BAUREKA.ONLINE - RESEARCH REPOSITORY, CATALOGUE AND ARCHIVE FOR ARCHITECTURAL HISTORY AND BUILDING ARCHAEOLOGY

Anke Naujokat*, Tobias Glitsch*, Felix Martin*, Hermann Schlimme**

*Rheinisch-Westfälische Technische Hochschule Aachen – Aachen, Germany.
**Technische Universität Berlin – Berlin, Germany.

Abstract

Building archaeology (Historische Bauforschung) is an internationally renowned methodology for investigating historic buildings in research, heritage preservation and planning. It is a hands-on process which documents the geometry, construction, materiality, usage, form, and meaning of buildings and makes it possible to trace the processes of their creation throughout the various layers of time embodied in their fabric. Despite the extensive use of digital tools within its workflows, building archaeology has not yet embraced the digital age as firmly as other disciplines. While the research community has long been interconnected via societies and conferences, there still is no digital platform for the overarching valorisation of scientific building documentations. The project baureka.online, a research data repository, catalogue and archive for building archaeology in the German-speaking area, seeks to provide an innovative answer to this desideratum.

Keywords

building archaeology, research data repository, open access, metadata model, research community activation, research data indexing

1. baureka.online – points of departure

To fully benefit from digital methods and instruments, the German speaking research community of historical building research or building archaeology (Historische Bauforschung) – hitherto only interconnected on a personal basis and via societies or recurring conferences and workshops¹ – urgently needs a comprehensive web-based data infrastructure allowing the discipline to efficiently exchange research data such as virtual models or digital surveys, to publish them in an appropriate manner and to ensure the permanent preservation and transmission of its digital output for future generations. The planned project baureka.online (2018) is aimed at setting up such an infrastructure for research data from historical building research or building archaeology and sees itself as an innovative answer to this desideratum. The project was conceived over the past two and a half years at the Chair of Architectural History at RWTH Aachen University in collaboration with the Department for History of Architecture and Urbanism at Technische Universität Berlin, involving various members of the community and several partners within and outside of academia.

In its final state, baureka.online hopes to address many of the unsolved challenges and unexploited potentials regarding the management of research data from building archaeology.

Firstly, the platform will provide a solution for the long-term digital preservation of the steadily growing amounts of data. Most readers will be familiar with the problem of finding a suitable repository that is able to guarantee sustainable long-term storage for their project data in order to make it accessible to future researchers. Today, the German Research Foundation (DFG) and other national and international funding bodies insist on elaborate preservation strategies before they hand out any grants. Nevertheless, in Germany – as in other European countries – so far there not only is no field-specific repository for data from building

¹ E.g. Koldewey-Gesellschaft, Gesellschaft für Bautechnikgeschichte, Arbeitskreis für Hausforschung or disciplinary forums like “Arbeitsgruppe Digitale Rekonstruktion”.


archaeology, perhaps apart from IANUS which was, however, developed by the German Archaeological Institute (DAI) and is thus reserved exclusively for research from the field of Classical Studies. An inventory compiled by the Network of Expertise in Long-Term Storage of Digital Resources also shows that there is currently no other obvious possibility for the long-term archiving of research data from historical building research (Neuroth, Oßwald, Scheffel, Strathmann, & Huth, 2010). It is only now that there is a chance to arrive at a convincing response – not least because Germany is currently developing a more comprehensive National Research Data Infrastructure (NFDI, 2020), in which baureka.online will be integrated.

A second challenge, which is clearly connected to the first one, lies in the problem of tracking down relevant data. All of us will have discovered that it is becoming more and more difficult to find one’s way through the large variety of existing databases, such as Ebidat, Denkmaldatenbank Bremen, Bayerischer Denkmalatlas, Rhineland’s KulaDig, Baden-Württemberg’s Datenbank Bauforschung/Restaurierung, Monumente-Online, Lineamenta, Bildarchiv Foto Marburg, Prometheus, AKBF or PHAROS and the equally large number of digital repositories, such as 3D-Icons, Edition Topoi, 3D Hop, Sketchfab, Massive Point Clouds for eScience, TopoRaZ, OpenInfra, Cisar, Patrimonium.net, MonArch or the Time Machine project (cfr. also Kuroczyński, Hauck, & Dvorak, 2016; Antonopoulou and Bryan, 2017; The Medieval Kingdom of Sicily; Visualizing Venice). This situation is mainly due to the fact that most of these resources either focus on specific places, periods and objects or remain limited to just one type of medium (such as drawings or photos) or one individual institution (such as a single heritage agency). Even harder to navigate are generic repositories like RADAR, GERDI, FIGSHARE or the storage solutions that various universities have set up for their own faculty (cfr. Naujokat, Glitsch, & Martin, 2017a, p. 30-31). How, then, when dealing with a historic building or construction that has been subject to prior documentation and research can one effectively locate existing data in a way similar to what is possible in disciplines like the geo and environmental or the social sciences that already have their national, field-specific repositories in, respectively, PANGAEA/KOMFOR and GESIS?

Finally, a third challenge arose when politicians declared their ambition to widely implement open-access standards in academia. Indeed, the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities now has almost 600 signatories (cfr: Deutsche Forschungsgemeinschaft, 2015; OECD, 2007; UNESCO, 2003; Wissenschaftsrat, 2012). It therefore seems quite likely that in future one will have to grant open access to the results of any research that has been publicly funded – a fact that might not only significantly hamper the potential to present scholarly research as books, but would also pose the question of what form publications in the field of building archaeology will take instead.

The same question of how the community is going to publish research in the future, however, also allows a closer look at the unexplored opportunities digitalisation offers to the discipline.

The first of those opportunities certainly consists in the rise of digital publication channels, which open up possibilities that far exceed the limits of traditional print publishing. For example, digital publication methods will finally provide adequate solutions for publishing 3D data. Furthermore, they also enable the publication of complete sets of primary data that, even in their raw state, can still be of great use to future researchers. Both trends will in turn encourage the wider discipline of architectural history to develop new, hybrid forms of publication in which, for example, a written research paper, published either online or in print, is accompanied by links to a complete digital dataset of CAD- or BIM-data that is accessible through a repository like that of baureka.online. In this way, it will finally become possible to directly join the interpretative results to the often huge body of data at their basis.

Other big opportunities will derive from the application of semantic web technologies and the use of metadata. By embracing new, quantitative analysis methods, these technologies could complement the current monographical approach of historical building research and provide ways to

---

2 Numerous universities in German-speaking countries have also adopted open access policies, among others TU Berlin, RWTH Aachen, TU München and the Universities of Heidelberg, Tübingen and Vienna.

3 CAD raw data e.g. as .dwg/.dxf; vector-based graphics as .svg, .pdf, .ai; pixel graphics as .tif, .jpg, .pdf, .raw, .png, .bmp, .cpt; tabular research data such as measurement protocols and tables as .csv; 3D data as .fbx, .stl, .wrl, .3ds, .3dm, .txt, .ptz, .obj, .ply or VR data as .urt, .ux, .usx.
tackle new, comparative research questions. And finally, they would create a completely new basis for cooperation and exchange within and beyond the community, for example by tapping into synergies between research, conservation and planning. Baureka.online has thus been conceived as a comprehensive web-based platform that will become a crucial prerequisite of a new culture of data exchange and publishing within the discipline and community of building archaeology.

2. Activation of the research community

If one considers the motivations of baureka as described before, it quickly becomes clear that the repository can reach its full potential only if it is actively used and supported by the discipline at large. While previous projects in the field of Digital Built Heritage or Digital Architectural History have always tended to focus on the development of innovative technologies rather than the interoperability of the respective systems, Baureka.online will therefore be a platform with as wide an appeal and as low a usability threshold as possible, while still providing forward-thinking functionalities. To achieve this objective, the input of the research community was sought and solicited from the very beginning. The initiative has thus been directly aligned with the research community's specific needs.

To this end, when the preparations for the project started in the autumn of 2016, one of the first steps was an online-survey amongst the various stakeholders within the field of historical building research (Naujokat, Glitsch, & Martin, 2017b). Answers not only showed that – with the vast majority of researchers storing their research data on their own computers, in their local networks or at best with their universities' computing services – there still is considerable uncertainty about how to permanently archive digital material and that – with most researchers obtaining other people's data only on a case by case basis on personal request – systematic reuse of material still remains extremely limited. The survey also demonstrated that among the respondents the interest in solutions which allow for a more structured, discipline-specific archiving of research data, facilitate data exchange or establish data provision as part of one's publication output is actually quite strong. Spurred by those observations, the baureka.online project has set itself the goal of making improvements at all those points.

4 The online survey was sent to all university departments in the German-speaking area that are involved in building research, as well as to heritage agencies, independent building archeologists and other institutions and foundations in the field of architectural history and monument preservation. Around 70 people took an active part in the survey.
Further talks with the research community were then undertaken in a day-long workshop in March 2017 (Fig. 1). On this occasion, together with approximately 40 participants the authors were able to verify the feasibility of the project, identify its possible pitfalls and collect additional opinions on the direction it would have to take. In addition, a steering committee was created which, as soon as the authors’ funding requests prove successful, will meet about once every six months to accompany and evaluate the development of the project. During the discussions it became clear that in addition to the original use cases, i.e. archiving, exchanging and publishing research data, there was also a rising demand for further tools aimed at improving communication within the discipline and at structuring collaborative workflows. At the same time, however, the participants reached the conclusion that in order to establish the platform as a central contact point for retrieving building research data it was essential to arrive at an operational system and a critical mass of data input as early in the project phase as possible.

3. baureka.online – structure

The authors therefore propose to divide the platform into five distinct modules - baureka.archive, baureka.index, baureka.publications, baureka.network and baureka.cloud - that at least to a certain extent can be implemented independently from each other in successive stages (Fig. 2). In the first stage, the project team is going to concentrate on the first two of those modules, which, in a sense, make up the core of the platform and therefore also directly inform its name. The platform’s title, after all, not only alludes to the Greek ‘heureka’, ‘I found it’, which obviously describes the project’s overall objective, i.e. allowing researchers to locate existing datasets on historic buildings, but also reflects the different parts of the platform, with “bau” standing for “Bauforschung” or “building archaeology”, “re” standing for “Repositorium” or “repository” and “ka” standing for “Katalog” or “catalogue”.

baureka.archive thus consists of the repository proper and allows the long-term storage and provision of research data. Its content will be limited to project-related primary material, i.e. data derived from original surveying and documentation work on historical buildings. The data will be submitted and made available on a simple upload-download platform, which will also accommodate future file formats. If certain quality standards are met, the platform is open not only to projects by established researchers and institutions, but also to contributions from younger scientists or even students. The data provider will be able to control access to and further use of its data by means of various license types. The individual datasets will then be displayed online in the form of thumbnails, metadata and short descriptions. And finally, depending on the license, an interested user will be able to download preview images or entire projects.

baureka.index, on the other hand, will provide a central reference catalogue for data in the field of historical building research. Both data stored directly in baureka.archive and data available from external repositories will be made searchable via a single, metadata-based web interface. In addition, baureka.index will also include references to restricted or older, analogue data sets in private or institutional hands and provide contact information that allows to directly request the material from its respective authors. The entries of baureka.index will in turn be searchable via suitable interfaces in national and international virtual libraries or research data infrastructures such as the Europeana, the German Digital Library or EUDAT.

In a second stage of the project once the core functionalities are in place, the platform can then be extended by three more modules. The module
baureka.publications will serve as an online data journal for historic building research. While in baureka.archive all records will already have received a digital object identifier that makes it possible to cite other people’s data or link own data with papers and books, in baureka.publications scientifically analysed and interpreted building research data and related articles will be able to undergo a peer-review process in order to underscore their status as a scientific contribution in their own right.

baureka.cloud is set to provide the community with a subject-specific exchange server, allowing research teams to access their data from different end devices and simplifying the joint processing of projects from different locations. Simultaneously, baureka.cloud is aimed at improving the subject-specific data management by, for example, assisting in the use of sustainable file formats and proposing suitable data structures, ultimately allowing the project data to be made available and archived in baureka.archive without having to go through a lengthy and costly revision workflow once the project has been completed. baureka.network, finally, will provide communication tools for the members of the discipline. The module should offer functions known from social networks, such as personal messages, discussion forums, appointment calendars and person-to-person data exchange.

In all of this, baureka.online will be developed with a view to the integration of both the more general techniques of the semantic web and the discipline-specific tools used for the processing and comprehensive analysis of 2D and 3D building research data, such as, among others, tools for searching for patterns in point clouds (De Luca, 2015) or for evaluating point clouds from wooden roof trusses (Pöchtrager, Styhler-Aydın, Döring-Williams, & Pfeifer, 2017; Döring-Williams, Styhler-Aydın, & Pöchtrager in this issue of SCIRES-IT). To this end, one of the central tasks will be to further develop specific metadata standards and exchange interfaces for the digital documentation of buildings. The project will therefore also tie in with the wider question of how to describe 3D resources in the field of cultural heritage, which is currently a much-debated subject of research (Hauck, & Kuroczyński, 2014; Blümel, 2013; Fernie, & D’Andrea, 2013; Münster, Pfarr-Harfst, Kuroczyński, & Ioannides (Eds.), 2016; Dudek, Blaise, De Luca, Bergerot, & Renaudin, 2015; Beetz, Blümel, Dietze, Fetahui, Gadiraju, Hecher, Krijnen, Lindlar, Tamke, Wessel, & Yu, 2016) and whose latest developments include plans to extend the .ifc 3D file format in such a way that it becomes possible to integrate point clouds (Beetz, & Krijnen, 2017).

Apart from an even more effective cooperation between researchers, conservationists and planners, the connections thus created within and between datasets will also permit the drafting of new, comparative research questions, the introduction of quantitative analysis methods and the development of topographical, epistemic or typological narratives that in a monographic approach have so far remained hidden. It will also be important to implement baureka.online in a way that allows it to become a forum for underlining and discussing the critical and analytical aspects in the interpretation of archival data, i.e. a forum, in which the work with baureka.online itself can be critically assessed.

4. baureka.online – schedule

In order to arrive at the intended benefits of baureka.online it is crucial to follow a cohesive schedule. In the following timetable, the project development is spread out over three years (Fig. 3). Overall, the authors have identified twelve work packages, of which some have to be completed at an early stage, some can only be tackled in the final phase of the development and some have to be considered on a continuous basis. At an early stage it will be important to re-evaluate and specify the objectives of baureka.online in order to articulate and communicate the platform policies. For the technical side of the platform, the most essential work package, which is set to become the main focus of the project’s first year, concerns the development and adoption of baureka’s metadata framework. Nevertheless, since the success of baureka as a whole will largely depend on the number of external archives and repositories to which it provides access, the project team will also use the first year to start acquiring and adapting external metadata

5 The Digital Object Identifier (DOI) is assigned by institutions belonging to the international non-profit organization DataCite. In Germany, the responsible body for allocating DOIs in the field of architecture is the Technische Informationsbibliothek TIB in Hanover.
sets and to conceive \textit{baureka.index}. In the second and third year, the authors will then not only step up the acquisition and adaptation of actual research data sets, but also address the technical development of \textit{baureka}'s core, its archive. Finally, towards the end of the three-year schedule attention will increasingly shift towards work packages like the design of \textit{baureka}'s front-end and the establishment of compelling strategies to make the platform sustainable. And of course, due to the nature of \textit{baureka.online}, the development proper of the platform will be accompanied along the whole project schedule by reports for and discussions with the community, for example via regular workshops.

For now, the authors are in the process of applying for public funding from the German Research Foundation, which particularly welcomes community-oriented projects like \textit{baureka}. In doing so, since the drivers behind \textit{baureka} up to now consist solely of architectural historians, the team have actively sought partnerships with experts in computing, data storage and data analysis, who have not only helped to better identify the central issues that need to be addressed within the project, but also given essential hints towards the definition of the work packages. While the authors had to learn about many technical aspects such as data modelling, metadata frameworks or the long-term storage of digital data, they also felt constantly rewarded and are particularly grateful to all those who kindly and willingly shared their knowledge. This was particularly helpful when it came to the question of the most recent standards of data science explicitly referred to in the funding guidelines of the German Research Foundation.

5. \textit{baureka.online} – project website

Similarly rewarding and challenging were the actions taken to activate the community and to discuss the use cases and objectives of \textit{baureka}. This is all the more true, because the community-based approach then also helped to meet an essential requirement of the application process, the duty to provide evidence for the necessity, practicality and successful outcome of the proposed project.

Since architects are trained to visualise concepts anyway, the authors in this context decided to design an interactive mock-up that shows the vision of \textit{baureka.online}. Such a visualisation did not merely help to communicate

![Implementation schedule for baureka.online.](image)

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\textbf{Year} & \textbf{Month} & \textbf{Day} & \textbf{Time} \\
\hline
1 & 1 & 1 & 12:00 PM \\
1 & 2 & 1 & 12:00 PM \\
1 & 3 & 1 & 12:00 PM \\
1 & 4 & 1 & 12:00 PM \\
1 & 5 & 1 & 12:00 PM \\
1 & 6 & 1 & 12:00 PM \\
1 & 7 & 1 & 12:00 PM \\
1 & 8 & 1 & 12:00 PM \\
1 & 9 & 1 & 12:00 PM \\
1 & 10 & 1 & 12:00 PM \\
1 & 11 & 1 & 12:00 PM \\
1 & 12 & 1 & 12:00 PM \\
2 & 1 & 1 & 12:00 PM \\
2 & 2 & 1 & 12:00 PM \\
2 & 3 & 1 & 12:00 PM \\
2 & 4 & 1 & 12:00 PM \\
2 & 5 & 1 & 12:00 PM \\
2 & 6 & 1 & 12:00 PM \\
2 & 7 & 1 & 12:00 PM \\
2 & 8 & 1 & 12:00 PM \\
2 & 9 & 1 & 12:00 PM \\
2 & 10 & 1 & 12:00 PM \\
2 & 11 & 1 & 12:00 PM \\
2 & 12 & 1 & 12:00 PM \\
3 & 1 & 1 & 12:00 PM \\
3 & 2 & 1 & 12:00 PM \\
3 & 3 & 1 & 12:00 PM \\
3 & 4 & 1 & 12:00 PM \\
3 & 5 & 1 & 12:00 PM \\
3 & 6 & 1 & 12:00 PM \\
3 & 7 & 1 & 12:00 PM \\
3 & 8 & 1 & 12:00 PM \\
3 & 9 & 1 & 12:00 PM \\
3 & 10 & 1 & 12:00 PM \\
3 & 11 & 1 & 12:00 PM \\
3 & 12 & 1 & 12:00 PM \\
\hline
\end{tabular}
\caption{Implementation schedule for \textit{baureka.online}.}
\end{table}
and promote the project's intentions within the community of building archaeologists, it also proved extremely useful in overcoming communicative obstacles that arose while trying to explain baureka to experts in computer science. The mockup can be tested online at https://baureka.online, which is currently a website for the communication and information on our initiative and which will gradually become the site of the actual platform (Fig. 4).

When visiting baureka's responsive and bilingual project-website, visitors are able to once more peruse the information presented in this article. The starting page provides an overview on the project of baureka.online with its modular sections. Visitors can take a closer look at the guiding principles for baureka and the work packages necessary for its realisation. Most recent information is listed in the sections "News" and "Downloads", in which all activities are recorded – often even providing additional material, such as various data sets from baureka's feasibility studies or images from its initial workshop in 2017.

Even if the project website remains a valuable resource for information on how the project develops, the best way towards understanding how baureka will function is to take a look at the mock-up. Here visitors are welcome to explore baureka's potentials and functionalities. Whenever the user scrolls down, a menu sticks to the upper edge of the browser window and allows the user to navigate the site. All intended modules of baureka can then be experienced as a preview of what the authors hope to develop. baureka.index allows the user to perform a mock search for various internal and external datasets. baureka.archive and baureka.publications demonstrate how the landing pages for raw and interpreted data sets might look like.

6. Conclusion

Overall, the development of the research data infrastructure baureka.online aims to improve the cooperation within the specialist community of historical building research and building archaeology in the German-speaking area and to create synergies between science and practice. Currently, there is no digital platform for the valorisation and overarching analysis of existing scientific documentations. The paper tried to explain the approach of baureka.online as a specific, user-oriented repository and research tool for the community that is being developed through a participatory design process. Following this path, the authors hope to be able to establish this urgently needed and desired piece of central digital infrastructure, which will make it possible to further propel the research methodology of building archaeology into the digital age and to efficiently position the discipline in an international context.

Fig. 4: Welcome page of baureka's interactive mock-up.
REFERENCES


AKBF. Retrieved from https://www.arthistoricum.net/netzwerke/akbf/


Baureka.online (2018). Retrieved from https://baureka.online


Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Retrieved from https://openaccess.mpg.de/Berliner-Erklärung


EUDAT. Retrieved from https://eudat.eu/
GERDI. Retrieved from http://www.gerdi-project.de/
IANUS. Retrieved from https://www.ianus-fdz.de/
KulaDig. Retrieved from https://www.kuladig.de/
Massive Point Clouds for eScience. Retrieved from https://www.esciencecenter.nl/project/massive-point-clouds-for-esciences/
MonArch. Retrieved from http://www.monarch.uni-passau.de/
Monumente-Online. Retrieved from https://www.monumente-online.de/


PANGAEA/KOMFOR. Retrieved from https://www.pangaea.de/ and https://www.komfor.net/


PHAROS. Retrieved from http://pharosartresearch.org/


RADAR. Retrieved from http://www.radar-project.org/


TopoRaZ. Retrieved from http://khi.phil-fak.uni-koeln.de/25891.html

