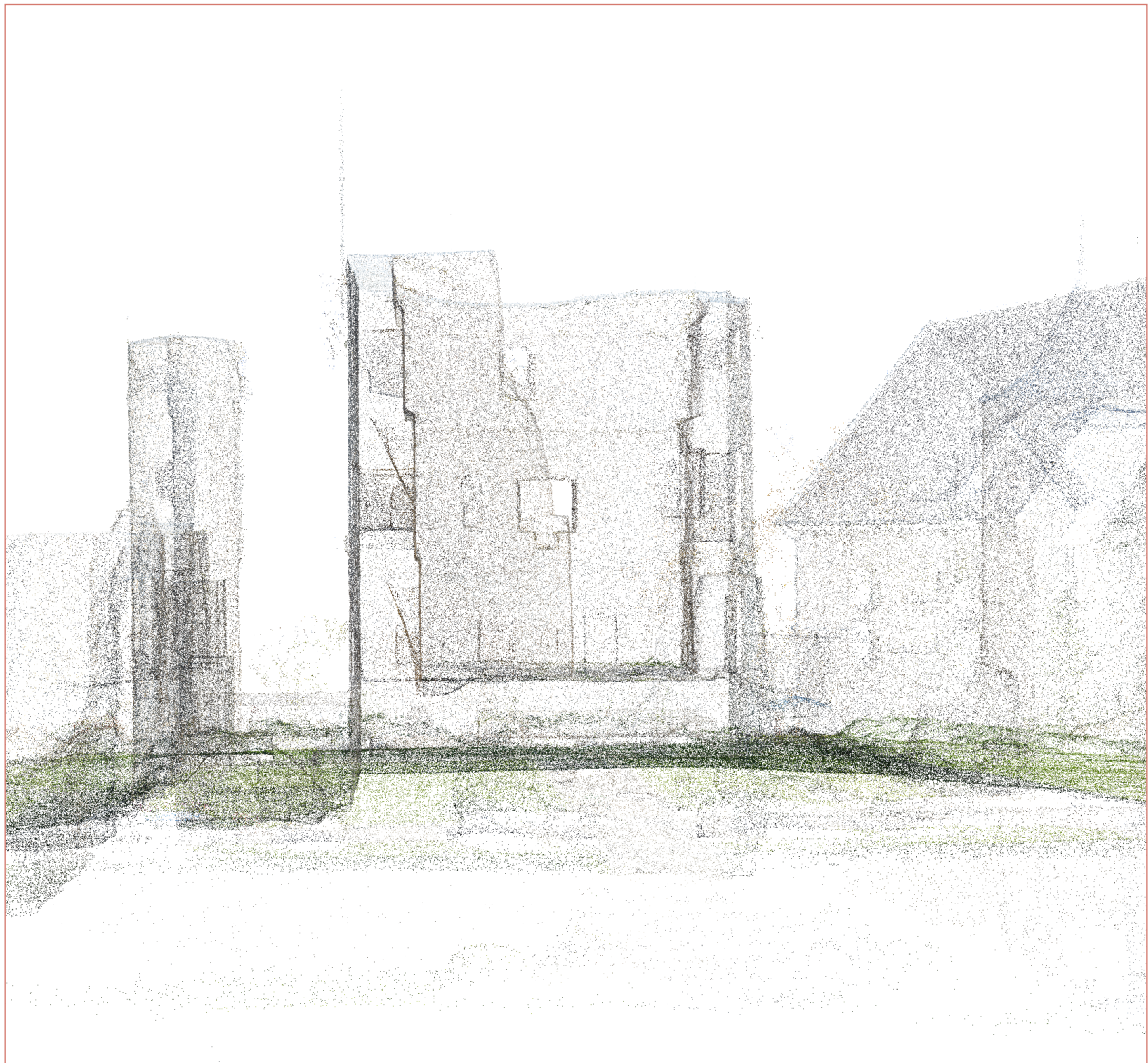




BUILD ON DATA

AUF DATEN BAUEN

Forschungsdaten in der Historischen Bauforschung und Denkmalpflege



Tagungsband zur Community-Tagung des DFG-Projekts baureka.online
am 4. und 5. Mai 2023 an der Technischen Universität Berlin

Abbildung Titelblatt

Sparse Point Cloud des Haus Heyden in Aachen-Richterich (Christian Klosterkötter / Felix Martin)

Impressum

Build on Data / Auf Daten bauen. Forschungsdaten in der Historischen Bauforschung und Denkmalpflege.

Eine Community-Tagung des DFG-Projekts baureka.online

© 2024 by Anke Naujokat and Sophie Helas

Herausgegeben von

Anke Naujokat, Sophie Helas

RWTH Aachen University

Lehrstuhl für Architekturgeschichte | Chair of Architectural History

Schinkelstraße 1

D-52062 Aachen

Germany

Layout und Grafik

Lara Draschoff, Felix Martin

Redaktion und Satz

Annika Dues, Sophie Helas, Svenja Kruijer

Founding

Diese Publikation wird von der Deutschen Forschungsgemeinschaft, Projektnummer 454194613, finanziert.



This work is licensed under Creative Commons Attribution 4.0 International.

Ausgenommen von der Lizenz sind Fotos, auf denen Personen erkennbar abgebildet sind, sowie alle Logos.

Verfügbar über das institutionelle Repository der RWTH Aachen University. DOI: 10.18154/RWTH-2024-04215

BUILD ON DATA

AUF DATEN BAUEN

Forschungsdaten in der Historischen Bauforschung und Denkmalpflege

Tagungsband zur Community-Tagung des DFG-Projekts baureka.online
am 4. und 5. Mai 2023 an der Technischen Universität Berlin

Herausgegeben von Anke Naujokat und Sophie Helas

In Gedenken an

Hermann Schlimme (1969-2023)



RWTHAACHEN
UNIVERSITY

FG | BSG
FACHGEBIET BAU- UND
STADTRAUGESCHICHTE



FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Gefördert durch
DFG Deutsche
Forschungsgemeinschaft

Inhalt

1.	Auf Daten bauen – Build on Data	6
2.	Auf Daten bauen – nicht ohne Forschungsdatenmanagement! <i>Matthias Razum</i>	10
3.	Das Forschungsdatenportal baureka.online. Die Vision <i>Anke Naujokat</i>	12
4.	Das Forschungsdatenportal baureka.online. Ein Fachrepositorium für die Historische Bauforschung <i>Nadine Marcinczik</i>	16
5.	Das Forschungsdatenportal baureka.online. baureka.index als zentraler Nachweiskatalog für Bauforschungsdaten <i>Tobias Glitsch</i>	22
6.	Madīnat al-Zahrā' <i>Heike Lehmann / Simon Trischberger</i>	28
7.	Forschungsdaten im Lehrkontext <i>Lukas Stampfer / Eva Kodžoman / Marina Döring-Williams</i>	38
8.	Die Datenbank Bauforschung/Restaurierung Baden-Württemberg <i>Claudia Mohn</i>	46
9.	Austausch und Standards für Daten und Metadaten photogrammetrischer Baufnahmen <i>Claudia Mächler / Andreas Noback</i>	56
10.	Forschungsdaten im Projekt „Synagogen-Gedenkbuch Hessen“ <i>Fani Gargova / Tilmann Gempp-Friedrich</i>	62
11.	Historischen Baubestand digital erfassen, modellieren, publizieren <i>Jörg Richter</i>	66

12.	Fehlende Bausteine in der GND: Normdaten zu Bauwerken	
	<i>Julia Rössel / Hanna-Lena Meiners</i>	70
13.	Das DFG-Projekt IDOVIR	
	<i>Markus Wacker</i>	78
14.	Coscine – Make Your Research Data FAIR	
	<i>Lukas C. Bossert</i>	82
15.	Bauforschung in der Lehre	
	<i>Melanie Kim-Lan Nguyen / Liang Song</i>	88
16.	3D Infrastructure for Digital Reconstructions	
	<i>Igor Bajena / Clemens Beck</i>	94
17.	Rechtsfragen in der Historischen Bauforschung	
	<i>Grischka Petri / Oliver Vettermann</i>	100
18.	Tagungsprogramm	104

Tobias Glitsch

5. Das Forschungsdatenportal baureka.online

baureka.index als zentraler Nachweiskatalog für Bauforschungsdaten

DOI: 10.18154/RWTH-2024-04231

baureka.index als plattformübergreifende Suchoberfläche

Mit seinem auf die Bedürfnisse des Fachs zugeschnittenen Aufbau, seinem spezifischen Metadatenschema und seiner Fokussierung auf die Daten nur dieser einen Disziplin stellt baureka.storage der Community der Historischen Bauforschung ein Repositorium zur Verfügung, das für die Datengeber*innen zahlreiche Vorteile gegenüber generischen oder hochschuleigenen Archivierungslösungen bietet. Zieht man darüber hinaus in Betracht, dass auch die Leitlinien und Best Practices des Forschungsdatenmanagements zunehmend empfehlen, bei der Veröffentlichung von Forschungsdaten fachspezifische Angebote zu bevorzugen (DFG, 2021, S. 1; RWTH Aachen, 2022), hat das Repositoriumsmodul von baureka.online daher das Potential, sich als zentraler Ort für die Ablage und Bereitstellung von Bauforschungsdaten zu etablieren und damit einen beträchtlichen Teil der für eine Nachnutzung relevanten Datensätze auf sich zu ziehen. Zugleich lässt jedoch das bei der Historischen Bauforschung ungewöhnlich breite Spektrum an Akteuren erwarten, dass gewisse Daten auch nach dem Betriebsbeginn von baureka.online weiterhin ihren Platz in anderen Speicherstrukturen finden werden. Zudem strebt baureka.online ausdrücklich nicht an, dass bereits anderswo verfügbare Daten auf der Plattform nochmals zweitveröffentlicht werden, sondern geht vielmehr davon aus, dass jene Daten auch weiterhin allein an ihrem bisherigen Ort verbleiben.

Auch wenn baureka.online bereits durch die Bereitstellung des Repositoriums und die dort mögliche Beschreibungstiefe entscheidend dazu beitragen wird, Forschungsdaten der Historischen Bauforschung besser auffindbar und besser nachnutzbar zu machen, dürfte eine Suche, die sich allein auf die repositoriumseigenen Bestände beschränkt, daher meist kein vollständiges Bild der verfügbaren Datensätze liefern. Um die Schwelle für die Nachnutzung von Daten möglichst niedrig zu halten, wäre jedoch eine Recherchelösung wünschenswert, die es den Datennehmer*innen erlaubt, sich an einem einzigen Ort und mit einer einzigen Suchanfrage einen gezielten Überblick über sämtliche zu einem bestimmten Thema existierenden Daten zu verschaffen. Idealerweise benötigt die Fachcommunity also ein Suchportal, das neben baureka.storage auch alle anderen infrage kommenden Datenablagen in die Suche miteinbezieht, das aus generischeren Speichersystemen aber lediglich die bauforschungsrelevanten Datensätze übernimmt und das darüber hinaus alle Datensätze, unabhängig von ihrer Provenienz, derart detailliert erschließt, dass auch komplexere Suchanfragen möglich sind.

baureka.online reagiert auf dieses Desiderat mit dem Modul baureka.index, das als zweites Standbein der Plattform zum eigentlichen Repositoriumsmodul, baureka.storage, hinzutritt. baureka.index soll dabei einerseits die Suche innerhalb von baureka.storage übernehmen, das als Repositorium somit gar keine separate Suchmaske mehr erhält. Andererseits soll baureka.index in der Suche aber auch das Material aus externen Quellen berücksichti-

gen und auf diese Weise als zentraler Nachweiskatalog für die Daten der Historischen Bauforschung dienen (Abb. 1).

Die Suchoptionen und die Erstellung der Datenerschließung

Die möglichen Suchanfragen orientieren sich zunächst einmal an der Struktur von baureka.storage und des zugehörigen Metadatenschemas (Göller / Marcinczik / Staschok, 2024). Dadurch, dass die dokumentierten Objekte darin zum einen unter dem Namen des Gesamtgebäudes verschlagwortet sind, sich zum anderen aber auch die speziell untersuchten Teilbereiche des Gebäudes angeben lassen, wird es baureka.index, erstens, erlauben, sowohl nach Forschungsdaten zu einem konkreten Objekt zu suchen als auch eine Recherche nach Daten zu einer bestimmten Art von Bauteil oder Architekturelement durchzuführen. Je nachdem, wie ausführlich die Objekte durch die Datengebenden beschrieben wurden, lassen sich die Suchergebnisse sodann, zweitens, auf einen bestimmten geographischen oder zeitlichen Kontext, auf bestimmte am Planungs- und Realisierungsprozess beteiligte Akteure sowie bestimmte beim Bau verwendete Materialien oder Konstruktionen einschränken. Schließlich wird, drittens, eine Filterung nach der Art des gesuchten Datenmaterials möglich sein, die Nutzer*innen der Plattform werden bei der Suche also gezielt angeben können, dass sie für ihre Zwecke z. B. speziell eine Fotodokumentation, einen zweidimensionalen Plansatz oder eine Punktwolke benötigen.

Während sich die dafür erforderlichen Informationen bei den in baureka.storage archivierten Datensätzen bereits im Uploadprozess mit abfragen und danach ohne größere Transformationsschritte der eigentlichen Suchfunktion zugrunde legen lassen, muss im Falle des Materials aus externen Quellen die Beschreibung der Daten jedoch zunächst aktiv importiert und gegebenenfalls sogar ergänzt werden. Je nach Charakter und Ablageort der externen Datenbestände erfordert deren Aufnahme in den Nachweiskatalog von baureka.online allerdings unterschiedliche Herangehensweisen und bringt eigene Herausforderungen mit sich.

Der automatisierte Metadateneingest aus fachspezifischen Archiven

Bei den Datenbeständen aus anderen fachspezifischen Archiven, dürfte es in der Regel möglich sein, einen automatisierten Ingest der das Material beschreibenden Metadaten vorzunehmen. Dies betrifft neben den Repositorien benachbarter Projektinitiativen wie des DFG-3D-Viewers (<https://dfg-viewer.de/dfg-3d-viewer>; <https://3d-repository.hs-mainz.de/>), des IDOVIR-Systems (<https://idovir.com/>) oder des Fachinformationsdienstes BAU-digital (<https://repo.fid-bau.de/>) insbesondere auch die – abgesehen von der Datenbank Bauforschung / Restaurierung Baden-Württemberg (<https://www.bauforschung-bw.de/>) allerdings noch nicht öffentlich zugänglichen – Datenablagen der Landesdenkmalämter. Ein vom Team des DFG-3D-Viewers initiiertes Vergleich der genutzten Metadatenschemata, der im Vorfeld eines gemeinsamen Workshops Ende 2022 von Vertreter*innen der diversen Projekte auf einem Miro-Board (https://miro.com/app/board/uXjVPDUEGMY=?share_link_id=777858896545) zusammengefasst wurde, sowie die Analyse der in der Datenbank Bauforschung / Restaurierung hinterlegten Informationen haben gezeigt, dass die Forschungsdaten auch dort in der Regel in einer für detaillierte Suchanfragen geeigneten Form und Tiefe erschlossen sind oder zumindest erschlossen werden könnten. Die Aufgabe besteht in

diesen Fällen also vor allem darin, die dortigen Beschreibungen auf das Metadatenschema von baureka.online abzubilden sowie eine entsprechende Importfunktion zu entwickeln, die die externen Plattformen regelmäßig auf Neuzugänge überprüft und die dort hinterlegten Metadaten selbständig in die für baureka.index benötigte Form überführt.

Die Rolle von Normdaten und kontrollierten Vokabularen

Die Transformation der Metadaten umfasst allerdings keinesfalls nur die Übertragung der Informationen in die entsprechenden Felder des Datenmodells von baureka.online. Wie in der Dokumentation des Metadatenschemas angedeutet, ist in baureka.online vorgesehen, Eintragungen, wo immer möglich, nicht über Freitexte, sondern anhand von Normdaten-identifiern vorzunehmen. Dies erlaubt es nicht nur, die Beschreibungen zu vereinheitlichen, bei der Suche auch Synonyme der Suchtermini miteinzubeziehen und die Treffermengen auf über- oder untergeordnete Begriffe und Entitäten auszudehnen oder zu verengen. Angesichts des Umstands, dass insbesondere die Normdatensätze zu Gebäuden und Ensembles zunehmend um Entity Facts zu Standort, Datierung, Entwurfsverfasser und grundsätzlichem Bautyp angereichert werden sollen (Rosenkötter, 2022, S. 8; DNB, 2023), ergibt sich damit vielmehr auch die Möglichkeit, Eigenschaften sogar dann für eine Suche heranzuziehen, wenn das entsprechende Metadatenfeld in der unmittelbaren Beschreibung des Datensatzes gar nicht ausgefüllt wurde oder im Metadatenschema der betreffenden Nachbarplattform nicht einmal existiert.

In baureka.online selbst spielen im Hinblick auf die übergreifende Durchsuchbarkeit des Datenmaterials insbesondere die im Beitrag zu baureka.storage genannten Normdateien und Vokabulare eine Rolle. Dagegen greifen die externen Archive zum Teil auf andere Normdateien und Vokabulare zurück oder nutzen – wie im Falle der Datenbanken der Denkmalämter – eigene Identifier und Begriffssammlungen. Der Metadatenimport nach baureka.index erfordert somit stets auch eine Überführung der auf der Basis alternativer Vokabulare hinterlegten Angaben in das System der von baureka.online genutzten Normdateien. Wo die dazu nötigen Konkordanzen zwischen den jeweiligen IDs noch nicht vorliegen oder wo die anderen Archive auf eine vokabular- beziehungsweise normdatenbasierte Beschreibung sogar vollständig verzichten, muss baureka.online daher zugleich daran mitwirken, die Äquivalenzen zwischen den jeweiligen Vokabularen zu klären, Werkzeuge zur Umwandlung der Freitextbegriffe in Vokabulareinträge zu entwickeln oder bei den externen Plattformen sogar eine Anpassung ihrer Beschreibungswerkzeuge zu erreichen.

Die Filterung nach Relevanz und die Einbindung von baureka.index in ein Suchökosystem

Trotz zahlreicher Überschneidungen in der Art der Datenbestände unterscheiden sich zumindest die Plattformen der aktuell entstehenden Nachbarprojekte allerdings in zentralen Punkten von baureka.online. So sind etwa in IDOVIR oder dem Repositorium des DFG-3D-Viewers zwar für gewisse Objekte durchaus auch Fotodokumentationen oder per Laserscan beziehungsweise per Structure-from-Motion generierte Punktwolken zu erwarten. Das am Gebäude erhobene Material dient dort aber primär als Begründung für bestimmte Rekonstruktionsentscheidungen, ist dem fertigen Modell daher tendenziell untergeordnet und kann sogar komplett fehlen. Das Datenarchiv des Fachinformationsdienstes BAUdigital ist sogar noch offener angelegt. Im Gegensatz dazu geht baureka.online als dezidiertes Reposi-

torium der Historischen Bauforschung stets von der im Feld erhobenen Dokumentation des Ist-Zustands aus, stellt diese auch bei der Suche ins Zentrum und beschränkt sich zudem keinesfalls allein auf dreidimensionale Daten.

Zusätzlich zum reinen Mapping der Metadaten muss die Ingestfunktion daher vor dem eigentlichen Import auch nochmals eine Filterung der fremden Einträge vorsehen, die sicherstellt, dass aus dem Material der externen Archive am Ende tatsächlich nur die bauforscherisch relevanten Datensätze in baureka.index übernommen werden und die Präzision der Suchtreffer nicht durch Einträge beeinträchtigt wird, die von vornherein aus dem von baureka.online abgedeckten Bereich herausfallen. Analog dazu müssen die Schnittstellen zu den eigenen Daten so konzipiert sein, dass sie es den Nachbarinitiativen ebenfalls erlauben, aus dem Bestand gezielt die für deren Anliegen interessanten Datensätze herauszugreifen und zu adressieren. In letzter Konsequenz erhält baureka.index damit das Potential, zu einem integralen Bestandteil eines größeren Verbunds von Recherchewerkzeugen zu werden, in dem für unterschiedliche Anwendungsgebiete jeweils eigene, gezielt auf die jeweiligen Anforderungen abgestimmte Sucheinstiege bestehen.

Die manuelle Katalogisierung externer Bestände und die Herausbildung einer Aggregatorfunktion

Im Gegensatz zu den Speicherlösungen der Nachbarprojekte sowie den Datenarchiven der Landesdenkmalämter, die ihre Metadatenschemata von vornherein auf fachspezifische Aspekte zuschneiden können, sind generische Repositorien dazu gezwungen, sich in ihrem Datenmodell auf wenige, allgemein verwendbare Felder zu beschränken. Der Umstand, dass sich dort Datensätze unterschiedlicher Fächer abwechseln, macht es auch deutlich komplizierter, das relevante Material dort überhaupt zu identifizieren. Sofern nicht die eigentlichen Datenpakete nochmals weitergehende Anmerkungen enthalten, für deren Extraktion dann aber ebenfalls erst einmal ein entsprechendes Verfahren entwickelt werden müsste, erscheint es in diesen Fällen daher kaum realistisch, die für eine Erschließung über baureka.index erforderlichen Informationen automatisiert einzulesen. Hinzu kommt, dass die Anzahl an generischen Repositorien in den vergangenen Jahren derart zugenommen hat, dass sich im Rahmen eines automatisierten Ingests eine vollständige Berücksichtigung aller infrage kommenden Speicherorte nicht mehr gewährleisten lässt.

Um dennoch auch die Datenbestände aus institutionellen Repositorien und anderen generischen Langzeitarchivierungsplattformen in baureka.index einbeziehen zu können, wird es baureka.online daher zusätzlich zu den automatisierten Ingestverfahren erlauben, externe Datensätze manuell zur Aufnahme in den Nachweiskatalog vorzuschlagen. Dazu wird es möglich sein, auch mit Projekten, bei denen die eigentlichen Daten nicht in baureka.storage aufgenommen werden sollen, zunächst den für den Datenupload konzipierten Workflow zu durchlaufen, am Ende jedoch anstelle eines oder mehrerer Datenpakete nur einen Link auf den externen Speicherort oder – bei sensiblem Material – nur eine Kontaktadresse zu hinterlegen (Abb. 1). Die Option, sich allein auf die Angabe einer Kontaktadresse zu beschränken, gestattet es in diesem Zusammenhang sogar, rein analoges Material aus der Zeit vor Einführung digitaler Werkzeuge und Arbeitstechniken mit in baureka.index aufzunehmen. Dabei hoffen wir darauf, dass gerade die an Universitäten und Hochschulen tätigen Bauforscher*innen zunehmend ein Eigeninteresse dafür entwickeln werden, auch ihrem bestehenden Material durch einen Eintrag in baureka.index eine zusätzliche Sichtbarkeit zu verleihen.

Insgesamt kann sich baureka.online auf diese Weise nicht nur zu einer Anlaufstelle für die Datenrecherche entwickeln, die gegenüber anderen digitalen Angeboten auf dem Gebiet der Baudokumentation einen erheblichen Mehrwert aufweist. Gerade für die über generische Repositorien verstreuten oder bisher undigitalisierten Datenbestände der Historischen Bauforschung kann baureka.online auch eine wichtige Aggregatorfunktion übernehmen, die es ihrerseits erlaubt, das Wissen um jenes ansonsten kaum auffindbare Material in die übergreifenden Systeme der NFDI oder der deutschen und europäischen digitalen Bibliotheken einzuspeisen.

Literatur

Sandra Göller / Nadine Marcinczik / Thalia Staschok: „Dokumentation des baureka-Metadatenschemas - Version 1.01“, *Zenodo*, 2024. <https://zenodo.org/records/11277405>

Internetquellen

DFG, 2021: DFG: Checkliste „Umgang mit Forschungsdaten“. URL: <https://www.dfg.de/resource/blob/174732/3c6343eed2054edc0d184edff9786044/forschungsdaten-checkliste-de-data.pdf> [Stand: 10.06.2024].

DNB, 2023: DNB: Erfassungs- und Verwendungshilfe EH-S-06-1. Bauwerke, Großplastiken, Grab- und Denkmäler. URL: <https://wiki.dnb.de/pages/viewpage.action?pageId=116426796> [Stand: 10.06.2024].

Rosenkötter, 2022: Martha Rosenkötter: Vortrag „Das GND-Anwendungsprofil Bauwerke. Sammlung bauwerksbezogener Anforderung an die GND“. URL: https://wiki.dnb.de/download/attachments/272246612/20221114_ForumBauwerkeAnwendungsprofil.pdf [Stand: 10.06.2024].

RWTH Aachen, 2022: RWTH Aachen: Blogbeitrag „FDM erklärt – Was ist eigentlich ein Repository?“. URL: <https://blog.rwth-aachen.de/forschungsdaten/2022/02/03/fdm-erklart-repositorium/> [Stand: 10.06.2024].

Abbildung

Abb. 1 Funktionsweise von baureka.index: Sammlung der Metadaten (oben), Trefferanzeige (links), Datenzugriff (unten) (Tobias Glitsch)

Tobias Glitsch
RWTH Aachen University
Lehrstuhl für Architekturgeschichte
Schinkelstraße 1
52062 Aachen
glitsch@ages.rwth-aachen.de
ORCID: 0000-0002-9911-2955

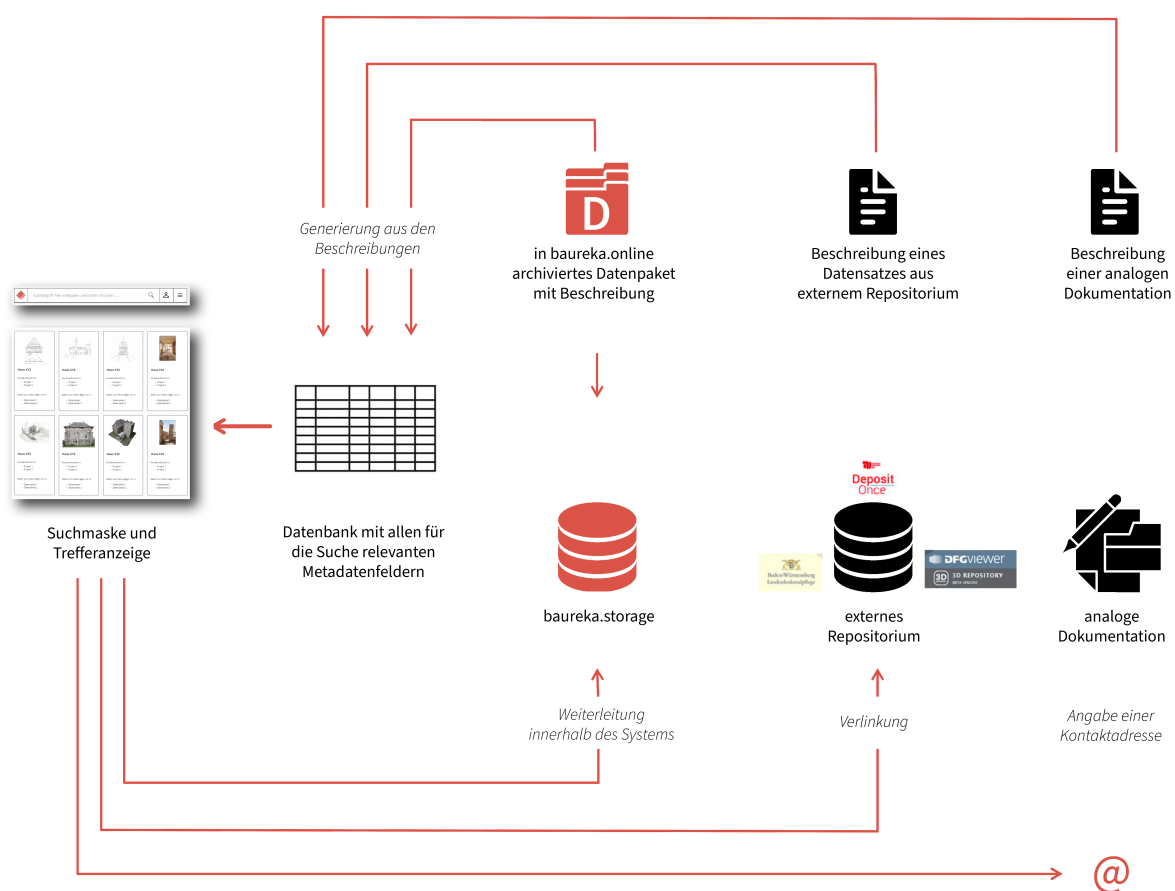


Abb. 1 Funktionsweise von baureka.index