

Managing and Annotating Historical Multimodal Corpora with the eHumanities Desktop

An outline of the current state of the LOEWE project 'Illustrations of Goethe's Faust'

Giuseppe Abrami¹, Michael Freiberg² and Paul Warner¹

¹ Johann Wolfgang Goethe-Universität
{abrami • P.Warner}@em.uni-frankfurt.de
Frankfurt am Main

² Frankfurter Goethe-Haus / Freies Deutsches Hochstift
mfreiberg@goethehaus-frankfurt.de
Frankfurt am Main

Abstract

Text corpora are structured sets of text segments that can be annotated or interrelated. Expanding on this, we can define a database of images as an iconographic multimodal corpus with annotated images and the relations between images as well as between images and texts. The Goethe-Museum in Frankfurt holds a significant collection of art work and texts relating to Goethe's Faust from the early 19th century until the present. In this project we create a database containing digitized items from this collection, and extend a tool, the ImageDB in the eHumanities Desktop, to annotate and provide relations between resources. This article gives an overview of the project and provides some technical details. Furthermore we show newly implemented features, explain the challenge of creating an ontology on multimodal corpora and give a forecast for future work.

1 Introduction

Art museum collections are widely varied. Factors such as finance, politics and aesthetics play a significant role in their formation. In any case, the act of assembling a collection is a historical process resulting in a group of art objects that are somehow related to each other, if in no other way than by being held in the same place. Often, a museum is devoted to a particular purpose, and this is the common ground. The collection of the Goethe-Museum in Frankfurt am Main¹ emphasizes the Age of Goethe and houses a significant collection of illustrations of Goethe's works. Among them are about two thousand drawings and prints of 'Faust' from the early 19th century until the present. Funded by the LOEWE² research focus "Digital Humanities"³ the project "Illustrations of Goethe's Faust", an interdisciplinary cooperation between the *Frankfurter Goethe-Haus* and the *Text Technology Lab*⁴ at the University of Frankfurt, aims to digitize the collection of Faust illustrations and to create a research environment for viewing and annotating the digital representations. The objective of the project is to build a framework which permits users to describe interpicture relations as indicated above and moreover relations between images and text.

This paper explains how an annotated image collection can be seen as a historical corpus and how it can be technically represented and automatically processed within the ImageDB, a module of the eHumanities Desktop⁵. The term corpus or text corpus is generally taken to mean a structured set of texts

¹<http://www.goethehaus-frankfurt.de/>

²National commitment to the Development of Scientific and Economic Excellence

³<http://www.digital-humanities-hessen.de/>

⁴<http://www.hucompute.org>

⁵<http://hudesktop.hucompute.org/>

subjected to annotations either defining text segments or the relations between them. Annotating the internal structure of "any sign aggregate (e.g., a sentence, a paragraph or a whole text) [...], we deal with intra-aggregate relations. [...] Alternatively, we may interrelate the aggregate as a whole with other aggregates and their segments. In this case we deal with inter-aggregate relations [...]"[4, p. 2]. From the point of view of semiotics an annotated image collection that links images and their constituents to the text and vice versa relates information of two different modalities. In this paper, we propose the following definition of a historical multimodal corpus: a set of intra- and interrelated signs of the same or a different modality where the relations can be annotated in a way that allows users to represent historical changes in the corpus. If we take the corpus of Faust illustrations as an example, the history of an art collection can be plotted - what it contained and when, what has been acquired or lost. Furthermore, it can be observed how the iconography of the artworks and the artists' interest for particular themes and motifs in the text change over time.

There are image database solutions offering the annotation of image interrelations (e.g. the commercial web-based solution ConedaKOR⁶) and the segmentation of images (e.g. HyperImage⁷). The combination of relation-based management and image segmentation has been worked out by Meta-Image[1]), which is integrated in the image archive Prometheus⁸. Other research projects like CLAROS⁹ or The WissKI Project¹⁰ developed effective techniques for information integration and retrieval. They rely on the CIDOC¹¹ Conceptual Reference Model (CRM)¹², an ontology for cultural heritage that will also be used in the course of our project. Although much work has been done in the field of image annotation, none of the projects, other than the eHumanities Desktop, provides as a whole the functionalities needed for the challenge of multimodal historical corpus management: a collaborative, web-based framework, management, segmentation and annotation of images, image-text-linking, multiple annotation layers, and ontology integration based on RDF.

The eHumanities Desktop is a platform independent web-based system which "allows users to upload, organize and share resources"[3]. These resources "can be [...] annotated and analyzed in various ways." [3] These ways include, for example, powerful analytical tools for grouping, searching, sorting, and filtering images or documents. Another powerful tool is the flexible and easy-to-use access permission system. This permits users to open or restrict access to their projects, and can be used to work on copyrighted material without infringement of the copyright. The ImageDB has been in existence for several years, and has been through several major iterations. It has significantly enhanced functionality compared to earlier versions, with changes incorporated based on extensive user feedback.

In section 2 we describe the technical structure of the eHumanities Desktop, in section 3 we explain the transforming of the existing metadata, in section 4 we give an overview on image segmentation and finally in section 5 we take a look at the ontological representation of images in the database.

2 The Technical Structure of Corpus Annotation

In order to build up a multimodal corpus, the integration of images, texts and their annotations in a complex network[6] is a central requirement. Typically, a large corpus is annotated by a research group rather than a single expert. Within the ImageDB, images are managed in repositories that are shared with users or groups. The data is organized in a graph-structure (Neo4J)¹³. We delimit generally two

⁶<http://www.coneda.net>

⁷<http://www.uni-lueneburg.de/hyperimage/hyperimage/>

⁸<http://www.prometheus-bildarchiv.de/>

⁹<http://www.clarosnet.org/>

¹⁰<http://wiss-ki.eu/node/23/>

¹¹International Committee for Documentation of the International Council of Museums (ICOM).

¹²<http://www.cidoc-crm.org/>

¹³<http://neo4j.org/>

kinds of objects in our data-structure: **authorities** and **resources** (figure 1).

Authorities represent a group or a user. Groups contains users and users belong to one or more groups. It is possible to have users without a group, but because most projects are collaborative, it makes sense to use groups primarily for sharing and restricting access to resources. Users can share their resources with other authorities.

Resources are represented primarily as documents¹⁴ especially, in our case, as images (*FileDocuments*). Resources can also be **repositories**, **annotations** or **annotation schema**. Figure 1 shows the structure of the master data. We can, when sharing, set permission levels on a resource (read, write, delete, grant). With repositories we can display a hierarchical structure similar to folders in a computer file system.[3]

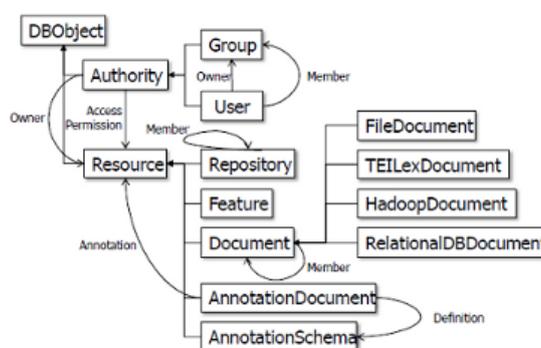


Figure 1: Class Diagramm of the master data[3]

Annotations based on **annotation schemas** are resources and are another important element in the data structure. The main challenge of the project is, after the digitalisation of the illustrations, to annotate the images to build a complete content reference. We can annotate resources based on an annotation schema. The schema is based on RDF¹⁵ and is very flexible. It is possible for the users who have the requisite permissions to create schemas or edit them to append or modify annotation fields. Each resource can then be annotated with any schema. It is also possible to annotate a resource with the same schema more than once.

3 Information Integration in Annotation Schemas

The extensible annotation architecture of the ImageDB enables knowledge representation on multiple levels. In our case, the Faust illustrations are going to be annotated with metadata and will be conceptualized in a formal ontology. On the first level the metadata lists the core information about an artwork, e.g. its title, who created it and when, the physical description, and so on. The second level is the integration of the annotated information in an ontology based on the CIDOC¹⁶ Conceptual Reference Model (CRM)¹⁷, a high-level ontology devised to describe "concepts and relations relevant to the

¹⁴Many kinds of documents are possible.

¹⁵<http://www.w3.org/TR/swbp-vocab-pub/>

¹⁶International Committee for Documentation of the International Council of Museums (ICOM).

¹⁷<http://www.cidoc-crm.org/>

documentation of cultural heritage"[8, p. 11]. Although there have been many attempts to establish a common metadata set for libraries, museums and scholars (e.g. Dublin Core), "the number of metadata vocabularies will continue to grow as individual communities seek to structure their own information for their own purposes".[2] At the early stage of the project, when a metadata standard had to be chosen, this situation cut both ways, offering a wider range of metadata standards, but also knowing that a decision could limit future interoperability. A metadata set satisfying the following constraints had to be found: XML-based syntax to ensure compatibility with the cataloguing system of the Goethe-Museum, the ability to describe objects of different kinds (e.g. drawings, books, manuscripts) and relations between them, the increasing acceptance and use in the museums sector and finally a structural affinity to the CIDOC CRM. The decision process led to the XML harvesting schema LIDO¹⁸ released in 2010 in version 1.0 by the *ICOM-CIDOC Working Group Data Harvesting and Interchange*. Being the summation of three widespread standards (CDWA Lite, museumdat and SPECTRUM) and specifically designed as an exchange format, it is intended to overcome the difficulties of sharing information with other portals or institutions. LIDO is multilingual, supports the "full range of descriptive information about museum objects"[5] and reserves a set of structural elements to express relations to other objects or thematic subjects. Most importantly, it implements the CIDOC CRM by using the concept of events: "the creation, collection, and use of an object are defined as events that have associated entities such as dates, places and actors".[5] For this reason, the later data mapping to the ontology will be much facilitated.

The second level of annotation, the knowledge representation with an ontological approach, will be the base of processing the whole of the images and annotations as a multimodal corpus. As Signore pointed out, the "metadata level cannot exploit the full richness of possible associations among different information items"[8, p. 28]. Keeping in mind all possible sorts of interrelations, a collection should be viewed as a totality rather than as a list of individual items. Furthermore, there are cultural (e.g. art historical, textual, social) contexts to be respected and annotated. By creating semantically defined relations, the ontological approach permits users to represent this richness of interconnections and knowledge accumulated by scholars. The integration of images and information in a network of semantic relations that is a multimodal annotated corpus, is the next task to be done within the project. Once accomplished, research in cultural studies can be supported by automatically processing corpora. Section 5 demonstrates how an historical development in iconography of Faust illustration could be interpreted by an intelligent agent (i.e., a computer program). Above all, we need to identify meaningful parts within the images in order to annotate the semantic content of an illustration. How this is technically realized will be explained in the next section.

4 Image Segmentation

To annotate more specifically and to connect parts of images with text (image-text-relation) we support the creation of subimages from other images. In order to allow precise annotations, it was necessary to support cropping images not only as rectangles. The first implementation of the crop-method in the ImageDB contained only the rectangle shape. For this project, we extended the functionality, so users are now able to crop images as rectangles, circles, ellipses and complex polygons.

After cropping, each new image becomes a new document owned by the current user. In addition the subimage has a back reference to its parent picture. As the subimage is an image itself it is possible to crop this subimage again to get finer detail for annotations. Each subimage, independent of the depth of cropping, knows its 'parentage', so to speak, and contains a reference to the original prototype image. Thus we make it possible to create a network of subimages within the prototype image and find inherent connections. By inverting the arrows in figure 2 one can see the way a subimage is linked to its 'parents'

¹⁸<http://www.lido-schema.org>

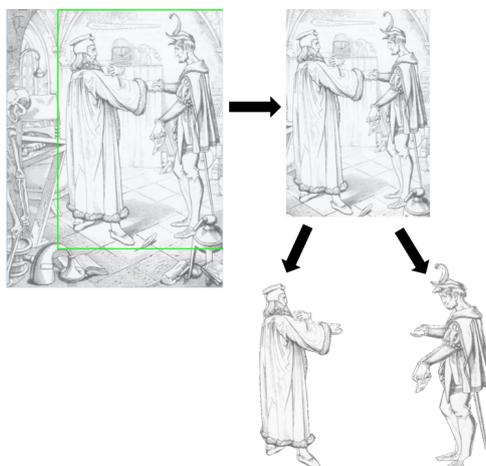


Figure 2: The crop sequence.

und finally the prototype image.

The selection of the shape will be done manually by the user, using the mouse. The first click marks the starting point of the shape. After the first click the user can see the shape according to the position of his cursor and can scale it while moving. After the last click, depending on the shape, the shape is finished and the image can be cropped. After finishing the shape-creation it is possible to move the whole shape by clicking on it. The shapes are displayed with a green border (see figure 4 for an example of a ellipse). This is realised with the JavaScript library `jsGraphics`¹⁹.

$$FORM : \langle Rectangle, Circle, Ellipse, Polygon \rangle$$

$$XYCoordinate : [x, y]$$

$$x, y \in \mathbb{N}$$

$$SHAPEFORM : XYCoordinate^*$$

$$MaxPoints(FORM) : Rectangle \rightarrow 2; Circle \rightarrow 2; Ellipse \rightarrow 3; Polygon \rightarrow n$$

For the polygon-shape there is no maximal count of points for the function (n). The user clicks around the selected shape to create the polygon. In this case the user must, on complex forms, make many points to create the shape. The cropped images are saved as new images with the backlink to the parent-image, but not with transparency. The image will be cropped with the minimal and maximal bounds of the selected shape, first as a rectangle (figure 4). The cropping is done by the HTML5 - Element: Canvas²⁰. By loading the image it is converted to a PNG. To create the transparency the shape will be drawn on the image and any pixels that are not part of the shape are deleted.(figure 3).

$$cropImage(FORM_{shape}) :$$

$$\min_{i=0}^n(XYCoordinate(x)), \min_{i=0}^n(XYCoordinate(y)), \max_{i=0}^n(XYCoordinate(x)), \max_{i=0}^n(XYCoordinate(y))$$

The method is illustrated in figure 4, using the ellipse as an example. The function requires four values to crop the image, the minimum and maximum x / y values of the shape: (1) the smallest x-value, (2) the smallest y-value, (3) the largest x-value and (4) the largest y-value.

¹⁹<http://www.walterzorn.de/jsgraphics/jsgraphics.htm>

²⁰<http://www.w3.org/TR/html5/the-canvas-element.html>



Figure 3: The cropped image, displayed in the ImageDB - Repository-View.



Figure 4: Schema to get the minimal and maximal values of the selected x/y - values.

After the shape is analysed, the picture will be cropped with a rectangle, then created with the values of the function. The values from the shape (*XYCoordinate**) will be saved as an annotation of the image. After that the image will be cropped *on the fly* when it is loaded in the selected shape-form (figure 3).

5 The Ontological Representation of Historical Changes

With the aid of the segmentation functionality of the ImageDB, the image will be divided into parts in order to single out particular motifs, persons and objects. These segments can be annotated with an iconographical topic, linked to the text or to another image. The ontological representation of these relations based on CIDOC CRM classes is illustrated in figure 5. The entities (e.g. E38 Image, E33 Linguistic Objects) are related with each other via properties (e.g. P67 refers to, P129 is about). For

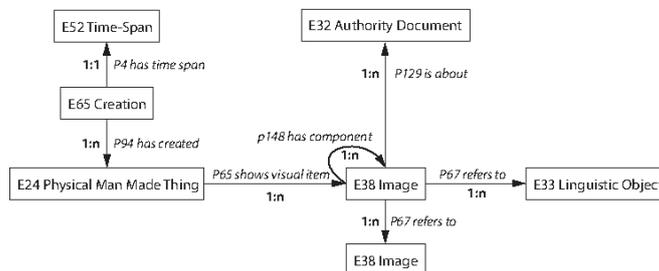


Figure 5: Ontological Representation of Interrelations

instance, an intrapictorial relation between image and sub-image is expressed by the triple "E38 Image *P148 has component* E38 Image". A link to the text would be "E38 Image *P67 refers to* E33 Linguistic Object". To refer to an iconographic theme (e.g. the poodle) of Faust illustration, an extensible topical thesaurus will be set up. It is represented as "E32 Authority Document", the CIDOC class that "comprises [...] documents that define terminology or conceptual systems for consistent use"[7]. Unlike text and image relations, it is linked by the property "P129 is about" since we are referring to a thematic subject and not to another artifact. The temporal dimension of the knowledge representation is gained by the event-based concept underlying the CIDOC CRM. Such an event is the creation or production of the object that brings into existence the material side of the artwork (E65 Creation *P94 has created* E24 Physical Man Made Thing). The limited extent of time in which this event happens is represented in the Entity E52. As a result of the annotation by means of an ontology, the image collection can be automatically processed as a historical corpus. Questions of visual traditions or iconographic changes over time can be analysed with an RDF reasoner or similar solution. If we take iconography as an

example, it has been noticed that early illustrations tend to be simple pictorial renditions of the text, in contrast to the expression on an abstract level sought by modern illustrators. The shift from a narrative to a psychological illustration could be proven by counting the text-image-links in combination with the knowledge of what has been illustrated - an object or a person. The more text items that have been illustrated, the closer an illustration is to the text. Seen from another side, it can show how the interest for a defined topic develops in the history of Faust illustrations. Such an automatic processing could lead to the discovery of relations that have as of yet gone unnoticed. Thus, when examined critically, the approach of this example could turn out to be too limited. In that case, reasoning over the ontology has to be refined in a sort of hermeneutic process.

6 Conclusion and Prospects

With the ImageDB we can upload, manage and share images and text documents to create a multimodal corpus and annotate it. Scalable access permissions allow users to share their projects without copyright infringement. To allow precise annotations, we created the possibility to crop images in varied forms. The cropped images have a recursive backlink to parent/ancestor images, including the prototype image. This functionality will be expanded in future development iterations. For instance, we plan eventually to include other multimodal content like audio and video files. The value of the applied ontology will become clearer as the size of the corpus grows. A platform combining all kinds of resources (e.g. letters, images, music, movies) with relevance to a literary work or to a specific period is an exciting vision, and a potential goal for the ImageDB. After having integrated the images and imported the corresponding metadata in the Image DB, the main goal is currently to prepare the annotation schema to connect images with the corresponding text from the Faust edition²¹.

References

- [1] Lisa Dieckmann, Anita Kliemann, and Martin Warnke. Meta-image: Forschungsumgebung für den bilddiskurs in der kunstgeschichte. *cms-journal. Computer- und Medienservice*, 35, 2012, S. 11-17, 2012.
- [2] Martin Doerr, Jane Hunter, and Carl Lagoze. Towards a core ontology for information integration. *Journal of Digital Information [online]*. April, 2006.
- [3] Rüdiger Gleim, Alexander Mehler, and Alexandra Ernst. SOA implementation of the eHumanities Desktop. In *Proceedings of the Workshop on Service-oriented Architectures (SOAs) for the Humanities: Solutions and Impacts, Digital Humanities 2012, Hamburg, Germany*, 2012.
- [4] Rüdiger Gleim, Paul Warner, and Alexander Mehler. eHumanities Desktop - An Architecture for Flexible Annotation in Iconographic Research. In *Proceedings of the 6th International Conference on Web Information Systems and Technologies (WEBIST '10), April 7-10, 2010, Valencia*, 2010.
- [5] ICOM-CIDOC Working Group Data Harvesting and Interchange. *LIDO - Lightweight Information Describing Objects*, 1.0 edition, 2010. <http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf>.
- [6] Alexander Mehler. Large Text Networks as an Object of Corpus Linguistic Studies. In Anke Lüdeling and Merja Kytö, editors, *Corpus Linguistics. An International Handbook of the Science of Language and Society*, pages 328–382. De Gruyter, Berlin/New York, 2008.
- [7] CIDOC CRM Special Interest Group (SIG). *Definition of the CIDOC Conceptual Reference Model*, 5.0.4 edition, 2011. http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf.
- [8] Oreste Signore. The semantic web and cultural heritage: Ontologies and technologies help in accessing museum information. In K. Robering (ed.): *Information Technology for the Virtual Museum*. Berlin, 2008.

²¹<http://faustedition.uni-wuerzburg.de>