Collection *Ægyptiaca Leodiensia* 9

Texts, Languages & Information Technology in Egyptology

Selected papers from the meeting of the Computer Working Group of the International Association of Egyptologists (Informatique & Égyptologie), Liège, 6-8 July 2010

Stéphane POLIS & Jean WINAND (eds.)

With the collaboration of Todd GILLEN

Presses Universitaires de Liège
2013

The MEKETREpository

A Collaborative Web Database for Middle Kingdom Scene Descriptions

Christian MADER, Bernhard HASLHOFER & Niko POPITSCH

Vienna

1. INTRODUCTION

Middle Kingdom (MK) tombs and tomb decorations offer a variety of complex and multi-layered information. However, comprehensive publications that deal with MK scene representations, iconography and scene development are still rare. There is especially a lack of literature performing comparative research on iconography in the MK. In 1922 Luise Klebs published the first assessment of MK representations[1] and in 1978 the last volume of Jacques Vandier's "Manuel d'archéologie égyptienne"[2] appeared. In this publication, he grouped various scenes according to their contents and tried to trace chronological developments in style and iconography. A large quantity of publications dealing with the art of the MK has appeared during the last 40 years, forming an excellent basis for further (comparative) research on scene iconography. However, this tremendous amount of literature poses great challenges to scientists and scholars in the domain of art-history, both in terms of gaining access to publications and keeping up with material being continuously published.

Within the scope of the MEKETRE project, we are developing a specialized software application that will enable scholars to describe MK scenes and scene fragments in a collaborative manner and provide comprehensive search and discovery mechanisms for accessing these items. This application will be referred to as the MEKETREpository, a digital repository of MK art items. It will allow Egyptologists to describe MK items in a structured way and aim at establishing vocabularies for that domain in order to support and improve communication among scholars. The repository and the collected data is already publicly accessible on the Web (http://www.meketre.org) and thus it seeks to make a valuable contribution to future (comparative) research on the MK.

2. METHODOLOGY

Information technology can support Egyptological research in various ways. In the MEKETREpository our focus is on the following four use cases (requirements):

1. *Die Reliefs und Malereien des mittleren Reiches. (VII.-XVII. Dynastie ca. 2475-1580 v. Chr.) Material zur ägyptischen Kulturgeschichte*, Heidelberg, 1922.
2. *Manuel d'archéologie égyptienne.* Tome IV. *Bas-reliefs et peinture. Scènes de la vie quotidienne. I^re partie: Les tombes*, Paris, 1964; Tome V. *Bas-reliefs et peinture. Scènes de la vie quotidienne. 2^ème partie: Élevage, chasse, pêche, navigation*, Paris, 1969; Tome VI. *Bas-reliefs et peinture. Scènes de la vie agricole à l'Ancien et au Moyen Empire*, Paris, 1978.

(1) Storage and retrieval of descriptive metadata for each item.
(2) Collaborative annotation of art items to stimulate cooperation between departments and individual researchers.
(3) Tools for developing and maintaining domain-specific vocabularies.
(4) Assignment of detailed bibliographical references to art items and their details.

When building software for a very specialized group of users, it is not only the implementation of these requirements that will at the end attract the user to the system. With the rise of the so-called "Web 2.0"[3] and wide patronage of websites like Youtube and Facebook, user-supplied content and social networking have attracted users who previously did not have any interest in using the Web. This gives us the opportunity and motivation to build an application that allows users to collaboratively collect and describe fragments of MK tomb decorations. Providing a Web application has advantages for inexperienced users because no local installation is required and the system is accessible from everywhere (when connected to the Internet) and from every device capable of running a Web browser (e.g. also from mobile phones). Since many users are already accustomed to using common Web applications, the technical competence for using the MEKETREpository application is expected to be quite low.

2.1. *Data Model*

The MEKETREpository enables the detailed description of two-dimensional *art items* (cf. Figure 1). Users can create new art items in the repository and may specify:

- the category this item belongs to (see §2.1.1);
- the tomb it belongs to;
- the current location (in situ or some other site, e.g., a museum);
- the position of the item in the tomb (plain language form possible);
- the execution style of the item (e.g., relief, painting, drawing).

Additionally, the user can specify detailed information about *tombs*:

- the necropolis where the tomb is located;
- the tomb number;
- the date;
- the tomb owner.

For both art items and tombs, it is possible to additionally specify the following information:

- a description in plain language;
- keywords (see §2.1.2);
- images depicting the item;
- annotations (see §2.1.3).

By providing this information, it is possible to connect art items to the tombs they originally belong to and describe them in plain language and terms taken from controlled vocabularies.

3. A definition and comprehensive explanation of the term can be found in O'Reilly, T. 2007. What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, in: *International Journal of Digital Economics* 65, p. 17-37.

2.1.1. Categorization

There exist many categorization schemes in the Egyptological domain, designed and used by different researchers. In the scope of the MEKETRE project, a new categorization scheme for MK art items will be developed. While this categorization scheme is envisioned to have an independent existence, it will also be linked and cross referenced to Egyptological schemes used elsewhere.

The MEKETRE categorization scheme is a taxonomy developed in a bottom-up, collaborative fashion by the researchers working on the MEKETRE project. Bottom-up means that whenever a researcher enters a new item and requires a new category that is not available in the current scheme, she is able to add this category easily. Depending on the scenes and activities depicted, items can be assigned to one or more categories. The categorization scheme is capable of describing the items contained in the repository but is no generic scheme for MK scenes. Such a generic scheme may be developed in a subsequent process.

2.1.2. Thesaurus

When describing an item, it is beneficial to use terms from controlled vocabularies in order to manage the available data more efficiently. By means of a thesaurus it will be possible to consider the semantic relationships between terms in search and retrieval. Search results could, for example, automatically include items that are tagged with more general terms than the one actually searched for and thus adjust the ranking of the results to produce more useful output. It is important to note that in a single research domain there may exist numerous thesauri, each developed by a different group of researchers, which in their most fundamental form are used to consistently describe items of research within a project. Researchers outside the group often don't have access to such a thesaurus and therefore are unclear about the meaning of its terms. Therefore it is essential to use existing standardized thesauri wherever possible to describe items. Hence, we use the "Multilingual Egyptological Thesaurus" (MET)[4] as a basis for the MEKETRE thesaurus. In cases where the MET does not provide appropriate terms, we allow users to add their terms and the relationships between those terms to a MEKETRE-specific thesaurus. By making this thesaurus publicly available on the Web and linking its entries to the MET it is expected to be a valuable contribution to the Egyptology domain and a complement to the MET.

2.1.3. The Concept of Annotations

In the user interface of the MEKETREpository it is possible to add so-called "details" to an art item or a tomb. A detail is a metadata description for one special aspect of an item. In the computer science community, the term "annotation" has been coined for that kind of architectural pattern. In the MEKETREpository application, each item can have an unrestricted number of annotations and annotations can also refer to each other. Thus it is possible to define relationships between different items like "this scene is contained in the picture of this tomb's wall" or classify parts of an image on a more detailed level.

We carefully designed the user interface of the MEKETREpository to facilitate the annotation of items and to allow for the quick addition of as many annotations to an item as the user desires. The screenshot depicted in Figure 1 illustrates the annotation of a scene item.

4. http://www.ccer.nl/apps/thesaurus/index.html.

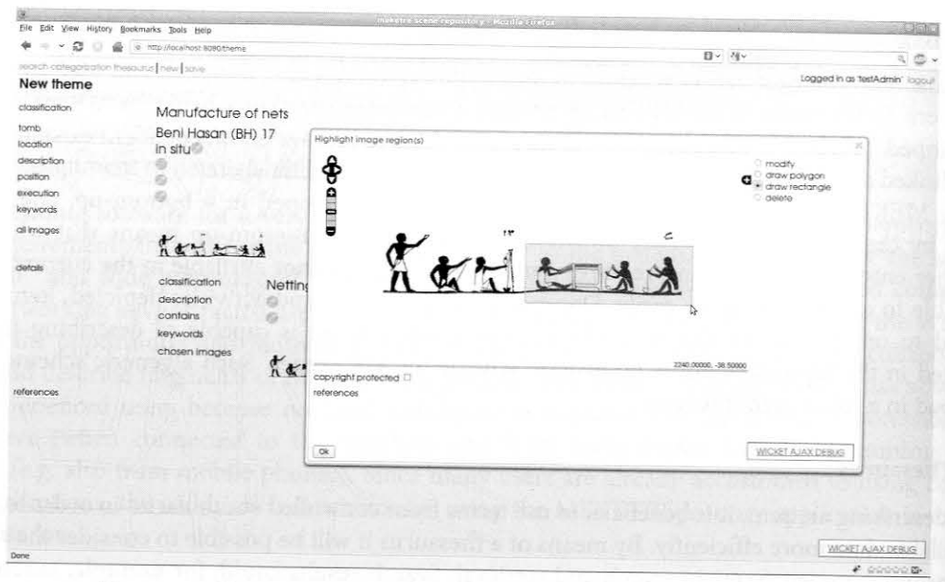

Figure 1. Adding an annotation to a scene item in the MEKETREpository[5]

2.2. *MEKETREpository Items as Linked Data on the Web*

The World Wide Web (most commonly just referred to as "the Web") was originally designed to support browsing through a large amount of interlinked documents. Although modern Web applications hide this fact to a great deal, working with the Web is essentially a sequence of sending a request for a document (i.e., entering an address or clicking a link) and getting an appropriate response (the actual document or some error message). The documents received by browsing the "traditional" Web are perfectly readable to human users but they are virtually useless for automated processing by computer systems. Efficient data processing requires structured typed data and the HTML markup language most Web pages are written in, does not fulfill this requirement.

Recently, Linked Data[6] has evolved as a method of exposing and linking structured data on the Web. It forms the foundation of the "Web Of Data". Just as the traditional Web serves HTML documents for human users, Linked Data is a method for serving machine-readable RDF data. A Linked Data entity holds data in a machine-readable format (RDF[7]). Furthermore, as the name implies, Linked Data resources are interlinked with resources from other sources, just as a usual Web page contains links to other pages[8]. Other than simple links between regular Web pages, links between RDF items (resources) are typed, i.e., each link has a given semantics and further describes the resource.

A Linked Data resource can describe virtually anything. It can be for example a future event, a concept, or even a feeling. More trivially it may describe an item of a specific domain, for example a person and its properties (e.g., name, age, birthday), a book (e.g., title, author, year of publication) or, in the case of the MEKETREpository, an art item, categorization scheme or thesaurus. These resources are published on the Web, identified by their unique address (URI – Universal Resource Identifier), and can easily be accessed and linked to other resources in the same or in other Linked Data sets. This

5. In this example, the user has selected the three people on the right of the scene in order to add a description of this particular detail. An unrestricted number of annotations of a scene are supported by our application.
6. http://www.w3.org/DesignIssues/LinkedData.html.
7. Resource Description Framework, a W3C standard model for publishing data on the Web. A good source for further reading is http://www.w3.org/RDF/, the complete specification can be found at http://www.w3.org/standards/techs/rdf#w3c_all.
8. An introduction can be found at http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial.

makes it possible, for instance, to link the Linked Data resource that represents an art item stored in the MEKETREpository with another online available resource (such as a Wikipedia article describing the discoverer of this art item) that is also exposed as a Linked Data resource (as is the case for Wikipedia articles in the DBpedia datasets).

In this fashion, data sets that are available as Linked Data become interlinked over time, forming a huge network of linked resources that can be exploited to learn about related information. Our application could, for example, follow a link to DBpedia, automatically retrieve biographical information about a particular person and display this data next to the depictions of the relevant art item.

2.2.1. Linked Data and the MEKETREpository

We intend to adopt the principles of Linked Data in our MEKETREpository and plan to (i) reuse data from existing Linked Data sources and (ii) publish the data available in MEKETRE as Linked Data on the Web. This will in turn allow other applications to reuse the data collected in MEKETRE by simply addressing the (interlinked) MEKETRE items by their URIs. More specifically we want to publish structured controlled vocabularies using the Simple Knowledge Organization System (SKOS[9]). For describing item metadata we will investigate the applicability of existing metadata standards such as Dublin Core (DC[10]), Friend-of-a-Friend (FOAF[11]), and others. One of the goals of the MEKETREpository application is to contribute to the Linked Data cloud[12] and provide interoperability with other data sources following the same standards.

2.3. *Copyright Issues*

Attaching media objects such as images to collected art items or tombs is useful for users to obtain a quick overview of the material or of interesting items of a search result. However, some pictures of MK scenes or scene fragments that should be added to the MEKETREpository's database are only available in books that are still protected by copyright. Hence the MEKETREpository allows the creation of items without any attached image. Such items are basically metadata descriptions that characterize a real-world art item but lack its depiction.

As an alternative, it is possible to tag an image as copyrighted when uploaded. This prevents image access for users who are not logged into the MEKETREpository application or do not have the right to view copyrighted material. As a general rule-of-thumb we propose to upload public domain material when available. Since any number of images can be uploaded, it is also possible to upload copyrighted material together with non-copyrighted material illustrating the same item.

2.4. *Accessing the Repository*

We designed the MEKETREpository for two different types of users. The first type relates to human users like researchers and students. Researchers will have read- and write-access to the repository and provide material along with a qualified description of the content. Students may browse the collected data for comparative research purposes without contributing to the repository. For both researchers and students we provide an easy-to-use Web application for accessing the stored items and performing their work.

Machine users represent the other type of user for the MEKETREpository: other systems connected to our repository via the Web with read-access to the stored data. Since we provide Linked Open Data (as described in §2.2), which is a novel approach for publishing machine-readable data,

9. http://www.w3.org/2004/02/skos/.
10. http://dublincore.org/.
11. http://www.foaf-project.org/.
12. To get an impression, see http://richard.cyganiak.de/2007/10/lod/.

our data can easily be queried and integrated with other data sets, potentially originating from domains outside of Egyptology.

2.5. *Long-term Archiving*

Preserving the collected data using a Long-Term Archiving Solution (LTAS) is another aim of the project. We are not going to develop a new solution, but focus on how to integrate the MEKETREpository with an already existing LTAS.

The University of Vienna hosts a digital asset management system with long-term archiving functions called PHAIDRA.[13] It is based on the popular Fedora Commons Repository Software[14] and can hold any kind of digital object, available worldwide around the clock with continual citability. PHAIDRA also uses metadata to store the content but its metadata standards are fixed and not easily tailorable to domain-specific needs. In the context of the MEKETREpository project, we use PHAIDRA as an additional storage solution. The data stored in the MEKETREpository is periodically replicated to PHAIDRA.

PHAIDRA is a solution for general archiving purposes whereas the MEKETREpository is custom-tailored to be used by Egyptologists. Since MEKETRE is an interdisciplinary project we are working hand-in-hand with our colleagues from the Institute of Egyptology to provide them with the tool they need to perform their research. The workflow in the MEKETREpository is optimized for finding, browsing and comparing scenes. Our strategy is to combine the intuitive user interface of the MEKETREpository with the long-term data archiving capabilities of PHAIDRA. Since the MEKETREpository also works with common metadata standards, the conversion to PHAIDRA datasets should be straightforward for most entries. The two systems will exist side-by-side and can be queried independently. It is important to note that the MEKETREpository is designed to exist on its own but we decided to additionally replicate the data to PHAIDRA to make use of an existing well-proven long-term archiving repository with relatively low effort expenditure.

2.6. *System Architecture*

Since the MEKETREpository application is expected to be used and maintained beyond the project's three-year limit, it is important to implement it using industry-standard components that are available under an open source license. The system is designed as a three-tier application consisting of a persistence layer, an intermediate service layer and a user interface layer. The main programming language is Java and the user interface layer is implemented using the Apache Wicket Web framework[15]. Operations on the data are passed to the service layer which in turn utilizes the persistence layer to save data to or obtain data from a relational database. The service layer is also responsible for providing Web services for other systems querying data. The persistence layer uses the Java Persistence API (JPA) to describe the data model. Hibernate is used as the JPA provider together with a MySQL database storing the actual data. Uploaded images of items and tombs are stored in the file system and are managed by an IIP (Internet Imaging Protocol[16]) Image Server which is a FCGI Application running on an Apache webserver. For fulltext search we make use of the Solr search server that runs as a separate process. Managing literature is done using an existing Web application (refbase[17]). However, by utilizing refbase's OpenSearch[18] interface, it is possible to edit literature

13. https://phaidra.univie.ac.at/.
14. http://fedora-commons.org/.
15. http://wicket.apache.org/.
16. The specification can be downloaded from http://iipimage.sourceforge.net/IIPv105.pdf.
17. http://www.refbase.net/.
18. http://www.opensearch.org.

references directly from the MEKETREpository's user interface. An overview of the system and the involved components is shown in Figure 2.

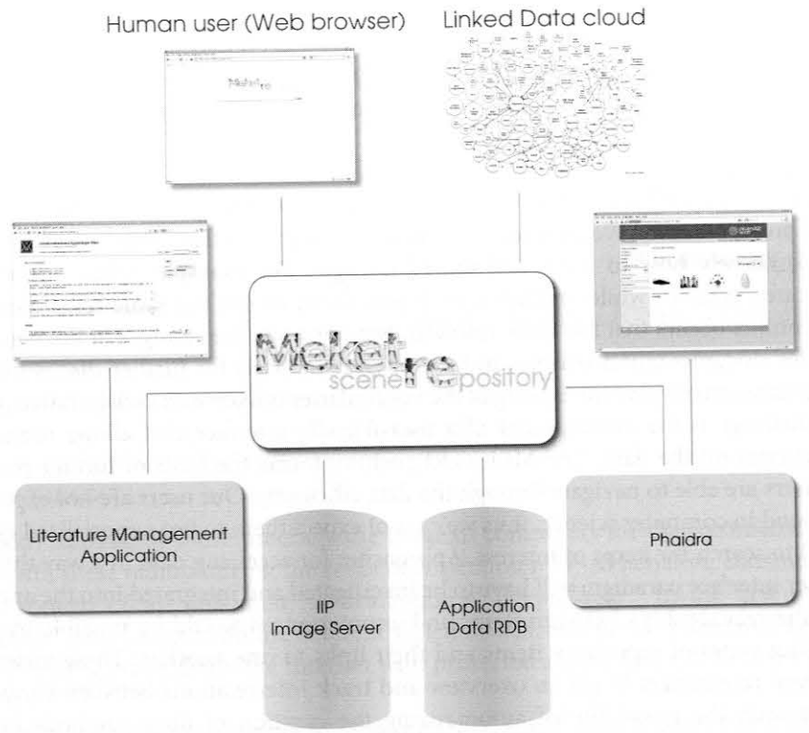

Figure 2. The MEKETREpository and its components[19]

3. RELATED WORK

The Oxford Expedition to Egypt (OEE)[20], which is affiliated academically to Linacre College, University of Oxford, created a scene details database. In the years from 2003 until 2006 the expedition collected data of scene details in Old Kingdom (OK) monuments. The database[21] went online in 2007 and is now freely available. It has been developed in close collaboration with the Archaeology Data Service (ADS)[22]. Similar to the categorization scheme developed during the work on the MEKETREpository's content, the OEE database uses a hierarchical scheme to organize the data into "Themes", "Scene types" and "Scene details". We plan to adopt and integrate already existing data modeling practices developed at ADS and give feedback in order to ultimately build a basis for publishing scene descriptions as Linked Data on the Web.

Scenes and their accompanying texts in OK tombs are also covered in the scope of the Leiden Mastaba Project (LMP)[23], also known as "MastaBase". In contrast to the MEKETREpository the data is not published directly on the Web for public access but purchasable on CD-ROM. There are also restrictions on the types of computer systems that may use this database, further restricting the possible user base of these data.

19. Both the Literature Management Application and Phaidra can be queried independently from the MEKETREpository using a Web browser.
20. http://www.oxfordexpeditiontoegypt.com/index.html.
21. http://ads.ahds.ac.uk/catalogue/archive/oee_ahrc_2006/.
22. http://ads.ahds.ac.uk/.
23. For a project summary see http://www.peeters-leuven.be/boekoverz.asp?nr=8170.

4. CONCLUSIONS AND FUTURE WORK

The MEKETREpository is a software solution capable of describing MK two-dimensional art items at an unrestricted level of detail. The contained data is published as Linked Data on the Web utilizing controlled vocabularies.

We aim at making it as easy as possible for scholars to enhance existing and develop new vocabularies. Since these vocabularies are published as Linked Data on the Web, it is essential to provide mappings to other existing vocabularies (e.g., DBpedia), so a strategy must be developed to create these mappings.

Since we expect the MEKETREpository to be used by more than one researcher concurrently, we will investigate possibilities for Web-based collaborative thesaurus editing. This brings up a whole bag of new challenges, namely how to track and record changes. For example, when one user deletes or reorganizes a categorization while another user is just about to use the same term/concept in a new annotation, a conflict occurs that has to be resolved. Furthermore, the changes in the vocabularies need to be tracked for documentation reasons and to provide the basis for further discussion. Annotating items is done collaboratively and the editing of the vocabularies is likewise a collaborative process.

Another challenge is the development of a user-friendly interface that allows users to formulate sophisticated queries on the data. The MEKETREpository forms the basis of further research, so it is essential that users are able to navigate through the data efficiently. Our users are not expected to have a strong background in computer science, thus we cannot expect them to use a specialized query language (e.g., SPARQL) to search for items of interest. Approaches for accessing data in a way that conforms to the Web 2.0 user interface paradigm will have to be investigated and integrated into the application.

Based on the relationships between items and annotations it would be possible to generate new visually appealing views of repository items and their links to one another. These views should help both scholars and researchers to get an overview and track interrelations between various items. We will further research the possibility of automatizing the creation of these methods of information visualisation.