Digital Reconstruction and Virtual Research Environments – A question of documentation standards

Piotr Kuroczyński

Herder Institute for Historical Research on East Central Europe – Institute of the Leibniz Association

Email: piotr.kuroczynski@herder-institut.de

Abstract. In view of proliferation of digital technologies and increasing popularity of digital 3D documentation of cultural heritage — being a result of both digitization of existing artefacts and computer-based 3D reconstruction of lost art and architecture — the implementation of reliable standards is of growing importance.

The paper considers potential benefits and inherent problems of digital reconstruction, which the Author understands as research and documentation tools. Some thoughts are devoted to the technological advancement and the risk of the digital obsolescence. The upcoming transition to the Web 3.0 is also considered. Concepts proposed for semantic data-modelling in the field of cultural heritage make it apparent that the future of e-documentation depends on the long-term availability and interoperability of data sets. Drawing on selected directives (e.g. London Charter) and projects (e.g. CARARE, 3D ICONS), current research in access to and understanding of 3D data sets is discussed.

The ongoing, joint German-Polish-Russian project called Virtual Reconstructions in Transnational Research Environments – the Web Portal: Palaces and Parks in Former East Prussia, offers the opportunity to present a prototype of a Virtual Museum. Exemplary in its interdisciplinary methodology, semantic indexing of sources and description of 3D objects, and of visualization based on interactive 3D models, the project is presented to encourage its critical discussion.

Keywords: Digital reconstruction, semantic modelling, documentation standards, virtual research environments

1. Introduction

The employment of computer-based visualization of destroyed or unbuilt architecture goes back to the end of the 1980s with the WINSOM computer model of the Old Minster[1] and the ASB BAUDAT computer model of Cluny III[2] amongst the pioneering projects. They embodied the earliest attempts to exploit digital (virtual) reconstruction for the documentation and communication of cultural heritage. By the 1990s the advancement of information technology, based on Moore's Law observed much earlier as the groundwork of the digital revolution, resulted in a rapid increase in the means of presentation and application of this kind of digital

---

1 BURRIDGE J. M., COLLINS B. M., GALTON B. N., HALBERT A. R. ET AL.: *The WINSOM solid modeller and its application to data visualization*, IBM Systems Journal, Vol. 28 (1989), No. 4, pp. 548–568.

2 CRAMER H., KOOB M.: *Cluny. Architektur als Vision*, Heidelberg 1993, pp. 58–103.

reconstruction. The still increasing computing power makes it possible to render computer-generated images ever more realistic. The game industry has introduced game engines and the interaction within immersive, virtual worlds. The automotive industry established the rematerialization of 3D data sets through Rapid Prototyping. The ubiquitous presence of the internet, in connection with new mobile devices and GPS, confronted us with various attempts at an augmented reality.

However, in the field of digital reconstruction of cultural heritage, these developments mainly help with cultural mediation, i.e. in historic exhibitions and film documentaries. Computer-generated picture sequences communicate complex architecture; their visual power results in a deeply emotional experience.[3] It seems that the rapid technological development has overshadowed the scholarly aspects and content requirements of digital reconstruction. Digital reconstruction does have a great potential, in particular for archaeological, architectural and art-historical research. For example, the holistic approach to the examination through 3D modelling has enormous advantages over standard representation through separate 2D drawings. In 3D reconstruction, spatial researchers are thus gaining a deeper insight into the field.

The main scholarly problem is a general problem of the digital revolution. Digital 3D data sets may be difficult to understand unless information about their meaning (metadata) and construction methods (provenance and paradata) is provided. In the context of the aforemention rapid technological development, the sluggish implementation of documentation standards and the common obsolescence of software and hardware, inevitably lead to what one may call a digital elephants' graveyard. Comprehensibility and long-term availability of information, both being essential in academic research contexts, are not ensured in 3D computer-based reconstruction of cultural heritage. A demand for "critical computer visualization"[4] has been growing for quite some time and in 2006 it was internationally recognized in the principles of the London Charter. By defining the need for paradata, the London Charter introduced to computer visualization an extension to provenance information, stipulating as follows: "Documentation of the evaluative, analytical, deductive, interpretative and creative decisions made in the course of computer-based visualization should be disseminated in such a way that the relationship between research sources, implicit knowledge, explicit reasoning, and visualization-based outcomes can be understood."[5]

2. Semantic data modelling

Libraries, archives and museums play a pioneering role in the field of e-documentation. These traditional custodians of knowledge and culture have endeavoured, since the beginning of the digital revolution, to transfer the principles of the analogue card index into the new era through digital metadata. Metadata exemplify the meaning (semantics) of digital data sets. Metadata of the sketching in technical digitalization (provenance), and paradata about the human intention and interpretation in the course of this process, ensure the accountability and quality of the data sets. The text-based Open-Source XML has been established as metadata format suitable for processing by the humans and machines. The semantic web, which marks the transition from the "Internet of

---

3 GRELLERT M.: *Immaterielle Zeugnisse. Synagogen in Deutschland. Potentiale digitaler Technologien für das Erinnern zerstörter Architektur*. Transcript-Verlag, Bielefeld, 2007.

4 GÜNTHER H.: *Kritische Computer-Visualisierung in der kunsthistorischen Lehre. In: Der Modelle Tugend. CAD und die neuen Räume der Kunstgeschichte*, Weimar 2001, pp.111–122.

5 DENARD H.: *A New Introduction to the London Charter*. In: Bentkowska-Kafel A., Baker D., Denard H. (eds.): Paradata and Transparency in Virtual Heritage, Ashgate, 2012, pp. 57–71. See also: htttp:// www.londoncharter.org/ (*accessed* August 12, 2014)

Documents” to the “Internet of Data”[6], offers the infrastructure for implementing more transparency, better visibility and availability of data.

In the field of documentation of cultural heritage, CIDOC-CRM[7] is the long-term attempt at establishing a „semantic glue“ between different sources of information. It defines and implements formal structures for the description of competences (entities) and relations (properties) in the documentation of cultural heritage. As the ISO-standard 21127:2006, it provides a general and extensive semantic frame. The advantage of CIDOC-CRM lies in its function as a reference ontology, which allows for the referencing of all sorts of heterogeneous information from the cultural field. The characteristic feature of CIDOC-CRM is its event-centric description. For example, the process of painting a portrait may be represented through the relations between objects, sources, persons, the time and place of the event (activity).

Semantic modelling requires an adequate qualification of 3D data sets, which, in turn, results in higher efforts when it comes to the creation and maintenance of information. From an economic standpoint, this can only be justified if semantic modelling benefits many users in a variety of applications. In this context, the virtual library EUROPEANA has to be mentioned. Its aim is to provide a broad audience with access to the cultural heritage of Europe in the form of images, texts, and audio and video data. Various subsequent European projects have preoccupied themselves with the semantic editing and contribution of content to this Linked Open Data Portal.

The EU project 3D-ICONS supplies EUROPEANA with 3D data sets. The 3D-ICONS is an ongoing project funded by the European Union. It is based on its predecessors, 3D-COFORM and CARARE. 3D-ICONS focuses on the implementation of a metadata schema for the documentation of the complete process of digitalization (including the provenance and paradata). It thus aims at a form of quality control of 3D data sets. The resulting CARARE 2.0 metadata schema has been adapted to the Europeana Data Model (EDM) and expanded by the inclusion of CRM_{dig}. CARARE 2.0 can be described as an application profile and extension of EDM, designed to cover the requirements of the archaeological and architectural heritage domains and of 3D digitization processes.[8]

CityGML, an XML-based format for the representation and exchange of virtual 3D city models also provides interesting points of reference for digital reconstruction of cultural heritage. It is based on the ISO standard 19100 and is implemented as an application ontology for Geography Markup Language (GML 3.1.1.).[9] Although it has been mainly conceived as a tool for public administration, urban and environmental planning, and the tourism industry, its five-step concept of Level-of-Detail (LoD) differentiates between different levels of information about geometry and themes, and is relevant to other subject areas.

6 BERNERS-LEE T., HENDLER J., LASSILA O.: *The Semantic Web*. In: Scientific American, May 2001

7 http://www.cidoc-crm.org/ (*accessed* August 12, 2014)

8 D’ANDREA A., FERNIE K.: *CARARE 2.0: a metadata schema for 3D Cultural Objects*. In: Proceedings of the 2013 Digital Heritage International Congress, pp. 137-143.

9 COX S., DAISY P., LAKE R., PORTELE C., WHITESIDE A.: *OpenGIS Geography Markup Language (GML3), Implementation Specification Version 3.1.0*, OGC Doc. No. 03-105r1, 2004.

3.1. Virtual reconstruction in transnational research environments

The international and interdisciplinary project tentatively called Virtual Reconstructions in Transnational Research Environments – the Web Portal: Palaces and Parks in Former East Prussia[10], examines the whole process of computer-based 3D reconstruction of lost architecture and destroyed interior decoration. The preliminary results are based on the digital reconstruction of two destroyed Baroque palaces, Schlodien (now Gładysze, Poland) and Friedrichstein (now Каменка, Russian Federation). New findings are concerned with the indexing of sources, documentation, semantic modelling, and visualization of 3D datasets in WebGL-technology.

The focus is on the development of a Cultural Heritage Markup Language (CHML)[11], a human and machine-readable XML Schema for the semantic annotation and integration of various metadata, including the provenance information and paradata, and the mapping (labelling) of geometrical, material, and lighting characteristics in the source text (for long-term archiving purposes).

The research-based, digital reconstruction of both palaces aims to combined earlier knowledge with new research. The knowledge portal that is to be implemented will be informed by cross-disciplinary research in the areas of architecture, art history, history, information technology, and knowledge transfer. In the medium term, the project aims to define standards for e-documentation and presentation of 3D datasets of destroyed architectural landmarks and artworks on the web. The scholarly and educational content of the prototype, open, virtual research environment for digital reconstruction is being designed by the project partners. Since the foundation of the working group Digital Reconstruction[12], the project is embedded in the Digital Humanities in the German-Speaking Region.

In brief, the project involves the preparatory task of identifying sources and the subsequent modelling of 3D objects. A web-based Content Management System (CMS) provides the interface to the annotation, linkage and storage of results. The CMS includes an application ontology for digital reconstruction. The CMS also generates the semantic database (Graph Database) and enables the export of 3D data sets to EUROPEANA, as well as the long-term preservation of the information about the 3D object in CHML Schema. The virtual research environment as a prototype Virtual Museum emerges thanks to the linkage of the CMS (Graph Database) to the binary data in the repository, as well as through the subsequent integration of interactive 3D models with the WebGL technology.

---

10 http://www.herder-institut.de/en/research-projects/current-projects/virtual-reconstructions-intransnational-research-environments-the-web-portal-palaces-and-parks-in-former-east-prussia.html, (*accessed* August 12, 2014)

11 HAUCK O., NOBACK A.: *CHML – Cultural Heritage Markup Language. Eine Auszeichnungssprache für das gebaute Weltkulturerbe*, Presentation at CeBit 2003. See http://archipelagus.de/lenya/archipelagus.de/live/chml-index/CeBit03.html, (accessed August 12, 2014).

12 http://www.digitale-rekonstruktion.info/uber-uns/ (*accessed* August 12, 2014)

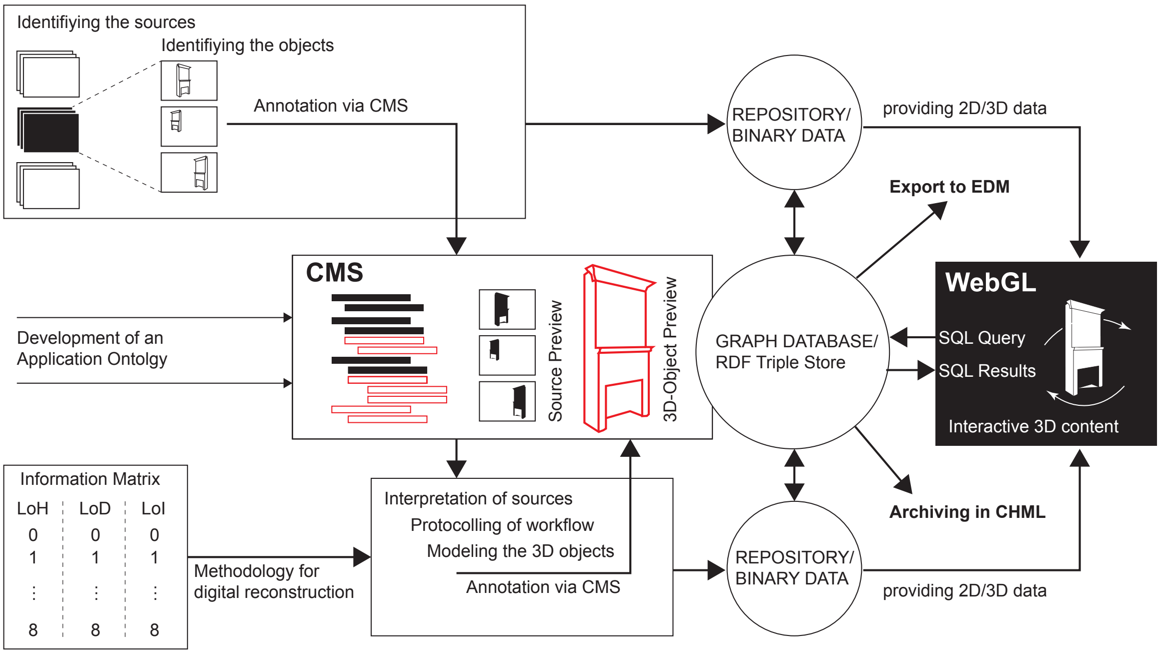

Fig. 1. Overview of the workflow and the project structure

3.2 Activities in digital reconstruction

The activity of identifying sources aims at structuralizing the existing sources, as well as identifying and indexing the objects that are to be reconstructed. In general, every source is entered into the CMS by a subject expert (e.g. an art historian) as an object. Images are primarily structured according to their geo-locative coverage in order to prepare a spatial orientation (mapping). Subsequently, the images are analysed and the objects which they depict are separately identified and annotated. For the indexing of coverage as well as for the object identification, Scalable Vector Graphics (SVG) is to be integrated into CMS. SVG is an XML-based vector image format for two-dimensional graphics which supports interactivity and animation. This method promises a smooth linkage of image sources and their content with the 3D data sets generated by the computer-based reconstruction.

The source indexation in CMS prepares for 3D modelling. The recording of spatial interpretation of the respective objects is based on the sources. The modeller records the process of digital reconstruction, so it is traceable within the digital 3D objects thus born.

The density of information and the interpretational margins are expressed on the basis of an enhancement of the Level-of-Detail concept under development.

3.3. Application ontology in digital reconstruction

The integral design feature of the e-documentation of cultural heritage under discussion is the Cultural Heritage Markup Language (CHML). The idea, which is to be developed further, was presented at the TU Darmstadt stand on CeBit in 2003.[13] The XML Schema of CHML combines the metadata, provenance and paradata, as well as the geometrical and material data, light and camera properties of 3D objects and scenes — all inline. Authority files and controlled

[13] cf. footnote 11

vocabularies (thesauri) are to be implemented. CHML guarantees the certification, interoperability and long-term preservation of the information contained in digital reconstruction.[14]

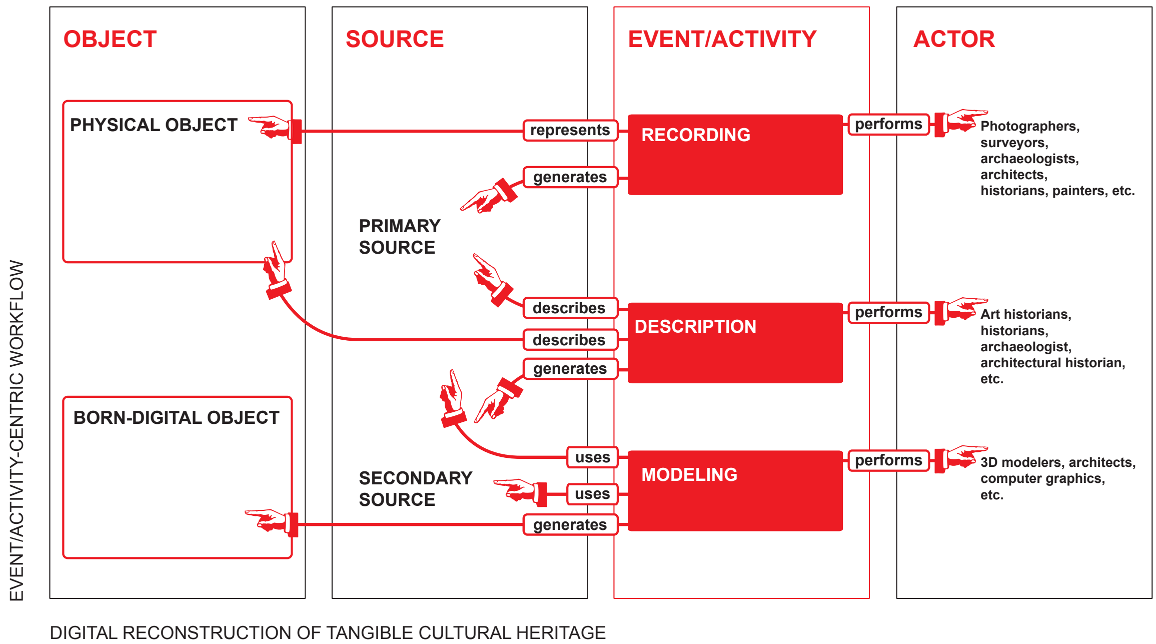

Fig. 2. Event-centric workflow described in CHML (Hauck O., Kuroczyński P., 2014)

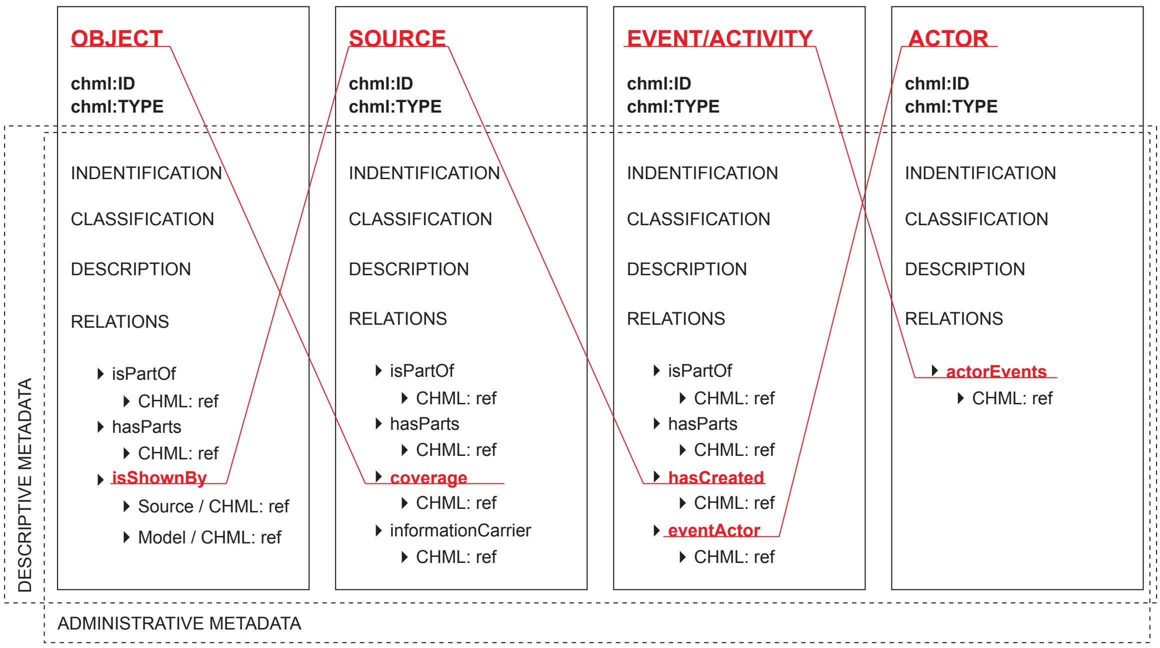

Fig. 3. CHML structure connecting the main themes (Hauck O., Kuroczyński P., 2014)

---

[14] KUROCZYŃSKI P., HAUCK O.: *Cultural Heritage Markup Language – How to record and preserve 3D assets of digital reconstruction*. In: 19th International Conference on Cultural Heritage and New Technologies, Vienna, 2014 (*submitted Paper*).

For an improved annotation of the CHML files in the process of computer-based 3D reconstruction, a custom Web Ontology Language (OWL) will be created on the basis of Erlangen-CRM (E-CRM). E-CRM is an application-related interpretation of CIDOC-CRM, which can establish a powerful virtual research environment for cultural heritage domains, inter alia in the ongoing WissKI-project.[15]

The application ontology for CHML, which is to be developed in cooperation with Erlangen-Nürnberg University, will be based on the CIDOC-CRM reference ontology. It will allow the storage of information in the triple-architecture of the Resource Description Framework (RDF), as well as mapping to CARARE 2.0 format and Europeana Data Model (EDM).

The action-based methodology employed by the subject expert and modeller of digital reconstruction corresponds to the event-based character of E-CRM. On the basis of respective activities, actors, sources and objects will be described and interlinked. The event-centric workflow described by CHML Schema and the interaction of the four main themes of CHML with their respective core elements are shown in the Figs. 2 and 3.

CHML includes a TYPE for the classification of the separate records. The CHML-TYPE is composed of four letters of the respective English description, e.g. SEPL stands for Section Plan. A connection to controlled vocabularies, such as the Getty AAT, is accounted for. Over the course of the project, the German, Polish and Russian equivalents are added to the already implemented English TYPE, with the aim of establishing a vocabulary for digital reconstruction of the architectural cultural heritage covered by the project.

3.4. Content Management System

The open source CMS Drupal of WissKI is semantically adapted to the four main themes of CHML. For the registration of the CHML main theme, Event/Activity, the project staff are using a free text input, which allows for subsequent keywording, automatic linkage with existing records and the creation of new objects, sources, activities and actors. Thereby, intuitive and easy access to the documentation of the working process, as well as to the linkage and creation of new records is provided. A deepening enhancement of metadata of the respective records can subsequently be carried out through the field-based input mask to the CHML themes. Here, all elements of descriptive and administrative metadata are listed, in addition to the obligatory input fields.

The Graph Database holds all data in the RDF Triple-Store on the web server. They can be issued out of the CMS in the form of relational graphs and/or triple paths. The binary data sets are stored separately on a repository and are linked to the semantic database. The repository holds all digitals of the text, audio, picture and video sources, as well as the 3D data sets of the reconstructed objects (FBX, OBJ and DAE COLLADA).

A Mapping of CHML Schema onto the Europeana Data Model will be tested. The long-term storage of the entire object-related information in CHML, including the metadata, provenance, and paradata, as well as the integration of geometrical, material, lighting and camera properties, through the embedded XML-format DAE COLLADA, is implemented.

The WebGL solution allows for the compilation of interactive 3D scenes (from repository; binary data) and the linkage of the 3D objects with the corresponding metadata (from Graph Database: RDF Triple-Store). 3D scenes can be constructed on demand, and the 3D objects, with the

---

15 SCHOLZ M., GOERZ G.: *WissKI: A Virtual Research Environment for Cultural Heritage*. In: De Raedt, Luc et al. (Ed.): 20th European Conference on Artificial Intelligence, ECAI 2012, Proceedings. Amsterdam: IOS Press.

appended metadata, can be retrieved and annotated. The virtual research environment thus becomes a Virtual Museum par excellence, in which every 3D object is documented in the same fashion as a museum object, and in which visitors/users are offered a tour of the reconstructed Baroque palaces, their respective context included.

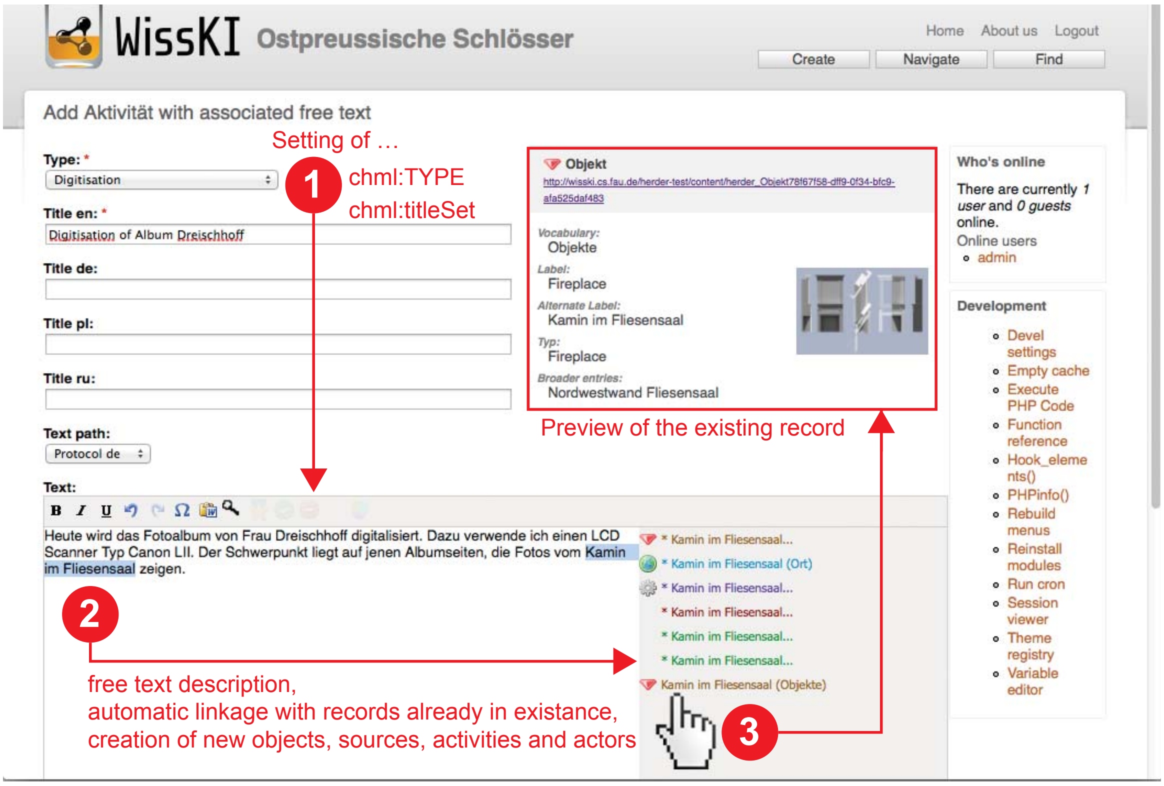

Fig. 4. Event-centric, free text annotation within CMS Drupal, referring to WissKI.

Fig. 5. Example of the integration of interactive semantic 3D objects in WebGL: exploring a room in Schlodien Palace, (1.) integrated position map with view perspective, (2.) visualization of sources linked in the repository, (3.) retrieval of metadata from the semantic Graph Database (Dworak D., 2014)

4. Conclusion

The readability and clarity of digital data can be ensured by descriptive metadata and semantic databases. For digital reconstruction of cultural heritage considered as a scholarly activity, this means the development of application ontologies, including the recognized standard reference ontologies, such as CIDOC-CRM. The request to provide the EUROPEANA with a diverse content in diverse database formats via mapping into EDM, has a politically driven impact. It is also an opportunity for research groups willing to create standardized metadata schemata for 3D Cultural Objects, such as CARARE 2.0.

Developing the CHML for digital reconstruction connects the sources used, the technical processes involved in creating digital objects (provenance), the human processes of understanding and interpreting the digitized sources (paradata) to the resulting born-digital 3D objects. CHML standardizes the methodology of computer-based 3D reconstruction of tangible cultural heritage and ensures the interoperability of data sets by referring to CIDOC-CRM. It ensures mapping the data sets into EDM and provides the described 3D content to linked open data environment of EUROPEANA. CHML stands also for long-term preservation of information through 3D-object-oriented records with an integrated inline description of geometry, materials, light and cameras in DAE COLLADA format.

The application of an Information Matrix – an enhancement of the Level-of-Detail concept – represents the granularity of information and the hypothesis within the 3D model. The development of the concept promises a new way of modelling and ensuring its transparency. Such critical approach to computer visualization is in line with the principles of the London Charter.

The proposed solution for the interactive integration of 3D objects in WebGL is under development. The early results are promising. The method enables the interactive connection between the 3D data from the repository and its descriptive metadata from the Graph Database. The engaged user gets the impression of a Virtual Museum, where visualized 3D objects are made available on demand, alongside their semantic metadata, and are ready to be edited and annotated.

It is anticipated that the described method of digital reconstruction and the semantic enrichment of 3D data sets doubles the efforts, but seems the price worth paying for creating more than just pretty pictures. It is the price of a scholarly approach and such values as transparency, consistency, certainty and access in the long-term.

5. Acknowledgements

The international, collaborative research project, Virtual Reconstructions in Transnational Research Environments – the Web Portal: Palaces and Parks in Former East Prussia, is funded by the Leibniz Association from 2013 to 2016. The project is coordinated by Herder Institute for Historical Research on East Central Europe in Marburg, Germany.