

Dr. Helen Wiedmaier

**Citizen Science in Stadtarchiven:  
Qualitätssicherung und -kontrolle bei  
ausgewählten Anwendungen**

**Eingereicht am:**

**22.03.2024**

Transferarbeit des 57. wissenschaftlichen Lehrgangs  
an der Archivschule Marburg

Betreuer an der Archivschule Marburg: Dr. Robert Meier

Betreuer des Ausbildungsarchivs: Dr. Andreas Neuburger

## **Inhalt**

1.	Einleitung .....	1
1.1.	Definitionen und begriffliche Abgrenzungen .....	2
1.2.	Forschungsliteratur .....	4
1.3.	Citizen Science im Archivwesen: Erschließung .....	5
1.4.	Citizen Science im Archivwesen: Chancen und Risiken .....	7
2.	Qualitätssicherung und -kontrolle bei Citizen Science: Das Projekt .....	12
2.1.	Methodik und Vorgehensweise .....	12
2.2.	Erster Projektabschnitt: Expert*inneninterviews .....	13
2.3.	Zweiter Projektabschnitt: Erschließungsarbeiten .....	15
3.	Auswertung des Projekts und der verwendeten Plattformen .....	18
3.1.	Microsoft Excel .....	18
3.2.	Daten-Eingabe-System (DES) .....	20
3.3.	Transkribus .....	23
4.	Fazit .....	27
5.	Zusammenfassung .....	29
6.	Quellenverzeichnis .....	29
7.	Literaturverzeichnis .....	29
8.	Anhang: Interviewfragen .....	35

## 1. Einleitung

Seit einigen Jahren kommt man an dem Thema „Citizen Science“ nicht mehr vorbei. Während dieser Ansatz früher vor allem im Bereich der Naturwissenschaften eine Rolle spielte, wird er inzwischen auch im geistes- und sozialwissenschaftlichen Bereich immer wichtiger.<sup>1</sup> Für Archive gibt es vielfältige Motivationen, um Citizen-Science-Projekte ins Leben zu rufen und zu unterstützen: Neben dem Versuch, damit neue Zielgruppen anzusprechen, offener zu werden, ein Bürger\*innenarchiv zu sein oder neue (meist digitale) Werkzeuge und Methoden zu testen, spielt oft der in vielen Archiven vorherrschende Personalmangel eine Rolle.<sup>2</sup> Wenn mittels Citizen-Science-Projekten Archivalien erschlossen werden, können dadurch Rückständen aufgearbeitet und Nutzer\*innen mehr erschlossene Archivalien zur Verfügung gestellt werden.

In vielen Archiven wird jedoch die Frage diskutiert, inwiefern die im Rahmen derartiger Projekte erarbeiteten Erschließungsdaten überhaupt übernommen werden können und wenn ja, ob man diese direkt in die Findbücher aufnehmen oder noch einer Qualitätskontrolle unterziehen sollte.<sup>3</sup> Diesen Fragen wird in dieser Arbeit nachgegangen. Dafür wurde ein zweistufiger Ansatz gewählt: Zum einen Interviews mit ausgewählten Expert\*innen und zum anderen die Umsetzung eines eigenen Erschließungsprojekts in Kooperation mit dem Verein für Computergenealogie sowie dem Stadtarchiv Leipzig. In beiden Fällen werden drei verschiedene Anwendungen zur Erschließung untersucht und vor allem hinsichtlich der Qualitätsüberprüfung und -sicherung analysiert.

Vor der eigentlichen Projektbeschreibung werden zunächst Begriffsdefinitionen sowie -abgrenzung thematisiert; anschließend erfolgt ein Forschungsüberblick. Das eigentliche Projekt mit den zwei Stufen – Experteninterviews und Erschließung – wird danach vorgestellt. Dabei werden nicht nur das Projekt selbst, sondern auch die verwendete Methodik sowie die beiden Projektpartner – der Verein für Computergenealogie sowie das Stadtarchiv Leipzig – thematisiert werden.

---

<sup>1</sup> Für die Entwicklung und Etablierung von Citizen Science sowie Public History ab den 1980ern vgl. Speit, *Public History*, S. 122–126.

<sup>2</sup> Vgl. Gränitz, *Schwarmwissen*, S. 1.

<sup>3</sup> Diese Frage ist bei vielen Citizen Science oder Crowdsourcing-Projekten zentral und ein häufiges Argument gegen deren Umsetzung; vgl. Patt, *Crowdsourcing*, S. 9f. Auch Balázs u.a., *Data Quality*, S. 141 vermerken: „(...) data quality issues are the Achilles heel of citizen science projects.“ Dies ist vor allem dann der Fall, wenn die generierten Daten weiterverwendet werden sollen; vgl. Ebd., S. 142. Da genau das das Ziel vieler Citizen-Science-Projekte in Archiven ist, ist die Frage nach der Qualitätskontrolle essenziell.

## 1.1. Definitionen und begriffliche Abgrenzungen

Wichtig ist zunächst, dass es sich zwar um ein Thema handelt, das aktuell vielfältig diskutiert wird, aber im Grunde um keine neue Entwicklung. Schon in vorangehenden Jahrhunderten wurden zentrale wissenschaftliche Erkenntnisse von „Laien“ gewonnen und publiziert;<sup>4</sup> allerdings nicht unter dem Schlagwort Citizen Science.<sup>5</sup> Während sich in früheren Jahrhunderten vor allem Einzelpersonen derart auszeichneten, ermöglicht heute der technische Fortschritt und die digitale Vernetzung einer wesentlich größeren Gruppe die Teilhabe an Forschung und die aktive Mitwirkung an verschiedenen Projekten. Auch der Austausch sowie die Erfassung von Daten ist durch das Internet und damit verbundene Tools wie etwa die Videokonferenz oder Clouddienste deutlich einfacher.<sup>6</sup>

Der Begriff Citizen Science wird häufig nicht klar von anderen, ähnlich besetzten Begriffen wie beispielsweise Crowdsourcing abgegrenzt. Gerade im deutschsprachigen Raum werden bei Projektbeschreibungen zudem oft andere Begriffe verwendet, da diese deutlich etablierter sind und eine längere Tradition aufweisen, etwa „Public History“ oder auch „ehrenamtliche Denkmalpflege.“<sup>7</sup> In der Forschung ist jedoch inzwischen sowohl im internationalen als auch im deutschsprachigen Raum der Terminus „Citizen Science“ üblich, obgleich er oft synonym zu „Crowdsourcing“ verwendet wird. Um in dieser Arbeit eine eindeutige Begrifflichkeit herstellen zu können, werden beide Begriffe nachfolgend definiert und einander gegenübergestellt.

Eine Definition, die häufig verwendet und inzwischen als maßgeblich angesehen wird,<sup>8</sup> ist jene aus dem Grünbuch Citizen Science Strategie 2020 für Deutschland, das 2016 im Rahmen des Bausteinprogramms „Bürger schaffen Wissen – Wissen schafft

---

<sup>4</sup> Ein oft genanntes Beispiel ist etwa Charles Darwin; vgl. Pettibone/ Ziegler, Citizen Science, S. 58. Andere Beispiele des 19. Jahrhunderts mit einem Fokus auf Kartographie führt Smolarski an; vgl. Smolarski, Missionskartograph, S. 71-87. Einen knappen Überblick über die Entwicklung von Citizen Science ausgehend von den Naturwissenschaften, über die Sozialwissenschaften bis hin zu den Geisteswissenschaften bietet Jaeger-Erben; vgl. Jaeger-Erben, Citizen Science, S. 46–50. Für die Begriffsgeschichte vgl. Pierroux/ Hetland/ Esborg, citizen science, S. 4–6.

<sup>5</sup> Die erste programmatische Verwendung des Begriffs erfolgte erst 1995; vgl. Irwin, Citizen Science.

<sup>6</sup> Vgl. Pettibone/ Ziegler, Citizen Science, S. 58. Für englischsprachige Definitionen vgl. Haklay u.a., Citizen Science, S. 14–23.

<sup>7</sup> Vgl. Pettibone/ Ziegler, Citizen Science, S. 65. Weitere gebräuchliche Begriffe sind etwa Bürger\*innenforschung, Bürger\*innenwissenschaft, civic science, community science, ehrenamtliche Forschung, scientific citizenship und partizipative Forschung; vgl. Jaeger-Erben, Citizen Science, S. 45. Die meisten Projekte verwenden jedoch entweder „Crowdsourcing“ oder eben „Citizen Science“; vgl. dazu auch den Literaturüberblick auf S. 4f.

<sup>8</sup> Eine Übersicht weiterer Definitionen trägt Fink in ihrer Arbeit zusammen; vgl. Fink, Citizen Science, S. 11f.

Bürger (GEWISS)“ zur Entwicklung von Citizen-Science-Kapazitäten publiziert wurde:

„Citizen Science beschreibt die Beteiligung von Personen an wissenschaftlichen Prozessen, die nicht in diesem Wissenschaftsbereich institutionell gebunden sind. Dabei kann die Beteiligung in der kurzzeitigen Erhebung von Daten bis hin zu einem intensiven Einsatz von Freizeit bestehen, um sich gemeinsam mit Wissenschaftlerinnen bzw. Wissenschaftlern und/oder anderen Ehrenamtlichen in ein Forschungsthema zu vertiefen. Obwohl viele ehrenamtliche Forscherinnen und Forscher eine akademische Ausbildung aufweisen, ist dies keine Voraussetzung für die Teilnahme an Forschungsprojekten. Wichtig ist allerdings die Einhaltung wissenschaftlicher Standards, wozu vor allem Transparenz im Hinblick auf die Methodik der Datenerhebung und die öffentliche Diskussion der Ergebnisse gehören.“<sup>9</sup>

Zusammenfassend lässt sich somit sagen, dass sich Citizen Science weder durch die Dauer oder Intensität auszeichnen, mit der sich Personen an Projekten beteiligen, noch durch die Laufzeit solcher Projekte, sondern durch die Einhaltung wissenschaftlicher Standards bei der ehrenamtlichen Arbeit von Personen, die mit der Wissenschaft nicht institutionell verbunden sind.

Vielen Definitionen beinhalten auch, dass die im Rahmen von Citizen-Science-Projekten gewonnenen Ergebnisse und Daten wieder in die Wissenschaft zurückgeführt werden. Dabei werden sie so verfügbar gemacht, dass sie weiterverarbeitet oder -analysiert werden können.<sup>10</sup> Als konkretes Beispiel werden etwa Erschließungsdaten angeführt, die bei Citizen-Science-Projekten von den Ehrenamtlichen erarbeitet werden.<sup>11</sup> Im Fokus stehen somit zum einen die Ehrenamtlichen und ihre Hinwendung zur Wissenschaft bei fehlender institutioneller Bindung und zum anderen die kreierte Forschungsdaten.

Eine etwas andere Ausrichtung lässt sich beim Begriff Crowdsourcing finden: Hier steht „das Auslagern von bisher in einem Unternehmen selbst erbrachten Leistungen auf eine große Anzahl von Menschen über das Internet“<sup>12</sup> im Vordergrund. Der Fokus liegt somit deutlich auf Projekten, die institutionell gebunden und Teil einer Wertschöpfungskette sind. Wissenschaftliche Standards werden in der Praxis zwar

---

<sup>9</sup> Bonn u.a., Grünbuch, S. 13. Knapper formuliert Jaeger-Erben; Jaeger-Erben, Citizen Science, S. 45: „(...) Citizen Science (...) beschreibt eine Form wissenschaftlicher Praxis, bei der nicht institutionell in der Forschung verankerte Bürgerinnen\* als aktiv Beteiligte in einen wissenschaftlichen Forschungsprozess eingebunden werden.“ Für das GEWISS-Projekt vgl. Kluttig, Citizen Science, S. 33–41.

<sup>10</sup> Smolarski/ Carius/ Prell, Citizen Science, S. 7f. Hier sind als Stichwörter Normdaten und Digital Humanities zu nennen; vgl. Ebd., S. 7f.

<sup>11</sup> Smolarski/ Carius/ Prell, Citizen Science, S. 7f.

<sup>12</sup> <https://www.duden.de/rechtschreibung/Crowdsourcing> (10.01.2024).

eingehalten, sind aber nicht Teil der Definition.<sup>13</sup> Stattdessen hängt der Begriff des Crowdsourcings eng mit jenem des „User Generated Contents“ zusammen, d.h. Inhalten, die von Personen generiert werden, die diese auch perspektivisch nutzen werden.<sup>14</sup> Huber/ Kansy/ Lüpold stellen in diesem Zusammenhang fest, dass der Begriff „User Generated Content“ den Fokus eher auf den Umgang mit den Inhalten legt, die von Benutzer\*innen generiert werden, während der Begriff „Crowdsourcing“ wiederum die Generierung der Inhalte umschreibt.<sup>15</sup> Man kann folglich Citizen Science von Crowdsourcing anhand der Projektgebundenheit unterscheiden: User Generated Content fließt im Rahmen von Crowdsourcing nicht in konkrete Forschungsprojekte, die Daten von Citizen Science hingegen schon.<sup>16</sup> Festzuhalten ist jedoch auch, dass die Begriffe sehr ähnlich sind und entsprechend oft bei Projekten und Publikationen synonym verwendet werden.<sup>17</sup>

## 1.2. Forschungsliteratur

Da sich Citizen-Science-Projekte heute vor allem im digitalen Raum finden lassen und dieser raschen Veränderungen unterworfen ist, werden vor allem neuere Werke aus den letzten zehn bis fünfzehn Jahren berücksichtigt. Einen Überblick über die ältere Literatur sowie internationale Entwicklungen fehlt bislang in der Forschung – vermutlich bedingt durch die Schnelllebigkeit des Themas in der digitalen Welt.<sup>18</sup>

Auffällig ist, dass sich wenige Arbeiten programmatisch mit Citizen Science auseinandersetzen, stattdessen überwiegen praxisbezogene Aufsätze und Abschlussarbeiten, die sich dezidiert mit Citizen Science oder Crowdsourcing in einem bestimmten Projektkontext – beispielsweise eingebunden in ein Museum, eine Bibliothek, ein Archiv oder auch in die Stadtentwicklung – befassen.<sup>19</sup> Auch Tagungs- und Sammelbände befassen sich vorrangig mit best-practice-Beispielen, aber auch ergänzend mit

---

<sup>13</sup> Vgl. dazu auch Bunge, Bibliotheken, S. 245.

<sup>14</sup> Vgl. Huber/ Kansy/ Lüpold, Crowdsourcing, S. 147.

<sup>15</sup> Vgl. Huber/ Kansy/ Lüpold, Crowdsourcing, S. 147.

<sup>16</sup> Vgl. Huber/ Kansy/ Lüpold, Crowdsourcing, S. 147.

<sup>17</sup> So beispielsweise Walter, Stormarn, S. 150 oder Becker, Citizen Science, S. 30–39. Berücksichtigt wurden in dieser Arbeit die neueren Definitionen. Ältere nennt Patt und verweist auf über 40 Definitionen für Crowdsourcing, die 2012 zusammengetragen wurden; vgl. Patt, Crowdsourcing, S. 5f. Wenn die Begriffe synonym verwendet werden, können die Ergebnisse zu Crowdsourcing in Archiven auf Citizen-Science-Projekte übertragen werden. Studien und Werke, bei denen dies der Fall ist, wurden in dieser Arbeit berücksichtigt und die Ergebnisse entsprechend übertragen.

<sup>18</sup> Für die Entwicklung des Begriffs vgl. Kapitel 1.1; besonders Fn. 4.

<sup>19</sup> Patt, Crowdsourcing; Fink, Citizen Science 2.0; Becker, Citizen Science, S. 30–39; Bunge, Wie viel Naturwissenschaft braucht die Bibliothek?, S. 241–250; Hirschbeck, Archive und Genealogen; Gränitz, Schwarmwissen;

einzelnen Fragestellungen zu beispielsweise der gesellschaftlichen Relevanz von Citizen Science.<sup>20</sup> Ein eigenes Thema mit Überschneidungen zu Citizen Science ist jenes der Ehrenamtlichen in Archiven. Dazu bietet der Aufsatz von Eibl, der auf ihrer 2013 verfassten Transferarbeit beruht, einen guten Überblick.<sup>21</sup>

Zu den drei Anwendungen, die in dieser Arbeit untersucht werden, existiert unterschiedlich viel Literatur. Während Transkribus im Kontext von Citizen-Science-Projekten vor allem im englischsprachigen Raum ausführlich thematisiert wird,<sup>22</sup> werden die anderen Plattformen eher stiefmütterlich behandelt. Excel wird hin und wieder erwähnt, eigene Publikationen zur Verwendung von Excel bei Citizen-Science-Projekten scheint es aber nicht zu geben. Die Eigenentwicklung des Vereins für Computergenealogie, das Dateneingabesystem (DES) wird vor allem in zahlreichen Vorträgen und kurzen Blogbeiträgen erläutert.<sup>23</sup> Aufsätze zu übergeordneten Fragestellungen und wissenschaftlichen Kontextualisierung des DES existieren hingegen nur wenige.<sup>24</sup> Dies lässt sich mit dem Zweck der Anwendung erklären: DES dient als praktisches Werkzeug, das speziell für das Erschließen serieller Quellen wie Kirchenbücher programmiert wurde. Somit bietet die Anwendung weniger Reibungsfläche als beispielsweise eine künstliche Intelligenz wie Transkribus, bei der neben praktischen Fragen auch moralische und solche nach der Zukunft der Erschließung verhandelt werden.

### 1.3. Citizen Science im Archivwesen: Erschließung

Im Archivwesen werden Citizen-Science-Projekte vor allem im Bereich der Erschließung umgesetzt.<sup>25</sup> Die Art und Weise, wie Ehrenamtliche bei derartigen Projekten

---

<sup>20</sup> Vgl. Bürger Künste Wissenschaft, hg. v. Kristin Oswald/ René Smolarski; Citizen Science in den Geschichtswissenschaften, hg. v. René Smolarski/ Hendrikje Carius/ Martin Prell; A History of Participation in Museums and Archives, hg. v. Per Hetland/ Palmyre Pierroux/ Line Esborg; The Science of Citizen Science, hg. v. Katrin Vohland u.a.;

<sup>21</sup> Vgl. Eibl, Engagement, S. 8–12.

<sup>22</sup> Für einen ausführlichen Literaturüberblick vgl. Nockels/ Gooding/ Ames/ Terras: Understanding, S. 367–392. Viele Aufsätze befassen sich mit konkreten Anwendungsfällen; vgl. beispielsweise Milioni, Transcription.

<sup>23</sup> Eine Auflistung einführender Vorträge und Publikationen findet sich unter [https://wiki.genealogy.net/DES#Publikationen.2FPr.C3.A4sentationen\\_zum\\_DES](https://wiki.genealogy.net/DES#Publikationen.2FPr.C3.A4sentationen_zum_DES) (14.02.2024). Häufig thematisieren Vorträge und Zoom-Meetings, die sich an die Freiwilligen wenden, das DES und erklären, wie die Erschließung praktisch funktioniert und was es zu beachten gibt. Vgl. exemplarisch die Vortragsankündigungen unter <https://www.compgen.de/2024/02/compgen-vortrag-zu-des-projekten-mit-transkribus-am-8-2-2024/> (14.02.2024) sowie <https://www.compgen.de/2023/10/online-treff-am-12-oktober-jeder-kann-mitmachen-bei-compgen-helft-mit-die-daten-einzugeben/> (14.02.2024).

<sup>24</sup> Vgl. Jesper, Daten-Eingabe-System, S. 28–32.

<sup>25</sup> Erschließung wird hier wie folgt definiert: „Erschließen bedeutet (...), Wissen und Sinn zu offenbaren und anderen Personen gegenüber zugänglich zu machen. (...) Erschließen ist das Bekanntmachen von sinnvollen Informationen, das Zugänglichmachen von Archivgut (...). Im Archivwesen wird unter Erschließung meist das Ordnen und Strukturieren sowie das Verzeichnen von Archivgut verstanden. (...)“

eingesetzt werden, lässt sich in drei Gruppen unterteilen: Erstens, indem sie bislang unerschlossene Bestände erschließen. Zweitens, indem sie bereits erschlossene Bestände ergänzen, etwa um Normdaten wie GND-Verknüpfungen oder Georeferenzdaten. Und drittens, indem sie mittels OCR-Texterkennung erschlossene Bestände überprüfen und korrigieren.<sup>26</sup> Die ersten beiden Einsatzmöglichkeiten beschränken sich nicht auf Textdokumente, sondern können auch andere Archivaliengruppen wie etwa Fotografien, Pläne oder Videos einschließen. Auch die Komplexität kann sehr divergent sein; allerdings muss die Begleitung umso enger sein, je komplexer die Erschließung ausfällt. Dementsprechend ist auch die Qualitätskontrolle bei komplexen Beständen und Aufgaben aufwändiger und zeitintensiver.<sup>27</sup>

Über die Frage, wie sich die „Qualität“ einer Erschließung bemessen lässt, beziehungsweise, welche Faktoren eine „gute“ Erschließung ausmachen, wurde viel debattiert. Inzwischen herrscht Konsens, dass eine gute Erschließung nicht zwingend tief, d.h. möglichst detailreich und folglich durch den hohen zeitlichen und personellen Aufwand teuer, und eine schlechte billig und oberflächlich ist. Heute bemisst sich die Qualität der Erschließung weniger an der Fülle der inhaltlichen Details, sondern an der Anschlussfähigkeit und Möglichkeit zur Weiterentwicklung sowie -verwendung der Daten.<sup>28</sup> Wenn alle Namen der in einer Urkunde erwähnten Personen aufgelistet, aber nicht bestimmten Menschen zugeordnet werden können, ist die Erschließung zwar tief, aber nicht mit weiteren Daten verknüpft. Ob eine genannte Person etwa noch in anderen Urkunden auftaucht, ist folglich eine Information, die fehlt.

Aktuell wird dafür plädiert, zum einen einheitliche, archivübergreifende Standards zu definieren und zu befolgen und zum anderen, zwischen verschiedenen Beständen und Archivalien zu differenzieren.<sup>29</sup> Kriterien für letzteres können beispielsweise das

---

Erschließung umfasst jedoch auch die Erstellung von Urkundenregesten, Editionen und Indices oder einer Bestandsübersicht – unterschiedliche normierte und gelehrte Formen zur Beschreibung von Archivgut;“ Joergens, *Erschließung*, S. 126. Erschließung gilt als archivische Kernkompetenz, deren Erlernung einen wichtigen Teil der archivischen Fachausbildung darstellt; vgl. Ebd., S. 126. Es gibt jedoch auch die Idee, das Archivieren selbst als kollektiven Prozess zu verstehen; vgl. Esborg, *Engaging*, S. 109–125. In der Praxis werden derartige Überlegungen deutlich seltener umgesetzt als partizipative Projekte im Bereich Erschließung.

<sup>26</sup> Vgl. Patt, *Crowdsourcing*, S. 10f. Zur zweiten Gruppen gehört auch das sog. „social tagging“, bei dem „kollektiv geteilter Inhalt mit Metadaten in Form von Schlagwörtern annotiert wird;“ vgl. Hirschbeck, *Archive*, S. 6.

<sup>27</sup> Vgl. Patt, *Crowdsourcing*, S. 11. Wichtig ist die Feststellung, dass Datenqualität für unterschiedliche Personen und Institutionen unterschiedliche Facetten beinhaltet; vgl. Balázs u.a., *Data Quality*, S. 139f. Während die einen vielleicht Wert auf möglichst vollständige Metadaten legen, wünschen sich die anderen vielleicht die Auflösung alter Ortsnamen in eine moderne Schreibweise in Findbücher und die dritten eine Georeferenzierung.

<sup>28</sup> Vgl. Haberer/ Rügge, *Überlegungen*, S. 130f.

<sup>29</sup> Vgl. Haberer/ Rügge, *Überlegungen*, S. 131f.

Interesse der Nutzer\*innen sein oder der Anspruch, keine unerschlossenen Archivalien zu haben. In diesem Fall wäre eine flache Erschließung einer tiefen vorzuziehen, da bei ersteren die Archivalien zumindest genutzt werden können. Optimal wäre es natürlich, wenn alle Archivalien bereits erschlossen wären und Zugänge innerhalb kürzester Zeit ebenfalls erschlossen werden könnten. Da knappe Ressourcen (sowie personell als auch finanziell) jedoch zu einer Priorisierung zwingen, strebt man in den meisten Archiven zumindest eine Verzeichnung und knappe Erschließung aller Bestände an und beschränkt sich bei der Tiefenerschließung zunächst auf jene Bestände, die von besonderem Interesse für die Nutzung sind – sei es aufgrund ihrer inhaltlichen Wertigkeit oder weil besonders viele Nutzer\*innen diese Bestände nachfragen.<sup>30</sup> Hierüber kann auch wieder Zeit gespart werden: sind Bestände gut erschlossen, reduziert sich die Zeit, die für Recherchen benötigt wird. Um diesen Anspruch umzusetzen, werden in vielen Archiven Citizen-Science-Projekte für die Erschließung genutzt.

#### **1.4. Citizen Science im Archivwesen: Chancen und Risiken**

Die meisten Archive, die Citizen-Science-Projekte ins Leben rufen,<sup>31</sup> erhoffen sich davon vielfältige Vorteile<sup>32</sup> für ihre Arbeit: Sie möchten Nutzer\*innen an sich binden, Arbeiten günstig auslagern und die Qualität von Erschließungen zu verbessern.<sup>33</sup> Bei gut geplanten und umgesetzten Projekten erfüllen sich diese Hoffnungen oft: Es werden kostensparend Aufgaben erfüllt, die sonst aufgrund fehlender personeller und finanzieller Ressourcen nicht zu bewältigen wären.<sup>34</sup> Bestände werden erschlossen und Rückstände können aufgearbeitet werden,<sup>35</sup> wodurch das Archiv eine Dienstleistung

---

<sup>30</sup> Vgl. Haberer/ Rügge, Überlegungen, S. 131f.

<sup>31</sup> Im Rahmen dieser Arbeit werden Citizen-Science-Projekte in den Blick genommen, die von Wissenschaftler\*innen – hier vor allem Archivar\*innen – initiiert werden. Selbstverständlich existiert auch eine Vielzahl an Projekten, die von den Bürger\*innen selbst ins Leben gerufen wurden.

<sup>32</sup> Generelle Vorteile von Citizen-Science-Projekten führen Pettibone u.a. auf; darunter beispielsweise die Inspiration und Generierung von Forschungsfragen durch das Einbringen neuer Ideen aus der Gesellschaft; die Möglichkeit zur Bearbeitung und Erschaffung großer Datensätze; die höhere Akzeptanz von Forschungsergebnisse durch das Mitforschen und die Evaluation von Bürger\*innen; die Überprüfung der praktischen Relevanz und Anwendbarkeit wissenschaftlicher Ergebnisse; der Perspektivwechsel für alle Beteiligten; die bessere Übertragung von Forschungsergebnissen in die Praxis durch frühzeitige Einbindung gesellschaftlicher Akteure; die Öffnung der Deutungshoheit der Wissenschaft; die Stärkung der Zivilgesellschaft; die Erhöhung der Bildung (auch durch Weiterqualifizierung) sowie generell Spaß haben und teilen; vgl. Pettibone u.a., Citizen Science, S. 8. In diesem Kapitel wird der Fokus jedoch auf die spezifischen Vorteile für Archive und Freiwillige, die sich an archivischen Projekten beteiligen, liegen.

<sup>33</sup> Vgl. Hirschbeck, Archive, S. 6f.

<sup>34</sup> Vgl. Gränitz, Schwarmwissen, S. 5.

<sup>35</sup> Vgl. Eibl, Engagement, S. 12.

kostenlos oder -günstig erhält.<sup>36</sup> Als Nebeneffekt werden mehr Bestände digitalisiert und folglich mehr Digitalisate zur Verfügung gestellt.<sup>37</sup>

Auch ermöglicht die Online-Erschließung die Teilhabe vieler Personen, da diese nicht mehr auf die Infrastruktur des Archivs vor Ort angewiesen sind.<sup>38</sup> Wenn es beispielsweise zehn Freiwillige gibt, aber vor Ort im Lesesaal nur drei Computer, ist das bei der klassischen, ortsgebundenen Erschließung ein Problem. Zudem können sich auch Personen beteiligen, die jenseits des eigentlichen Einzugsgebiets leben oder aufgrund einer Krankheit oder ihres Alters nicht mehr physisch das Archiv aufsuchen können.<sup>39</sup> Auch die freie Zeiteinteilung durch die Unabhängigkeit von Lesesaalöffnungszeiten sowie die Möglichkeiten zur digitalen Vernetzung sind von Vorteil – sowohl für das Archiv als auch für die Nutzer\*innen.<sup>40</sup>

Gleichzeitig verbessern derartige Projekte auch die Außenwirkung der beteiligten Archive und Verbände, da sie Innovationsfähigkeit signalisieren und somit als Werbung fungieren.<sup>41</sup> Citizen Science fügt sich somit in den Trend, Archive partizipativer zu gestalten.<sup>42</sup>

Auch die Bestände werden bekannter,<sup>43</sup> wodurch – zusammen mit der als positiv wahrgenommenen Möglichkeit der Partizipation – mehr Nutzer\*innen generiert werden.<sup>44</sup> Durch die Angebote zur Partizipation können weiterhin Bindung und Loyalität der Nutzer\*innen zum Archiv gestärkt werden.<sup>45</sup> Durch die Zusammenarbeit mit Bürger\*innen können die Archive darüber hinaus auf einen Wissens- und Kompetenzpool zugreifen,<sup>46</sup> der jenen der eigenen Mitarbeiter\*innen ergänzt.

---

<sup>36</sup> Vgl. Hirschbeck, *Archive*, S. 8f.

<sup>37</sup> Vgl. Hirschbeck, *Archive*, S. 10.

<sup>38</sup> Vgl. Eibl, *Engagement*, S. 11. Weiterhin ist die Digitalisierung ein wichtiger Baustein der Bestandserhaltung, da die Originale nicht mehr im Lesesaal vorgelegt werden müssen. Ebenso bleibt der Bezug zwischen Quelle und Transkription im digitalen Raum erhalten; vgl. Moeller/ Müller, *Heimatsforscher*, S. 100.

<sup>39</sup> Vgl. Eibl, *Engagement*, S. 11.

<sup>40</sup> Vgl. Eibl, *Engagement*, S. 11.

<sup>41</sup> Vgl. Gränitz, *Schwarmwissen*, S. 5.

<sup>42</sup> Vgl. Theimer, *Partizipation*, S. 9. Dieser wurde zunächst nicht durch die eigene Motivation der Archive ausgelöst, sondern war Resultat der veränderten und vor allem deutlich abnehmenden Nutzer\*innenzahlen; vgl. Ebd., S. 6-10. Erst im Laufe der Zeit entwickelte sich das Selbstverständnis der Archive weg vom Ort, der Wissen hütet und den Zugang beschränkt ermöglicht, hin zu einem offenen Zentrum der Begegnung, des Erwerbs und der Vermittlung von Wissen; vgl. Ebd., 9f. Die abnehmenden Nutzer\*innenzahlen haben vielfältige Ursachen, wobei zwei der wichtigsten Punkte der vereinfachte Zugang zu Informationen mittels des Internets sowie die zunehmende Fülle an konkurrierenden Angeboten (etwa bei der Recherche nach Vorfahren) darstellen; vgl. Ebd., S. 9f.

<sup>43</sup> Vgl. Hirschbeck, *Archive*, S. 8f.

<sup>44</sup> Vgl. Gränitz, *Schwarmwissen*, S. 5f.

<sup>45</sup> Vgl. Hirschbeck, *Archive*, S. 8f. Weiterhin können Archive durch die aktive Einbindung der Nutzer\*innen besser deren Wünsche erkennen und auf diese eingehen; vgl. Gränitz, *Schwarmwissen*, S. 5f.

<sup>46</sup> Vgl. Becker, *Citizen Science*, S. 34.

Aber nicht nur die Archive profitieren von Citizen-Science-Projekten, auch die Freiwilligen können ihr Wissen und ihre Kompetenzen erweitern.<sup>47</sup> Verschiedene Studien untersuchen die Motivation der Freiwilligen, sich bei derartigen Projekten zu engagieren, und die daraus resultierende Vorteile für diese. Häufig genannte motivierende Faktoren sind etwa das Bedürfnis, anderen zu helfen, die eigenen Forschungen (gerade im Bereich der Genealogie) voranzutreiben, zu entspannen oder mit spannenden Archivalien arbeiten zu können.<sup>48</sup> Ein weiteres Argument, dass gerade von Genealog\*innen, die im Verein für Computergenealogie tätig sind, angeführt wird, ist der Wunsch, dem eigenen Verein etwas zurückzugeben.<sup>49</sup>

Neben der Motivation sollte auch in den Blick genommen werden, welche Vorteile Bürger\*innen durch die Teilnahme erlangen, da immer wieder betont wird, dass die Freiwilligen nicht ausgenutzt werden dürfen, sondern alle Seiten – sowohl die Freiwilligen als auch die Archive und eventuell beteiligten Vereine – profitieren sollten.<sup>50</sup> Die 2018 erschienene Studie von Phillips/ Porticella/ Conostas/ Bonney untersucht im Detail, welche Vorteile Personen erlangen, die sich bei Citizen-Science-Projekten aktiv engagieren und bezieht auch ältere Studien mit ein. Sie gelangen zum Schluss, dass die Personen vor allem neues Wissen und Kompetenzen erwerben sowie verschiedenste Fähigkeiten erlernen und/oder ausbauen.<sup>51</sup> Zudem steigerte sich ihr Interesse an wissenschaftlichen Themen sowie damit verbundener Personen wurden positiver wahrgenommen.<sup>52</sup>

Zusammengefasst lässt sich sagen, dass sowohl für die Archive als auch die Freiwilligen, die sich an Citizen-Science-Projekten beteiligen, zahlreiche Vorteile entstehen. Zuvor wurde jedoch bereits angesprochen, dass Freiwillige nicht ausgenutzt werden dürfen – hierin liegt eine reale Gefahr und ein potentieller Nachteil für die Bürger\*innen, die sich beteiligen.

---

<sup>47</sup> Vgl. Hirschbeck, Archive, S. 8f.

<sup>48</sup> Vgl. Hirschbeck, Archive, S. 7.

<sup>49</sup> Vgl. Hirschbeck, Archive, S. 7.

<sup>50</sup> Vgl. Eibl, Engagement, S. 12.

<sup>51</sup> Phillips/ Porticella/ Conostas/ Bonney, Framework, S. 6.

<sup>52</sup> Phillips/ Porticella/ Conostas/ Bonney, Framework, S. 6. Für eine ausführliche Erklärung zu den einzelnen Bereichen, in denen sich Kompetenzen, Fähigkeiten, Motivationen und Ansichten von Teilnehmenden an Citizen-Science-Projekten änderten beziehungsweise verbesserten vgl. Ebd., S. 7–11. Es existieren zahlreiche weitere Studien, die die Personen untersuchen, die sich in Citizen-Science-Projekten engagieren. Einen guten Literaturüberblick bietet Reinsone; vgl. Reinsone, Searching, S. 187f. Dazu auch Land-Zandstra/ Agnello/ Gültekin, Participants, S. 243–259.

Zumeist werden sich Personen, die das Gefühl haben, ausgebeutet zu werden, von den Projekten abwenden und sich nicht länger beteiligen.<sup>53</sup> Es besteht jedoch auch die Gefahr, dass einige Personen trotzdem weiter für die Citizen-Science-Projekte arbeiten, da sie sich beispielsweise verpflichtet fühlen. Ebenfalls darf nicht vergessen werden, dass die Freiwilligen teilweise Interna erfahren, mit sensiblen Daten arbeiten oder Zugriff auf Digitalisate erhalten, die nicht veröffentlicht werden dürfen. Hier besteht das Risiko, dass diese weitergeben und unbefugten Personen zugänglich gemacht werden.<sup>54</sup> Es muss allerdings auch betont werden, dass dieses Risiko ebenfalls bei regulären Archivmitarbeiter\*innen besteht. Hier helfen Verträge und klare Absprachen, um das rechtliche Risiko zu minimieren.

Durch die fehlende Entlohnung besteht weiterhin die Gefahr, dass sich niemand für das Projekt meldet,<sup>55</sup> den Freiwilligen die Motivation fehlt<sup>56</sup> beziehungsweise Anreizstrukturen erst aufwendig geschaffen werden müssen.<sup>57</sup> Hier kann allerdings durch gute Absprache und Betreuung das Risiko minimiert oder zumindest frühzeitig erkannt werden.

Wie zuvor erwähnt, werden Citizen-Science-Projekte oft als kostengünstig angepriesen, allerdings können die Kosten bei schlecht geplanten Projekten teilweise nicht kalkulierbar sein oder zu teuren Fehlinvestitionen führen. Dies kann beispielsweise geschehen, wenn Archivalien digitalisiert werden, mit denen später nicht gearbeitet wird, da etwa Freiwillige oder technische Voraussetzungen fehlen.<sup>58</sup> Zudem muss berücksichtigt werden, dass auch personelle Ressourcen von Seiten des Archivs kostenintensiv sein können. Dies ist gerade dann der Fall, wenn das Fachpersonal die Freiwilligen intensiv betreuen muss.<sup>59</sup>

Auch kann das Archiv Kernkompetenzen verlieren, wenn diese an Externe ausgelagert werden, was wiederum zu internen Widerständen führen kann.<sup>60</sup> Zu bedenken ist

---

<sup>53</sup> Vgl. Becker, Citizen Science, S. 34. Zur Frage, ob Crowdsourcing als Ausbeutung verstanden werden kann, ausführlich Kidd, die betont, dass es bei der Beantwortung stark auf den Kontext ankommt: Profitieren die Freiwilligen in einer anderen, nicht monetären Form von ihrer Arbeit? Dies können etwas die Möglichkeiten zur Beteiligung an wissenschaftlichen Prozessen sein. Wichtig ist ihrer Meinung auch, sich den ethnischen Fragen zu stellen, die in diesem Kontext auftreten und zu überlegen, wie Citizen-Science-Projekte Beteiligung, Wissens- und Kompetenzzuwachs sowie Selbstermächtigung statt Ausbeutung fördern können; vgl. Kidd, Citizen participation, S. 73–89.

<sup>54</sup> Vgl. Hirschbeck, Archive, S. 8f.

<sup>55</sup> Vgl. Hirschbeck, Archive, S. 8f.

<sup>56</sup> Vgl. Gränitz, Schwarmwissen, S. 6.

<sup>57</sup> Vgl. Becker, Citizen Science, S. 34.

<sup>58</sup> Vgl. Hirschbeck, Archive, S. 8f.

<sup>59</sup> Vgl. Gränitz, Schwarmwissen, S. 6.

<sup>60</sup> Vgl. Becker, Citizen Science, S. 34.

allerdings, dass zwar oft die Erschließungstätigkeiten im Rahmen von Citizen-Science-Projekten kritisiert wird, dabei jedoch übersehen wird, dass in vielen Archiven seit Jahren reichlich Archivgut von Freiwilligen/Ehrenamtlichen erschlossen wird. Teilweise handelt es sich um ehemalige Mitarbeiter\*innen, teilweise um engagierte Einzelpersonen.<sup>61</sup> Diese Tätigkeiten werden weniger publik als die Erschließungsarbeit in Citizen-Science-Projekten, über die deutlich intensiver berichtet wird. Erfahrungsberichte zeigen, dass derartige Erschließungen in vielen Archiven unverzichtbar sind und eben nicht zu einem Personalabbau oder Abfluss von Wissen oder Kompetenzen führt.<sup>62</sup> Entsprechend ist davon auszugehen, dass bei Citizen-Science-Projekten im Bereich der Erschließung das Risiko von internen Widerständen und Personalabbau zwar existiert, es aber kaum zu derartigen Problemen kommt. Schon alleine, da – wie zuvor angeführt – auch für die Betreuung von Freiwilligen Personal benötigt wird.

Ein großes Problem stellt die Schnittstellenproblematik sowie die Qualitätssicherung dar: Wenn die Daten in einem Citizen Science-Projekt und nicht direkt im Archiv mit der Archivinformationssoftware erschlossen werden, besteht die Gefahr, dass die Daten aufgrund fehlender Schnittstellen anschließend nicht übernommen werden können.<sup>63</sup> Auch eine fehlende oder fehlerhafte Qualitätssicherung kann zu ungeeigneten Daten führen,<sup>64</sup> etwa wenn sich Freiwillige nicht an vorgegebene Regeln (wie beispielsweise an die Erschließungsrichtlinien) halten wollen oder können. Letzteres kann vorkommen, wenn die Regeln unklar formuliert sind oder nicht ausreichend zwischen den Freiwilligen, den Projektkoordinator\*innen sowie den beteiligten Archiven abgestimmt wurden.<sup>65</sup> Eine gute Qualitätssicherung sowie -kontrolle ist folglich unabdingbar für das Gelingen eines Citizen Science-Projekts und – wie zuvor ausgeführt – das größte Risiko.

---

<sup>61</sup> Vgl. Eibl, Engagement, S. 9.

<sup>62</sup> Vgl. dazu auch den Erfahrungsbericht von Lippik, die schildert, wie im Stadtarchiv Mühlacker seit Jahren Ehrenamtliche archivische Kernaufgaben erledigen; vgl. Lippik, Ehrenamtliche, S. 13-17.

<sup>63</sup> Vgl. Hirschbeck, Archive, S. 10f.

<sup>64</sup> Vgl. Hirschbeck, Archive, S. 11. Ein weiteres Risiko, das verschiedentlich angesprochen wird, ist jenes der abnehmenden Nutzer\*innen durch mehr online zur Verfügung stehenden Digitalisaten; vgl. beispielsweise Ebd., S. 10. Da inzwischen in den meisten Archiven die Zahl der Nutzer\*innen nicht nur anhand der Personen im Lesesaal, sondern auch anhand derjenigen, die online auf die Archivalien zugreifen, gezählt werden, ist dieses Argument hinfällig.

<sup>65</sup> Vgl. Balázs u.a., Data Quality, S. 145f.

## **2. Qualitätssicherung und -kontrolle bei Citizen Science: Das Projekt**

Diese Arbeit soll die Frage beantworten, ob im Rahmen von Citizen-Science-Projekten gewonnene Erschließungsdaten von Archiven übernommen werden können und wenn ja, ob eine Qualitätskontrolle von Nöten ist. Qualität meint in diesem Kontext akkurat, vollständig und relevant.<sup>66</sup> Akkurat bedeutet wiederum, dass erstens der Datenkontext verdeutlicht wird, d.h. Metadaten generiert, bereitgestellt und gesichert werden. Zweitens, dass die Daten gültig und verlässlich sind. Drittens, dass ihre langfristige Nutzung und Weiterverwendung gesichert wird, d.h. früh über Schnittstellen und kompatible Systeme nachgedacht wird. Und viertens, dass die rechtliche Lage eindeutig ist, d.h. klar ist, wem die Daten gehören, wer Zugriff erhält und wie mit ihnen gearbeitet werden kann.<sup>67</sup> Zentral ist die Frage, wie und ob die Qualität der Daten während des Eingebens und danach geprüft werden – dass eine Prüfung stattfinden muss, ist allgemeiner Konsens und zentral für das Gelingen des Projekts.<sup>68</sup>

### **2.1. Methodik und Vorgehensweise**

Um diese Fragen beantworten zu können, ist ein Blick auf die von den Ehrenamtlichen verwendeten Plattformen unabdingbar. Citizen-Science-Projekte nutzen eine Vielzahl von Plattformen, Anwendungen und Vorgehensweisen, die sich nicht nur von Projekt zu Projekt, sondern teilweise auch innerhalb des Projekts unterscheiden. Im Rahmen dieser Arbeit werden verschiedenen Anwendungen beleuchtet, die aktuell zur Erschließung eingesetzt werden. Dabei handelt es sich sowohl um Anwendungen, die online – beispielsweise über ein Programm im Webbrowser – oder offline verwendet werden. Auch Anwendungen, die auf künstlicher Intelligenz beruhen, werden in den Blick genommen.

Bei der offline Anwendung wurde Excel ausgewählt, da es sich um eine gängige Anwendung handelt, die auch außerhalb von Citizen-Science-Projekten häufig zum Erschließen genutzt wird und mit den meisten Archivinformationssystemen kompatibel ist. Bei den online Anwendungen werden sowohl die Eigenentwicklung des Vereins für Computergenealogie, das Daten-Eingabe-System (= DES), als auch Transkribus, eine Anwendung zum Transkribieren von Druck- und Handschriften berücksichtigt. Ersteres bietet sich an, da es eine häufig verwendete Anwendung ist, die vor allem

---

<sup>66</sup> Vgl. Balázs u.a., Data Quality, S. 140.

<sup>67</sup> Vgl. Balázs u.a., Data Quality, S. 143.

<sup>68</sup> Vgl. Balázs u.a., Data Quality, S. 147f.

vom Verein für Computergenealogie genutzt wird. Da das Erschließungsprojekt in Kooperation mit diesem Verein durchgeführt wird, erscheint es sinnvoll, auch ihre Anwendung zu untersuchen. Als dritte Anwendung wurde Transkribus ausgewählt, um der Frage zu begegnen, welche Rolle KI heute im Rahmen von Erschließungsprojekten spielt. Transkribus ist besonders geeignet für eine derartige Untersuchung, da es die weltweit am meisten genutzte Anwendung bei der Erkennung von (historischen) Handschriften ist.<sup>69</sup>

## **2.2. Erster Projektabschnitt: Expert\*inneninterviews**

Um diese Anwendungen zu untersuchen, wurden in einem ersten Schritt Personen befragt, die mit diesen Anwendungen arbeiten. Hierfür wird ein qualitativer Ansatz verwendet, bei dem die inhaltlichen Fragen im Vordergrund stehen und keine quantitativen Fragen (wie etwa jene nach der Anzahl der Nutzer\*innen einer Anwendung) beantwortet werden. Repräsentativität wurde folglich nicht in einem statistischen, sondern einem inhaltlichen Sinne realisiert.<sup>70</sup>

Als Methode für den ersten Teil dieser Arbeit wurde die Konzeption und Umsetzung von halboffenen, semistrukturierten Expert\*inneninterviews ausgewählt. Bei derartigen Interviews sind die Fragen, die den Interviewten gestellt werden, festgelegt und für alle befragten Personen identisch. Während die Fragen festgelegt sind, sind es die Antwortmöglichkeiten hingegen nicht. Die Interviewten haben folglich die Möglichkeit, frei zu antworten.<sup>71</sup>

Es handelt sich dabei nicht um narrative Interviews, bei denen die Strukturierung durch die Interviewten vorgegeben wird, sondern um semistrukturierte Leitfrageninterviews, bei denen die Struktur durch die Forschenden bestimmt wird.<sup>72</sup> Entsprechend der drei Grundprinzipien qualitativer Forschung wurden die Interviewfragen gestaltet: Offen, d.h. es sind keine Ja/Nein-Antworten möglich; kommunikativ, d.h. bei den

---

<sup>69</sup> Stand 2022; vgl. Nockels/ Gooding/ Ames/ Terras: *Understanding*, S. 368. Selbstverständlich existieren weitere Anwendungen, bei zur Erschließung in Citizen-Science-Projekten verwendet werden; beispielsweise im Zusammenhang mit der Initiative #everynamecounts der Arolsen Archives; vgl. Hirschbeck, *Archive*, S. 37f. oder die Anwendung von FamilySearch; vgl. Ebd., S. 40-42. Ursprünglich wurde auch FactGrid, ein Projekt des Forschungszentrums der Universität Erfurt, für eine Untersuchung ausgewählt. Da jedoch nur ein Interview zu Stande kam und der Umfang der Arbeit eine Analyse von vier Anwendungen nicht zuließ, wurde darauf verzichtet. Zu erwähnen ist allerdings die lesenswerte Einschätzung von Simons zur Nutzung der Anwendung bei Citizen-Science-Projekten; vgl. Simons, *Selbstverständlichkeit*, S. 242–265.

<sup>70</sup> Vgl. Misoch, *Interviews*, S. 2.

<sup>71</sup> Vgl. Misoch, *Interviews*, S. 13.

<sup>72</sup> Vgl. Misoch, *Interviews*, S. 65.

Formulierungen wurde darauf geachtet, dass diese verständlich sind, sowie prozesshaft, d.h. die Fragen berücksichtigen Veränderungen sowie die Rolle der Umwelt.<sup>73</sup> Letzteres bezieht sich etwa auf die Rolle der Einbettung einer Erschließungsanwendung in ein übergeordnetes Projekt.

Der Begriff der Expert\*inneninterviews bezieht sich auf die Personen, mit denen die Interviews geführt werden.<sup>74</sup> Personen können als Expert\*innen bezeichnet werden, wenn sie „für ein bestimmtes Wissensgebiet aufgrund ihrer eigenen Aktivität in diesem Bereich fundierte Auskunft geben können.“<sup>75</sup> Diese Aktivität kann, muss aber nicht hauptberuflich erfolgen. Wichtig ist lediglich, dass diesen Personen Expert\*innenstatus „für einen bestimmten Wissens/Erkenntnisbereich zugesprochen werden kann.“<sup>76</sup> In dieser Arbeit wird die Rolle des/der Expert\*in Personen zugeschrieben, die eine langjährige Erfahrung mit Citizen-Science-Projekten und vor allem mit Anwendungen, die zur Erschließung verwendet werden, vorweisen könne. Dabei handelt es sich zum einen um Personen, die konzeptionell an diesen Anwendungen beteiligt sind, etwa indem sie diese selbst programmiert haben oder in die Programmierung/Planung involviert waren, und zum anderen um Personen, die selbst aktiv an der Erschließung von Archivalien mittels dieser Anwendungen mitwirken oder mitgewirkt haben. Sie stehen „nicht als individuelle Personen im Vordergrund (...), sondern als Funktionsvertreter und somit als spezielle Wissensträger. Damit rücken subjektive Konstruktionen und Sinngebungen in den Hintergrund und institutionalisierte Wissensbestände und Problemlösungen stehen im Vordergrund.“<sup>77</sup> Es wurde darauf geachtet, dass pro Anwendung mindestens zwei Interviews geführt wurden. Die Fragen wurden größtenteils schriftlich beantwortet und nur in zwei Fällen wurde das Interview mündlich durchgeführt. Die Ergebnisse der Interviews fließen anonymisiert in diese Arbeit ein<sup>78</sup> und lassen Rückschlüsse auf die Qualitätssicherung und -kontrolle bei den einzelnen Anwendungen sowie – daraus abgeleitet – deren Eignung für Citizen-Science-Projekte zu.

---

<sup>73</sup> Vgl. Misoch, Interviews, S. 66f.

<sup>74</sup> Vgl. Misoch, Interviews, S. 119.

<sup>75</sup> Misoch, Interviews, S. 120.

<sup>76</sup> Misoch, Interviews, S. 120.

<sup>77</sup> Misoch, Interviews, S. 120.

<sup>78</sup> Vgl. dazu auch die ethischen Grundprinzipien für sozialwissenschaftliche Untersuchungen mittels Interviews; Misoch, Interviews, S. 18-20.

### 2.3. Zweiter Projektabschnitt: Erschließungsarbeiten

Um die Aussagen der Expert\*innen zu überprüfen und weitere Daten als Ergänzung zu den Interviews zu sammeln, wurde ein eigenes Erschließungsprojekt in Kooperation mit dem Stadtarchiv Leipzig sowie dem Verein für Computergenealogie initiiert und durchgeführt: Die Erschließung von Mieter\*innenverzeichnisse aus dem 19. Jahrhundert. Da das Gelingen eines Citizen Science-Projekts von der Vorarbeit abhängt, wird hier in Kürze dargelegt, welche Vorüberlegungen erfolgten: Das Projekt wurde so gestaltet,<sup>79</sup> dass

1. dessen Thema für Freiwillige von Interesse ist;
2. die Fragestellung bzw. das Thema mittels Citizen Science bearbeitet werden kann;
3. die benötigten Ressourcen (u.a. personell und zeitlich) vorhanden sind;
4. eine langfristige Ergebnissicherung gewährleistet ist;
5. geklärt wird, was nach Projektende mit den Erschließungsdaten geschieht;
6. die rechtlichen Rahmenbedingungen geklärt sind und
7. Evaluationskriterien festgelegt werden.

Da es sich um ein Erschließungsprojekt von überschaubarer Größe handelt, ist es gut für ein Citizen Science-Projekt geeignet. Das Thema ist für eine große Zielgruppe von Interesse, da sich zum einen Genealog\*innen für derartige Verzeichnisse begeistern. Durch Mieter\*innenverzeichnisse können sie erforschen, wo ihre Vorfahren gelebt haben, welchen Berufen sie nachgegangen sind und in welcher (räumlichen) Beziehung sie zu ihren Familienmitgliedern oder Bekannten standen.<sup>80</sup> Zudem sind Genealog\*innen nicht nur eine sehr große, sondern auch eine oft innovative und technikaffine Nutzer\*innengruppe, die sich häufig in Citizen-Science-Projekten engagiert und folglich Erfahrung bei der Online-Erschließung von Archivalien aufweist.<sup>81</sup>

Zum anderen ist das Projekt für Studierende der Universität Halle sowie verschiedene Multiplikatoren interessant, da es an bereits bestehende Initiativen anknüpft. Seit einigen Semester existiert eine von Prof. Dr. Georg Fertig, Professor an der Universität

---

<sup>79</sup> Vgl. Pettibone u.a., Citizen Science, S. 9. Der Fragebogen wurde hier paraphrasiert und auf jene Fragen gekürzt, die zu geisteswissenschaftlichen Citizen-Science-Projekten passen. Fragen, die auf naturwissenschaftliche Projekte abzielen, werden für diese Arbeit nicht berücksichtigt (wie etwa das Vorhandensein eines Labors). Mit der Notwendigkeit von digitalem Spezialwissen befassen sich inzwischen auch Forscher\*innen; vgl. beispielsweise Lemaire/ Rommelfanger, Berücksichtigung, S. 69–90.

<sup>80</sup> Nach Aussage von Prof. Dr. Georg Fertig führte räumliche Nähe teilweise auch zu Eheschließungen. Die Informationen zu den Wohnorten lassen somit auch Rückschlüsse darauf zu, wie sich spätere Ehepaare kennenlernten.

<sup>81</sup> Vgl. Becker, Citizen Science, S. 31.

Halle sowie Vorsitzendem des Vereins für Computergenealogie, geleitete Seminarreihe.<sup>82</sup> In dieser werden historische Daten zur Leipziger Stadtgeschichte erschlossen und mit dreidimensionalen, digitalen Modellen verbunden. Konkret wurden bislang wohnungs- und stockwerksgenaue Daten über Mieter\*innen und Hausbesitzer\*innen erschlossen und diese verschiedentlich digital verknüpft sowie dreidimensionale Modelle erstellt, die diese visualisieren.<sup>83</sup> Grundlage waren verschiedene Verzeichnisse, die Informationen zu den Häusern und Wohnungen enthielten, darunter etwa Brandkataster und Schuldverzeichnisse, die im Zusammenhang mit der kommunalen Schuldentilgung der Stadt Leipzig in der ersten Hälfte des 19. Jahrhunderts entstanden.<sup>84</sup>

Für das Erschließungsprojekt dieser Arbeit wurde der Bestand 0008 (Ratsstube), Nr. 3171 bis Nr. 3203 digitalisiert.<sup>85</sup> Davon wurden bereits für vergangene Seminare einige Bücher erschlossen (darunter Nr. 3186 und Nr. 3187), allerdings anhand von privaten Fotografien und nicht Digitalisaten. Im Rahmen dieses Projekts wurde mit StadtAL 0008 Nr. 3190 ein umfangreiches Buch für die Erschließung ausgewählt. Das Mieter\*innenverzeichnis beinhaltet vor allem die tabellarische Auflistung aller Mieter\*innen im Universitätsgebäude samt Hausnummern und dem jährlichen Mietzins der Jahre 1842 bis 1853.<sup>86</sup> Die handschriftlichen Einträge wurden von verschiedenen Schreibern verfasst und beinhalten auch Nachträge und Korrekturen. Erschlossen wurde es in erster Linie mit der Anwendung Excel sowie ausgewählte Seiten mit Transkribus.

Durch die zuvor benannten Anknüpfungsmöglichkeiten lassen sich auch einige der anderen Fragen beantworten: Die Erschließungsdaten können nicht nur von Genealog\*innen zur Recherche verwendet werden, sondern auch von Studierenden, um weitere Projekte mit Bezug zur Leipziger Stadtgeschichte umzusetzen.<sup>87</sup> Die

---

<sup>82</sup> Für die Beschreibungen im Kommentierten Vorlesungsverzeichnis vgl. für das Sommersemester 2022 „Das Leipzig-Projekt, Phase I – digitale Methoden für die urbane Sozialgeschichte“, für das Wintersemester 2023/24 „Das Leipzig-Projekt, Phase II“ und für das Sommersemester „Das Leipzig-Projekt, Phase III“ (zu finden unter [https://studip.uni-halle.de/dispatch.php/course/details?sem\\_id=15442f4cd15b1f054a2cddb495cf2556&again=yes](https://studip.uni-halle.de/dispatch.php/course/details?sem_id=15442f4cd15b1f054a2cddb495cf2556&again=yes) (16.02.2024)).

<sup>83</sup> Eine Übersicht zu den bislang umgesetzten Projekten findet sich unter <https://database.factgrid.de/wiki/FactGrid:Leipzig> (16.02.2024). Dazu gehört auch das vielbeachtete Hebammenverzeichnis sowie die visualisierten Umzüge vom und zum Gewandgässchen.

<sup>84</sup> Für die französischen Kontribution und Schuldentilgung in Leipzig vgl. Beachy, Soul, S. 137–221.

<sup>85</sup> Vgl. <https://recherche-stadtarchiv.leipzig.de/actaproweb/archive.xhtml> (die Digitalisate sind unter 1. Städtische Körperschaften und deren nachgeordnete Einrichtungen -> 1.2. Stadtverwaltung -> 1.2.1. Anfänge bis um 1830 -> 0008 Ratsstube -> Titelakten -> 43. Stadtschulden, Kommunalablagen, -abgaben -> 43.2. Französische Kontribution zu finden).

<sup>86</sup> [https://recherche-stadtarchiv.leipzig.de/actaproweb/document/Vz\\_5c85ade0-19da-43ac-a9fa-a8fcb8cd1749](https://recherche-stadtarchiv.leipzig.de/actaproweb/document/Vz_5c85ade0-19da-43ac-a9fa-a8fcb8cd1749) (28.02.2024).

<sup>87</sup> Die Erschließungsdaten, die im Rahmen des praktischen Projekts entstanden, werden perspektivisch in FactGrid gespeist. Im FactGrid wird es anschließend möglich sein, sich beispielsweise alle

Erschließungsdaten werden somit nach Projektende weiterverwendet und dienen als Grundlage für weitere Seminare und stadtgeschichtliche Projekte. Zudem besteht die Überlegung, die Erschließungsdaten in das Findbuch des Stadtarchivs zu übernehmen – allerdings nur, wenn diese qualitativ hochwertig genug sind und in einer Form vorliegen, in der sie übernommen werden können.<sup>88</sup> Damit kann auch die Frage nach der Evaluation beantwortet werden: das Projekt ist dann erfolgreich, wenn die Erschließungsdaten am Ende in einer Form vorliegen, die nachhaltig ist, mit der weitergearbeitet werden kann, deren Qualität gesichert ist und die von anderen Stakeholdern wie eben dem Leipziger Stadtarchiv übernommen werden können.

Durch die Kooperation mit Prof. Dr. Fertig und dem Leipziger Stadtarchiv ist zudem gewährleistet, dass die benötigten Ressourcen vorhanden sind: Das Archiv finanzierte die Digitalisierung der Archivalien, mit denen in diesem Projekt gearbeitet wurde, und Prof. Dr. Fertig stelle den Kontakt zu verschiedenen Personen her, die sich an der Erschließung beteiligten.<sup>89</sup>

---

Angehörigen eines bestimmten Berufsstands oder zusammengehörige Familiengruppen anzeigen zu lassen. Perspektivisch sollen alle Häuser- und Mieter\*innenverzeichnisse erschlossen werden, die die Leipziger Innenstadt betreffen. Geplant ist weiterhin eine Verbindung mit der Plattform „Altes Leipzig“, die verschiedene archivalische Quellen nutzt, um zum einen digitale 360°-Panoramen der Stadt zwischen 1880 und 1920 zu erstellen und zum anderen ein stadtgeschichtliches Portal zu betreiben. Bei dem virtuellen Rundgang lassen sich nicht nur die animierten Straßen erkunden, sondern auch Informationen zu einzelnen Gebäuden abrufen. Hier könnte eine Verbindung zu dem Projekt hergestellt werden, indem die Informationen um jene bereichert werden, die im Projekt erarbeitet wurden. Denkbar wäre beispielsweise, dass beim Mausklick auf einzelne Gebäude direkt die ehemaligen Bewohner\*innen angezeigt werden. Geplant ist bei „Altes Leipzig“ auch die Erstellung von Familienmatrizen, die detaillierte Informationen zu einzelnen Personen, ihren beruflichen Stationen, Wirkungsorten und der Veränderung ihrer familiären Situation aufzeigen. Eine beispielhafte Matrix findet sich unter [https://www.altes-leipzig.de/media/galerie\\_detail/2020\\_struktur03.jpg](https://www.altes-leipzig.de/media/galerie_detail/2020_struktur03.jpg) (12.01.2024). Auch hier zeigt sich die Anschlussfähigkeit direkt: Die Daten, die in den Mietsverzeichnissen aufgenommen wurden, eignen sich ideal, um derartige Matrizen zu ergänzen oder überhaupt erst zu erstellen. Für die Projektseite auf Bürger schaffen Wissen vgl. <https://www.buergerschaffenwissen.de/projekt/altes-leipzig> (12.01.2024). Für die Website sowie das digitale Panorama vgl. <https://www.altes-leipzig.de/virtuelles-leipzig/> (12.01.2024) sowie <https://www.altes-leipzig.de/pano/1894/A1/A1.html> (12.01.2024). Geplant ist auch der Abgleich der erhobenen Daten mit dem Holzmodell der Stadt, das 1823 von Johann Christoph Merzdorf im Maßstab von 1:390 gefertigt wurde. Bereits jetzt ist ein 3D-Modell online verfügbar, mit dem allerdings nur wenige Informationen verknüpft sind. Für das Modell vgl. <https://www.leipzig.de/bauen-und-wohnen/bauen/geodaten-und-karten/historisches-3d-stadtmodell-leipzig> (27.02.2024). Die Verknüpfung mit den ehemaligen Bewohner\*innen der einzelnen Gebäude ist ein spannendes Projekt, das allerdings nicht mehr 2024 realisiert werden wird.

<sup>88</sup> Hieran zeigt sich erneut die Relevanz dieser Arbeit, die nach Qualitätssicherung sowie -kontrolle fragt und dafür verschiedene Anwendungen in den Blick nimmt.

<sup>89</sup> Moeller/ Müller betonen, dass Forschungsergebnisse vor allem dann aufschlussreich sind, wenn sie auf der tatsächlichen Zusammenarbeit mit Freiwilligen beruhen; vgl. Moeller/ Müller, Heimatforscher, S. 95.

### 3. Auswertung des Projekts und der verwendeten Plattformen

Im Rahmen dieser Arbeit wurden drei Anwendungen ausgewählt und mittels Experteninterviews sowie einem praktischen Erschließungsprojekt untersucht. Während die Interviews alle Anwendungen abdecken, wurde im Projekt aus Zeitgründen nur mit zwei der Anwendungen – Excel und Transkribus – gearbeitet. Die Ergebnisse werden in den Kapiteln zu den einzelnen Anwendungen präsentiert und mit bereits bestehenden Studien und Untersuchungen kombiniert.

#### 3.1. Microsoft Excel

Die erste Anwendung, die ausgewählt wurde, ist Microsoft Excel, eine Software, die zum Anlegen und Bearbeiten von Tabellen genutzt wird und weit verbreitet ist. Excel kann sowohl offline als auch in Form von GoogleSheets oder ExcelOnline im Browser verwendet werden. Während bei der offline Nutzung Lizenzgebühren anfallen, ist die online Nutzung aktuell (noch) kostenlos. Eine Anmeldung mit eigenem Konto ist jedoch bei allen drei vorgestellten Varianten zwingend erforderlich.<sup>90</sup> Der Export der Daten ist problemlos im csv-, txt- oder XML-Format möglich,<sup>91</sup> wodurch die Erschließungsdaten gut in die gängigen Archivischen Fachinformationssysteme übernommen werden kann. Erforderlich ist lediglich das Anlegen und die passende Benennung der einzelnen Spalten.<sup>92</sup>

Vor der Einführung von Citizen Science-Anwendungen, die speziell für die Erschließung konzipiert sind, wurde in vielen Orten mit der offline-Version von Excel gearbeitet.<sup>93</sup> Das damalige Verfahren stellte sich jedoch als umständlich und fehleranfällig dar: Das Hauptproblem war, dass Fehler oft erst nach tausenden Einträgen bemerkt wurden.<sup>94</sup> Auch die Absprachen gestalteten sich schwierig, da man nicht sehen konnte, welche Person was genau wie bearbeitet und es folglich teilweise zu doppelten

---

<sup>90</sup> <https://www.microsoft.com/de-de/microsoft-365/excel?market=de> (23.02.2024). Für ExcelOnline fallen keine Kosten an, es enthält aber etwas weniger Funktionen als das klassische Excel mit Lizenzgebühren. Diese sollten allerdings bei Citizen-Science-Projekten nicht ausschlaggebend sein, da es sich vor allem um Unterschiede bei der Darstellung von Diagrammen und ähnlichem handelt. Ein Beispiel: Während man diese bei Excel in jeder gewünschten Farbe erstellen kann, muss man sich bei ExcelOnline auf vorgefertigte Designs beschränken.

<sup>91</sup> <https://support.microsoft.com/de-de/office/importieren-oder-exportieren-von-textdateien-txt-oder-csv-5250ac4c-663c-47ce-937b-339e391393ba> (23.02.2024) und <https://support.microsoft.com/de-de/office/exportieren-von-xml-daten-0b21f51b-56d6-48f0-83d9-a89637cd4360> (23.02.2024). Wenn die Spalten entsprechend angepasst werden, können die Daten auch direkt von Excel in ActaPro übernommen werden; vgl. Gränitz, Schwarmwissen, S. 41f. Das Leipziger Stadtarchiv nutzt ActaPro.

<sup>92</sup> Vgl. Interview E; vgl. Interview F.

<sup>93</sup> Vgl. Interview A.

<sup>94</sup> Vgl. Interview A. Vgl. dazu auch Zedlitz, Daten-Eingabe-System, S. 28, der ebenfalls dieses Problem hervorhebt.

Erschließungen kam. Als letzter Punkt wurde der Versand der Daten angesprochen, der früher mittels CD-Versand erfolgte und somit umständlich war sowie Kosten verursachte.<sup>95</sup> Inzwischen besteht zwar die Möglichkeit, auch per Mail oder Drittanbieter große Datenmengen zu versenden, die anderen beiden Probleme existieren jedoch weiterhin. Es zeigt sich folglich, dass Excel in der offline-Version heutzutage wenig geeignet für Erschließungen mittels Citizen-Science-Projekten ist. Eine Qualitätskontrolle erfolgt, wenn überhaupt, nur zeitlich versetzt und wenn Fehler gefunden werden, besteht die Gefahr, dass diese tausende Einträge betreffen, wodurch das Projekt ungeplant deutlich länger dauern könnte.

Etwas anderes gestaltet es sich bei den online-Versionen. Wenn online gearbeitet wird, können mehrere Personen gleichzeitig an einer Tabelle arbeiten; benötigt wird lediglich eine stabile Internetverbindung.<sup>96</sup> Der Vorteil ist hierbei, dass gemeinsam an einer Tabelle gearbeitet werden kann.<sup>97</sup> Besonders positiv wurde in den Interviews die Übersichtlichkeit hervorgehoben und die Möglichkeit, benötigte Anpassungen oder Konfigurationen selbst vornehmen zu können,<sup>98</sup> da dies bei vielen anderen Anwendungen nur ein Administrator erledigen kann.

Auch die gegenseitige Überprüfung und Qualitätskontrolle ist einfach: Man kann jederzeit (bei Bedarf in Echtzeit) auf die Erschließungsdaten zugreifen und diese korrigieren. Dies wurde auch im Erschließungsprojekt so umgesetzt und führte zu weitestgehend fehlerfreien Einträgen. Bei Leseunsicherheiten kann zudem das betreffende Wort markiert werden, damit die Personen, die die Einträge kontrollieren, direkt sehen, welche Stellen unklar sind und entsprechend ein besonderes Augenmerk auf diese legen können. Wichtig ist in diesem Fall, dass alle beteiligten Personen über gute Lesekenntnisse verfügten, da kein Administrator eine Endabnahme der Ergebnisse beaufsichtigt.

Filterfunktionen können zudem die Darstellung übersichtlicher gestalten, die sonst – gerade bei großen Datenmengen – schnell unübersichtlich werden kann.<sup>99</sup> Bei sehr großen Datenmengen, etwa bei Millionen von Adressbucheinträgen, stößt die Anwendung jedoch an ihre Grenzen und es kann zu deutlichen Verzögerungen oder

---

<sup>95</sup> Vgl. Interview A. Zudem erfolgte die Publikation der Daten immer erst nach Abschluss des Projekts. Bei anderen Anwendungen wie etwa dem DES können die Daten bereits während der Erschließung bzw. Bearbeitung eingesehen und durchsucht werden; vgl. Nicola, Quellen, S. 68.

<sup>96</sup> <https://www.google.de/intl/de/sheets/about/#features> (24.02.2024).

<sup>97</sup> Vgl. Interview E; vgl. Interview F.

<sup>98</sup> Vgl. Interview E; vgl. Interview F.

<sup>99</sup> Vgl. Interview E; vgl. Interview F.

Problemen bei der Darstellung kommen.<sup>100</sup> Zudem benötigt die Anwendung mit den einzelnen Spalten relativ viel Platz und man benötigt entweder einen sehr großen Bildschirm oder direkt zwei, da man sowohl das Digitalisat der Quelle als auch die Anwendung sehen sollte.<sup>101</sup> Wenn zwischen beidem gesprungen wird, erhöht sich das Risiko, beispielsweise in der Zeile zu verrutschen und fehlerhaft zu erschließen.

Neben der aktuellen Version werden frühere Versionen gespeichert, was einen gewissen Schutz gegen Vandalismus bietet. Wenn eine Person alle Einträge löscht, kann auf eine frühere Version zurückgegriffen werden. Dennoch sind regelmäßige Sicherungen von Vorteil. Hier sind gute Absprachen von Nöten, da es keine Administrator\*innen gibt, die derartige Aufgaben routinemäßig erledigen.

Zusammengefasst lässt sich sagen, dass Excel in der offline-Version wenig geeignet für Citizen-Science-Projekte erscheint, da die Qualitätskontrolle deutlich zeitversetzt erfolgt und auch die Absprachen fehleranfällig sind. Bei den online-Versionen sieht es hingegen anders aus: wenn die Freiwilligen über gute Lesekompetenzen verfügen, gut zusammenarbeiten können und die Anzahl der Einträge nicht in die Millionen gehen, ist die Anwendung gut geeignet für derartige Erschließungsprojekte. Wichtig sind regelmäßige Sicherungen und eine gute Absprache mit den beteiligten Archiven, um die Daten am Ende in das Findbuch übernehmen zu können.

### **3.2. Daten-Eingabe-System (DES)**

Als zweite Anwendung wurde die Eigenentwicklung des Vereins für Computergenealogie, das Daten-Eingabe-System (kurz DES) untersucht. Hierbei handelt es sich um eine Anwendung, bei der direkt auf dem Digitalisat und nicht in einem separaten Fenster erschlossen wird.<sup>102</sup> Besondere Fähigkeiten oder eine vertiefte Einarbeitung in die Anwendung sind nicht notwendig: die Freiwilligen klicken auf den Zeilenanfang und können direkt mit der Eingabe auf dem Digitalisat beginnen.<sup>103</sup>

Wenn die Freiwilligen Mitglieder im Verein für Computergenealogie sind, können sie zudem einfach ihren bereits bestehenden Zugang nutzen, um sich anzumelden und müssen sich nicht neu registrieren.<sup>104</sup> Nach dem Anmelden werden den Freiwilligen eine Liste aller verfügbaren Projekte mit einer Prozentangabe angezeigt, die den

---

<sup>100</sup> Zedlitz, Daten-Eingabe-System, S. 28. Wie die praktische Projektphase zeigte, sind mehrere tausend Einträge unproblematisch. Es kommt weder zu Fehlern bei der Darstellung noch zu Verzögerungen.

<sup>101</sup> Zedlitz, Daten-Eingabe-System, S. 28.

<sup>102</sup> Zedlitz, Daten-Eingabe-System, S. 28.

<sup>103</sup> Zedlitz, Daten-Eingabe-System, S. 29.

<sup>104</sup> Zedlitz, Daten-Eingabe-System, S. 29.

Fortschritt dokumentiert und motivierend wirken kann.<sup>105</sup> Die Freiwilligen können an mehreren Projekten parallel arbeiten oder kurz vor Abschluss ihres Wunschprojekts sich direkt über das nächste informieren,<sup>106</sup> was dazu beiträgt, sie längerfristig an Projekte und die Plattform zu binden. Die Anbindung an den Verein für Computergenealogie trägt dazu bei, dass sich viele Personen an einem Projekt beteiligen und – durch deren Interesse an Genealogie und Handschriften – meist das benötigte Fachwissen für eine Erschließung mitbringen.<sup>107</sup>

Für jedes Projekt werden projektspezifische Erschließungsrichtlinien und verschiedene Einführungen zur Verfügung gestellt, die gewährleisten (sollen), dass die Erschließungen systematisch ablaufen und Freiwillige ausreichend Unterstützung erhalten.<sup>108</sup> Oft werden bei umfangreichen Projekten ergänzende Onlineschulungen angeboten.<sup>109</sup>

Hierin besteht direkt der erste Vorteil der Anwendung: die Einarbeitung in das Projekt wird gut unterstützt. Die Erschließung direkt auf dem Digitalisat bedeutet zudem, dass ein kleiner Bildschirm ausreicht, um gut erfassen zu können und sowohl die Freiwilligen als auch die Administrator\*innen auf einen Blick sehen, ob alles erfasst wurde.<sup>110</sup>

Das bedeutet jedoch auch, dass die Vorbereitung deutlich aufwendiger ist und das einfache Digitalisat nicht ausreicht. Stattdessen muss ein Projekt angelegt, die Digitalisate in Eingabefelder unterteilt und grundlegende Entscheidungen hinsichtlich der aufzunehmenden Informationen getroffen werden.<sup>111</sup> Dafür benötigt man Administrator\*innen mit Fachwissen und Zeit. Das Projekt muss somit mit einer längeren Vorlaufzeit geplant werden als vergleichbare Projekte, die eine andere Anwendung (wie beispielsweise Excel) verwenden. Hier haben die interviewten Expert\*innen auch Verbesserungsbedarf gesehen: Das Anlegen der Projekte sollte perspektivisch leichter werden.<sup>112</sup>

---

<sup>105</sup> <http://des.genealogy.net/start/selectProject> (12.02.2024).

<sup>106</sup> Zedlitz, Daten-Eingabe-System, S. 29.

<sup>107</sup> Vgl. Gränitz, Schwarmwissen, S. 40.

<sup>108</sup> [https://wiki.genealogy.net/Kartei\\_Leipziger\\_Familien#Editionsrichtlinien](https://wiki.genealogy.net/Kartei_Leipziger_Familien#Editionsrichtlinien) (14.02.2024).

<sup>109</sup> <https://www.compgen.de/2023/10/online-treff-am-12-oktober-jeder-kann-mitmachen-bei-compgen-helft-mit-die-daten-einzugeben/> (14.02.2024).

<sup>110</sup> Zedlitz, Daten-Eingabe-System, S. 28f. Die Seite, die von einer bestimmten Person erschlossen wird, wird für diese eine gewisse Zeit lang reserviert. Würden mehrere Personen auf einer Seite arbeiten, könnte das zum einen die Motivation senken und zum anderen zu technischen Problemen führen; vgl. Ebd., S. 30.

<sup>111</sup> Zedlitz, Daten-Eingabe-System, S. 31.

<sup>112</sup> Vgl. Interview A.

Dennoch überwiegen für die interviewten Expert\*innen die Vorteile: Da die Eingabe online erfolgt, können die Administrator\*innen direkt regulierend eingreifen und die Erschließung im Blick zu behalten.<sup>113</sup> Auch Fortschritte sind leicht erkennbar und können beispielsweise an das beteiligte Archiv weitergemeldet werden. Wer Administrator\*in wird, bestimmt der Verein für Computergenealogie.<sup>114</sup> Dabei ist nicht von Bedeutung, wie viel die gewählten Personen in der Vergangenheit erschlossen haben, sondern das genaue Kennen der Erschließungsrichtlinien, gewissenhaftes Arbeiten und gute paläografische Fähigkeiten.<sup>115</sup> Betreut werden die einzelnen Projekte je nach Umfang und Komplexität von einem bis fünf Administrator\*innen,<sup>116</sup> wobei bei Kooperationen mit Archiven gerne Fachpersonal aus diesen als Administrator\*innen aufgenommen werden. Durch deren Fachexpertise ist ein besonderes Maß an Qualitätskontrolle gewährleistet.<sup>117</sup>

Die Rolle der Administrator\*innen beschränkt sich jedoch nicht auf das Anlegen und Begleiten eines Projekts, sondern geht darüber hinaus, indem sie in ein elaboriertes, mehrstufiges System zur Fehlerermittlung und -verbesserung eingebunden sind: Zunächst wird jede fertig erschlossene und indizierte Seite Korrektur gelesen und bei Bedarf verbessert.<sup>118</sup> Die Person, die erschlossen hat, kann auch direkt Leseunsicherheiten markieren, die dann den Administrator\*innen in einer Listenansicht als zu korrigierende Einträge angezeigt werden.<sup>119</sup> Darüber hinaus werden die Einträge systematisch nach Abweichungen von den Erschließungsrichtlinien durchsucht.<sup>120</sup> Erschließungen von Personen, die erstmalig an einem Projekt beteiligt sind, werden besonders genau überprüft.<sup>121</sup> Weiterhin können sowohl während der Eingabe als auch danach Fehler an die Administrator\*innen gemeldet werden.<sup>122</sup>

---

<sup>113</sup> Vgl. Interview A. Vgl. Dazu Zedlitz, Daten-Eingabe-System, S. 28.

<sup>114</sup> Vgl. Hirschbeck, Archive, S. 21.

<sup>115</sup> Vgl. Hirschbeck, Archive, S. 25.

<sup>116</sup> Vgl. Hirschbeck, Archive, S. 25.

<sup>117</sup> Vgl. Hirschbeck, Archive, S. 25f.

<sup>118</sup> Vgl. Hirschbeck, Archive, S. 20.

<sup>119</sup> Vgl. Moeller/ Müller, Heimatforscher, S. 107. Diese Funktion wurde in einem Interview als besonders positiv bewertet; vgl. Interview A.

<sup>120</sup> Vgl. Zedlitz, Daten-Eingabe-System, S. 31. Typische Fehler sind beispielsweise bei der Erschließung von Adressbüchern die Eintragung von „Witwe“ als Beruf; vgl. Ebd., S. 31.

<sup>121</sup> [https://docs.genealogy.net/des/de/15-16.html#typische\\_fehler](https://docs.genealogy.net/des/de/15-16.html#typische_fehler) (13.02.2024).

<sup>122</sup> Vgl. Hirschbeck, Archive, S. 21. Das DES wird nicht nur zum Erschließen, sondern auch zum Recherchieren verwendet. Nach der abgeschlossenen Erschließung werden vor allem Fehler gemeldet, die von Nutzer\*innen bemerkt wurden; vgl. Zedlitz, Daten-Eingabe-System, S. 30. Auch Verbesserungsvorschläge können eingereicht und von den Administrator\*innen überprüft werden; vgl. Hirschbeck, Archive, S. 24. Das Recherchieren ist eine Funktion, die das DES von anderen Anwendungen zur Erschließung abhebt und von Genealog\*innen oft verwendet wird; vgl. Interview B.

Dieses Vorgehen wurde auch von den interviewten Expert\*innen als besonders vorteilhaft hervorgehoben: Die mehrstufige Prüfung sichert die Qualität und die Möglichkeit, beim Recherchieren noch eventuell vorhandene Fehler zu melden, steigert die Qualität zusätzlich.<sup>123</sup> In normalen Findbüchern, wie sie die meisten Archive anbieten, gibt es selten derartige Funktionen. Man kann davon ausgehen, dass nur wenige Nutzer\*innen die Mühe auf sich nehmen, einzelne Fehler, die ihnen beim Recherchieren auffallen, an das Archiv zu melden. Eine integrierte Meldefunktion senkt die Hemmschwelle und verbessert langfristig die Qualität der Einträge.

Auch der Export in anderen Archivinformationssysteme ist dank csv-Export möglich. Gerade durch die strukturierte Angabe mittels vorgefertigter Felder auf dem Digitalisat ist sichergestellt, dass die Erschließung standardisiert erfolgt und die Daten ohne komplizierte Nacharbeiten unmittelbar nutzbar sind.<sup>124</sup> Ein Excel-Export ist ebenfalls eine Option; dafür sind jedoch im Vorfeld manuelle Anpassungen von Nöten, bevor die Daten beispielsweise in ActaPro übernommen werden, da die von DES erzeugten Excel-Tabellen nicht kompatibel mit den ActaPro-Eingabemasken sind.<sup>125</sup> Ein weiteres Vorteil ist die Kompatibilität mit Transkribus.<sup>126</sup>

### **3.3. Transkribus**

Die Anwendung „Transkribus“ entstand aus einem von der EU finanzierten Projekt und wird inzwischen von der Europäischen Genossenschaft READ-COOP SCE betrieben.<sup>127</sup> Diese Genossenschaft wurde 2019 gegründet und besitzt über hundert Mitglieder aus Europa, den USA und Kanada, darunter neben zahlreichen Universitäten und Archiven auch den Verein für Computergenealogie.<sup>128</sup> Aktuell sind über 100.000 Benutzer\*innen registriert, die über 40 Millionen Dateien hochgeladen und über 15.000 Modelle zur Erkennung von Handschriften trainiert haben.<sup>129</sup> Die dabei entstandenen Trainingsdaten stellen weltweit das umfangreichste Datenset zur automatischen Erkennung historischer Handschriften dar. Die Anwendung erkennt inzwischen neben Buchstaben und Wörtern auch Strukturen wie beispielsweise Tabellen<sup>130</sup> –

---

<sup>123</sup> Vgl. Interview A.

<sup>124</sup> Vgl. Moeller/ Müller, Heimatforscher, S. 101.

<sup>125</sup> Vgl. Gränitz, Schwarmwissen, S. 41; vgl. Interview A.

<sup>126</sup> Vgl. Interview B.

<sup>127</sup> Vgl. Nockels/ Gooding/ Ames/ Terras: Understanding, S. 368.

<sup>128</sup> Mühlberger/ Siegl/ Scharr, Crowdsourcing, S. 226.

<sup>129</sup> <https://readcoop.eu/de/transkribus/> (26.02.2024).

<sup>130</sup> Vgl. Mühlberger/ Siegl/ Scharr, Crowdsourcing, S. 226.

unverzichtbar für viele serielle Quellen wie beispielsweise Sterberegister.<sup>131</sup> Während Transkribus in den Anfängen eine Anwendung war, die auf den Computer heruntergeladen werden musste, kann sie inzwischen auch in einer etwas reduzierten Version im Webbrowser verwendet werden.<sup>132</sup>

Die Anwendung ist teilweise kostenpflichtig, dabei aber sehr transparent: Einzelpersonen können 100 handgeschriebene oder 600 gedruckte Seiten pro Monat kostenlos erkennen lassen, während Institutionen und Personen mit Bezug zu Schulen oder Universitäten eine monatliche Gebühr zahlen müssen. Die Preise sind abhängig von der Anzahl der Seiten, die von Transkribus verarbeitet werden sollen. Je mehr Seiten, desto teurer.<sup>133</sup>

Im Rahmen von Citizen-Science-Projekten wird Transkribus inzwischen regelmäßig eingesetzt.<sup>134</sup> Der Grundgedanke ist, dass Freiwillige nicht mehr Archivalien abtippen, sondern diese in Transkribus einspeisen und lediglich die von der Anwendung erstellten Daten überprüfen und verbessern. Durch Exportmöglichkeiten und Schnittstellen zu anderen Anwendungen wie beispielsweise DES können die Daten weiterverarbeitet werden<sup>135</sup> und mit weiterführenden Informationen oder Normdaten verknüpfen werden.

Um die Eignung der Anwendung für Citizen-Science-Projekte zu untersuchen, wurden im Rahmen dieser Arbeit mehrere Seiten aus dem Verzeichnis StadtAL 0008 Nr. 3190 bei Transkribus hochgeladen. Die handschriftlichen Einträge auf diesen Seiten wurden mittels der automatischen Texterkennung transkribiert und mit jenen Erschließungen abgeglichen, die in Excel erschlossen wurden. Es wurde aus Zeitgründen davon abgesehen, ein eigenes Modell zu trainieren und stattdessen mit Modellen gearbeitet, die Transkribus kostenfrei zur Verfügung stellt. Es zeigte sich, dass bei den Mieter\*innenverzeichnissen die Überschriften in Fraktur fehlerfrei erkannt wurden, es allerdings bei der Erkennung der handschriftlichen Einträge zu zahlreichen Fehlern kam. Gängige Wörter wie „Frau“ erkannte die Anwendung ohne Probleme und auch bei Zahlen sowie besonders groß und deutlich geschriebenen Wörtern (meist Nachnamen) war die Erfolgsquote gut. Auch Strukturelemente wurden meistens erkannt und die Inhalte den richtigen Tabellenzeilen oder -spalten zugeordnet.

---

<sup>131</sup> Vgl. Interview B.

<sup>132</sup> Mühlberger/ Siegl/ Scharr, Crowdsourcing, S. 226.

<sup>133</sup> <https://www.transkribus.org/plans> (26.02.2024).

<sup>134</sup> Für ein Beispiel vgl. Milioni, Transcription.

<sup>135</sup> <https://www.compgen.de/2024/02/compgen-vortrag-zu-des-projekten-mit-transkribus-am-8-2-2024> (26.02.2024).

Die Fehlerquote war hingegen bei zahlreichen anderen Wort- und Eintragstypen sehr hoch, vor allem bei Namen, Abkürzungen und Orten, was besonders bei seriellen Quellen ein Problem darstellt, da diese in erster Linie aus den zuvor genannten Elementen bestehen. Oft wurden zwar nur einzelne Buchstaben falsch erkannt, aber gerade bei Eigennamen ist eine fehlerfreie Erkennung zentral, da sonst Personen und Orte falsch benannt werden und folglich nicht korrekt identifiziert oder verknüpft werden können. Große Schwierigkeiten traten auch bei dem Erkennen von Worten auf, die nicht mit dunkler Tinte geschrieben wurden. Sowohl Bleistift als auch rote Tinte stellten die Anwendung vor Schwierigkeiten und sorgten zum Teil für erhebliche Fehler.

Zu bedenken ist zwar, dass es sich um ein nicht speziell für diese Handschrift trainiertes Modell handelt, aber andere Studien kommen zu einem ähnlichen Ergebnis: In vergleichbaren Projekten lieferte die automatische Erkennung selbst mit eigens trainierten Modellen mittels Transkribus oft die schlechtesten Ergebnisse.<sup>136</sup>

In den Interviews mit Expert\*innen wurde jedoch deutlich, dass Transkribus trotzdem von diesen als eine geeignete Anwendung für Citizen-Science-Projekte angesehen wird – sofern sie mit einer weiteren wie beispielsweise DES oder irgendeiner Form der Qualitätskontrolle kombiniert wird. Der Grund dafür liegt in der Geschwindigkeit, mit der die KI transkribieren kann. Selbst wenn sie zahlreiche Fehler produziert, ist es oft (vor allem im Vergleich zu Freiwilligen, die alte Handschriften nur langsam lesen) deutlich schneller, mehrere Seiten von Transkribus erschließen zu lassen und anschließend zu überprüfen sowie zu korrigieren.<sup>137</sup> Auch die Möglichkeit, Randnotizen, Spalten und andere Strukturelemente automatisch erkennen zu lassen, spart deutlich Zeit.<sup>138</sup> Ebenfalls zeitsparend ist die Funktion, eigene Skripte zu schreiben.<sup>139</sup>

Um die Fehlerquote zu senken, lohnt es sich laut einer der Interviewten, Zeit in die Suche nach einem passenden Modell zu investieren. Da inzwischen eine Vielzahl existiert (nicht nur für Handschriften in deutscher Sprache, sondern auch für englische,

---

<sup>136</sup> Als Beispiel kann etwa das Projekt „Hallische Heiratsgeschichten“ angeführt werden, das ebenfalls in Kooperation mit dem Verein für Computergenealogie umgesetzt wurde. Ermittelt wurde eine Fehlerquote von 5%, die sich vor allem auf Orts-/Eigennamen sowie Berufsbezeichnungen bezog und als deutlich zu hoch angesehen wurde, da gerade Eigennamen fehlerfrei erkannt werden müssen; vgl. Moeller/Müller, *Heimatsforscher*, S. 106.

<sup>137</sup> Vgl. Interview B; vgl. Interview C.

<sup>138</sup> Vgl. Interview B.

<sup>139</sup> Vgl. Interview B. In Adressbüchern wird meist ein Nachname aufgeführt und darunter die Liste aller Familienangehörigen. Mittels eines Skripts kann für die Erschließung jeder Vorname automatisch dem passenden Nachnamen zugewiesen werden, sodass man nicht etwa „Müller, Anna Marie Ludwig“, sondern „Müller, Anna; Müller, Marie; Müller, Ludwig“ als Datensatz erhält.

holländische, französische usw.), müssen nur in Ausnahmefällen eigene Modelle trainiert werden.<sup>140</sup>

Zudem gibt es hilfreiche Zusatzfunktionen wie jene, Orte als solche zu kennzeichnen und zu georeferenzieren, was besonders für die weitere Arbeit mit den Daten hilfreich sein kann. Ein neues Modul ermöglicht es, sich das Digitalisat und die Erschließung platzsparend und übersichtlich nebeneinander anzeigen zu lassen, wodurch ein guter Abgleich möglich ist.<sup>141</sup> Ähnlich wie bei DES sinkt somit das Risiko, einzelne Zeilen zu überspringen.

In den Interviews wurden allerdings auch mehrere Nachteile angesprochen: Zum einen die regelmäßigen Änderungen – sowohl, was die Oberfläche der Anwendung als auch ihre Funktionen und vor allem die Bezahlmodelle betrifft. Hier müssen Freiwillige sich regelmäßig neu einarbeiten. Weiterhin kann Transkribus schnell teuer werden. Wenn ein Freiwilliger ungeduldig ist und etwa mehrfach auf das Feld zum Transkribieren klickt, verarbeitet das die Anwendung als mehrfache Anfrage und zieht entsprechend „Credits“ (die digitale Währung) ab.<sup>142</sup> Zudem ist die Qualität der Vorlage entscheidend: scheint etwa die Rückseite einer Seite durch, liefert Transkribus schlechte Ergebnisse.<sup>143</sup>

Neben diesen Punkten, die die grundlegende Qualität der Erschließung betreffen, muss auch die anschließende Qualitätskontrolle und -sicherung beachtet werden. In dem Projekt, für das eine der Interviewten aktuell erschließt, gibt es keine Kontrolle durch eine zweite Person. In vorherigen Projekten wurden hingegen nach dem Vier-Augen-Prinzip gegenseitig Korrektur gelesen.<sup>144</sup> Festzuhalten ist in jedem Fall, dass es bei Transkribus – ähnlich wie bei Excel und im Gegensatz zu DES – keine implementierte Funktion zur Qualitätskontrolle gibt. Auch das „Mitlesen“ ist nicht direkt in der Anwendung möglich, da die jeweiligen Erschließungen an eine Person gebunden sind. Um die Erschließungsdaten zu teilen, müssen sie erst exportiert werden. Das ist zwar durch Schnittstellen und die Möglichkeit eines xml-Exports oder eines Exports in Excel und darüber im csv-Format problemlos möglich,<sup>145</sup> verzögert aber den Prozess

---

<sup>140</sup> Vgl. Interview B. Perspektivisch sollen Modelle für Notenblätter trainiert werden – auch an solche Sonderformate wird gedacht; vgl. Ebd.

<sup>141</sup> Vgl. Interview B; vgl. Interview C.

<sup>142</sup> Vgl. Interview B; vgl. Interview C.

<sup>143</sup> Vgl. Interview C.

<sup>144</sup> Vgl. Interview B.

<sup>145</sup> Vgl. <https://help.transkribus.org/de/dokumente-und-bilder-herunterladen> (27.02.2024); vgl. Interview B; vgl. Interview C.

des Korrekturlesens im Vergleich zu einer direkten Kontrolle der Daten in der jeweiligen Anwendung.

Die Interviewten betonten vor allem die Notwendigkeit einer Schnittstelle zu DES, um mit den Daten im Rahmen von Citizen-Science-Projekten sinnvoll weiterarbeiten zu können. In einer früheren Version von Transkribus existierte diese Funktion, in der Browser-Version der Anwendung hingegen (noch) nicht.<sup>146</sup> Die alleinige Verwendung von Transkribus wurde von den Expert\*innen als wenig geeignet für Citizen-Science-Projekte beurteilt.<sup>147</sup> Dieser Einschätzung ist zuzustimmen, da eine einfache Möglichkeit der Qualitätskontrolle fehlt – schon alleine, da die Erschließungsdaten personen-gebunden sind und nicht innerhalb der Anwendung geteilt oder gemeinsam bearbeitet werden können. Die Kombination von Transkribus mit DES wäre eine gute Möglichkeit, um die Qualitätssicherung und -kontrolle verbessern beziehungsweise überhaupt sicherstellen zu können.

#### **4. Fazit**

Die vorliegende Arbeit kombinierte eine praktische Erschließung in Kooperation mit dem Verein für Computergenealogie sowie dem Stadtarchiv Leipzig mit Expert\*inneninterviews, um die Frage zu beantworten, ob die im Rahmen von Citizen-Science-Projekten gewonnenen Erschließungsdaten qualitativ hochwertig sind und von Archiven übernommen werden können. Es wurde deutlich, dass die Qualität stark von der verwendeten Anwendung sowie der darin verankerten Qualitätskontrolle abhängt. Eine KI wie Transkribus liefert zwar schnelle Ergebnisse, diese sind jedoch oft fehlerhaft. Da die Erschließungsdaten innerhalb der Anwendung personen-gebunden sind, ist eine Teilung und somit gegenseitige Kontrolle nach dem Vier-Augen-Prinzip nicht möglich. Eine Qualitätskontrolle durch Administrator\*innen oder ähnliches ist ebenfalls nicht implementiert. Während die praktische Erschließung einiger Seiten aus StadtAL 0008 Nr. 3190 die hohe Fehlerquote unterstrich, wurde die Anwendung hingegen in den Expert\*inneninterviews positiv hervorgehoben – allerdings nur, wenn sie in Kombination mit einer anderen Anwendung wie dem DES verwendet wird. Hier eröffnen Schnittstellen die Möglichkeit, die Geschwindigkeit von Transkribus mit der elaborierten Qualitätssicherung und -kontrolle von DES, der Eigenentwicklung des Vereins für Computergenealogie, zu kombinieren.

---

<sup>146</sup> Vgl. Interview B; vgl. Interview C.

<sup>147</sup> Vgl. Interview B; vgl. Interview C.

Wie dargelegt wurde, werden die Erschließungsdaten, die mit DES erhoben werden, mehrfach von Administrator\*innen überprüft und bei Bedarf korrigiert und auch Nutzer\*innen können auf Fehler aufmerksam machen. Diese Anwendung ist für alle Projekte in Kooperation mit dem Verein für Computergenealogie geeignet und erzeugt Erschließungsdaten, die aufgrund ihrer hohen Qualität ohne Bedenken von Archiven übernommen werden können. Auch die paläografischen Kenntnisse der meisten Genealog\*innen sind positiv zu bewerten und sichern die Qualität der Erschließungen. Exportmöglichkeiten in Archivinformationssysteme wie ActaPro erleichtern zudem die Übernahme der Daten.

Für Projekte, die weniger interessant für Genealog\*innen sind oder nicht auf das DES zurückgreifen wollen, eignet sich auf Microsoft Excel in der online-Version gut. Wie dargelegt wurde, ist von der offline-Version eher abzuraten. In der online-Version kann hingegen eine gegenseitige Qualitätskontrolle erfolgen – wichtig sind allerdings zum einen gute Absprachen und zum anderen Freiwillige, die über sehr gute paläografische Fähigkeiten sowie Kenntnisse über den jeweiligen Archivalientyp verfügen. Dies ist von Bedeutung, da eine übergeordnete Kontrolle durch Administrator\*innen fehlt. Im praktischen Erschließungsprojekt zeigte sich, dass die Qualität der Erschließungsdaten sehr gut ist, wenn die zuvor genannten Punkte beachtet werden. Auch diese Daten können bedenkenlos von einem Archiv übernommen werden. Stichpunktartige Kontrollen sind jedoch vor der Übernahme der Daten in jedem Fall sinnvoll.

Zusammengefasst lässt sich sagen, dass Citizen-Science-Projekte bei der Wahl passender Anwendungen gut geeignet sind, qualitativ hochwertige Erschließungsdaten zu erzeugen, die von Archiven übernommen werden können. Zudem können sie mit verschiedensten Projekten verknüpft werden und so auch das Interesse anderer Bürger\*innen an Geschichte und dem Archiv wecken. Anzuführen ist beispielsweise das Projekt „Altes Leipzig“ oder die Pläne rund um das Holzmodell der Stadt von 1823. Die ehrenamtliche Arbeit von Freiwilligen sollte nicht als Gefahr für die Kernkompetenzen eines Archivs gesehen werden, sondern als wertvolle Ergänzung. Von Citizen-Science-Projekten können sowohl die Freiwilligen als auch beteiligte Vereine und Partnerinstitutionen wie Museen und Archive profitieren. Derartige Projekte können neben Erschließungsdaten auch mehr Kooperationen zwischen verschiedenen Fächern, Einrichtungen und Institutionen sowie eine größere Bürger\*innennähe hervorbringen.

## **5. Zusammenfassung**

Im Rahmen dieser Arbeit wurde untersucht, ob Erschließungsdaten, die bei Citizen-Science-Projekten gewonnen werden, qualitativen Anforderungen genügen und von Archiven ohne Bedenken übernommen werden können. Dafür wurden mittels semi-strukturierter Expert\*inneninterviews und einem eigenen Erschließungsprojekt in Kooperation mit dem Verein für Computergenealogie sowie dem Stadtarchiv Leipzig drei Anwendungen untersucht, die aktuell bei derartigen Projekten häufig zum Einsatz kommen: Microsoft Excel, das Daten-Eingabe-System und Transkribus.

Es wurde deutlich, dass sich Anwendungen, bei denen nicht gemeinsam an einem Dokument gearbeitet werden kann bzw. die keine implementierte Qualitätskontrolle aufweisen, nur eingeschränkt für Citizen-Science-Projekte eignen, bei denen die generierten Erschließungsdaten von Archiven übernommen werden sollen. Anwendungen mit diesen Funktionen erzeugen hingegen bei einer durchdachten Projektplanung Erschließungsdaten von einer hohen Qualität, die ohne Bedenken von Archiven übernommen werden können.

## **6. Quellenverzeichnis**

StadtAL 0008 Nr. 3190.

## **7. Literaturverzeichnis**

Altes Leipzig. Das ALTE LEIPZIG, unter: <https://www.altes-leipzig.de/pano/1894/A1/A1.html> (12.01.2024).

Altes Leipzig. Virtuelles Leipzig, unter: <https://www.altes-leipzig.de/virtuelles-leipzig> (12.01.2024).

Ansicht in Matrixanzeige, unter: [https://www.altes-leipzig.de/media/galerie\\_detail/2020\\_struktur03.jpg](https://www.altes-leipzig.de/media/galerie_detail/2020_struktur03.jpg) (12.01.2024).

Balázs, Bálint u.a.: Data Quality in Citizen Science, in: *The Science of Citizen Science*, hg. v. Katrin Vohland u.a., Cham 2021, S. 139–157.

Beachy, Robert: *The Soul of Commerce. Credit, Property, and Politics in Leipzig, 1750–1840* (Studies in Central European History 34), Leiden/ Boston 2005.

Becker, Denny: Citizen Science in Archiven. Möglichkeiten und Grenzen von Crowdsourcing bei der archivischen Erschließung von Fotografien, in: *ABI Technik* 40/1 (2020), S. 30–39.

- Bonn, Aletta u.a.: Grünbuch Citizen Science Strategie 2020 für Deutschland, Berlin 2016.
- Bunge, Eva: Wie viel Naturwissenschaft braucht die Bibliothek? Scientific Literacy und Citizen Science in Öffentlichen Bibliotheken, in: Öffentliche Bibliothek 2030. Herausforderungen – Konzepte – Visionen, hg. v. Petra Hauke, Bad Honnef 2019, S. 241–250.
- Bürger schaffen Wissen. Altes Leipzig, unter: <https://www.buergerschaffenwissen.de/projekt/altes-leipzig> (12.01.2024).
- Crowd-sour-cing, das, unter: <https://www.duden.de/rechtschreibung/Crowdsourcing> (10.01.2024).
- Dokumente und Bilder herunterladen, unter: <https://help.transkribus.org/de/dokumente-und-bilder-herunterladen> (27.02.2024).
- Eibl, Sabine: Freiwilliges Engagement in Archiven. Ergebnisse einer Transferarbeit, in: Der Archivar 72/1 (2019), S. 8–12.
- Esborg, Line: Engaging disenfranchised publics through citizen humanities projects, in: A History of Participation in Museums and Archives. Traversing Citizen Science and Citizen Humanities, hg. v. Per Hetland/ Palmyre Pierroux/ Line Esborg, London/ New York 2020, S. 109–125.
- FactGrid:Leipzig, unter: <https://database.factgrid.de/wiki/FactGrid:Leipzig> (16.02.2024).
- Fink, Kira: Citizen Science 2.0. Zukunftsstadt 2030+ Gelsenkirchen, Dortmund 2020.
- Finke, Peter: Citizen Science und die Rolle der Geisteswissenschaften für die Zukunft der Wissenschaftsdebatte, in: Bürger Künste Wissenschaft. Citizen Science in Kultur und Geisteswissenschaften, hg. v. Kristin Oswald/ René Smolarski, Gutenberg 2016, S. 31–56.
- Google Sheets. Datengestützte Entscheidungen treffen, unter: <https://www.google.de/intl/de/sheets/about/#features> (24.02.2024).
- Gränitz, Alina: Schwarmwissen. Die Einsatzmöglichkeiten von Crowdsourcing am Beispiel des Stadtarchivs Herne, Potsdam 2022.
- Haberer, Stephanie/ Rügge, Nicola: Grunderschliessung als Ersterfassung – und weiter? Überlegungen zu einer angemessenen Erschließungstiefe, in: Der Archivar 73/2 (2020), S. 129–133.

- Haklay, Mordechai (Muki) u.a.: What Is Citizen Science? The Challenges of Definition, in: *The Science of Citizen Science*, hg. v. Katrin Vohland u.a., Cham 2021, S. 13–33.
- Hirsch, Volker/ Linde, Roland: Juden in Westfalen und Lippe. Bilanz eines Crowdsourcing-Projekts zur Tiefenerschließung von Personenstandsregistern des 19. Jahrhunderts, in: *Zwischen populären Praktiken und akademischer Forschung. Jahrbuch für Geschichte des ländlichen Raumes* 18 (2021), S. 274–280.
- Hirschbeck, Georg: *Archive und Genealogen – Zusammenarbeit und Nutzen durch Crowdsourcing?*, Potsdam 2021.
- Historisches 3D-Stadtmodell Leipzig, unter: <https://www.leipzig.de/bauen-und-wohnen/bauen/geodaten-und-karten/historisches-3d-stadtmodell-leipzig> (27.02.2024).
- Huber, Christian J./ Kansy, Lambert/ Lüpold, Martin: Crowdsourcing in Archiven. Ein Werkstattbericht, in: *Der Archivar* 73/02 (2020), S. 145–149.
- Jaeger-Erben, Melanie: Citizen Science, in: *Handbuch Transdisziplinäre Didaktik*, hg. v. Tobias Schmohl/ Thorsten Philipp/ Johanna Schabert, Bielefeld 2021, S. 45–55.
- Joergens, Bettina: Erschließung – eine Einleitung zum Themenschwerpunkt, in: *Der Archivar* 73/2 (2020), S. 126–128.
- Junkers, Günther: Online-Treff am 12. Oktober: „Jeder kann mitmachen bei CompGen: Helft mit, die Daten einzugeben“, unter: <https://www.compgen.de/2023/10/online-treff-am-12-oktober-jeder-kann-mitmachen-bei-compgen-helft-mit-die-daten-einzugeben/> (14.02.2024).
- Kartei Leipziger Familien, unter: [https://wiki.genealogy.net/Kartei\\_Leipziger\\_Familien#Editionsrichtlinien](https://wiki.genealogy.net/Kartei_Leipziger_Familien#Editionsrichtlinien) (14.02.2024).
- Kidd, Jenny: Infrastructures that democratize? Citizen participation and digital ethics, in: *A History of Participation in Museums and Archives. Traversing Citizen Science and Citizen Humanities*, hg. v. Per Hetland/ Palmyre Pierroux/ Line Esborg, London/ New York 2020, S. 73–89.
- Kluttig, Thekla: Die Citizen Science Strategie 2020 für Deutschland und die Archive, in: *Kompetent! – Archive in der Wissensgesellschaft*. 86. Deutscher Archivtag in Koblenz (Tagungsdokumentation zum Deutschen Archivtag 21), Fulda 2018, S. 33–41.
- Land-Zandstra, Anne/ Agnello, Gaia/ Gültekin, Yaşar Selman, Participants in Citizen Science, in: *The Science of Citizen Science*, hg. v. Katrin Vohland u.a., Cham 2021, S. 243–259.

- Lemaire, Marina/ Rommelfanger, Yvonne: Berücksichtigung von Data-Literacy-Kompetenzen bei der Planung und Durchführung von Crowdsourcing-Projekten mit Bürgerwissenschaftlerinnen und -wissenschaftlern, in: Citizen Science in den Geschichtswissenschaften. Methodische Perspektive oder perspektivlose Methode?, hg. v. René Smolarski/ Hendrikje Carius/ Martin Prell, Göttingen 2023, S. 69–90.
- Lippik, Marlis: Ohne Sie geht gar nichts – wenn Ehrenamtliche reguläre Mitarbeiter ersetzen müssen. Praxisbericht aus einem kleinen Stadtarchiv, in: Der Archivar 72/1 (2019), S. 13–17.
- Martin-Luther-Universität Halle-Wittenberg, Stud.IP, unter: [https://studip.uni-halle.de/dispatch.php/course/details?sem\\_id=15442f4cd15b1f054a2cddb495cf2556&again=yes](https://studip.uni-halle.de/dispatch.php/course/details?sem_id=15442f4cd15b1f054a2cddb495cf2556&again=yes) (16.02.2024).
- Microsoft 365. Microsoft Excel, unter: <https://www.microsoft.com/de-de/microsoft-365/excel?market=de> (23.02.2024).
- Microsoft Unterstützung. Exportieren von XML-Daten, unter: <https://support.microsoft.com/de-de/office/exportieren-von-xml-daten-0b21f51b-56d6-48f0-83d9-a89637cd4360> (23.02.2024).
- Microsoft Unterstützung. Importieren oder Exportieren von Textdateien (TXT oder CSV), unter: <https://support.microsoft.com/de-de/office/importieren-oder-exportieren-von-textdateien-txt-oder-csv-5250ac4c-663c-47ce-937b-339e391393ba> (23.02.2024).
- Milioni, Nikolina: Automatic Transcription of Historical Documents. Transkribus as a Tool for Libraries, Archives and Scholars, Uppsala 2020.
- Misoch, Sabina: Qualitative Interviews, Berlin/ Boston <sup>2</sup>2019.
- Moeller, Katrin/ Müller, Moritz: Heimatforscher, Citizen Science und/oder Digital History? Organisationsformen und Qualitätssicherung zwischen Wissenschaft und bürgerwissenschaftlicher Community, in: Citizen Science in den Geschichtswissenschaften. Methodische Perspektive oder perspektivlose Methode?, hg. v. René Smolarski/ Hendrikje Carius/ Martin Prell, Göttingen 2023, S. 91–108.
- Nicola, Susanne: Historische Quellen gemeinsam erschliessen. Daten-Eingabe-System, in: Familienforschung (2019/2020), S. 68–69.
- Nockels, Joe/ Gooding, Paul/ Ames, Sarah/ Terras, Melissa: Understanding the application of handwritten text recognition technology in heritage contexts.

- A systematic review of Transkribus in published research, in: *Archival Science* (2022), S. 367–392.
- Nr. 3190 – Hauptbuch über die Beiträge zum Stadtschuldentilgungsfonds 1842-1853. Vollansicht Verzeichnungseinheit Nr. 3190, unter: [https://recherche-stadtarchiv.leipzig.de/actaproweb/document/Vz\\_5c85ade0-19da-43ac-a9fa-a8fcb8cd1749](https://recherche-stadtarchiv.leipzig.de/actaproweb/document/Vz_5c85ade0-19da-43ac-a9fa-a8fcb8cd1749) (28.02.2024).
- Oswald, Kristin/ Smolarski, René: Einführung. Citizen Science in Kultur und Geisteswissenschaften, in: *Bürger Künste Wissenschaft. Citizen Science in Kultur und Geisteswissenschaften*, hg. v. Kristin Oswald/ René Smolarski, Gutenberg 2016, S. 9–27.
- Patt, Gregor: *Crowdsourcing bei Urkunden und Briefen. Kooperative Erschließung von Altbeständen in der digitalen Welt*, Marburg 2015.
- Pettibone, Lisa u.a.: *Citizen Science für alle. Eine Handreichung für Citizen Science-Beteiligte*, Berlin 2016.
- Pettibone, Lisa/ Ziegler, David: Citizen Science. Bürgerforschung in den Geistes- und Kulturwissenschaften, in: *Bürger Künste Wissenschaft. Citizen Science in Kultur und Geisteswissenschaften*, hg. v. Kristin Oswald/ René Smolarski, Gutenberg 2016, S. 57–69.
- Phillips, Tina/ Porticella, Norman/ Conostas, Mark/ Bonney, Rick: A Framework for Articulating and Measuring Individual Learning Outcomes from Participation in Citizen Science, in: *Citizen Science: Theory and Practice* 3/2 (2018), S. 1–19.
- Pierroux, Palmyre/ Hetland, Per/ Esborg, Line: Traversing citizen science and citizen humanities. Tacking stitches, in: *A History of Participation in Museums and Archives. Traversing Citizen Science and Citizen Humanities*, hg. v. Per Hetland/ Palmyre Pierroux/ Line Esborg, London/ New York 2020, S. 3–23.
- Reinsone, Sanita: Searching for deeper meanings in cultural heritage crowdsourcing, in: *A History of Participation in Museums and Archives. Traversing Citizen Science and Citizen Humanities*, hg. v. Per Hetland/ Palmyre Pierroux/ Line Esborg, London/ New York 2020, S. 186–207.
- Rose, Michael: CompGen-Vortrag zu DES-Projekten mit Transkribus am 8.2.2024, unter: <https://www.compgen.de/2024/02/compgen-vortrag-zu-des-projekten-mit-transkribus-am-8-2-2024/> (14.02.2024).
- Simons, Olaf: Keine Selbstverständlichkeit: Citizen Science auf der FactGrid Wikibase-Plattform, in: *Citizen Science in den Geschichtswissenschaften*.

- Methodische Perspektive oder perspektivlose Methode?, hg. v. René Smolarski/ Hendrikje Carius/ Martin Prell, Göttingen 2023, S. 242–265.
- Smolarski, René/ Carius, Hendrikje/ Prell, Martin: Citizen Science in den Geschichtswissenschaften aus methodischer Perspektive. Zur Einführung, in: Citizen Science in den Geschichtswissenschaften. Methodische Perspektive oder perspektivlose Methode?, hg. v. René Smolarski/ Hendrikje Carius/ Martin Prell, Göttingen 2023, S. 7–21.
- Smolarski, René: Der Gothaer Missionskartograph. Ein historisches Beispiel für „Crowdsourcing“ und „Citizen Science“ im 19. Jahrhundert, in: Bürger Künste Wissenschaft. Citizen Science in Kultur und Geisteswissenschaften, hg. v. Kristin Oswald/ René Smolarski, Gutenberg 2016, S. 71–87.
- Speit, Sina: Public History und historische Grundlagenforschung. Das Projekt „Die Geschichte der Landesministerien in Baden und Württemberg in der Zeit des Nationalsozialismus“, in: Bürger Künste Wissenschaft. Citizen Science in Kultur und Geisteswissenschaften, hg. v. Kristin Oswald/ René Smolarski, Gutenberg 2016, S. 119–137.
- Stadt Leipzig, Stadtarchiv, Tektonik, unter: <https://recherche-stadtarchiv.leipzig.de/actaproweb/archive.xhtml> (11.02.2024).
- Theimer, Kate: Partizipation als Zukunft der Archive, in: Der Archivar 71/1 (2018), S. 6–12.
- Transkribus. Historische Dokumente mit KI entschlüsseln, unter: <https://readcoop.eu/de/transkribus> (26.02.2024).
- Verein für Computergenealogie. Projektauswahl, unter: <http://des.genealogy.net/start/selectProject> (12.02.2024).
- Welche Erfassungsfehler werden zu Anfang häufig gemacht?, unter: [https://docs.genealogy.net/des/de/15-16.html#typische\\_fehler](https://docs.genealogy.net/des/de/15-16.html#typische_fehler) (13.02.2024).
- Walter, Sarah: Stormarn schafft Wissen. Erschließung der Bildsammlungen des Kreisarchivs Stormarn durch interessierte Bürger\*innen, in: Der Archivar 73/02 (2020), S. 150–154.
- Was ist das DES, unter: <https://docs.genealogy.net/des/de/> (13.02.2024).
- Zedlitz, Jesper: Das Daten-Eingabe-System (DES), in: Computergenealogie 4 (2013), S. 28–32.
- Zedlitz, Jesper: DES, unter: [https://wiki.genealogy.net/DES#Publikationen.2FPr.C3.A4sentationen\\_zum\\_DES](https://wiki.genealogy.net/DES#Publikationen.2FPr.C3.A4sentationen_zum_DES) (14.02.2024).

Irwin, Alan: Citizen Science. A study of people, expertise and sustainable development 1995.

## **8. Anhang: Interviewfragen**

### **Fragen zur Person & zum Projekt**

Wie lange erschließen Sie bereits Archivalien im Rahmen von Citizen-Science-Projekten?

Wie lange erschließen Sie bereits Archivalien mit der Anwendung, die Sie aktuell benutzen?

Für welches Projekt erschließen Sie aktuell Archivalien?

Seit wann sind Sie in dieses Projekt involviert und wie sind Sie dazu gekommen?

Wie sind Sie in das Projekt eingebunden bzw. welche Rollen füllen Sie aus?

### **Fragen zur Anwendung**

Welche Anwendung nutzen Sie für die Erschließung der Archivalien?

Welche Vorteile bietet die Anwendung aus Ihrer Sicht?

Welche Nachteile bietet die Anwendung aus Ihrer Sicht?

Haben Sie in der Vergangenheit andere Anwendungen genutzt? Wenn ja, weshalb sind Sie auf die aktuelle Anwendung umgestiegen?

Wo sehen Sie Verbesserungsbedarf bei der von Ihnen genutzten Anwendung?

Wie sieht die Qualitätssicherung bei Ihren Erschließungen aus?

Wie erfolgt der Datenimport und -export in/aus der Anwendung?

Hatte die verwendete Plattform/Anwendung Einfluss auf Ihre Entscheidung, sich in dem Projekt zu engagieren?