

MSG Moderne Stadtgeschichte, Bd. 55/1 (2024), 116-123

DOI: 10.60684/msg.v55i1.35

Rita Gudermann

Leibniz-Institut für Gesellschafts- und Raumforschung (IRS)

Paul Perschke

Leibniz-Institut für Gesellschafts- und Raumforschung (IRS)

Öffentliches Engagement bei der Erschließung von Archivalien. Citizen Science-Ansätze im Archiv

MSG Moderne Stadtgeschichte

ISSN: 2941-6159 online

<https://moderne-stadtgeschichte.de>

Dieses Werk steht unter der [Lizenz Creative Commons Namensnennung 4.0 International](#).

Bei den Abbildungen sind eingeschränkte Lizenzformen möglich, Weiternutzungsrechte sind gesondert abzuklären.

© Rita Gudermann/Paul Perschke 2024



Rita Gudermann / Paul Perschke

Öffentliches Engagement bei der Erschließung von Archivalien. Citizen Science-Ansätze im Archiv

1. Einleitung: Entstehung und Ziele des Projekts „CitizenArchives“

Citizen Science-Projekte innerhalb der historisch arbeitenden Disziplinen konzentrierten sich bisher vorrangig auf die Erforschung historischer Sachverhalte, weniger auf die alltägliche Arbeit der Archive. Dabei bieten bürgerwissenschaftliche Ansätze gerade bei der Erschließung von Archivalien vielversprechende Möglichkeiten. Als Beispiel eines Projekts, das derartige Ansätze mit neuen technischen Verfahren kombiniert, sei im Folgenden das Projekt „CitizenArchives“ vorgestellt.

Im Projekt „CitizenArchives“ sollten Citizen Science-Ansätze mit Technologien des Semantic Web verbunden werden, um Archivbestände mit Hilfe kostenfreier Ressourcen von Nutzer*innen in geordneten, IT-gestützten Prozessen zu digitalisieren, zu erschließen und der Öffentlichkeit zur Verfügung zu stellen. Anhand der Bestände der Wissenschaftlichen Sammlungen zur Architektur- und Planungsgeschichte der DDR und mit Hilfe verschiedener Pilotnutzer*innen wurden Prozesse und ein Produkt zur einfachen Erschließung und Veröffentlichung von digitalisierten Beständen kleinerer, spezialisierter Archive entwickelt, deren personelle und finanzielle Möglichkeiten dies häufig nicht zulassen. Von wesentlicher Bedeutung waren dabei Ressourcen wie Engagement, Wissen und Zeit interessierter und fachkundiger Nutzer*innen – Wissenschaftler*innen oder Zeitzeug*innen –, die insbesondere bei der Verschlagwortung und Beschreibung von Beständen halfen.

Im Folgenden wird das für ein gesellschaftswissenschaftliches Forschungsinstitut ungewöhnliche Hybrid-Projekt aus Wissenschaft und Wirtschaft vorgestellt, das die Wissenschaftlichen Sammlungen des Leibniz-Instituts für Raumbezogene Sozialforschung (IRS) in Erkner bei Berlin im Dezember 2021 gemeinsam mit dem Unternehmen Programmfabrik GmbH umsetzten. Finanziert wurde das zweijährige Projekt „CitizenArchives“ aus der Förderlinie „Innovations-

programm für Geschäftsmodelle und Pionierlösungen (IGP)“ des Bundesministeriums für Wirtschaft und Klimaschutz (BMWK). Erklärtes Ziel des Vorhabens war die „Entwicklung eines Citizen-Science- und Semantic-Web-basierten Verfahrens zur Digitalisierung und Erschließung der Bestände kleiner Archive“.¹

2. Projektbeteiligte

Die Wissenschaftlichen Sammlungen zur Bau- und Planungsgeschichte der DDR am IRS sind ein kleines, in ihren Beständen jedoch stetig wachsendes Spezialarchiv mit großer Bedeutung auch für die Stadtgeschichte der DDR.² Neben Unterlagen ehemaliger Bau- und Planungsinstitutionen bewahrt das Archiv mittlerweile über 140 Vor- und Nachlässe von Persönlichkeiten aus Stadtplanung, Architektur, Kunst, Forschung und Lehre auf. Dementsprechend vielfältig ist das Sammlungsgut: Neben persönlichen Unterlagen und Planungsdokumenten besteht ein Großteil der Sammlungen aus Bildmaterial wie Zeichnungen und Plänen sowie mehreren zehntausend Fotos in Form von Abzügen, Negativen und Dias. Ein weiterer, immer wichtiger werdender Aspekt ist zudem die Archivierung bereits digital vorliegender Bestände wie Forschungsdatensammlungen, Websites oder Datenbanken. Für eine rasche und vollständige Digitalisierung, Erschließung und Nutzbarmachung von Beständen fehlen dem Archiv jedoch langfristig die personellen und finanziellen Mittel. Damit stehen die Sammlungen des IRS exemplarisch für unzählige kleine, spezialisierte Sammlungen und Archive, die zwar einen wichtigen Beitrag zur Bewahrung von Kulturgut leisten, häufig aber nur über eine bescheidene Ausstattung verfügen. So bleiben viele kulturelle und historische Schätze zwangsläufig ungehoben und unzugänglich für Bildung, Forschung und Öffentlichkeit.

Der Projektpartner Programmfabrik ist ein Berliner Softwarehaus, das seit 2003 die webbasierte Software „Easydb“ (2023 umbenannt in „FYLR“) zur Verwaltung digitaler Medieninhalte wie zum Beispiel Bildern, Videos, Musik und Text-Dateien entwickelt. Neben diesem technischen Know-how bringt Programmfabrik langjährige Wissenschafts- und Kulturprojekterfahrung mit. Zu den Kunden des Unternehmens gehören Archive sowie Forschungs- und Bildungseinrichtungen im gesamten deutschsprachigen Raum. Die Wissenschaftlichen Sammlungen erwarben „Easydb“ im Jahre 2020 als einen Baustein ihrer neuen digitalen Infrastruktur.

¹ Projektantrag igp03-0757 vom 29.01.2021 im Rahmen der Fördermittellinie des BMWK, „Innovationsprogramm für Geschäftsmodelle und Pionierlösungen“, dritte Ausschreibungsrunde, Innovationen im Bereich Bildung und Informationszugang mit hohem „sozialem Impact“.

² Vgl. Onlineauftritt: <http://stadt-raum-geschichte.de> [21.04.2024].

3. Rahmenbedingungen

Als Projektpartner boten sich die Wissenschaftlichen Sammlungen nicht nur wegen ihrer relativ jungen Bestände an, sondern auch aufgrund ihres über Jahrzehnte aufgebauten Netzwerks von Citizens, mit deren Hilfe die Mechanismen und Abläufe in einer auf Archive zugeschnittenen Citizen Science getestet und optimiert werden sollten. Die Teilhabe und Mithilfe von Nutzer*innen eröffnete die Möglichkeit, Sammlungsgut schon in der frühen Phase der Digitalisierung und Erschließung öffentlich zugänglich zu machen. Dabei sollte allerdings auf die Moderation und Qualitätskontrolle der gewonnenen Inhalte und Daten durch Archivmitarbeiter*innen nicht verzichtet werden. Solch eine Qualitätskontrolle wurde insbesondere durch Mitautor Paul Perschke, wissenschaftlicher Mitarbeiter der Sammlungen, geleistet. Die dazu notwendigen Prozesse wurden im Rahmen des Projekts vor allem durch die Wissenschaftlichen Sammlungen herausgebildet und beschrieben und anschließend durch von Programmfabrik entwickelte zusätzliche Plug-ins für die Software „Easydb“ technisch unterstützt und vereinfacht.

Eine Besonderheit der Wissenschaftlichen Sammlungen ist ihre enge Vernetzung mit der stadtgeschichtlichen Forschung im Forschungsschwerpunkt „Zeitgeschichte und Archiv“ des IRS. Nicht selten werden Bestände mit Blick auf laufende Forschungsarbeiten im Hause eingeworben oder bevorzugt zugänglich gemacht. Doch nicht nur von Seiten der Wissenschaft werden Anfragen an die Wissenschaftlichen Sammlungen gerichtet: Informationsgesuche stammen auch aus der Praxis, etwa Fragen von Architekt*innen oder Ingenieur*innen nach Bauzeichnungen oder der Verwendung bestimmter Materialien in historischen Gebäuden aus dem gesamten Gebiet der früheren DDR.

Eine weitere Besonderheit ist die enge Bindung zwischen Archiv und den Akteur*innen der Stadt- und Architekturgeschichte aus der Zeit der DDR. Regelmäßig finden „Werkstattgespräche“ statt, bei denen es immer wieder zu angeregten und auch sehr kontroversen Diskussionen kommt. Debatten entzünden sich dabei häufig an dem von den Akteur*innen erhobenen Anspruch, die eigentlichen und einzigen Expert*innen für ihr Thema zu sein. Damit stellen sie die Kompetenz von Wissenschaftler*innen in Frage, die in der Regel jünger sind, keine praktischen Erfahrungen in der Architektur und Planung haben und häufig auch nicht in der DDR sozialisiert wurden. Die Wissenschaftler*innen sind somit in ganz besonderer Weise herausgefordert, ihr historisches Wissen und ihre Methodenkenntnis darzustellen und zu verteidigen.

4. Vorgehensweise

Entwickelt wurde ein Verfahren, bei dem den Citizens über ein Web-Portal noch unerschlossene Bilder präsentiert werden, die sie dann über eine Online-Oberfläche mit Metadaten versehen können. Die Partizipationsmöglichkeiten der Citizens waren hierbei zunächst beschränkt auf das Ausfüllen weitgehend vorgegebener Masken, um die Einheitlichkeit der entstehenden Metadaten zu gewährleisten. Allerdings gibt es auch Datenfelder für freie Kommentare und Hinweise. Die meisten Beitragenden blieben zudem anonym. Hier soll in Zukunft weiter daran gearbeitet werden, eine lebendige Community rund um die Architektur- und Planungsgeschichte der DDR an die Wissenschaftlichen Sammlungen zu binden und in die im Institut geleistete Forschungsarbeit zu integrieren.

Um die von den Wissenschaftlichen Sammlungen des IRS und dem Unternehmen Programmfabrik entwickelten Technologien und Erfahrungen auch anderen kleinen Archiven zugutekommen zu lassen, wurde anfangs das an der Universität Oxford gehostete, kostenlose Tool „Zooniverse“ genutzt, das bisher überwiegend für naturwissenschaftliche Projekte eingesetzt wird. Zwar war es auf diese Weise möglich, sehr schnell eine funktionierende Plattform für die Dateneingabe aufzusetzen und über einen geheimen Link für Interessierte zu öffnen. Allerdings erwies sich das Tool aufgrund einiger Einschränkungen dann doch für den geplanten Anwendungszweck als suboptimal. So war es beispielsweise nicht möglich, den Citizens eine Thumbnail-Übersicht aller zu beschreibenden Bilder anzubieten, um schnell jene auswählen zu können, zu denen sie einen Beitrag leisten konnten und wollten. Um das Projekt auf „Zooniverse“ für die Allgemeinheit freizugeben und zu veröffentlichen, hätte es zudem in Gänze aus dem Deutschen ins Englische übersetzt werden müssen, nur um es anschließend wieder ins Deutsche zurück zu übersetzen; was bei den rein deutschsprachigen Inhalten des Projekts als nicht besonders sinnvoll erschien. Für andere bürgerwissenschaftliche Projekte im Archiv, etwa solche aus dem naturwissenschaftlichen Bereich, mag diese Hürde aber nicht besonders hoch sein.

In der Endphase des Projekts wurde zudem ein freiberuflicher Entwickler damit beauftragt, Algorithmen und ein Plug-in für die Anwendung von Semantic Web-Technologien zu entwickeln, mit dem die von den Citizens erstellten Inhalte bewertet, KI-unterstützt angereichert beziehungsweise als unzuverlässig markiert werden können. Diese Art der technischen Kontrolle sollte also die Arbeit der Archivmitarbeiter*innen unterstützen und erleichtern.

Der in diesem Zusammenhang entwickelte „IRS AI-Validator“ ist ein konfigurierbares Werkzeug zur semantischen Textanalyse. Der Validator prüft nicht

nur die Rechtschreibung in den User-Einträgen, sondern entdeckt auf der Basis des ChatGPT-Sprachmodells auch Eingaben, die Fehler oder Fehlinformationen enthalten oder unmittelbar als Spam einzuordnen sein könnten. Er erkennt unzulässige bedeutungssteigernden Markierungen (Kapitälchen, mehrere Rufzeichen, Emojis ...) sowie „Sentiments“ (positive, negative Konnotation). Des Weiteren findet er Wörter, die auf einer „Blacklist“ stehen („böse Wörter“) oder solche auf einer „Greylist“, die unbeabsichtigte oder beabsichtigte Falscheingaben nach sich ziehen könnten. Außerdem werden auffällig gewordene Nutzer-Pseudonyme (zu häufige, zu kurzfristige Eingaben, vorherige Fehleingaben) erkannt und es kommen Verfahren des „Rate Limiting“ (Blockierung von Nutzern, Bots oder Anwendungen, die eine Website übermäßig beanspruchen oder missbrauchen) und „Sandboxing“ (Freigabe von User-Beiträgen erst nach Qualitätscheck) zum Einsatz. Entsprechende User-Eingaben werden durch das Plug-in mit farbigen Tags markiert und einem Online-Redakteur (im Projekt: Paul Perschke) vorgelegt, der dann über den Erhalt oder die Löschung der Eingaben entscheidet.

Um die positiven Ergebnisse des nur auf zwei Jahre angelegten und nicht verlängerbaren Projekts zu perpetuieren, wurden einige zunächst simple Eingabefelder in das neu entstandene Online-Portal der Wissenschaftlichen Sammlungen (<http://stadt-raum-geschichte.de>) integriert, um unabhängig von weiteren Fördermaßnahmen auf den bisherigen Ergebnissen aufbauen zu können. Der weitere Ausbau dieser Infrastruktur bietet verschiedene Möglichkeiten der tieferen Integration von bürgerwissenschaftlichen Eingaben.

5. Projektergebnisse

Trotz der nicht immer erfolgreichen technischen Experimente waren die Erträge bereits während der Laufzeit des Projekts ganz erstaunlich. Nicht nur stieß es auf ein außerordentlich großes Publikumsinteresse, das sich unter anderem bei der Präsentation des Projekts im Rahmen der Langen Nacht der Wissenschaften in Berlin in den Jahren 2022 und 2023 bemerkbar machte. Für viele ehemalige Bürger*innen der DDR war der Kontakt mit den historischen Fotos von Orten und Gebäuden ihrer Vergangenheit ein sehr emotionales Erlebnis („In der Straße habe ich gewohnt, so sah das damals aus!“), woraus sich schließen lässt, dass ein solcher Citizen Science-Ansatz eine stark identitätsbildende Wirkung entfalten kann. Auch ließen sich nach entsprechenden Presseaktivitäten der Abteilung Wissenschaftskommunikation des IRS sehr hohe Klickraten messen: Eine Pressemitteilung mit der Bitte, bei der Erschließung der Bilder zu helfen, wurde von mehr als 35 Zeitungen und Medienportalen in der ganzen Bundesrepublik aufgegriffen und führte innerhalb weniger Tage zu 40.000 Zu-

griffen auf die Zooniverse-Seiten. Pressereaktionen auf die Eröffnung des neuen Portals stadt-raum-geschichte.de mit dem erstmaligen Online-Zugriff auf die Wissenschaftlichen Sammlungen führten zu etwa zehntausend neuen Besucher*innen, 50.000 Zugriffen und Tausenden von ausgefüllten Metadatenfeldern. Allerdings flaute das Interesse nach einer Weile immer sehr schnell wieder ab. Die daraus gewonnene Erkenntnis lautet, dass Pressearbeit sinnvollerweise ein solches Projekt möglichst langfristig und regelmäßig begleiten sollte.

Die von den Citizens in der Folge geleisteten Beiträge waren in aller Regel überaus hilfreich. Auch und gerade, wenn es „nur“ um Ortskenntnis ging, also etwa die Benennung von Straßen oder Gebäuden in einer Stadt, war der Nutzen ganz außerordentlich hoch. Schließlich bieten diese grundlegenden Informationen, einmal den Digitalisaten zugefügt, Anhaltspunkte zur weiteren Recherche und Erschließung. Darüber hinaus zeigt eine Vielzahl entsprechender Eingaben von Freiwilligen, dass auch die persönlichen Erinnerungen und Erläuterungen zur Alltagsgeschichte einen wertvollen Beitrag zur Einordnung von Bildern leisten können.

Aufgrund des unerwartet hohen Ansturms war es, anders als zunächst geplant, auch möglich, die verschiedenen Motive mehrfach mit Metadaten zu versehen. Der Abgleich dieser Daten, ebenso wie die Kontrolle der Ergebnisse durch die Projektmitarbeiter*innen, ergab eine hohe Konsistenz der Eingaben. Auch kam es zu ganz unverhofften Beiträgen und Erkenntnissen, etwa, wenn ein Bürger sich die Mühe machte, alle Automarken auf einer Sammlung von Auto-Motiven minutiös zu benennen.

6. Herausforderungen und weitere Entwicklung

Die Herausforderungen des Projekts waren weniger wissenschaftlicher als organisatorischer und technischer Natur und erforderten von den Projektmitgliedern eine große Bereitschaft, sich mit den Denkansätzen der jeweils anderen Beteiligten zu beschäftigen. Für die Wissenschaftler*innen bedeutete dies die Einarbeitung in Themen wie Datenmodelle, Sicherheitskonzepte und semantische Verfahren, für die Mitarbeiter*innen des IT-Unternehmens die Auseinandersetzung mit dem „langen wissenschaftlichen Atem“ und der Freude an der Diskussion von Details. Insgesamt wurde die Zusammenarbeit jedoch als sehr bereichernd empfunden.

Im Ergebnis entstand ein vom Unternehmen Programmfabrik zur Verfügung gestelltes Open-Source-Produkt in Form einer zentralen Wissens- und Kommunikationsplattform mit verschiedenen Plug-ins, die jedes Archiv befähigt, seine Nutzer*innen in ähnlicher Weise an der Digitalisierungs- und Erschließungsarbeit zu beteiligen. Dem Förderziel des Fördermittelgebers, nämlich gemein-

nützige und kommerzielle Interessen miteinander zu verbinden, wurde damit entsprochen.

Mit Hilfe des entwickelten Verfahrens und der entstandenen Infrastruktur wurde es möglich, Schätze zugänglich zu machen, die insbesondere für die Stadtgeschichte von großem Wert sind. Zugleich wird ein wichtiger Beitrag zum Forschungstransfer geleistet. Die Wissenschaftlichen Sammlungen, bereits seit vielen Jahren aktiv in der Einbindung von Citizens in ihre Arbeit, sind damit in gewisser Weise Teil der Bewegung der „Third Mission“ der Wissenschaft, also der Aktivitäten, die weit über Forschung und Lehre hinausgehen.³ Wesentlich war hierbei der Beitrag der Abteilung Wissenschaftskommunikation des Instituts („WUK“), denn das große Publikumsinteresse musste immer wieder durch Pressemitteilungen angefacht werden. Dadurch ergab sich aber auch für die „WUK“ eine Win-Win-Situation, gelang es so doch, ein breites Publikum für die Arbeit des gesamten Instituts zu interessieren.

7. Fazit

Abschließend lässt sich vermuten, dass Ansätze wie der hier dargestellte Selbstverständnis und Profil von Archiven stark verändern werden.⁴ Angesichts knapper werdender Ressourcen und eines großen öffentlichen Interesses ist diese Entwicklung nicht von Schaden. Es bleibt anhand weiterer Praxisbeispiele zu prüfen, für welche Typen von Archiven sich dieser stark technisch geprägte Ansatz am besten eignet.

Auch liegt in der Einbindung von Citizens eine gewisse Gefahr, denn je mehr Kontrolle die Gedächtnisinstitutionen über die Erschließung ihres Materials abgeben, desto notwendiger werden die Vorkehrungen zur Kontrolle oder zumindest Kommentierung von Citizen-Eingaben. Die dargestellten KI-Ansätze können dabei nur in gewissen Grenzen helfen; die Qualitätsprüfung durch das Ar-

³ Zur Third-Mission vgl. Siegfried Balleis, Raus aus dem Elfenbeinturm. Third Mission für Universitäten und Hochschulen, in: Politische Studien 73:504, 2022, S. 52-59; sowie Judith Welker u. a., Inwieweit erfüllen Hochschulen ihre „Third Mission“? Sichtweisen der regionalen Bevölkerung auf Hochschulen und ihre Angebote, Frankfurt 2023.

⁴ Vgl. zu weiteren Citizen Science-Projekten u. a. folgendes Themenheft: Bürgerschaftliches Engagement in Archiven, in: Archivar 72:1, 2019; sowie Christian J. Huber/Lambert Kansy/Martin Lüpold, Crowdsourcing in Archiven. Ein Werkstattbericht, in: Archivar 73:2, 2020, S. 145-149; Sandra Leibner, Citizen Science-Projekt des Historischen Archivs des Bayerischen Rundfunks zur Erschließung rundfunkhistorischer Fotobestände: Ein Erfahrungsbericht, in: Archivar 74:1, 2021, S. 31-34. „Crowdsourcing: Konzeptionelle Überlegungen für den Einsatz in Archiven: DFG-Projekt „Digitalisierung und Entwicklung neuer Nutzungsmöglichkeiten von archivalischen Fotobeständen“ des Landesarchivs Baden-Württemberg“, 2015, https://www.archivschule.de/uploads/Forschung/Digitalisierung/Ergebnisse/LABW-08_Crowdsourcing_Grundsatzkonzept_2.pdf [21.04.2024].

chivpersonal bleibt notwendig. Als wichtig erscheint es hier, den einmal angestoßenen Dialog mit den Citizens fortzuführen und weitere bürgerwissenschaftliche Verfahren zu erproben, etwa die Etablierung eines digitalen Forums oder die Durchführung von „Hackathons“ zur Massenerschließung von Beständen.

Insgesamt aber lassen sich die Möglichkeiten und Chancen sehr positiv bewerten, gerade mit Blick auf klein- und mittelstädtische Archive. Und sind die Bestände erst einmal erschlossen und zugänglich, steht auch der weiteren Partizipation der Bürger*innen an der Erforschung der Geschichte ihrer Stadt nichts mehr entgegen.

Rita Gudermann, Dr., ist Wissenschaftlerin am Leibniz-Institut für Gesellschafts- und Raumforschung (IRS). Sie promovierte an der Freien Universität Berlin am Lehrstuhl für Wirtschafts- und Sozialgeschichte des Fachbereichs Wirtschaftswissenschaften zum Thema "Landwirtschaftliche Bodenverbesserung in Brandenburg und Westfalen im 19. Ihre Schwerpunkte sind Agrar-, Umwelt-, Raum- und Digitalgeschichte des 19. und 20. Jahrhunderts". Darüber hinaus verfügt sie über langjährige Berufserfahrung als IT-Beraterin im Bereich DAM- und ERP-Systeme und ist Leiterin des Projekts zur Verbesserung der digitalen Infrastruktur der wissenschaftlichen Sammlungen des IRS.

rita.gudermann@leibniz-irs.de

Paul Perschke, MA, ist Stadthistoriker und bearbeitete als wissenschaftlicher Mitarbeiter zwischen 2021 und 2023 das Projekt „CitizenArchives“ in den Wissenschaftlichen Sammlungen zur Bau- und Planungsgeschichte der DDR am Leibniz-Institut für Raumbezogene Sozialforschung. Derzeit ist er dort Dokumentar mit dem Schwerpunkt digitale Bestandserschließung und zugleich Mitarbeiter im Drittmittelprojekt „Crafting Entanglements: Afro-Asian Pasts of the Global Cold War (CRAFTE)“. In dem Projekt, das die afro-asiatischen Verflechtungen während des Kalten Krieges mit seinen Akteuren, Praktiken und ihre alltäglichen Interaktionsorten untersucht, ist er verantwortlich für den Aufbau einer Projektdatenbank.

paul.perschke@leibniz-irs.de