Bypassing metadata

A research platform for 3D models, images and textual information with a new approach to filtering

Cindy KRÖBER, Friedrich-Schiller-Universität Jena, Germany
Aaron PATTEE, Ludwig-Maximilians-Universität München, Germany
Sander MÜNSTER, Friedrich-Schiller-Universität Jena, Germany
Jonas BRUSCHKE, Julius-Maximilians-Universität Würzburg, Germany
Ronja UTESCHER, Friedrich-Schiller-Universität Jena, Germany

Keywords: *Metadata, filtering, georeferencing, multimodality, AI*

CHNT Reference: Kröber, C., Pattee, A., Münster, S., Bruschke, J. and Utescher, R. (2022). 'Bypassing metadata - A research platform for 3D models, images and textual information with a new approach to filtering, in CHNT Editorial board. *Proceedings of the 27th International Conference on Cultural Heritage and New Technologies, November 2022.* Heidelberg: Propylaeum.
DOI: xxxxxxx.

Introduction

Photographs and other images as well as textual information and documents serve as important source materials and provide a foundation for many subject-orientated and theory-based investigations within historical studies; e.g., architectural studies, art history, and cultural studies. When investigating urban development, visual media and additional information are consulted in order to analyze buildings, draw stylistic comparisons, document changes in the building structure or – in general – to answer open questions regarding the history of a certain building, quarter, or city. Finally, in case of a complete loss of physical traces, historical pictures can form the basis for different kinds of reconstruction efforts. A lot of images and information are needed to provide a thorough picture of a building necessary for a reconstruction, and the search for suitable images and texts can prove challenging. On the one hand, there is a vast amount of data available online and on the other hand, filtering – usually based on metadata entries – does not often illicit the desired results or is simply avoided altogether.

Issues connected to metadata

The present inquiry standard is primarily concentrated on embedded metadata, such as the author, the date of an object or photograph, a more or less exact localization, a description of the topic or depicted matter and, of course, several keywords that classify the object or information in various aspects (Kohle, 2013, p. 23). Unfortunately, not all repositories and institutions adhere to a standard for metadata. The metadata is very heterogenous even within the same repository. This can partly be attributed to the lack of an approved standard as well as rushed digitization processes. Descrip-

tions can be flawed or missing, names are misspelled, and information on the digitization is completely neglected. Entries for 'times' are usually not specified and can relate to the acquisition time, digitization time, or sometimes even the time of a display at an exhibition.

Many researchers do not trust filtering and are rather hesitant when it comes to reducing search results. They prefer to get as many hits as possible to make sure nothing of importance gets left out (Kröber, 2021).

Recent approaches try to avoid possible flawed metadata by searching and filtering with the help of *Artificial Intelligence* (AI) providing advanced keyword tags and object recognitions. Metadata can become unnecessary for a search when photographs are geo-referenced within a digital map or 3D city model. Buildings depicted in the photographs can thus be directly linked to places or the virtual architectural models (F Maiwald, Henze, Bruschke, & Niebling, 2019).

Research project HistKI

This paper presents a work in progress of the ongoing, multi-institutional HistKI (Historische Künstliche Intelligenz, or Historical Artificial Intelligence) project seeking to integrate 2D images, 3D models as well as text into a web application with an interface that supports searching and filtering, amongst other features. The project is a collaboration of researchers from across Germany specializing in Art History, Computer Linguistics, Computer Vision, and Information Sciences.

Research platform 4D Browser

A prototypical web application called 4D Browser[1] presents historical images of the city of Dresden in a virtual 3D city model (see Fig. 1).

A timeline introduces the *Fourth Dimension* (4D) and provides information on the development of the city by filtering photos and 3D models according to any selected point in time. The features support image searching and filtering, data analysis, data visualization based on acquisition habits, and contextualization via linking photos and their spatial-temporal locations within the model. The currently accessible collection consists of 3,633 historical images and is intended to include textual resources for further research and analysis. Therefore, all three entities (text, 2D images, and 3D models) will be interconnected or linked, and additional tools and functionalities will support further research. The idea of the 4D Browser is directly connected to research questions for art and architectural heritage studies. Users were involved from the onset of the development of the 4D Browser interface and functionalities, and will continue be part of the quality and usability assessment (Kröber, Hammel, Schade, Filz, & Dewitz, 2021).

[1] https://4dbrowser.urbanhistory4d.org/

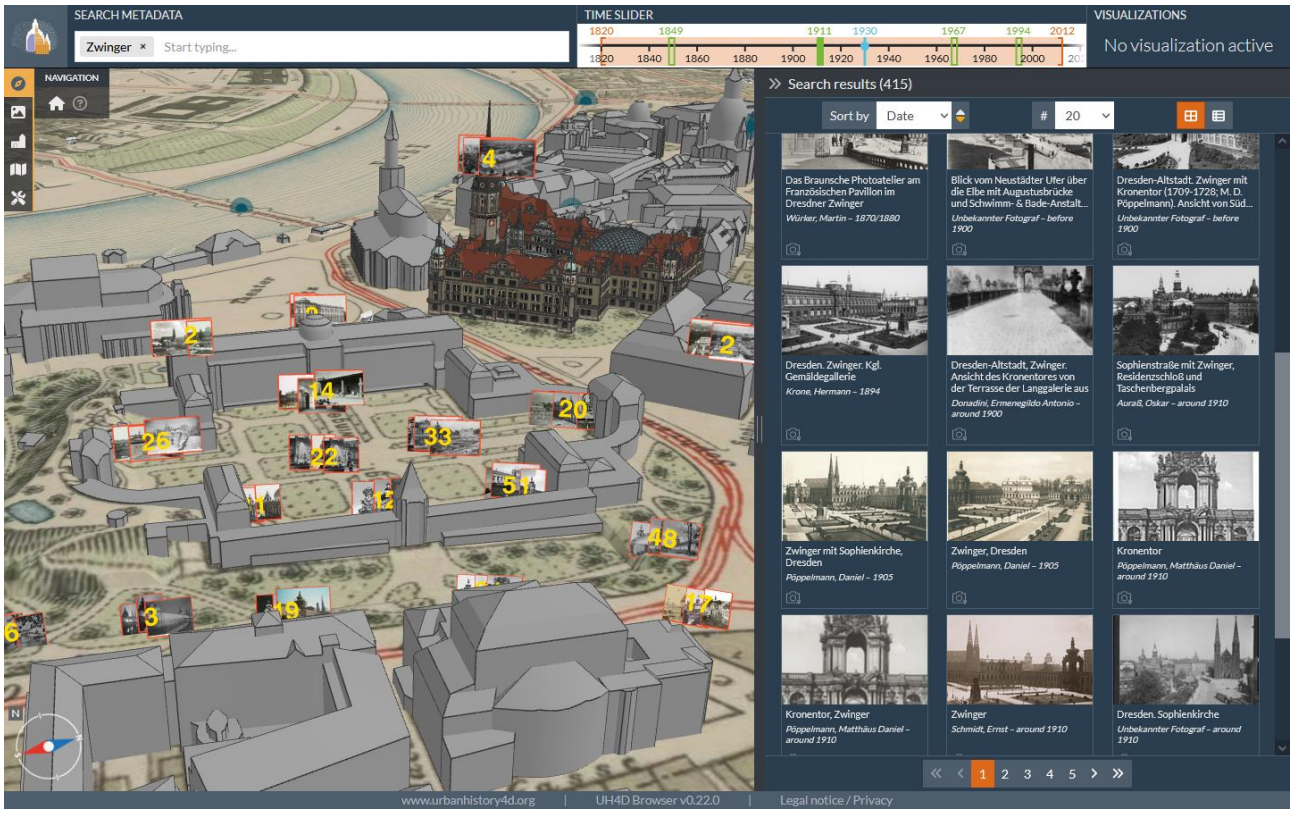

Fig. 1. Interface of the 4D Browser

Prospects for art history research

The 4D Browser has the potential of becoming a new interface for repositories or at least inspire new designs and functionalities for them. The available functions help with gathering, cataloguing, and locating a large number of sources. The timeframe adds even more support by allowing a scholarly examination of several stages of urban development, something that is difficult to achieve without digital solutions. A whole catalogue of mid-ranged questions can be derived from the work with the 4D Browser that supports and pushes scientific tasks forward.

Besides large-scale changes due to wars, earthquakes or simple city reorganizations, also smaller observations are relevant for art historians and building researchers, e.g., changes of building details like the roofing, decorations or façade renovations. Another setting for art history research is the perception of the city and its stock of historic and modern buildings based on the analysis of photography. This may even give insight into the interests, self-image, and intellectual attitude of the photographer.

Specifically, within the realms of architectural art history and archaeology, the primary sources are the material remains and in the case of a building, it is the building itself (Großmann, 2010). As for buildings that no longer exist, architectural historians must rely upon image data for accurate interpretations of the building phases, progression of construction, and the exhibited motifs ornamenting the building. Employing AI-based methodologies for this purpose will buttress architectural investigations of former buildings, in which existing image and text sources can be connected and examined chronologically. This effectively exhumes destroyed buildings from the bog of immateriality by gathering all known evidences of their previous existence. Through the eyes of a contemporary it will be possible to evaluate which buildings dominated and formed the idea of city.

Sources from urban history are quite challenging to work with. The view of the city they represent is mostly internal and therefore biased. Some sources are more objective, e.g., construction plans or mapping photographs with coordinate points. Written descriptions of a city, engravings or artistic photographs show a much more personal attitude towards the cityscape. Researchers are therefore responsible for contextualizing them, taking these biases into account. Art historians are aware of this problem and objectivize results - still predominantly a manual process (Schneider & Stegers, 2002). As a result, these more laborious studies consider only narrow time spans or locations. Computerized tools enable easier access to the sources and wider research foci, creating more opportunities for utilization, and making a multi-focus analysis of the urban tissue possible. This leads to new questions: How do buildings and cities change over time? In which contexts, such as political or formal developments, does a historical cityscape evolve? What similarities can be found between objects in terms of construction standards and requirements, building codes, regional, temporal or personal tastes and styles? Furthermore, what connotations does a building have? Does the number of pictures taken of a specific building change over time? Are there connections to other buildings or urban spaces? Which interactions of architecture with other artistic genres, inscriptions or infrastructural facilities can be found? Which buildings are likely to form the architectural backdrop for social events, demonstrations, or celebrations?

New approach to searching and filtering

The 4D Browser supports a simple query using a search bar and corresponding metadata entries. But the advanced searching and filtering functionalities also take advantage of the fact that all images are georeferenced—meaning that they are placed at their location of acquisition (camera position). In the beginning, the georeferencing was accomplished manually but now the use of Machine Learning (ML) approaches support a (semi-)automatic reconstruction of the camera position (Ferdinand Maiwald, Lehmann, & Lazariv, 2021). The image is then projected onto the 3D city model, allowing the pixels to correspond with the 3D buildings. This makes it possible detect pictures with (parts of the) buildings even when nothing is mentioned in the metadata (see Fig. 2).

In order to search for texts, the current approach will be extended to include a specific terminology (based on Wikidata IDs and the Getty Art and Architecture Thesaurus IDs) and annotate the 3D models accordingly. This will help with image segmentation as well as the discovery of textual paragraphs connected to a certain image, providing more information and context.

The approach for interconnecting or linking text, images, and 3D models will serve as a first step to improve searching and filtering without relying solely upon metadata. It will also be a starting point for further research and source criticism as a complex and foundational technique within the humanities via a multimodal, AI-based approach.

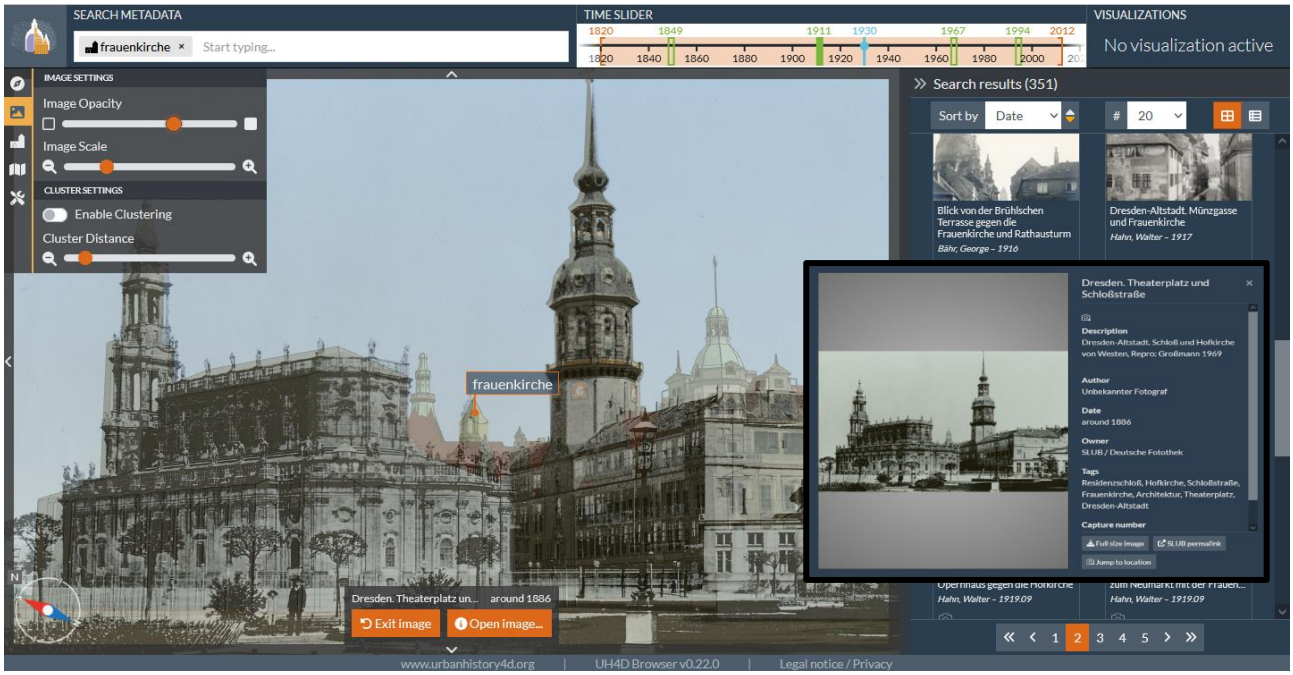

Fig. 2. Successful detection of an image with the dome of the Frauenkirche based on the correspondence of image projection and 3D model

Funding

The research upon which this paper is based is part of the research project HistKI which has received funding from the German Federal Ministry of Education and Research (BMBF) under grant identifier 01UG2120A.

References

Großmann, G. U. (2010). *Einführung in die historische und kunsthistorische Bauforschung*. Wissenschaftliche Buchgesellschaft, Darmstadt.

Kohle, H. (2013). Digitale Bildwissenschaft. Hülsbusch.

Kröber, C. (2021). German Art History Students' Use of Digital Repositories: An Insight. Paper presented at the 16th International Conference, iConference 2021, Beijing, China. DOI: 10.1007/978-3-030-71305-8_14

Kröber, C., Hammel, K., Schade, C., Filz, N., & Dewitz, L. (2021). User Involvement for Application Development: Methods, Opportunities and Experiences from Three Different Academic Projects. Paper presented at the Workshop on Research and Education in Urban History in the Age of Digital Libraries. DOI: 10.1007/978-3-030-93186-5_3

Maiwald, F., Henze, F., Bruschke, J., & Niebling, F. (2019). Geo-information technologies for a multimodal access on historical photographs and maps for research and communication in urban history. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences. DOI: 10.5194/isprs-archives-XLII-2-W11-763-2019

Maiwald, F., Lehmann, C., & Lazariv, T. (2021). Fully Automated Pose Estimation of Historical Images in the Context of 4D Geographic Information Systems Utilizing Machine Learning Methods. ISPRS International Journal of Geo-Information, 10(11), 748. DOI: 10.3390/ijgi10110748

Schneider, R., & Stegers, R. (2002). Glück, Stadt, Raum in Europa: 1945-2000: Springer.