EDEN – An Epigraphic Web Database of Ancient Inscriptions

Martin Scholz[1], Marvin Holdenried[2], Boris Dreyer[2], and Guenther Goerz[1]

[1] *Digital Humanities Research Group*
[2] *Chair of Ancient History*
University of Erlangen-Nürnberg
{firstname}.{lastname}@fau.de

Keywords: semantic web, archaeology, 3D, ancient history, ontologies, cidoc crm

1 Introduction

The Epigraphische Datenbank Erlangen-Nürnberg / Epigraphic Database Erlangen-Nürnberg (EDEN)[1] is an evolving online database of ancient Greek inscriptions from cities of today's western Turkey: Metropolis (in Ionia), Magnesia ad Maeandrum and Apollonia ad Rhyndacum.

The development of EDEN started 2012 at the University of Erlangen-Nürnberg in close collaboration between ancient historians, archaeologists and computer scientists. Contentwise, the database is maintained by the Chair of Ancient History. The Digital Humanities Research Group is responsible for the technical framework. The cooperation assures the rapid evolution of the database despite the tight budget.

At the time of writing, the database contains nearly 600 Greek an Latin inscriptions of Hellenistic and Roman emperial period enriched with metadata.

[1]see *wisski.cs.fau.de/eden* and [HRS14]

2 The WissKI System

The technical base for EDEN is the virtual research environment WissKI[2] that is being developed and maintained by the Digital Humanities research group of the University of Erlangen-Nürnberg, the Germanic National Museum in Nuremberg and the Zoological Research Museum Alexander Koenig in Bonn (all located in Germany).

The WissKI system is built upon the content management system (CMS) Drupal[3] and, as such, is purely web-based.

Like Wikipedia, the web interface serves as the one and only point of access to the database. Data acquisition, editing and presentation are done solely via the internet, usually and preferably with a web browser.

As a full-featured CMS, WissKI also offers functionality for managing web sites, from creating simple web pages like an "About us" section to fine-grained user roles and permissions. EDEN, for example, currently knows four roles: unauthenticated users (the public), authenticated users, editors, and administrators; each having different access rights.

WissKI allows data to be semantically enriched using Semantic Web techniques.[4] With RDF[5], knowledge is represented in triples[6], forming a loosely linked net composed of nodes and edges instead of rather rigidly defined records. Through the use of URIs, RDF has innate facilities to combine one's own local data with other data world-wide. Furthermore, OWL[7] ontologies provide flexible and modular ways of categorizing the data.

However, practitioners in both fields are still mostly unfamiliar with net-based visualization of data and a lot of state-of-the-art visualisation techniques do not scale well when applied to larger datasets. Also, the complexity of an ontology with hundreds of classes and properties constitutes a major obstacle for those not familiar with the ontology's structure.

In order to provide users with a well-known interface, WissKI hides the net-based data storage and the ontology's complexity and simulates a record-oriented approach. This is done by defining mappings between paths through the data, so-called ontology paths, and fields of a record.[8]

[2] See *wiss-ki.eu* and [SG12]

[3] *http://www.drupal.org*

[4] *http://www.w3.org/2001/sw/wiki/Main_Page*, an introduction to Semantic Web and the techniques mentioned hereafter can be found in [HKR10]

[5] Resource Description Framework. RDF is the standard metadata format of the Semantic Web; see *http://www.w3.org/RDF/*

[6] An RDF triple can be seen as a simple proposition of the form *subject predicate object*

[7] *http://www.w3.org/OWL/*

[8] A more detailed description of the path approach can be found in [FR12].

3 Construction and Structure of the Epigraphic Database

The digital, web-based nature of EDEN together with the applied techniques offer many new possibilities. The following sections describe two design issues of the database that illustrate these possibilities.

3.1 One Database for Multiple Disciplines

Contentwise, the EDEN database is edited by ancient historians from University of Erlangen-Nürnberg in cooperation with and archaeologists from Erlangen, Vienna and Turkey. One aim of the initiators is to compile the different views and requirements of both ancient historians and archaeologists into one database, building a bridge between the disciplines and hopefully leading to synergy effects.

EDEN contains information crucial for both disciplines: the textual or immaterial part of the inscription — which is the main focus of ancient historians — is addressed as detailed as the physical aspects of the inscription carriers (stones, coins, . . .), which are essential for archaeologists. This combination is invaluable for both disciplines, since the inscription and its carrier must always be analyzed together.

The flexible data schema allows for extending EDEN quite easily. New categories of metadata can be added almost on the fly and "marginal pieces of information" can be enriched with new metadata, which make it possible to shift or multiply the focus of the database. EDEN already has undergone such focus creation: starting with two foci, the inscription and its carrier, the formerly marginal person and place references are gaining more and more focus. E.g. descriptions and geospatial information have been and still are being added, making each person and place a hub for accessing related inscriptions, etc. This means that the database can be readily adopted to support new research questions.

3.2 Knowledge (re)presentation

As pointed out, the underlying WissKI system uses RDF and OWL for knowledge representation. EDEN defines a domain ontology that is based on the Erlangen CRM[9], an OWL DL implementation of the CIDOC CRM[10], and extends it with its own specialized classes and properties where needed.

The information of each object of interest (inscription, carrier, person, place, etc.) is assembled on a dedicated page. The layout is based on Wikipedia: The center displays free text while tabular data and images are shown on the right side. Where reasonable, metadata and pieces of free text referring to other objects are displayed as links, so that it is easy to browse the database from record to record (as known from Wikipedia). There are various types of free text in EDEN. An entry for an inscription contains at

[9] *http://erlangen-crm.org*
[10] see *http://cidoc-crm.org* and [CDG^{+}11]

least the original greek or latin transcription. However, most entries are accompanied by further elements, such as translations into German and English, basic information on finding place, editions and measurements and a traditional scientific commentary. Each of the text segments is a object in its own right and may be reused in other records. The same applies tabular metadata, where field autocompletion lets the user select existing entries. Inconsistencies due to copy-paste errors or spelling variants are thus avoided and the data gets interconnected more tightly. WissKI also features a graphical editor that allows for semantic annotation of the free text segments. The annotations are interwoven with the text and RDF triples may be created from them. EDEN currently annotates inscriptions, persons, places, publications and dates and automatically adds the links between these entities to the tabular data.

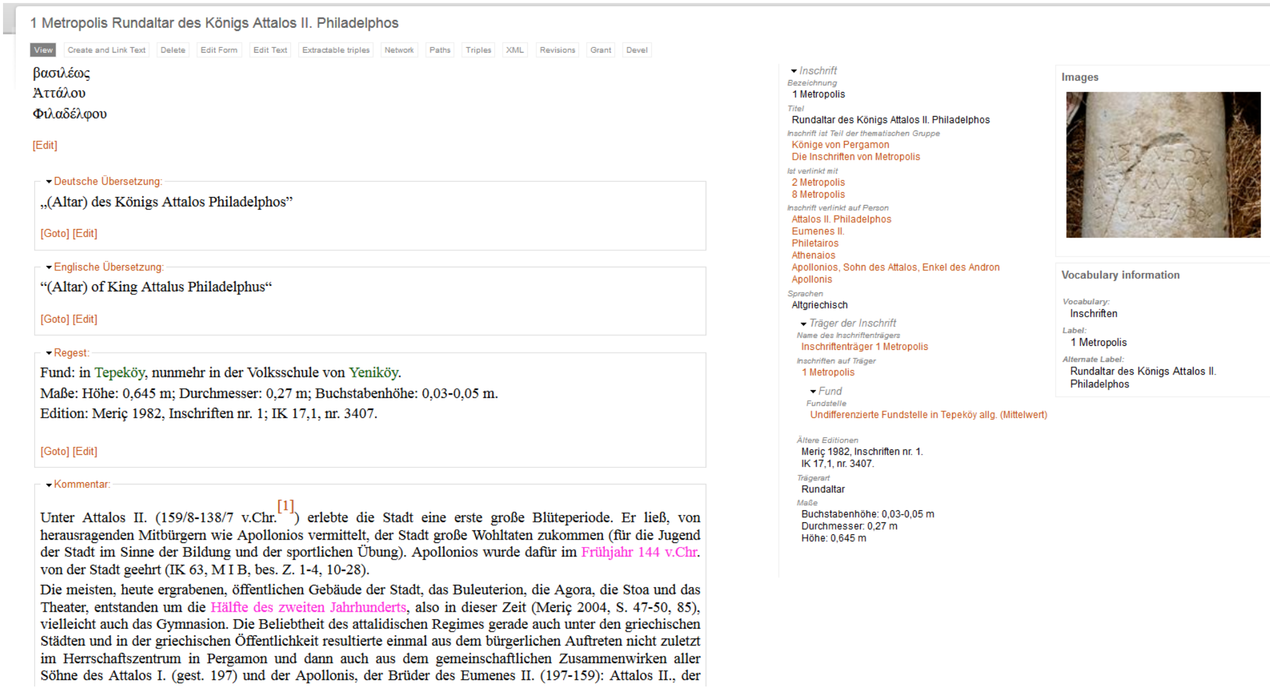

Figure 1: Inscription 1 Metropolis, *Rundaltar des Königs Attalos II. Philadelphos*, *http://wisski.cs.fau.de/eden/content/ecrm_E34_Inscription01e6*: The information describing the inscription is assembled to a record.

4 From a Database to a Virtual Excavation Site

In 2012, 3D scans of various excavation sites in Metropolis in Ionia have been produced. Geometries of large buildings like a bath complex, a stairway and many others have been assembled (as surfaces, not mere point clouds). The goal is to tightly integrate the 3D models with the epigraphic database, in order to eventually enable users to hover through a joint virtual version of the Metropolis excavation site. In the following paragraphs we want to provide a short résumé of the work done so far, the challenges encountered and possible solutions.

WissKI is a web-based system and therefore the 3D models have to be rendered on web pages. A couple of years ago this was quite an obstacle, as there were no standardized ways to use the efficient rendering features of the graphics card through a web browser. In 2011, the specification of WebGL[11] closed that gap. WebGL enables efficient rendering of three dimensional content on web pages and is widely supported nowadays. WissKI uses the three.js library[12] that builds on top of WebGL and provides a simple and yet powerful API with lots of functionalities, like predifined motion controls and support for different file formats. Various code examples make it easy to set up a scene. Those with some expertise in 3D models may use plenty of parameters for lights, colors, materials, etc. for fine-tuning. The layout of EDEN was altered such that the 3D model, if existent, is shown in the center in the same place as free text.

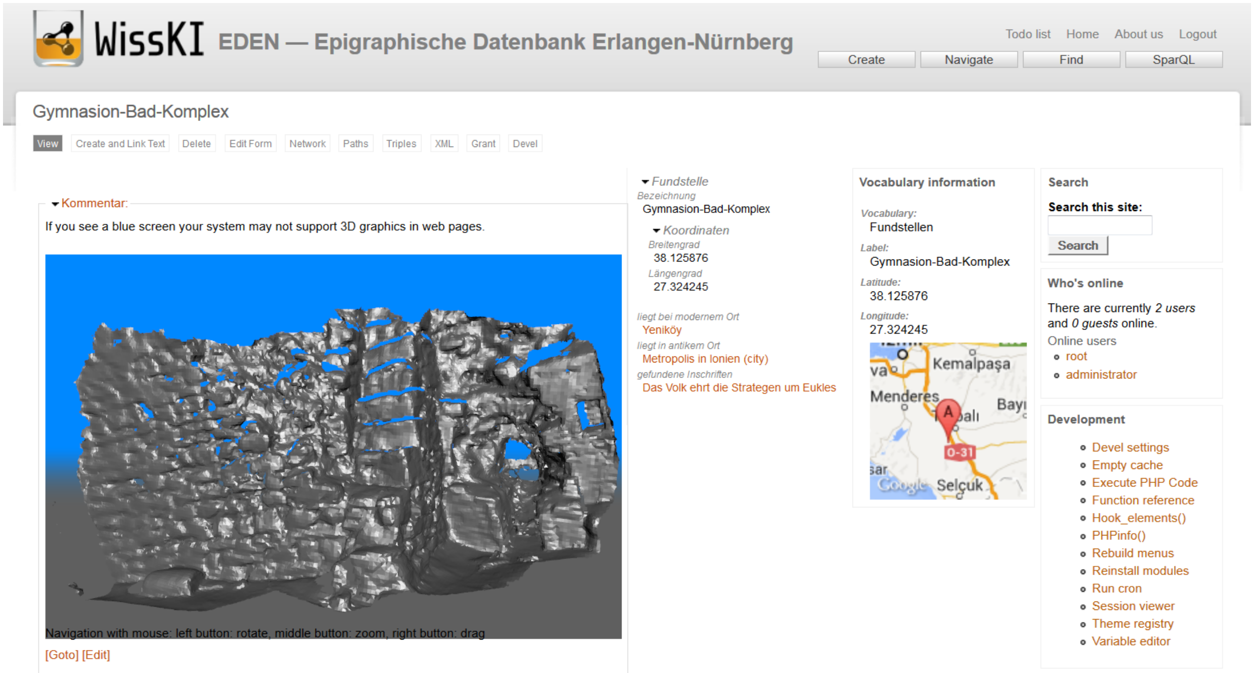

Figure 2: Find spot *Gynmasium Bath Complex with view on the fireplace*, *http://wisski.cs.fau.de/eden/content/eden_Fundstelle1d9aecfe-d714-4d44-a4b8-8670da214cd0*: The 3D model complements the semantically enriched tabular metadata.

The original geometries are of variable size, some containing several million vertices, leading to files with sizes much bigger than 100 MB. Such file sizes, however, are not acceptable in web applications like EDEN; downloading such files would take far too long

[11] see *http://webgl.org*. The first version of the specification from 2011 can be found at *https://www.khronos.org/registry/webgl/specs/1.0.0/*

[12] *http://threejs.org*

and furthermore, the mere size could make weaker clients crash. In a first approach, we therefore reduced the file size to some kilobytes, sacrificing the high resolution. Although the sites' overall shape is still intact, detailed surface structures are not retained. This, of course, is far from optimal for displaying inscriptions. For a better visual experience with acceptable load time delays we therefore plan to integrate and experiment with three techniques:

1. Loading geometries on demand and dynamically adjusting the level of detail of objects, starting with coarse-grained geometries and loading finer ones if needed. This is commonly used to reduce complexity of objects far away from the user's perspective. In our case it can also help to reduce internet traffic and loading delays.

2. Bump maps are used to render smaller surface structures on a coarse-grained mesh. In our case we thus could recover the details lost in the downsampling process mentioned above. Bump maps are typically stored as normal 2D bitmaps and encode the surface information much denser than a mesh, again reducing traffic.

3. Textures may show smaller surface structures like scratches or even inscriptions, too, by simulating light and shadows.

The 3D models are not only nice visualizations, but also a way to access the semantically enriched information in EDEN. Thus, the question arises how and how much should the 3D data be intermingled with the semantically enriched data? In the long run, we would like to annotate parts of 3D models in the same way we annotate texts with named entities. These parts, then, should be highlighted, showing web links to their references. But where should these annotations be stored? One option would be to store it directly in the mesh, as it is done for textual annotations. The other option is to point into the mesh, by using techniques like the selectors in the Open Annotation Collaboration model[13].

This has a lot of technical issues. But also imposes challenges on the user interface, e.g. ergonomically selecting a subset of a mesh — only with the capabilities of modern browsers. In this case, there seem to be solutions at hand soon: The German research project "Inschriften im Bezugsystem des Raumes" e.g. is developing such a mechanism by combining 3D mesh and 2D bitmap information.[14]

Since recently, the developers of WissKI at the University of Erlangen-Nürnberg have joined forces with the research project "Virtuelle Rekonstruktion von Barockschlössern im ehemaligen Ostpreußen"[15] of the Herder-Institut, which uses WissKI for virtually reconstructing palaces in former East Prussia. This cooperation may be fruitful for EDEN as well.

In conclusion, the way to a virtual 3D excavation site seems laborious, as out-of-the-box tools and techniques are still missing, but nevertheless within reach.

[13] See *http://www.openannotation.org/spec/core/specific.html#Selectors*

[14] See *http://www.spatialhumanities.de/ibr/technologie/genericviewer.html*

[15] *https://www.herder-institut.de/forschung-projekte/laufende-projekte/virtuelle-rekonstruktion-von-barockschloessern-im-ehemaligen-ostpreussen.html*

References

[CDG+11] Nick Crofts, Martin Doerr, Tony Gill, Stephen Stead, and Matthew Stiff. Definition of the CIDOC Conceptual Reference Model — Version 5.0.4, 2011.

[FR12] Mark Fichtner and Vincent Ribaud. Paths and shortcuts in an event-oriented ontology. In Juan Manuel Dodero, Manuel Palomo-Duarte, and Pythagoras Karampiperis, editors, *Metadata and Semantics Research*, volume 343 of *Communications in Computer and Information Science*. Springer Berlin Heidelberg, 2012.

[HKR10] Pascal Hitzler, Markus Krötzsch, and Sebastian Rudolph. *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC, 2010.

[HRS14] Marvin Holdenried, Charlotte Roueché, and Martin Scholz. Digital epigraphy in archaeological context: The case of metropolis, magnesia & apollonia. In Boris Dreyer, editor, *Orient & Occident in Antiquity*, volume 1 (Die Surveys im Hermos- und Kaystrostal und die Grabungen an den Thermen von Metropolis (Ionien) sowie am Stadion von Magnesia am Mäander. Neue Methoden und Ergebnisse), pages 163–183. Berlin — Münster, 2014.

[SG12] Martin Scholz and Guenther Goerz. WissKI: A Virtual Research Environment for Cultural Heritage. In Luc De Raedt, Christian Bessière, Didier Dubois, Patrick Doherty, Paolo Frasconi, Fredrik Heintz, and Peter J. F. Lucas, editors, *ECAI*, volume 242 of *Frontiers in Artificial Intelligence and Applications*, pages 1017–1018. IOS Press, 2012.