application programming interface to make it possible for third-party programs to make use of the system's functionality.

Attribute: in XML, an attribute specifies a property of an element. In e.g. `<lg type="motto" lang="lat">`, the *lg* element has a type attribute (value 'motto') and a lang (language) attribute.

Cocoon: a system for publishing XML documents on the web. See http://cocoon.apache.org/.

CSS: Cascading Style Sheets. See http://www.w3.org/Style/CSS/. Cascading Style Sheets contain instructions for the rendering of HTML and XML documents in browsers, determining e.g. fonts and font sizes, size and location of text areas, margins, etc. To be distinguished from XSLT stylesheets.

Diigo: a web annotation tool. Diigo can highlight fragments of web pages and add sticky notes to the pages. See http://www.diigo.com/.

DocBook: an XML markup vocabulary, similar to TEI. DocBook defines a tagset especially suitable for technical documentation. See http://www.docbook.org/.

DOM: Document Object Model. A convention for representing and interacting with elements in an HTML or XML document (Wikipedia).

DTD: Document Type Definition. An XML schema language. A DTD specifies rules for the elements and attributes that can occur in a class of XML documents. An example of such a rule could be: books contain chapters which contain paragraphs, and chapters must be numbered.

EAD: Encoded Archival Description . See http://www.loc.gov/ead/. An XML markup vocabulary for encoding archival finding aids.

Element: in XML, a constituent of a document that is delimited by opening and closing tag, that can contain text and/or sub-elements. The example shows one *lg* element (linegroup) and 3 *l* elements (verse lines):
```
<lg>
  <l>Ama à Dios de coraçon, </l>
  <l>Paraque se satisfaga</l>
  <l>Que amor con amor se paga. </l>
</lg>
```

Extension: in Firefox, a program installed within the web browser that extends the browser's functionality.

Feature structures: 'a general purpose data structure which identifies and groups together individual features, each of which associates a name with one or more values' (Burnard & Bauman 2007: ch. 18).

Firefox: a web browser, a program that is used to view web pages. See http://www.mozilla.com/firefox/.

Gecko: Firefox 'layout engine', the component of the browser which handles the display of web pages.

Google Notebook: a web annotation tool. See http://www.google.com/notebook.

HTML: Hyper Text Markup Language. The language used to define web pages.

Internet Explorer (IE): a web browser, a program that is used to view web pages. Made by Microsoft.

Java: an open source programming language. See http://java.com/.

Markup: a set of annotations to text that describe how it is to be structured, laid out, or formatted (Wikipedia). Typically, markup and text are interspersed.

Mozilla: Organisation that develops the Firefox browser and other open source programs.

MySQL: an open source database system. See http://www.mysql.com/.

Notefish: a web annotation tool. See http://www.notefish.com/.

OAI-PMH: Open Archive Initiative – Protocol for Metadata Harvesting. A standard for how sites can fetch ('harvest') metadata about objects from other sites. See http://www.openarchives.org/.

ODD: an XML markup vocabulary for defining (among other things) TEI extensions. See Burnard & Rahtz (2004).

Ontology: a model of the sorts of objects, with their properties and relations, that exist in a given domain. In the Semantic Web, ontologies play a role similar to that of XML Schemas in XML: they describe the elements that can occur in classes of RDF documents. OWL and RDF Schema are examples of ontologies.

OWL: Web Ontology Language (*sic*). A language for defining ontologies used in RDF documents. See http://www.w3.org/TR/owl-features/.

OWLDoc: a tool that generates documentation for OWL ontologies. See http://www.co-ode.org/downloads/owldoc/.

PHP: a scripting language originally designed for producing dynamic web pages (Wikipedia). See http://www.php.net/.

P4: TEI version released in 2002, the first to be based on XML. See http://www.tei-c.org/Guidelines/P4/index.xml.

P5: TEI version released in 2007. Major new release. See http://www.tei-c.org/Guidelines/P5/.

Plain text: unstructured text, text without markup.

Query language: a language used to formulate queries, e.g. to request a selection from a database

Querystring: the part of a Uniform Resource Locator (URL) that contains data to be passed to web applications (Wikipedia). The query string consists of the part of the URL that follows the question mark: `http://server/path/program?query_string`. The query string can be used to send requests or commands to the web application.

RAP: RDF API for PHP. An Application programming interface used to handle RDF data in PHP programs. See http://www.seasr.org/wp-content/plugins/meandre/rdfapi-php/doc/index.html.

RDF: Resource Description Framework. A framework used for modelling information that is implemented in web resources. RDF models can be expressed in XML or in syntaxes. RDF is used in many Semantic Web applications. See http://www.w3.org/RDF/.

RDF Schema: a light-weight ontology language used in conjunction with RDF. See http://www.w3.org/TR/rdf-schema/.

RDQL: an RDF query language. See http://www.w3.org/Submission/RDQL/. RDQL has been superseded by SPARQL.

Relax NG: an XML Schema language. The preferred schema language of the TEI. See http://www.relaxng.org/.

Rich text: text that uses formatting such as italics, bold, bulleted lists, different fonts, etc.

SKOS: Simple Knowledge Organization System. A W3C standard in development to represent knowledge organization systems (such as thesauri) in RDF. See http://www.w3.org/2004/02/skos/.

SPARQL: a query language for RDF, standardised by W3C. See http://www.w3.org/TR/rdf-sparql-query/.

Stand-off markup or annotation: markup or annotation that is not present in the text, at the location of the phenomenon that is being annotated, but elsewhere. The stand-off markup will then use a pointing mechanism to refer to the location that it belongs to.

SVG: Scalable Vector Graphics. An XML-based file format for describing two-dimensional vector graphics (Wikipedia). See http://www.w3.org/Graphics/SVG/. SVG can be used to render e.g. bar charts.

Tag, tagging: label, labelling. The labelling can be applied to objects such as web sites or newspaper articles, but tagging can also refer to the process of text encoding. In that case, and in the context of this thesis, the tags delimit XML elements.

TEI: Text Encoding Initiative. See http://www.tei-c.org. 'A consortium which collectively develops and maintains a standard for the representation of texts in digital form' (from the website). The Emblem Project Utrecht followed the TEI Guidelines in encoding the emblems.

Triple: a statement in RDF, consisting of subject, predicate and object.

URL: Uniform Resource Locator. A web address, 'specifies where an identified resource is available and the mechanism for retrieving it' (Wikipedia).

W3C: World Wide Web Consortium. See http://www.w3.org/. 'Develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential' (from the website).

Web 2.0: 'a perceived second generation of web development and design', that 'encapsulates the idea of the proliferation of interconnectivity and interactivity of web-delivered content' (Wikipedia).

Wiki: 'a collection of Web pages designed to enable anyone with access to contribute or modify content, using a simplified markup language' (Wikipedia).

Wysiwyg: What You See Is What You Get. Used to describe software, e.g. word processors, where a document as shown on the screen closely resembles the final output.

XHTML: XML version of HTML.

XML: eXtensible Markup Language. See http://www.w3.org/TR/2004/REC-xml-20040204/. A formalism that describes how tags can be added to text in order to describe properties of that text. By itself, XML does not describe which tags (markup vocabulary) to use; in the context of the humanities, the TEI has proposed such a markup vocabulary.

XML Schema: XML Schemas describe which elements and attributes can occur in a certain class of XML documents. They can also describe what combinations of elements and attributes are allowed. Relax NG and DTD are examples of XML Schemas.

Zotero: a reference manager, developed as a Firefox extension. See http://www.zotero.org/.

Literature

Adams, Alison (2003), *Webs of Allusion. French Protestant Emblem Books of the Sixteenth Century* (Travaux d'humanisme et Renaissance 0082-6081, 378; Genève: Droz).

Agosti, Maristella, Benfante, Lucio, and Orio, Nicola (2003), 'IPSA: A Digital Archive of Herbals to Support Scientific Research', *Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access* (Lecture Notes in Computer Science, 2911; Heidelberg: Springer), 253-64.

Agosti, Maristella, Ferro, Nicola, and Orio, Nicola (2005), 'Annotating Illuminated Manuscripts. An Effective Tool for Research and Education', *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries* (New York: ACM Press), 121-30.

**Agosti, Maristella*, et al.* (2004), 'Annotations in Digital Libraries and Collaboratories. Facets, Models and Usage', *Research and Advanced Technology for Digital Libraries: 8th European Conference, ECDL 2004, Bath, UK, September 12-17, 2004. Proceedings* (Lecture Notes in Computer Science, 3232; Heidelberg: Springer), 244-55.

Alciato, Andrea (1531), *Emblematum Liber* (Augsburg: Steyner).

Alciato, Andrea (2007), 'Alciato at Glasgow', <http://www.emblems.arts.gla.ac.uk/alciato/>, accessed 2009-04-23.

Antoniou, Grigoris and Van Harmelen, Frank (2004), *A Semantic Web Primer* (Cooperative Information Systems; Cambridge (Ma); London: MIT Press).

**Arko, Robert A.*, et al.* (2006), 'Using Annotations to Add Value to a Digital Library for Education', *D-Lib Magazine,* 12 (5).

**Atkinson, Paul*, et al.* (2001), *Handbook of Ethnography* (Sage).

Audenaert, Neil (2008), 'Breaking the Frame: Decomposition and Juxtaposition of Visually Complex Documents', *Joint Conference on Digital Libraries* (Pittsburgh).

Bannasch, Bettina (2007), *Zwischen Jakobsleiter und Eselsbrücke. Das 'bildende Bild' im Emblem- und Kinderbilderbuch des 17. und 18. Jahrhunderts* (Berliner Mittelalter- und Frühneuzeitforschung, 3; Göttingen: V&R unipress).

Barasch, Moshe (1987), *Giotto and the Language of Gesture* (Cambridge Studies in the History of Art; Cambridge, London, New York: Cambridge University Press).

Bargeron, David M. , Gupta, Anoop, and Bernheim Brush, Alice Jane (2001), 'A Common Annotation Framework', (Redmond, WA: Microsoft Corporation, Microsoft Research).

Barnett, Dene (1987), *The Art of Gesture. The Practices and Principles of 18th Century Acting* (Heidelberg: Carl Winter).

Bath, Michael (1994), *Speaking Pictures* (London: Longman).

Bath, Michael (2005), 'Christopher Harvey's *School of the Heart*', in Peter M. Daly (ed.), *Emblem Scholarship. Directions and Developments. A Tribute to Gabriel Hornstein* (Imago Figurata. Studies, 5b; Turnhout: Brepols), 1-23.

Battestin, Martin C. (1981), 'A Rationale of Literary Annotation. The Example of Fielding's Novels ', *Studies in Bibliography,* 34, 1-22.

Bauer, Barbara (1982), 'Das Bild als Argument. Emblematische Kulissen in den Bühnenmeditationen Franciscus Langs', *Archiv für Kulturgeschichte,* 64, 79-170.

Bernat Vistarini, Antonio and Cull, John T. (2008), 'The Emblem in Spain: History and Characteristics', in Peter M. Daly (ed.), *Companion to Emblem Studies* (20; New York: AMS Press), 347-68.

Berners-Lee, Tim, Hendler, James, and Lassila, Ora (2001), 'The Semantic Web', *Scientific American,* 284 (5), 28-37.

Bertrand, Paul (2006), 'Telma. Naissance d'un centre de digital humanities en France', *L'IRHT, avenir d'une tradition* (Paris: IRHT).

Bird, Steven and Liberman, Mark (2001), 'A Formal Framework for Linguistic Annotation', *Speech Communication,* 33 (1-2), 23-60.

Black, Max (1962), *Models and Metaphors. Studies in Language and Philosophy* (Ithaca [etc.]: Cornell University Press).

Black, Max (1979), 'More about Metaphor', in A Ortony (ed.), *Metaphor and Thought* (Cambridge: Cambridge University Press), 19-43.

Bloemendal, Jan (2002), 'Een emblematicus en zijn inspiratie. De bronnen van Otto Vaenius' Amoris divini emblemata, Antwerpen 1615', *Tijdschrift voor Nederlandse taal- en letterkunde,* 118 (4), 273-87.

Bocconi, Stefano, Nack, Frank, and Hardman, Lynda (2005), 'Using Rhetorical Annotations for Generating Video Documentaries', *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, 1070-73.

Bodard, Gabriel and Garcés, Juan (2009), 'Open Source Critical Editions. A Rationale', in Marilyn Deegan and Kathryn Sutherland (eds.), *Text Editing, Print, and the Digital World* (Ashgate Press), 83-98.

Boeije, H. (2005), *Analyseren in kwalitatief onderzoek. Denken en doen* (Boom onderwijs).

Bolzoni, Lina (2004), 'Emblemi e arte della memoria. Alcune note su invenzione e ricezione', in Sagrario López Poza (ed.), *Florilegio de Estudios de Emblematica. A Florilegium of Studies on Emblematics. Actas del VI Congreso Internacional de The Society for Emblem Studies. Proceedings of the 6th International Conference of the Society for Emblem Studies. A Coruña, 2002* (Ferrol: Sociedad de Cultura Valle Inclán), 15-31.

Boot, Peter (2004a), 'Accessing Emblems using XML. Digitisation Practice at the Emblem Project Utrecht', in Sagrario López Poza (ed.), *Florilegio de Estudios de Emblematica. A Florilegium of Studies on Emblematics. Actas del VI Congreso Internacional de The Society for Emblem Studies. Proceedings of the 6th International Conference of the Society for Emblem Studies. A Coruña, 2002* (Ferrol: Sociedad de Cultura Valle Inclán), 191-97.

Boot, Peter (2004b), 'Beyond the Digital Edition. Tools for Emblem Research', in Mara Wade (ed.), *Digital Collections and the Management of Knowledge. Renaissance Emblem Literature as a Case Study for the Digitization of Rare Texts and Images* (DigiCULT), 121-30.

Boot, Peter (2005), 'Advancing digital scholarship using EDITOR', *Humanities, Computers and Cultural Heritage. Proceedings of the XVI international conference of the Association for History and Computing 14-17 September 2005* (Amsterdam: Royal Netherlands Academy of Arts and Sciences), 43-48.

Boot, Peter (2006a), 'Decoding Emblem Semantics', *Literary and Linguistic Computing,* 21 supplement 1, 15-27.

Boot, Peter (2006b), 'A SANE Approach to Annotation in the Digital Edition', *Jahrbuch für Computerphilologie,* 8, 7-28.

Boot, Peter (2007), 'A Mirror to the Eyes of the Mind. Metaphor in Otto van Veen's Amoris Divini Emblemata (Antwerp 1615)', in Agnès Guiderdoni-Bruslé and Ralph Dekoninck (eds.), *Emblemata Sacra. Rhétorique et herméneutique du discours sacré dans la littérature en images. The Rhetoric and Hermeneutics of Illustrated Sacred Discourse* (Turnhout: Brepols), 291-304.

Boot, Peter (2008a), 'Playing and displaying love. Theatricality in Otto van Veen's *Amoris divini emblemata* (Antwerp 1615)', *Emblematica,* 16, 339-64.

Boot, Peter (2008b), 'An OWL-based Index of Emblem Metaphors', *Digital Humanities 2008* (Oulu).

Boot, Peter (2008c), 'Emblem Book Digitisation. State of Affairs, Options and Challenges', *EDITIO*.

Boot, Peter (2008d), 'Facilitating Emblem Research at the Emblem Project Utrecht.' in Lina Bolzoni and Silvia Volterrani (eds.), *Con parola brieve e con figura. Emblemi e imprese fra antico e moderno* (Pisa: Edizioni della Normale), 619-30.

Boot, Peter (2009), 'Towards a TEI-based encoding scheme for the annotation of parallel texts', *Lit Linguist Computing,* 24 (3), 347-61.

Boot, Peter (forthcoming 2009), 'Digital annotation of emblem books using EDITOR', *Early Modern Literary Studies*.

Boot, Peter and Stronks, Els (2002), 'Digitale edities en letterenonderzoek, een verkenning', *Neerlandistiek.nl,* 02 (08).

Boot, Peter and Stronks, Els (2003), 'Digitaal onderzoek naar Cats' sinne- en minnebeelden. Ingrediënten van een succesformule', *Nederlandse letterkunde,* 8 (1), 24-40.

Bradley, Jean-Claude (2008a), 'Enhancing Scientific Communication through Open Notebook Science', *Scholar2Scholar Symposium* <http://usefulchem.blogspot.com/2008/04/scholar2scholar-tomorrow.html>, accessed 2009-04-23.

Bradley, Jean-Claude, Owens, Kevin, and Williams, Antony (2008), 'Chemistry Crowdsourcing and Open Notebook Science', <http://precedings.nature.com/documents/1505/version/1>, accessed 2009-04-23.

Bradley, John (2003), 'Finding a Middle Ground between 'Determinism' and 'Aesthetic Indeterminacy'. A Model for Text Analysis Tools', *Literary and Linguistic Computing,* 18 (2), 185-207.

Bradley, John (2004), 'Tools to augment scholarly activity. An architecture to support text analysis', in Dino Buzzetti, Guliano Pancaldi, and Harold Short (eds.), *Augmenting Comprehension: Digital Tools and the History of Ideas* (Office for Humanities Communication Publication, 17; London: OHC), 19-47.

Bradley, John (2005), 'Documents and Data. Modelling Materials for Humanities Research in XML and Relational Databases', *Literary and Linguistic Computing,* 20 (1), 133-51.

Bradley, John (2006), 'A Response to the B2C "Cultural Hegemony" in Humanities Computing. Note Manager', *Digital Humanities* (Paris).

Bradley, John (2008b), 'Thinking about interpretation. Pliny and scholarship in the humanities', *Lit Linguist Computing,* 23 (3), 263-79.

Bradley, John (2008c), 'Pliny: A model for digital support of scholarship', *Journal of Digital information,* 9 (26).

Bradley, John and Short, Harold (2005), 'Texts into Databases. The Evolving Field of New-style Prosopography', *Literary and Linguistic Computing,* 20 (Suppl. 1), 3-24.

Bradley, John and Vetch, Paul (2007), 'Supporting Annotation as a Scholarly Tool. Experiences from the Online Chopin Variorum Edition', *Literary and Linguistic Computing,* 22 (2), 225-41.

Bradshaw, Shannon and Light, Marc (2007), 'Annotation Consensus. Implications for Passage Recommendation in Scientific Literature',

Proceedings of the 18th Conference on Hypertext and Hypermedia, 209-16.

Brandhorst, Hans (2004), 'Using Iconclass for the Iconographic Indexing of Emblems', in Mara Wade (ed.), *Digital Collections and the Management of Knowledge. Renaissance Emblem Literature as a Case Study for the Digitization of Rare Texts and Images* (DigiCULT), 29-44.

Bray, Tim, Paoli, Jean, and Sperberg-McQueen, C. M. (1998), 'Extensible Markup Language (XML) 1.0. W3C Recommendation 10 February 1998', <http://www.w3.org/TR/1998/REC-xml-19980210>, accessed 2009-04-17.

Bray, Tim*, et al.* (2004), 'Extensible Markup Language (XML) 1.0 (Third Edition). W3C Recommendation 04 February 2004', <http://www.w3.org/TR/2004/REC-xml-20040204/>, accessed 2009-04-17.

Brickley, Dan and Guha, Ramanathan V. (2004), 'RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation 10 February 2004', <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>, accessed 2009-02-16.

Brush, A. J. B.*, et al.* (2001), 'Robust Annotation Positioning in Digital Documents', *Proceedings of the SIGCHI conference on Human factors in computing systems* (Redmond, WA), 285-92.

Burghart, Marjorie (2004), 'Annotation collaborative d'un corpus de documents médiévaux: outils pour l'analyse de la structure et du contenu des sermons de Jacques de Voragine', *Le Médiéviste et l'ordinateur,* 43. <http://lemo.irht.cnrs.fr/43/43-11.htm>.

Burnard, Lou and Rahtz, Sebastian (2004), 'Relax NG with Son of ODD', *Extreme Markup Languages* <http://www.tei-c.org/cms/Talks/extreme2004/paper.html>.

Burnard, Lou and Bauman, Syd (2007), 'P5: Guidelines for Electronic Text Encoding and Interchange', <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>, accessed 2009-03-20.

Burnard, Lou and Sperberg-McQueen, C. M. (eds.) (2001-2004), *TEI P4. Guidelines for Electronic Text Encoding and Interchange. XML-Compatible Edition* (TEI Consortium).

Buschhoff, Anne (2004), *Die Liebesemblematik des Otto van Veen. Die Amorum Emblemata (1608) und die Amoris Divini Emblemata (1615)* (Bremen: Hauschild).

Buzzetti, Dino and McGann, Jerome (2006), 'Critical Editing in a Digital Horizon', in John Unsworth, Lou Burnard, and Katherine O'Brien O'Keeffe (eds.), *Electronic Textual Editing* (New York: MLA), 53-73.

Buzzetti, Dino and Rehbein, Malte (2008), 'Towards a Model for Dynamic Text Editions', *Digital Humanities 2008* (Oulu), 78-81.

Cadiz, J. J., Gupta, Anop, and Grudin, Jonathan (2000), 'Using Web Annotations for Asynchronous Collaboration around Documents', *CSCW '00: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (ACM), 309-18.

Carletta, Jean*, et al.* (2003), 'The NITE Object Model Library for Handling Structured Linguistic Annotation on Multimodal Data Sets', *Proceedings of the EACL Workshop on Language Technology and the Semantic Web (3rd Workshop on NLP and XML, NLPXML-2003)*.

Carletta, Jean*, et al.* (2005), 'A Generic Approach to Software Support for Linguistic Annotation using XML', in G. Sampson and D. McCarthy (eds.), *Corpus Linguistics: Readings in a Widening Discipline* (London: Continuum International Publishing Group).

Cats, Jacob (1618), *Silenus Alcibiadis sive Proteus* (Middelburg: Hans van der Hellen).

Cavazza, Marc and Pizzi, David (2006), 'Narratology for Interactive Storytelling: A Critical Introduction', *Technologies for Interactive Digital Storytelling and Entertainment* (Lecture Notes in Computer Science), 72-83.

Cheung, K.*, et al.* (2007), 'SCOPE-A Scientific Compound Object Publishing and Editing System'.

Chisik, Yoram and Kaplan, Nancy (2006), 'The Social Life of Books in the Humane Library', *Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries*, 312-13.

Ciula, Arianna, Spence, Paul, and Vieira, Jose Miguel (2008), 'Expressing Complex Associations in Medieval Historical Documents: The Henry III Fine Rolls Project', *Lit Linguist Computing,* 23, 311-25.

Clements, R. J. (1960), *Picta Poesis. Literary and Humanistic Theory in Renaissance Emblem Books* (Temi e Testi, 6; Roma: Edizioni di Storia e Letteratura).

Cohen, Daniel J. (2007), 'Zotero: Social and Semantic Computing for Historical Scholarship', *Perspectives Online,* 45 (5).

Colclough, Stephen (2004), 'R R, A Remarkable Thing or Action: John Dawson as Reader and Annotator', *Variants,* 2/3, 61-78.

Committee on Scholarly Editions, MLA (2007), 'Guidelines for Editors of Scholarly Editions', <http://www.mla.org/resources/documents/rep_scholarly/cse_guidelines>, accessed 2009-04-23.

Covarrubias Orozco, Sebastián (1610), *Emblemas Morales* (Madrid).

Craig, Hugh (2004), 'Stylistic Analysis and Authorship Studies', in John Unsworth, Ray Siemens, and Susan Schreibman (eds.), *A Companion to Digital Humanities* (Blackwell).

Czejdo, Bogdan D., Biguenet, Jonathan, and Biguenet, John (2007), 'Metaphor Modeling on the Semantic Web', *Lecture Notes in Computer Science,* 4512, 97.

D'Iorio, Paolo (ed.), (2000), *HyperNietzsche. Modèle d'un hypertexte savant sur Internet pour la recherche en sciences humaines. Questions philosophiques, problèmes juridiques, outils informatiques.* (Paris: Presses Universitaires de France).

Daly, Peter M. (1979), *Emblem Theory. Recent German Contributions to the Characterization of the Emblem Genre* (Wolfenbütteler Forschungen, 9; Nendeln: KTO Press).

Daly, Peter M. (1980), *The European Emblem. Towards an Index Emblematicus* (Waterloo, Ontario: Wilfrid Laurier UP).

Daly, Peter M. (1986), 'Directions in Emblem Research. Past and Present', *Emblematica,* 1 (1), 159-74.

Daly, Peter M. (1990), *The Index of Emblem Art Symposium. Papers from the McGill Symposium "An Index of Emblem Art"* (AMS Studies in the Emblem, 6; New York: AMS Press).

Daly, Peter M. (1998), *Literature in the Light of the Emblem. Structural Parallels between the Emblem and Literature in the Sixteenth and Seventeenth Centuries* (Toronto: University of Toronto Press).

Daly, Peter M. (2002), *Digitizing the European Emblem. Issues and Prospects* (AMS studies in the emblem, 15; New York, N.Y.: AMS Press).

Daly, Peter M. (2004), 'Digitizing the European Emblem: Issues, Problems and Prospects', in Sagrario López Poza (ed.), *Florilegio de Estudios de Emblemática. A Florilegium of Studies on Emblematics. Actas del VI Congreso Internacional de The Society for Emblem Studies. Proceedings of the 6th International Conference of the Society for Emblem Studies. A Coruña, 2002* (Ferrol: Sociedad de Cultura Valle Inclán), 49-64.

Daly, Peter M. (2007), 'Digitising Dutch Love Emblems', in Els Stronks and Peter Boot (eds.), *Learned Love* (DANS Symposium Publications, 1; The Hague: DANS), 183-200.

Daly, Peter M. and Young, Alan R. (2002), *The Emblems of George Wither and Gabriel Rollenhagen. A Collection of Emblemes*, eds Antonio Bernat Vistarini and Tamás Sajó (Studiolum).

Daly, Peter M. (ed.), (1985-1998), *Index Emblematicus* (Toronto: University of Toronto Press).

Daly, Peter M. (ed.), (2008), *Companion to Emblem Studies* (AMS studies in the emblem, 20; New York, NY: AMS Press).

Damiano, Rossana, Lombardo, Vincenzo, and Pizzo, Antonio (2005), 'Formal Encoding of Drama Ontology', *Virtual Storytelling* (Lecture Notes in Computer Science), 95-104.

David, Amos and Robert, Charles Abiodun (2006), 'AMIE: An annotation model for information research', *International Conference on Computers in Education (ICCE)* (Hiroshima).

De Jongh, Eddy (1976), *Tot Lering en Vermaak. Betekenissen van Hollandse genrevoorstellingen uit de zeventiende eeuw.* (Amsterdam: Rijksmuseum Amsterdam).

Dietz, Feike (2007), '[Review of] Learned love. Proceedings of the Emblem Project Utrecht Conference on Dutch Love Emblems and the Internet (November 2006);' *Nederlandse letterkunde,* 12 (4), 283-85.

Dimler, G. Richard (2007), *Studies in the Jesuit Emblem* (AMS Studies in the Emblem, 18; New York, NY: AMS Press).

Eggert, Paul (2005), 'Text-encoding, Theories of the Text, and the 'Work-Site'', *Literary and Linguistic Computing,* 20 (4), 425-35.

Elam, Keir (1980), *The Semiotics of Theatre and Drama*, ed. Terence Hawkes (New Accents; London; New York: Methuen).

Evert, Stefan and Voormann, Holger (2002), 'NITE Query Language', <http://www.ltg.ed.ac.uk/NITE/documents/NiteQL.v2.1.pdf>, accessed 2009-04-23.

Farzan, R. and Brusilovsky, P. (2005), 'Social Navigation Support Through Annotation-Based Group Modeling', *User Modeling 2005* (Lecture Notes in Computer Science, 3538; Berlin, Heidelberg: Springer), 463-72.

Ferrari, Stéphane (1996), 'Using Textual Clues to Improve Metaphor Processing', *Proceedings of the 34th Conference on Association for Computational Linguistics*, 351-53.

Ferrer, Daniel (2004), 'Towards a Marginalist Economy of Textual Genesis', *Variants,* 2/3, 7-18.

Finley, Moses I. (1985), *Ancient history. Evidence and models* (London: Chatto & Windus).

Fitzpatrick, Kathleen (2007), 'CommentPress: New (Social) Structures for New (Networked) Texts', *Journal of Electronic Publishing,* 10 (3). <http://hdl.handle.net/2027/spo.3336451.0010.305>, accessed 2009-04-23.

Flanders, Julia (1997), 'Trusting the Electronic Edition', *Computers and the Humanities,* 31 (4), 301-10.

Flynn, Christopher (s.a.), 'Defoe's Review: An annotated and annotatable edition of *A Review of the State of the British Nation* 1704-1713', <http://www.defoereview.org/>, accessed 2009-04-23.

Forster, Leonard W. (1969), *The Icy Fire. Five Studies in European Petrarchism* (London: Cambridge University Press) XVI, 203 , 5 pl. ill., facs., portr. 23 cm.

Frommholz, Ingo*, et al.* (2003), 'Document-Centered Collaboration for Scholars in the Humanities–The COLLATE System', *Proc. 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2003)*, 434-45.

Gazan, Rich (2008), 'Social Annotations in Digital Library Collections ', *D-Lib Magazine,* 14 (11/12).

Gelderblom, Arie Jan (2000), 'Binnen en buiten. Symboliek in de emblemen van Jan Luyken', *Jaarboek van de Maatschappij der Nederlandse Letterkunde, 1999* (Leiden: Maatschappij der Nederlandse Letterkunde), 18-35.

Gelderblom, Arie Jan (2003), 'Leerzaam huisraad, vol van vuur', *De steen van Alciato. Literatuur en visuele cultuur in de Nederlanden. Opstellen voor prof. dr. Karel Porteman bij zijn emeritaat. Literature and visual culture in the low countries. Essays in honour of Karel Porteman* (Leuven: Peeters), 901-21.

Gelderblom, Arie Jan, Paasman, A. N., and Steenbeek, J. W. (eds.) (1996), *Luyken, Jan. Duytse lier* (Amsterdam: Amsterdam University Press).

Genette, Gérard (1982), *Palimpsestes. La littérature au second degré* (Collection Poétique; Paris: Éditions du Seuil).

Gertz, Michael and Sattler, Kai-Uwe (2003), 'Integrating Scientific Data through External, Concept-Based Annotations', *Proceedings of the VLDB 2002 Workshop EEXTT and CAiSE 2002 Workshop DTWeb on Efficiency and Effectiveness of XML Tools and Techniques and Data Integration over the Web-Revised Papers*, 220-40.

Gervás, Pablo*, et al.* (2006), 'Narrative Models: Narratology Meets Artificial Intelligence', <http://www1.uni-hamburg.de/story-generators/Birte/NarrativeModels_LREC06_Workshop.pdf>, accessed 2009-04-23.

Giovio, Paolo (1555), *Dialogo dell'imprese militari et amorose* (Roma: Antonio Barre).

Goodman, Nelson (1976), *Languages of Art. An Approach to the Theory of Symbols* (2 edn.; Indianapolis: Hackett).

Graham, David (1991a), 'Putting Old Wine in New Bottles. Emblem Books and Computer Technology', *Emblematica,* 5, 271-85.

Graham, David (1991b), 'The Emblematic Hyperbook. Using HyperCard on Emblem Books', in Paul Delany and George Landow (eds.), *Hypermedia and Literary Studies* (MIT Press), 273-86.

Graham, David (2004), 'Three Phases of Emblem Digitization. The First Twenty Years, the Next Five', in Mara Wade (ed.), *Digital Collections and the Management of Knowledge. Renaissance Emblem Literature as a Case Study for the Digitization of Rare Texts and Images* (DigiCULT), 13-18.

Graham, David (2005), 'Emblema multiplex. Towards a Typology of Emblematic Forms, Structures and Functions', in Peter M. Daly (ed.), *Emblem Scholarship. Directions and Developments. A Tribute to Gabriel Hornstein* (Imago Figurata. Studies, 5b; Turnhout: Brepols), 131-58.

Greene, Thomas M. (1982), *The light in Troy. Imitation and discovery in Renaissance poetry* (Elizabethan club series, 7; New Haven [etc.]: Yale University Press).

Grosso, Paul*, et al.* (2003), 'XPointer Framework. W3C Recommendation 25 March 2003', <http://www.w3.org/TR/2003/REC-xptr-framework-20030325/>, accessed 2009-04-23.

Grove, Laurence (2005), 'Emblems with Speech Bubbles', in Alison Saunders and Peter Davidson (eds.), *Visual Words and Verbal Pictures. Essays in Honour of Michael Bath* (Glasgow Emblem Studies. Special Volume.; Glasgow: Department of French, University of Glasgow), 89-103.

Guarino, Nicola (1998), 'Formal Ontology and Information Systems', in Nicola Guarino (ed.), *Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998* (Amsterdam: IOS Press), 3-15.

Guiderdoni-Bruslé, Agnès (2004a), 'La métamorphose des objets dans les emblèmes sacrés. Les marqueurs du processus emblématique', in Sagrario López Poza (ed.), *Florilegio de Estudios de Emblematica. A Florilegium of Studies on Emblematics. Actas del VI Congreso Internacional de The Society for Emblem Studies. Proceedings of the 6th International Conference of the Society for Emblem Studies. A Coruña, 2002* (Ferrol: Sociedad de Cultura Valle Inclán), 431-38.

Guiderdoni-Bruslé, Agnès (2004b), 'L'ame amante de son Dieu by Madame Guyon (1717). Pure Love Between Antwerp, Paris and Amsterdam, at the Crossroads of Orthodoxy and Heterodoxy', in Arie Jan Gelderblom, Jan L. De Jong, and Marc Van Vaeck (eds.), *Intersections* (3; Leiden; Boston: Brill), 297-318.

Gut, Ulrike*, et al.* (2004), 'Querying Annotated Speech Corpora', *Speech Prosody 2004, International Conference*.

Gyford, Phil (s.a.), 'The Diary of Samuel Pepys. Daily Entries from the 17th Century London Diary', <http://www.pepysdiary.com/>, accessed 2009-01-16.

Heckscher, William S. and Wirth, Karl-August (1959), 'Emblem, Emblembuch', *Reallexikon zur Deutschen Kunstgeschichte* (v), col. 85-228.

Heckscher, William S. and Sherman, Agnes B. (1995), *Emblematic Variants. Literary Echoes of Alciati's Term 'Emblema'. A Vocabulary Drawn from the Title Pages of Emblem Books* (AMS Studies in the Emblem, 11; New York: AMS Press).

Heinsius, Daniël (1601), *Quaeris quid sit amor* (Leiden).

Henkel, Arthur and Schöne, Albrecht (1976), *Emblemata. Handbuch zur Sinnbildkunst des XVI. und XVII. Jahrhunderts* (Stuttgart: Metzler).

Henkel, Arthur and Schöne, Albrecht (1996), *Emblemata. Handbuch zur Sinnbildkunst des XVI. und XVII. Jahrhunderts. Taschenausgabe.* (Stuttgart: Metzler).

Henning, Andreas (1999), 'Vision und Text - Bildliche Inszenierung der Gestik um 1600', *diss.sense. Zeitschrift für Literatur und Kommunikation,* Gestik. <http://www.diss.sense.uni-konstanz.de/gestik/henning-and-1a.htm>, accessed 2009-04-23.

Hockey, Susan (1996), 'Creating and Using Electronic Editions', in Richard Finneran (ed.), *The Literary Text in the Digital Age* (Ann Arbor, [Mich.]: University of Michigan Press), 1-21.

Hockey, Susan (2000), *Electronic Texts in the Humanities: Principles and Practice* (Oxford University Press).

Höltgen, Karl Josef (1986), *Aspects of the Emblem* (Problemata Semiotica; Kassel: Edition Reichenberger).

**Hooft, P. C*., et al.* (1983), *Emblemata amatoria = afbeeldinghen van minne = emblèmes d'amour : (Amsterdam, 1611)* (Nijhoffs Nederlandse klassieken; Leiden: Martinus Nijhoff) VIII, 222.

Höpel, Ingrid (1987), *Emblem und Sinnbild. Vom Kunstbuch zum Erbauungsbuch* (Frankfurt am Main: Athenäum).

Hrachovec, Herbert and Köhler, Dieter (2002), 'Digital Hermeneutics: A New Approach to Wittgenstein's Nachlass', in Rudolf Haller and Klaus Puhl (eds.), *Wittgenstein und die Zukunft der Philosophie. Eine Neubewertung nach 50 Jahren* (Wien: öbv & hpt), 151-59.

Hrachovec, Herbert, Köhler, Dieter, and Pichler, Alois (2004), 'Wittgenstein MS115 in APE', <http://wab.aksis.uib.no/wab_115ape.page>, accessed 2009-04-23.

Hüllen, Werner (1989), 'In the beginning was the gloss', in G. James (ed.), *Lexicographers and their works* (Exeter: University of Exeter), 100-16.

Hugo, Herman (1624), *Pia Desideria* (Antwerp: Hendrick Aertssens).

Huitfeldt, Claus and Sperberg-McQueen, C. M. (2003), 'TexMECS. An Experimental Markup Meta-Language for Complex Documents', <http://decentius.aksis.uib.no/mlcd/2003/Papers/texmecs.html>, accessed 2009-02-27.

Hummelen, W. M. H. (1992), 'Het tableau vivant, de "toog", in de toneelspelen van de rederijkers', *Tijdschrift voor Nederlandse Taal- en Letterkunde,* 108, 193-222.

Hunter, Jane (2006), 'Scientific Publication Packages – A Selective Approach to the Communication and Archival of Scientific Output', *International Journal of Digital Curation,* 1 (1), 33-52.

Huygen, Pieter (1689), *De beginselen van Gods koninkryk in den mensch uitgedrukt in zinnebeelden.* (Amsterdam: Wed. P. Arentsz).

**Hyvonen, Eero*, et al.* (2007), 'Semantic Kalevala—Accessing Cultural Content Through Semantically Annotated Stories', *Proceedings of the Cultural Heritage on the Semantic Web*.

**Hyvonen, Eero*, et al.* (2006), 'CultureSampo—Finnish Culture on the Semantic Web: The Vision and First Results', *Proceedings of the 12th Finnish Artificial Intelligence Conference STeP*, 25-36.

Ide, Nancy and Romary, Laurent (2003), 'Outline of the International Standard Linguistic Annotation Framework', *Proceedings of ACL,* 3, 1-5.

Ide, Nancy and Romary, Laurent (2004), 'A Registry of Standard Data Categories for Linguistic Annotation', *Proceedings of 4th International Conference on Language Ressources and Evaluation* (Lisbon), 135-38.

Importantia (2004), *De Nieuwe Bijbelvertaling cd-rom* (Dordrecht: Importantia Publishing).

Indurkhya, Bipin (1987), 'Approximate Semantic Transference. A Computational Theory of Metaphors and Analogies', *Cognitive Science: A Multidisciplinary Journal,* 11 (4), 445-80.

Isaac, Antoine and Summers, Ed (2009), 'SKOS Simple Knowledge Organization System Primer. W3C Working Draft 17 March 2009', <http://www.w3.org/TR/2009/WD-skos-primer-20090317/>, accessed 2009-04-23.

ISO TC 37/SC 4 (2006), 'Language Resource Management -- Feature Structures -- Part 1: Feature Structure Representation', (Geneva: ISO).

Junnila, M., Hyvonen, E., and Salminen, M. (2006), 'Describing and Linking Cultural Semantic Content by Using Situations and Actions', *Proceedings of the 12th Finnish Artificial Intelligence Conference STeP*, 14.

Juola, P. (2008), 'Killer Applications in Digital Humanities', *Literary and Linguistic Computing,* 23 (1), 73.

**Kahan, José*, et al.* (2001), 'Annotea: An Open RDF Infrastructure for Shared Web Annotations', *World Wide Web Conference 10*.

Kernodle, George R. (1944), *From Art to Theatre. Form and Convention in the Renaissance* (Chicago: University of Chicago Press).

Kintsch, W. (2000), 'Metaphor Comprehension. A Computational Theory', *Psychonomic Bulletin & Review,* 7 (2), 257-66.

Kirchner, Gottfried (1970), *Fortuna in Dichtung und Emblematik des Barock. Tradition und Bedeutungswandel eines Motivs* (Metzler Studienausgabe; Stuttgart: Metzler).

Klyne, Graham and Carroll, Jeremy J. (2004), 'Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation 10 February 2004', <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>, accessed 2009-04-23.

Köhler, Dieter (2006), 'Persistent Links for the Internet. Fundamentals and Implementation', *Literary and Linguistic Computing,* 21 (Suppl. 1), 77-86.

Koivunen, Marja-Riitta (2005), 'Annotea and Semantic Web Supported Collaboration', *Workshop on User Aspects of the Semantic Web*.

Kranich, Nancy (2004), 'The Information Commons. A Public Policy Report', (New York School of Law. Brennan Center for Justice. Free Expression Policy Project).

Kuiper, Willem (2009), 'Wiki-wijs', *Neder-L,* 2009, 0903.57. <http://www.neder-l.nl/bulletin/2009/03/090357.html>, accessed 2009-04-15.

Lacasse, Germain (2004), 'Le cinéma direct en 1906', *Nouvelles « vues » sur le cinéma québécois. ,* Hiver 2004 (1). <http://cinema-quebecois.net/01_hiver_2004/parler_lacasse_02b.htm>, accessed 2009-04-23.

Lagoze, Carl and Van de Sompel, Herbert (2008), 'ORE User Guide. Primer', <http://www.openarchives.org/ore/1.0/primer.html>, accessed 2009-04-23.

Lakoff, George and Johnson, Mark (1980), *Metaphors We Live By* (Chicago; London: University of Chicago Press).

Lancashire, Ian (1995), 'Early Books, RET Encoding Guidelines, and the Trouble with SGML', *The Electronic Scriptorium* <http://www.ucalgary.ca/~scriptor/papers/lanc.html >, accessed 2009-04-23.

Landow, George P. (2006), *Hypertext 3.0. Critical Theory and New Media in an Era of Globalization* (Parallax; Baltimore, MD [etc.]: Johns Hopkins University Press).

Langendoen, D. Terence and Simons, Gary F. (1995), 'A Rationale for the TEI Recommendations for Feature-Structure Markup', *Computers and the Humanities,* 29, 191-209.

Lavagnino, John (1995), 'Reading, Scholarship, and Hypertext Editions', *Text*, 109-24.

Lavagnino, John (2006), 'When Not to Use TEI', in Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth (eds.), *Electronic Textual Editing* (New York: MLA), 334-38.

Leerintveld, Ad M. Th. (2008), '[Review of] Learned love. Proceedings of the Emblem Project Utrecht Conference on Dutch Love Emblems and the Internet (November 2006)', *De zeventiende eeuw,* 24 (1), 138-39.

Leggett, John J. and Shipman III, Frank M. (2004), 'Directions for hypertext research: exploring the design space for interactive scholarly communication', *Proceedings of the fifteenth ACM conference on Hypertext & hypermedia*, 2-11.

Leuf, Bo and Cunningham, Ward (2001), *The Wiki Way. Quick Collaboration on the Web* (Boston, MA: Addison-Wesley Longman Publishing Co., Inc.).

Library of Congress (2003), 'METS: An Overview & Tutorial', <http://www.loc.gov/standards/mets/METSOverview.html>, accessed 2009-04-23.

Liu, Yin and Smith, Jeff (2008), 'A Relational Database Model for Text Encoding', *CH Working Papers,* A.43. <http://www.chass.utoronto.ca/epc/chwp/CHC2007/Liu_Smith/Liu_Smith.htm>, accessed 2009-04-23.

Lönneker, Birte and Eilts, Carina (2004), 'A Current Resource and Future Perspectives for Enriching WordNets with Metaphor Information', *2nd*

International Conference of the Global Word-Net Association, Brno, Czech Republic, 2004 (GWC-2004), 157-62.

López Poza, Sagrario (2000), 'Los libros de emblemas como tesoros de erudición auxiliares de la inventio', in Rafael Zafra and José Javier Azanza (eds.), *Emblemata Aurea. La emblemática en el arte y la literatura del Siglo de Oro*, 263-69.

López Poza, Sagrario (2004-2009), 'Poliantea', <http://rosalia.dc.fi.udc.es/Poliantea/index.jsp>, accessed 2009-04-23.

López Poza, Sagrario and Fernández Vales, Sandra Maria (2009), 'Catálogo de ediciones digitales de libros de emblemas y obras afines accesibles en Internet', (updated Enero 2009) <http://rosalia.dc.fi.udc.es/emblematica/CatalogoDEBOW.pdf>, accessed 2009-04-23.

López Poza, Sagrario*, et al.* (1996), 'Acceso a un base de datos de literatura emblemática a través de Internet', *Boletín de la red nacional de I+D, RedIRIS,* (38), 55-63.

Lucía Megías, José Manuel ([In print]), 'La edición crítica hipertextual. Hacia la superación del incunable del hipertexto', *Editar en el siglo XXI: nuevos retos de la edición* (Jaén).

Luijten, Hans (1996), *Jacob Cats. Sinne- en minnebeelden*, 3 vols. (Monumenta Literaria Neerlandica; Den Haag: Constantijn Huygens Instituut).

Luyckx, K., Daelemans, W., and Vanhoutte, E. (2006), 'Stylogenetics: Clustering-based stylistic analysis of literary corpora', <http://www.edwardvanhoutte.org/pub/2006/LREC06.pdf>, accessed 2009-04-23.

Luyken, Jan (1671), *Duytse lier* (Amsterdam).

Luyken, Jan (1678), *Jesus en de ziel* (Amsterdam: P. Arentsz.).

Luyken, Jan (1694), *Het menselyk bedryf* (Amsterdam).

Machan, Tim William (1994), 'Chaucer's Poetry, Versioning, and Hypertext', *Philological Quarterly,* 73 (3), 299-316.

MacMullen, W. J. (2005), 'Annotation as Process, Thing, and Knowledge. Multi-Domain Studies of Structured Data Annotation', *ASIST Annual Meeting* (Chapel Hill: School of Information and Library Science, University of North Carolina at Chapel Hill).

Manning, John (2002), *The Emblem* (London: Reaktion Books).

Marshall, Catherine C. (1998a), 'Toward an ecology of hypertext annotation', *HYPERTEXT '98: Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space---structure in hypermedia systems* (Pittsburgh, Pennsylvania: ACM), 40-49.

Marshall, Catherine C. (1998b), 'The Future of Annotation in a Digital (Paper) World', <http://www.csdl.tamu.edu/~marshall/uiuc-paper-complete.pdf >, accessed 2009-04-23.

Marshall, Catherine C. and Shipman III, Frank M. (2003), 'Which Semantic Web?' *Proceedings of the fourteenth ACM conference on Hypertext and hypermedia* (New York: ACM), 57-66.

Marshall, Catherine C. and Brush, A. J. B. (2004), 'Exploring the relationship between personal and public annotations', *Digital Libraries, 2004. Proceedings of the 2004 Joint ACM/IEEE Conference on*, 349-57.

Marshall, Catherine C., Shipman III, Frank M., and Coombs, J. H. (1994), 'VIKI: spatial hypertext supporting emergent structure', *Proceedings of the 1994 ACM European conference on Hypermedia technology*, 13-23.

Martin, J. H. (1990), *A computational model of metaphor interpretation* (Academic Press Professional, Inc. San Diego, CA, USA).

Marx, Vivien (2009), 'Wiki-Based Annotation Takes Off in 2008, But Some Say Data-Mining Tools are Lacking', *BioInform*.

Mason, Z. J. (2004), 'CorMet: a computational, corpus-based conventional metaphor extraction system', *Computational Linguistics,* 30 (1), 23-44.

McCarty, Willard (2004), 'Modelling: A Study in Words and Meanings', *A Blackwell Companion to Digital Humanities* (Oxford: Blackwell's).

McCarty, Willard (2005), *Humanities Computing* (Basingstoke: Palgrave Macmillan).

McCarty, Willard (2007), 'Beyond retrieval? Computer science and the humanities. Plenary lecture for the CATCH Midterm Event', *CATCH Midterm Event* (The Hague).

McGann, Jerome (s.a.), 'An Introduction to D.G. Rossetti', <http://nines.org/exhibits/6>, accessed 2009-04-23.

McGann, Jerome and Samuels, Lisa (2004), 'Deformance and Interpretation', *radiant textuality. literature after the world wide web* (New York, Basingstoke: Palgrave Macmillan), 105-35.

McGuinness, Deborah L. (2002), 'Ontologies Come of Age', in Dieter Fensel and Jim Hendler (eds.), *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential* (MIT Press), 171-95.

McGuinness, Deborah L. and Van Harmelen, Frank (2004), 'OWL Web Ontology Language. Overview. W3C Recommendation 10 February 2004 ', <http://www.w3.org/TR/owl-features/>, accessed 2009-04-23.

Meisner, Daniel (1623-1631), *Thesaurus Philo-Politicus. Das ist: Politisches Schatzkästlein guter Herren und bestendiger Freund* (Frankfurt am Main: Eberhard Kieser).

Miedema, Hessel (1968), 'The term 'Emblema' in Alciati', *Journal of the Warburg and Courtauld Institutes,* 31, 234-50.

Moorman, K. and Ram, A. (1996), 'The role of ontology in creative understanding', *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*.

Moss (1996), *Printed commonplace-books and the structuring of Renaissance thought* (Oxford [etc.]: Clarendon Press).

Murray-Rust, Peter and Rzepa, Henry S. (2004), 'The Next Big Thing: From Hypermedia to Datuments', *Journal of Digital Information,* 5 (1).

Noë, Hilde (2001), 'Register', <http://www.ua.ac.be/main.aspx?c=*RUUSBROEC&n=10224>, accessed 2009-04-23.

Nöth, W (1995), *Handbook of Semiotics* (Bloomington, Indianapolis: Indiana University Press).

Nowviskie, Bethany (2005), 'COLLEX. Semantic collections & exhibits for the remixable web', (Nines), 17.

Nowviskie, Bethany and McGann, Jerome (2005), 'Nines. A federated model for integrating digital scholarship', (Nines), 40.

Nunn, Remmel (2008), 'Crossroads: A New Paradigm for Electronically Researching Primary Source Documents', *Journal of Electronic Publishing,* 11 (2). <http://hdl.handle.net/2027/spo.3336451.0011.207>, accessed 2009-04-23.

O'Donnell, James J. (1992) *The Confessions of Augustine: An Electronic Edition* [online text], Stoa Consortium <http://www.stoa.org/hippo/>

Obendorf, Hartmut (2006), 'The Indirect Authoring Paradigm–Bringing Hypertext into the Web', *Journal of Digital information,* 5 (1).

Opitz, Andrea (2004), 'Indexing Emblem Books on the Internet - the Opportunities offered by TEI', in Mara Wade (ed.), *Digital Collections and the Management of Knowledge. Renaissance Emblem Literature as a Case Study for the Digitization of Rare Texts and Images* (DigiCULT), 71-78.

Palmer, Carole L. (2004), 'Thematic Research Collections', in Susan Schreibman, Ray Siemens, and John Unsworth (eds.), *A Companion to Digital Humanities* (Oxford: Blackwell), 348-65.

Pasanek, Brad (2006), 'The Mind is a Metaphor. A Database of Eighteenth-Century Metaphors of Mind', <http://mind.textdriven.com/db/search.php>, accessed 2009-04-23.

Pasanek, Brad and Sculley, D. (2006), 'Mining Millions of Metaphors', *Colloquim on the Digital Humanities* (Chicago).

Pasin, M. and Motta, E. (2006), 'Paving the way towards the e-humanities: a Semantic Web approach to support the learning of philosophy', *European Semantic Web Conference*.

Peil, Dietmar (2008), 'The Emblem in the German-Speaking Regions', in Peter M. Daly (ed.), *Companion to Emblem Studies* (20; New York: AMS Press), 187-221.

Pierazzo, Elena (2008), 'Learned Love. Proceedings of the Emblem Project Utrecht Conference on Dutch Love Emblems and the Internet (November 2006) Els Stronks and Peter Boot (eds)', *Literary and Linguistic Computing,* 23 (4), 495-97.

Piez, Wendell (2007), 'Markup schema as theory', *TEI@20*. <http://www.lib.umd.edu/dcr/events/teiconference/program.html#theory>, accessed 2009-04-19.

Plaisant, C., et al. (2006), 'Exploring Erotics in Emily Dickinson's Correspondence with Text Mining and Visual Interfaces', *Proceedings of the Joint Conference on Digital Libraries* (New York: ACM), 141 - 50.

Porteman, Karel (1977), *Inleiding tot de Nederlandse emblemataliteratuur* (Groningen: Wolters-Noordhoff).

Porteman, Karel (1983), 'Inleiding', in Karel Porteman (ed.), *Emblemata amatoria. Afbeeldinghen van minne. Emblèmes d'amour* (Reprint of: (Amsterdam, 1611) edn.; Leiden: Martinus Nijhoff).

Porteman, Karel (1996a), 'Introduction', in Karel Porteman (ed.), *Vaenius, Amorum Emblemata* (Emblem book facsimile series; Aldershot: Scolar press), 1-20.

Porteman, Karel (1996b), *Emblematic Exhibitions (affixiones) at the Brussels Jesuit College (1630-1685). A Study of the Commemorative Manuscripts* (Turnhout: Brepols).

Porteman, Karel and Smits-Veldt, Mieke B. (2008), *Een nieuw vaderland voor de muzen* (Amsterdam: Prometheus).

Prätor, Klaus (2007), 'Topologie und Navigation. Zur Bewegung in elektronischen Editionen', in Annette Seil (ed.), *Editionen - Wandel und Wirkung* (Beihefte zu Editio, 25; Tübingen: Niemeyer).

Praz, Mario (1975), *Studies in Seventeenth-Century Imagery* (2nd. edition considerably increased. Offset reprint of the edition published in 1964. edn., Sussidi Eruditi, 16; Roma: Edizioni di Storia e Letteratura).

Prud'Hommeaux, Eric and Seaborne, Andy (2008), 'SPARQL Query Language for RDF. W3C Recommendation 15 January 2008', <http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>, accessed 2009-04-23.

Ramsay, Stephen (2003), 'Special Section: Reconceiving Text Analysis: Toward an Algorithmic Criticism', *Literary and Linguistic Computing,* 18 (2), 167-74.

Ramsay, Stephen (2005), 'In praise of pattern', *TEXT TECHNOLOGY,* 14 (2), 177.

Ramsay, Stephen (2007), 'Algorithmic Criticism', in Ray Siemens and Susan Schreibman (eds.), *Blackwell Companion to Digital Literary Studies* (Oxford: Blackwell), 477-91.

Rawles, Stephen (2004), 'A Spine of Information Headings for Emblem-Related Electronic Resources', in Mara Wade (ed.), *Digital Collections and the Management of Knowledge. Renaissance Emblem Literature as a Case Study for the Digitization of Rare Texts and Images* (DigiCULT), 19-28.

Renear, A. H. (2004), 'Text Encoding', *A Companion to Digital Humanities*, 218-39.

Richards, Ivor Armstrong (1936), *The Philosophy of Rhetoric* (New York, London: Oxford University Press).

Robinson, Peter (2003), 'Where we are with electronic scholarly editions, and where we want to be', *Jahrbuch für Computerphilologie,* 5, 125-46.

Robinson, Peter (2005), 'Current issues in making digital editions of medieval texts—or, do electronic scholarly editions have a future?' *Digital Medievalist,* 1 (1).

Rollenhagen, Gabriel (1611-1613), *Nucleus emblematum selectissimorum* (Arnhem).

Rommel, T. (1996), '"So violent a metaphor." Adam Smith's metaphorical language in the Wealth of Nations', *Joint International Conference ALLC/ACH* (Bergen).

Roodenburg, Herman (2004), *The Eloquence of the Body. Perspectives on Gesture in the Dutch Republic* (Studies in Netherlandish Art and Cultural History, 6; Zwolle: Waanders).

Russell, Daniel S. (1985), *The emblem and device in France* (French Forum monographs 59; Lexington, Ky.: French Forum).

Russell, Daniel S. (1995), *Emblematic structures in Renaissance French culture* (University of Toronto romance series 0082-5336 71; Toronto [etc.]: University of Toronto Press).

Russell, Daniel S. (2008), 'The Emblem in France and French-Speaking Countries', in Peter M. Daly (ed.), *Companion to Emblem Studies* (20; New York: AMS Press), 155-86.

Saavedra Fajardo, Diego (1640), *Idea de un príncipe politico cristiano* (München).

Sahle, Patrick (2003), 'Vom editorischen Fachwissen zur digitalen Edition. Der Editionsprozeß zwischen Quellenbeschreibung und Benutzeroberfläche', *Fundus. Forum für Geschichte und ihre Quellen,* 2 (Quellen und Quellenedition im neuen Medienzeitalter).

Saunders, Alison (1988), *The sixteenth-century French emblem book a decorative and useful genre* (Travaux d'Humanisme et Renaissance, 224; Genève: Droz).

Saunders, Alison (2000), *The seventeenth-century French emblem. A study in diversity* (Travaux du Grand Siècle 18; Genève: Droz).

Schellner, K., et al. (2003), 'CULTOS: Towards a World-Wide Digital Collection of Exchangeable Units of Multimedia Content for Intertextual Studies', *Proc. of the Conference on Distributed Multimedia Systems (DMS 2003)* (Vienna: Dept. of Computer Science and Business Informatics, University of Vienna).

Schilling, Michael (1979), *Imagines Mundi. Metaphorische Darstellungen der Welt in der Emblematik* (Mikrokosmos, 4; Frankfurt am Main [etc.]: Lang).

Scholarly Publishing Office (2007), 'University Publishing in a Digital Age. Social Commentary on the Ithaka Report', <http://scholarlypublishing.org/ithakareport/>, accessed 2009-04-23.

Scholz, Bernhard F. (1979), 'Semantik oder Ontologie. Zur Bestimmung des 'Verhältnisses zur Wirklichkeit' in der neueren Emblemtheorie', *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte,* 53 (3 (okt)), 362-77.

Scholz, Bernhard F. (1984), 'Overwegingen bij de tekstsoort 'embleem' en bij het genre van Roemer Visschers 'Sinnepoppen'', *Vorm en funktie in tekst en taal: bundel opstellen verschenen ter gelegenheid van de voltooiing van het honderdste deel van het Tijdschrift voor Nederlandse Taal- en Letterkunde; red. D.M. Bakker ... [et al.]:* (Leiden: Brill), 73-106.

Scholz, Bernhard F. (1993), 'Marginalia bij het Boek der natuur', *Feit en fictie,* 1 (2), 51-74.

Scholz, Bernhard F. (2002), *Emblem und Emblempoetik. Historische und systematische Studien* (Allgemeine Literaturwissenschaft, 3; Berlin: Erich Schmidt Verlag).

Schöne, Albrecht (1993), *Emblematik und Drama im Zeitalter des Barock* (München: Verlag C. H. Beck).

**Schreiber, Guus*, et al.* (2001), 'Ontology-Based Photo Annotation', *IEEE Intelligent Systems,* 16 (3), 2-10.

Schreibman, Susan (2003), 'The Text Ported', *Literary and Linguistic Computing,* 17 (1), 77-87.

Sculley, D. and Pasanek, Bradley M. (2008), 'Meaning and mining: the impact of implicit assumptions in data mining for the humanities', *Lit Linguist Computing,* 23 (4), 409-24.

Seaborne, Andy (2003), 'RDQL - A Query Language for RDF. W3C Member Submission 9 January 2004', <http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/>, accessed 2009-04-23.

Sebastián López, Santiago (2001), *La mejor emblemática amorosa del barroco* (Collección Sielae: Sociedad de Cultura Valle Inclán).

Sherman, William H. (2008), *Used books. Marking readers in Renaissance England* (Material texts; Philadelphia, PA: PENN, University of Pennsylvania Press).

Shillingsburg, Peter L. (2006), *From Gutenberg to Google: Electronic Representations of Literary Texts* (Cambridge: Cambridge University Press).

Shipman III, Frank M. and Marshall, Catherine C. (1999), 'Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems', *Computer Supported Cooperative Work (CSCW),* 8 (4), 333-52.

**Shipman III, Frank M*., et al.* (2001), 'The visual knowledge builder: a second generation spatial hypertext', *Proceedings of the twelfth ACM conference on Hypertext and Hypermedia*, 113-22.

**Shipman III, Frank M*., et al.* (2003), 'Identifying Useful Passages in Documents Based on Annotation Patterns', *Research and Advanced Technology for Digital Libraries: 7th European Conference, Ecdl 2003, Trondheim, Norway, August 17-22, 2003: Proceedings* (Lecture Notes in Computer Science, 2769; Berlin, Heidelberg), 101-12.

Siemens, Ray (with the TAPoR community) (2005), 'Text Analysis and the Dynamic Edition? A Working Paper, Briefly Articulating Some Concerns with an Algorithmic Approach to the Electronic Scholarly Edition', *CH Working Papers,* A.36.

Silcox, Mary V. (2008), 'The Emblem in the United Kingdom and America', in Peter M. Daly (ed.), *Companion to Emblem Studies* (20; New York: AMS Press), 369-91.

Sinclair, Stéfan (2005), 'Toward Next Generation Text Analysis Tools: The Text Analysis Markup Language (TAML)', *CH Working Papers,* A.38. <http://www.chass.utoronto.ca/epc/chwp/Casta02/Sinclair_casta02.htm>, accessed 2009-04-23.

**Smeulders, A. W. M*., et al.* (2000), 'Content-based image retrieval at the end of the early years', *IEEE Transactions on pattern analysis and machine intelligence,* 22 (12), 1349-80.

Smit, W. A. P. (1957), 'The Emblematic Aspect of Vondel's Tragedies as the Key to their Interpretation', *Modern Language Review,* 52, 554.

Smith, Michael K., Welty, Christopher, and McGuinness, Deborah L. (2004), 'OWL Web Ontology Language Guide. W3C Recommendation 10 February 2004', <http://www.w3.org/TR/2004/REC-owl-guide-20040210/>, accessed 2009-04-23.

Smith, Paul J. (2006), *Het schouwtoneel der dieren. Embleemfabels in de Nederlanden (1567-ca. 1670)* (Zeven Provinciënreeks, 24; Hilversum: Verloren).

Smulyan, Susan (1999), 'Everyone A Reviewer? Problems and Possibilities in Hypertext Scholarship', *American Quarterly,* 51 (2), 263-67.

Sperberg-McQueen, C. M. (1996), 'Textual Criticism and the Text Encoding Initiative', in Richard Finneran (ed.), *The Literary Text in the Digital Age* (Ann Arbor, [Mich.]: University of Michigan Press), 37-61.

Sperberg-McQueen, C. M. (2009), 'How to teach your edition how to swim', *Lit Linguist Computing,* 24, 27-39.

Steinhart, Eric Charles (2001), *Logic of Metaphor: Analogous Parts of Possible Worlds* (Dordrecht: Kluwer Academic Publishers).

Stronks, Els (2007a), 'Churches as Indicators of a Larger Phenomenon. The Religious Side of the Dutch Love Emblem.' in Els Stronks and Peter Boot

(eds.), *Learned Love* (DANS Symposium Publications, 1; The Hague: DANS).

Stronks, Els (2007b), 'The Emblem Project Utrecht as a knowledge site', in Els Stronks and Peter Boot (eds.), *Learned Love* (DANS symposium Publications, 1; The Hague: DANS), 151-55.

Stronks, Els (2008), 'The Emblem and the Low Countries', in Peter M. Daly (ed.), *Companion to Emblem Studies* (20; New York: AMS Press), 267-89.

Stronks, Els and Boot, Peter (2007a), 'The Dutch love emblem on the Internet: an introduction', in Els Stronks and Peter Boot (eds.), *Learned Love* (DANS Symposium Publications, 1; The Hague: DANS), 1-9.

Stronks, Els and Boot, Peter (eds.) (2007b), *Learned Love* (DANS symposium Publications, 1; The Hague: DANS).

Suderman, Jan (1724), *De godlievende ziel vertoont in zinnebeelden door Herman Hugo en Otto van Veen met dichtkunstige verklaringen van Jan Suderman* (Amsterdam: Henrik Wetstein).

Swartjes, Ivo and Theune, Mariët (2006), 'A Fabula Model for Emergent Narrative', *Technologies for Interactive Digital Storytelling and Entertainment*, 49-60.

Szőnyi, György E. (2003), 'The 'Emblematic' as a Way of Thinking and Seeing in Renaissance Culture', *e-Colloquia,* 1 (1).

The Orwell Prize (s.a.), 'Orwell Diaries', <http://orwelldiaries.wordpress.com/>, accessed 2009-04-23.

Thøfner, Margit (2002), '"Let your desire be to see God": Teresian Mysticism and Otto van Veen's Amoris Divini Emblemata', *Emblematica,* 12, 83-104.

Tilstra, Johan (2007), 'The technical backbone of the Emblem Project Utrecht', in Els Stronks and Peter Boot (eds.), *Learned Love* (DANS symposium Publications, 1; The Hague: DANS), 175-81.

Treffers, Bert (1995), *Een hemel op aarde. Extase in de Romeinse barok* (Nijmegen: SUN).

Tuffield, M. M., Millard, D. E., and Shadbolt, N. R. (2006), 'Ontological Approaches to Modelling Narrative', *Proceedings of 2nd AKT DTA Symposium* (Aberdeen: AKT).

Tummarello, G., *et al.* (2008), 'A proposal for textual encoding based on semantic web tools', *Online Information Review,* 32 (4), 467-77.

Unsworth, John (1997a), 'Documenting the reinvention of text: the importance of imperfection, doubt, and failure', *Journal of Electronic Publishing,* 3 (2).

Unsworth, John (1997b), 'Creating Digital Resources: the Work of Many Hands', *Digital Resources for the Humanities* (Oxford).

Unsworth, John (2000), 'Scholarly Primitives: what methods do humanities researchers have in common, and how might our tools reflect this', *Humanities Computing: formal methods, experimental practice* (London).

**Uren, Victoria S*., et al.* (2003), 'Scholarly publishing and argument in hyperspace', *Proceedings of the twelfth international conference on World Wide Web*, 244-50.

Van Dalen-Oskam, Karina and Van Dijk, Suzan (2006), 'De Vaderlandsche Letteroefeningen online! Nieuw onderzoeksperspectief: de contemporaine receptie van vergeten schrijfsters', *Mededelingen van de Stichting Jacob Campo Weyerman,* 29, 3-18.

Van den Berg, Jörgen (1995), 'Subject retrieval in pictorial information systems', *Electronic Filing, Registration, and Communication of Visual Historical Data. Abstracts for Round Table no 34 of the 18th International Congress of Historical Sciences* (Copenhagen), 21–29.

Van Gemert, Lia (1991), 'Renaissancetragedies: emblemen op het toneel?' *De zeventiende eeuw,* 7 (2), 101-12.

Van Horik, René (2007), 'Data Curation', in Kasja Weenink, Leo Waaijers, and Karen Van Godtsenhoven (eds.), *A DRIVER'S Guide to European Repositories. Five studies of important Digital Repository related issues and good Practices* (Amsterdam: Amsterdam University Press).

Van Hulle, Dirk and Van Mierlo, Wim (2004), 'Reading Notes: Introduction', *Variants,* 2/3.

Van Raemdonck, Bert (ed.), (2004), *Virginie Loveling. In Oorlogsnood. Virginie Lovelings dagboek 1914-1918. Tekstkritische editie door Bert Van Raemdonck* (Gent: CTB).

Van Stipriaan, René (2001), 'Een warenhuis, geen pakhuis', *Taal en computer: congresnummer; [red.: Peter Debrabandere ... et al.]. Speciaal nr. van: Ned. nu* (49), 15-19.

Van Stipriaan, René (2003), 'Waar zit de winst? Een digitale bibliotheek voor de hele neerlandistiek. Enkele overwegingen bij de inrichting van de DBNL ', *Spiegel der Letteren,* 45 (4), 403-16.

Van Vaeck, Marc (1993), *Adriaen van de Vennes "Tafereel van de belacchende werelt" (Den Haag, 1635)*, 3 vols. (Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde).

Van Veen, Otto (1615), *Amoris Divini Emblemata* (Antwerp: Martinus Nutius, Joannes Meursius).

Van Veen, Otto (2006), *Amorum Emblemata (1608) - a web edition* (Utrecht: Emblem Project Utrecht).

Van Zundert, Joris and Van Dalen-Oskam, Karina (2005), 'The e-Laborate Project and the Usability of Another Textual Paradigm', *ACH/ALLC Conference 2005* (Victoria).

Vanhoutte, Edward (1999), 'Where is the editor? Resistance in the creation of an electronic critical edition', *Human IT,* (1).

Vanhoutte, Edward (2007), 'Traditional editorial standards and the digital edition', in Els Stronks and Peter Boot (eds.), *Learned Love* (DANS Symposium Publications, 1; The Hague: DANS), 157-74.

Vinken, Pierre J. (1960), 'Aan d'overkant, is 't zalig land', *Levende talen,* 205, 319-27.

Vinken, Pierre J. (1993), 'The founding of the Heckscher School in the Netherlands', in Karl Selig (ed.), *Polyanthea. Essays on art and literature in honor of William Sebastian Heckscher* (The Hague: Van der Heijden), 35-36.

Visscher, Roemer (1614), *Sinnepoppen* (Amsterdam: Willem Jansz).

Vondel, Joost van den (1613), *Gulden Winckel* (Amsterdam: Pers).

Vondel, Joost van den (1617), *Vorsteliicke warande der dieren* (Amsterdam: Pers).

Waller, R. (2003), 'Functionality in digital annotation: Imitating and supporting real-world annotation', *Ariadne,* 35 (30), 2005-04.

Warncke, Carsten-Peter (2004), *Théâtre d'Amour. Vollständiger Nachdruck der kolorierten "Emblemata amatoria" von 1620* (Köln: Taschen).

Warncke, Carsten-Peter (2005), *Symbol, Emblem, Allegorie. Die zweite Sprache der Bilder* (Köln: Deubner Verlag für Kunst, Theorie und Praxis).

Weise, Georg and Otto, Gertrud (1938), *Die religiösen Ausdrucksgebärden des Barock und ihre Vorbereitung durch die italienische Kunst der Renaissance* (Schriften und Vorträge der Württembergischen Gesellschaft der Wissenschaften. Geisteswissenschaftliche Abt., 5; Stuttgart: W. Kohlhammer).

Whitney, Geffrey (1586), *A Choice of Emblemes, and other devises, For the moste part gathered out of sundrie writers, Englished and Moralized. And divers newly devised* (Leiden: Plantijn).

Wielemaker, J., Schreiber, A. Th, and Wielinga, B. J. (2003), 'Supporting Semantic Image Annotation and Search', in S. Handschuh and S. Staab (eds.), *Annotation for the Semantic Web* (Amsterdam: IOS Press), 147-55.

Wiering, Frans (2009), 'Digital Critical Editions of Music: A Multidimensional Model', in Tim Crawford and Lorna Gibson (eds.), *Modern Methods for Musicology. Prospects, Proposals, and Realities* (Ashgate), 23-45.

Wither, George (1635), *A collection of Emblemes, Ancient and Moderne* (London: Henry Taunton).

Zafrin, Vika (2006), 'RolandHT [as|and] Corpus Study', *Literary and Linguistic Computing,* 21 (Supplementary issue), 157-67.

Zafrin, Vika (2007), 'RolandHT'. <http://wordsend.org/rht/xml/index.php>, accessed 2009-04-23.

Zafrin, Vika and Armstrong, Guyda (2005), 'Towards the electronic Esposizioni: the challenges of the online commentary', *Digital Medievalist,* 1 (1). <http://www.digitalmedievalist.org/article.cfm?RecID=1>, accessed 2009-04-23.

Zittrain, Jonathan*, et al.* (2006), 'H2O Playlist'. <http://h2obeta.law.harvard.edu/home.do>, accessed 2009-04-23.

Zöllner-Weber, Amelie (2005), 'Formale Repräsentation und Beschreibung von literarischen Figuren', *Jahrbuch für Computerphilologie - online,* 7.
Zöllner-Weber, Amelie (2008), 'Noctua literaria : a computer-aided approach for the formal description of literary characters using an ontology', (Universität Bielefeld).

Acknowledgements

This book, like others, did not come about in a vacuum. Many people contributed to it, each in his or her own way, and I would like to take the opportunity to thank the most important among those contributors. First the thesis advisors. To the fact that the thesis spans two disciplines I owe the luxury of having had two, Arie Gelderblom and Frans Wiering. I wonder whether I haven't sometimes been a rather self-willed student. I thank Arie and Frans for leaving me room to pursue my own ideas while still correcting me when there was a need to do so. I also thank them for their many, many comments on larger and smaller matters, and for their belief in the thesis in times when I doubted the final outcome. Moreover, I thank Arie for broaching the idea of writing a thesis; I thank Frans for introducing me to my present employer, the Huygens Institute. I thank my supervisor, Lia van Gemert, who over the years, even at a distance, was a stimulating presence. Without her intervention in the final stages, the book would not be finished by now.

A very special word of thanks should go to Els Stronks. As a pioneering explorer of the potential of information technology for the humanities, Els created an environment where I could use and deepen both recently acquired knowledge of Renaissance literature and the IT knowledge from my day job. More importantly, without her trust in what I could do, I might never have decided to leave that day job (which is the second best decision I ever took). I might still live in the intellectual slumber of financial computing. Our joint work on the Emblem Project Utrecht (EPU) was to be my passport into the international world of emblem digitisation. I thank that group of colleagues, too numerous to mention by name, that has made me feel very welcome.

Our work on the EPU also led to a magic season when job offers seemed to arrive every few weeks. The Huygens Institute was one of those that believed I had something to offer them. I thank former director Henk Braakhuis who, when I indicated I'd rather be a researcher than an IT policy officer, made available funding for a joint PhD project with Utrecht University. I also thank the Research Institute for History and Culture at Utrecht University for providing the other half of the necessary funds. It has not always been easy to balance work on the thesis and work on other projects at the Huygens Institute. However, if it was hard to strike that balance, it was so because the other projects were

interesting and the colleagues a pleasure to work with. I thank them for the many expressions of interest in the thesis, and sometimes of concern: they have meant much to me. Of the many colleagues who spring to mind, I will mention only two: former colleague Hanneke van Kempen suggested the idea of the EDITOR project, which developed into an important subject for this thesis, and Herman Brinkman helped me put flesh on the idea.

As to contributors from the personal sphere of life, I thank all those friends and family members who have regularly expressed interest in the thesis' progress. I want to thank my parents especially: that I could write a thesis that spans the disparate fields of technology and the humanities must be due in part to the liberal education they provided. And though this thesis arrives rather late in life, I know they have always been certain I could do this.

Finally: shortly after starting work on this thesis, one day I took a train to Hengelo. That was the best decision I ever took. Hanna, your support over these years has been more important than anything else.

Peter Boot

Curriculum Vitae

Peter Boot was born in 1961 in Alphen aan den Rijn, the Netherlands. After completing secondary education (VWO) in 1979 he studied Mathematics in Leiden. He obtained his Masters degree in 1986. After fulfilling his military service he went to work at the ICT department of a large insurance company, where he worked until 2000, lastly as a database consultant. From 1996 onwards he studied Dutch Language and Culture in Utrecht, where he specialised in Older Dutch Literature. He obtained his Masters degree in 2003 'cum laude'. He helped create the Emblem Project Utrecht and the international network that the project is embedded in.

Since 2003 he has been employed at the Huygens Institute, where he has worked as a humanities computing consultant and researcher. The research project that resulted in the present thesis was co-financed by the Huygens Institute – KNAW and the Research Institute for History and Culture (Utrecht University). Since 2008 the main project he has been working on is a digital edition of the letters of Vincent van Gogh.