

ARCHIV

theorie & praxis

KÜNSTLICHE INTELLIGENZ IM ARCHIV

Einfluss von Bildformaten auf
die KI-basierte Analyse von
digitalisierten Dokumentbeständen

Alexa trifft Archiv. Der aktuelle Stand
des Einsatzes von KI in Archiven

Letting AI loose in an archive.
Technology to manage or to manage
with

DATENSCHÄTZE

ACTAPRO[•]

*Die Datenmanagement-Software
für **Archive**.*



Digitale Archivierung

- Eigenes OAIS-Modul
- Schnittstellen zu DA-NRW und DIMAG
- DMS-Anbindung

Standards, Schnittstellen

- ISDIAH, ISAD(G), EAD, EAD DDB, METS
- Offene Schnittstellen
- Vollständige XML-Datenhaltung und Unicode-Unterstützung

Service

- Begleitende Beratung
- Individuelle Anpassungen
- Konsequente Pflege und Weiterentwicklung



Kennenlern-Produktdemos:

Per QR-Code-Scan oder Linkeingabe gelangen Sie zu den Terminen.

produkt demos.startext.de



IM GRIFF

Archivierung digitaler Unterlagen

*Antworten auf die
digitale Herausforderung.*

Übernehmen

- Automatische Übernahme digitaler Massenakten
- Bewertung und Übernahme von Dateisammlungen
- Webseitenarchivierung

Bewahren

- Erkennung und Validierung von Dateiformaten
- Automatische Umwandlung in Archivformate
- Standardkonforme Dokumentation aller Bearbeitungsschritte

Nutzen

- Automatische Erstellung von Nutzungskopien
- Integration von Wasserzeichen
- Recherche in Metadaten und Volltextsuche in den Dateien

startext GmbH
Riemenschneiderstr. 11 • 53175 Bonn
Telefon: +49 228 959 96-0
info@startext.de • www.startext.de

startext

Datenschätze im Griff

INHALT

EDITORIAL	5
KÜNSTLICHE INTELLIGENZ IM ARCHIV	
Bettina Joergens/Tobias Krafft: Künstliche und andere Intelligenz im Archiv.	
Annäherungen zwischen Methoden und Praxis in Archiven einerseits und KI-Technologien	6
Isto Huvila: Letting AI loose in an archive. Technology to manage or to manage with	12
Elena Williams: Alexa trifft Archiv. Der aktuelle Stand des Einsatzes von KI in Archiven	16
Luciana Duranti: I Trust AI. The Fifth Phase of the InterPARES Project	20
Martin Vogel: Künstliche Intelligenz im Archiv. Herausforderungen und Chancen	24
Oliver Tüselmann/Fabian Wolf/Tim Raven/Gernot A. Fink: Einfluss von Bildformaten auf die KI-basierte Analyse von digitalisierten Dokumentbeständen	29
Andreas Neuburger: Die Zähmung des Chaos. Perspektiven zur KI-gestützten Erschließung von Entschädigungsakten im Landesarchiv Baden-Württemberg	34
Florian Probst/Thomas Reich: Das Digitalisat als Datenquelle. Ein Workflow für die Nutzung bereits digitalisierter Archivbestände	38
Ruth von dem Bussche/Jasmin Hartmann: Die KI-basierte Erschließung von Archivbeständen für die Provenienzforschung in NRW	43
Miriam Eberlein: Bild-Erschließung mit KI. Sechs Vorschläge für die Nutzung KI-generierter Metadaten	49
INNOVATION UND ERFAHRUNG	54
Standardbasierte Aussonderung elektronischer Gerichtsverfahrensakten über das digitale Zwischenarchiv des Bundes (L. Bannert/I. Müller/N. Trossen) • Die Bergungserfassung im Historischen Archiv mit Rheinischem Bildarchiv Köln. Entwicklungen seit dem Archiveinsturz im Jahr 2009 (S. Lurz/S. Theemann) • „... denn sie wissen nicht, wonach sie suchen sollen.“ Ein Plädoyer für die Erweiterung archivpädagogischer Methoden zugunsten angehender Historiker*innen (N. Pfeuffer) • Das „Wiesbadener Modell“ zur Untersuchung historisch belasteter Straßennamen. Ein Zwischenfazit aus archivischer Sicht (P. Quadflieg) • Ein neues Urteil zum Verhältnis von Wissenschaftsfreiheit versus postmortalem Persönlichkeitsschutz (W. H. Stein)	
TAGUNGSBERICHTE	72
Kölner Forum „Jenseits des Hype – Anwendung von KI in Archiven“ (H. Kannengießer/F. Klein) • Entsäuerung und Kaltlagerung. Bestandsaufnahme und Perspektive des Originalerhalts – Workshop am Landesarchiv Sachsen-Anhalt vom 18. bis 19. Juni 2024 (R. Lusiardi) • Das Internationale Archivsymposium 2024 (S. Rudolf)	
LITERATURBERICHTE	77
MITTEILUNGEN UND BEITRÄGE DES LANDESARCHIVS NRW	81
Transformation der Wiedergutmachung. Ein neues mehrjähriges Erschließungs- und Digitalisierungsprojekt in der Abteilung Rheinland (D. Hecken) • Zwischen „Infohappen“ und „Blockbuster“ – Neue Wege der elektronischen Behördenberatung im LAV NRW (M. Schlemmer)	
MITTEILUNGEN UND BEITRÄGE DES VdA	86
Heute, morgen, übermorgen – Archive im Fokus: 91. Deutscher Archivtag 2024 in Suhl • Fachgruppen 7 und 8 • Berichte aus dem Verband: Landesverband Berlin • Landesverband Sachsen • Gedenken	
PERSONALNACHRICHTEN	99
NACHRUFE	103
KURZINFORMATIONEN	104
VORSCHAU	105

EDITORIAL

Liebe Leserinnen und Leser, liebe Kolleginnen und Kollegen,

das Thema Künstliche Intelligenz (KI) oder Artificial Intelligence (AI) findet seit geraumer Zeit nicht nur große Beachtung in den Medien, sondern nach und nach auch Eingang in unser tägliches Leben. Firmen wie Open AI, Microsoft und Google bringen immer neue generative KI-Modelle auf den boomenden Markt. Zukunftsweisende Prozesse ohne die Einbeziehung von KI sind in vielen Bereichen mittlerweile undenkbar. Angesichts dieser Entwicklungen stellt sich auch für die Archive, nicht zuletzt vor dem Hintergrund enger personeller und finanzieller Ressourcen, die Frage, ob und auf welche Weise auch hier KI-Technologien zum Einsatz gebracht werden können.

Aufgrund der Aktualität und Dynamik, die in der Entwicklung der KI zu beobachten ist, hat der Beirat der Zeitschrift ARCHIV. theorie & praxis beschlossen, dem Thema mit dem vorliegenden Heft einen eigenen Schwerpunkt zu widmen. National wie international wirkende Autorinnen und Autoren nähern sich in der vorliegenden Ausgabe dem Phänomen KI nicht nur definitorisch, sondern auch in der Perspektive praktischer Umsetzbarkeit im Archiv. Die Einleitung in das spannende und zum Teil kontrovers diskutierte Thema haben Bettina Joergens vom Landesarchiv NRW und Tobias Krafft von der Trusted AI GmbH verfasst.

Unter den Beiträgen der Rubrik „Innovation und Erfahrung“ des vorliegenden Heftes sei besonders auf den Bericht von Lutz Bannert, Ines Müller und Nils Trossen über ein Verfahren zur standardbasierten Aussonderung elektronischer Gerichtsverfahrensakten über das digitale Zwischenarchiv des Bundes sowie auf den Überblick über die Bergungserfassung im Historischen Archiv mit Rheinischem Bildarchiv Köln seit dem Archiveinsturz im März 2009 von Selina Lurz und Sven Theemann hingewiesen. Ebenfalls sehr empfehlenswert ist der aufschlussreiche Beitrag von Wolfgang Hans Stein über ein neues Urteil zum Verhältnis von Wissenschaftsfreiheit und postmortalem Persönlichkeitsschutz.

Wie gewohnt finden Sie im vorliegenden Heft auch eine Reihe von Tagungsberichten, und zwar über das neue Format „Kölner Forum“, das im August 2024 erstmals im Stadtarchiv Köln stattgefunden und sich ebenfalls mit dem Thema KI befasst hat, über einen Workshop am Landesarchiv Sachsen-Anhalt zur Entsäuerung und Kaltlagerung als Bestandsaufnahme und Perspektiven des Originalerhalts sowie schließlich über das Internationale Archivsymposium, das jährlich mit begrenzter Teilnehmerschaft aus Belgien, Luxemburg, den Niederlanden und den staatlichen und kommunalen Archiven in Nordrhein-Westfalen und Rheinland-Pfalz durchgeführt wird. Im Mai 2024 fand die Tagung in Arnheim statt und fragte nach der Zukunftsfähigkeit der Archive im Zeichen der tiefgreifenden Transformationsprozesse, mit denen sich Archivarinnen und Archivare in den vergangenen Jahren intensiv befassen mussten, vor allem mit Digitalität, Klimawandel und Naturkatastrophen.

Aus dem Landesarchiv NRW wird über ein mehrjähriges Erschließungs- und Digitalisierungsprojekt der Abteilung Rheinland zur Transformation der Wiedergutmachung berichtet. Außerdem werden neue Formate der seit mehreren Jahren erfolgreich laufenden Behördenberatung des Landesarchivs NRW zur elektronischen Schriftgutverwaltung vorgestellt.

Schließlich berichtet der VdA ausführlich über den 91. Deutschen Archivtag, der im Oktober 2024 in Suhl veranstaltet wurde. Daneben wird auch aus den Landesverbänden Berlin und Sachsen berichtet.

Wir wünschen Ihnen eine interessante Lektüre und eine gute Zeit.

*Herzlichst, Kathrin Pilger, in Verbindung mit Ralf Jacob,
Bettina Joergens, Frank M. Bischoff, Torsten Musial und Bernhard Homa*

KÜNSTLICHE UND ANDERE INTELLIGENZEN IM ARCHIV

ANNÄHERUNGEN ZWISCHEN METHODEN UND PRAXIS IN ARCHIVEN EINERSEITS UND KI-TECHNOLOGIEN

von *Bettina Joergens und Tobias Krafft*

EINLEITUNG

Der Titel des 1. Kölner Forums, „Jenseits des Hype – Anwendung von KI in Archiven“¹ war der Versuch, praxisnah und methodisch reflektiert einige erste KI-Projekte im Archivwesen im deutschsprachigen Raum zu diskutieren. Am Ende des Jahres 2024 kann – nichtsdestotrotz – ein Hype konstatiert werden. Neben zahlreichen Projekten und Praxistests wird im Archivwesen auch viel über verschiedene KI-Technologien, über Qualitätssicherung, über Datenmanagement und Datenschutz sowie archivische Prozesse diskutiert, so etwa beim 1. Kölner Forum am 28. August 2024, beim Deutschen Archivtag in Suhl² und bei einer internen Fachtagung des Landesarchivs NRW im Oktober 2024 sowie bei zahlreichen anderen archivfachlichen Veranstaltungen, wie etwa bei der Frühjahrstagung der Medienarchive im April 2024³ oder beim 7. Berliner Archivtag⁴. Auch schon vorher gab es erste Aktivitäten und Publikationen zur Frage, wie Archive KI nutzbar machen können: So fanden 2022 bereits die interdisziplinär ausgerichteten „EDV-Tage“ in Regensburg zum Thema „Intelligente Prozesse & KI in Museum und Archiv“ und die Tagung „Offene Archive“ in Koblenz statt, bei der Elisabeth Klindworth und Benjamin Rosemann über „Künstliche Intelligenz im Archiv“ referierten⁵. Und in den Blogs „Archivalia“ (unter Hypotheses)⁶ und Sax Archiv⁷ werden die Einsatzmöglichkeiten von KI im Archivwesen erklärt, es werden Tools und Praxisbeispiele vorgestellt und u. a. Fragen des Datenschutzes besprochen.⁸ Wenn man jetzt schon von einem ersten Eindruck vom Einsatz von KI im Archivwesen sprechen kann, fällt zumindest Stand heute auf, dass sich Archive meist um KI-gestützte Anwendungen zur Erschließung und Text- und Formularerkennung bemühen, so auch das Saarländische Landesarchiv in seinem Pilotprojekt zur Handschriftenerkennung.⁹

Insofern kann dieses Themenheft als eine Art Zwischenbilanz gelesen werden, wobei hierfür bei weitem nicht alle Diskussionen, Aktivitäten, Projekte und Veranstaltungen zum Thema „KI im Archivwesen“ rezipiert werden können. In dieser Ausgabe finden sich Beiträge, die beim Kölner Forum diskutiert wurden, so der von Elena Williams, Andreas Neuburger, Gernot Fink mit Oliver Tüselmann, Fabian Wolf und Tim Raven, Martin Vogel, Jasmin Hartmann und Ruth von dem Bussche-Hünnefeld sowie Miriam Eberlein. Darüber hinaus konnten Florian Probst und Thomas Reich gewonnen werden, die ihr Projekt beim Internen Archivtag des Landesarchivs vorstellten. Luciana Duranti, Professorin für Archivtheorie und -wissenschaft an der Universität British Columbia in Vancouver in Kanada, und Isto Huvila, Professor für Informationswissenschaften an der Universität Uppsala in Schweden, bereichern außerdem diesen Themenschwerpunkt mit einer internationalen und archiv- bzw. informationswissenschaftlichen Perspektive. Auch die Vorträge beim Deutschen Archivtag in der Sektion „KI, Machine Learning und Algorithmen – Technologische Revolution im Archivwesen“ von Maria Schmalenbach (Landesbetrieb IT.NRW) zu „Ist das schon KI? – Einsatzszenarien in der öffentlichen Verwaltung“, Benjamin Rosemann (Landesarchiv Baden-Württemberg) über „Archive: Orte des Maschinellen Lernens? KI im praktischen Einsatz bei der Aufbereitung und Nutzung archivischer Daten“ und Esther Lemmerz (Bundesarchiv, Koblenz) zu der Frage, „Welche Möglichkeiten bietet uns KI? Erste Erfahrungen aus dem Bundesarchiv“, gaben wichtige Impulse. Diese werden jedoch im Tagungsband des Deutschen Archivtags veröffentlicht. Alle Veranstaltungen waren stark interdisziplinär besetzt, so wie auch die Beiträger*innen dieses Schwerpunkts aus der Perspektive verschiedener Fach- und Denkrichtungen schreiben.

Im Folgenden wird zunächst 1) eine Begriffsdefinition versucht. 2) geht es um ‚lernende Technik‘, Big Data und den Praxisdurst in den Archiven, 3) die Folgen der Informationen speichernden Technologie (Datenschutz) und 4) um das Mensch – Maschine bzw. Archivar*in – Maschine-Verhältnis.

1) „KÜNSTLICHE INTELLIGENZ“ – VERSUCH EINER DEFINITION

Der Begriff „Künstliche Intelligenz“ (kurz KI) oder englisch „Artificial Intelligence“ (AI) ist bereits wegen des Intelligenz-Begriffs umstritten, da noch nicht einmal klar definierbar ist, was (menschliche) Intelligenz ist. „Konsens ist“, laut Manuela Lenzen, „dass Intelligenz mit Flexibilität und Lernen zu tun hat, mit der Fähigkeit, sich auf wechselnde Anforderungen einzustellen“.¹⁰ Laut der Informatik-Professorin Katharina Zweig ist der „Begriff Künstliche Intelligenz ... mehrfach besetzt: Auf der einen Seite meint man damit das Forschungsfeld, das Methoden entwickelt, um Computer Dinge machen zu lassen, für die Menschen Intelligenz brauchen würden. Auf der anderen Seite nennt man die Methoden selbst auch manchmal ‚Künstliche Intelligenz‘. Und nicht zuletzt, als wäre nicht alles schon verwirrend genug, nennt man insbesondere auch die Software, die diese Methoden verwendet, ebenfalls ‚Künstliche Intelligenz‘.“¹¹ Im aktuellen Diskurs dagegen hat „es sich eingebürgert, Programme, die mit einem Verfahren des maschinellen Lernens arbeiten, vor allem mit dem sogenannten Deep Learning auf Künstlichen Neuronalen Netzen ... als ‚eine KI‘ zu bezeichnen“, so Lenzen.¹²

Allerdings wird in der Literatur schnell deutlich, dass es bei KI nicht mehr nur um neue „Tools“ der Informationstechnologien geht. Vielmehr ist sie zu einem umfassenden Konzept geworden, dessen genaue Grenzen und Verwirklichungswege unklar bleiben.¹³ Isto Huvila bringt dies auf den Punkt: Auch wenn es naheliegend erscheint, die zentralen Herausforderungen von KI in der Archiv- und Schriftgutverwaltung vor allem als technische zu betrachten – sowohl direkt als auch durch ihre Verflechtungen mit zunehmend komplexen gesellschaftlichen Prozessen –, liegen diese in Wirklichkeit häufig in ganz anderen Bereichen.¹⁴ Laut Huvila hat der Begriff „Artificial Intelligence“ inzwischen seine Funktion als eindeutiger Referenzpunkt verloren.¹⁵

2) ‚LERNENDE TECHNIK‘, BIG DATA UND PRAXISDURST DER ARCHIVE

„Künstliche Intelligenz“ beschreibt i. d. R. Technologien aus der Informationstechnik, die mit Algorithmen arbeiten, Prozesse automatisieren, Muster erkennen und z. T. dabei lernen. In vielerlei Hinsicht ist es das, was Computer seit jeher machen. Daher ist es relevant, die angewandten Technologien und deren Funktionsweisen zu benennen, um die Prozesse zu verstehen, um zu entscheiden, welche archivischen Aufgaben der KI-Technologie überlassen werden und welche von Expert*innen erledigt werden müssen.

Mit den gesteigerten Rechnerleistungen und der Fähigkeit von Computern, enorm große Mengen an Daten (Big Data) zu verarbeiten, kann die als „Generative Künstliche Intel-

lizenzen“ bezeichnete Technologie ihre Möglichkeiten erst wirksam werden lassen: Bereits 1843 hat die „erste Computerprogrammiererin der Welt“, Ada Lovelace, vorhergesagt: „Alles, was sich mit Logik beschreiben lässt, werde die Technologie eines Tages reproduzieren können. Weil diese Dinge erkennbare Muster aufweisen, sei die Maschine zukünftig imstande, Musik zu komponieren und Gedichte zu schreiben“, so Miriam Meckel und Lea Steinacker in ihrem Buch „Alles überall auf einmal“.¹⁶ Etwa ein Jahrhundert später, in den 1950er Jahren, entwickelten sich zwei grundlegende Ansätze innerhalb der KI-Forschung: Der erste Ansatz konzentrierte sich auf die symbolische KI, die auf explizit definierten Regeln und logischen Ableitungen basiert. Der zweite Ansatz fokussierte sich auf künstliche neuronale Netzwerke, die auf datengetriebenem Lernen beruhen und Muster sowie Zusammenhänge direkt aus den Daten erkennen können. Zunächst, so Meckel und Steinacker weiter, schien die Symbolische KI ‚das Rennen‘ zu machen. Heute wissen wir, dass das Feld zunehmend von der Generativen KI dominiert wird, die auf neuronalen Netzwerken basiert. Eine Vielzahl von Anwendungen wie ChatGPT, Anwendungen zur maschinellen Übersetzung sowie Techniken wie Maschinelles Lernen, Deep Learning und Reinforcement Learning (RFL) nutzen diese. Generative KI ermöglicht die Erstellung neuer Inhalte, wie Texte, Bilder oder Musik, basierend auf bestehenden Datenmustern.¹⁷

Warum der Weg von den 1950er Jahren bis heute lang war und durch zwei „KI-Winter“ führte, liegt auf der Hand: Die „Generative KI“ ist – wie gesagt – datenbasiert. Sie benötigt nicht nur eine schier unvorstellbare Menge an Daten, sondern es müssen Rechner in der Lage sein, diese Daten zu verarbeiten. „Mit dem World Wide Web begann die Ära von

¹ Diese Veranstaltung wurde vom Historischen Archiv der Stadt Köln mit Rheinischem Bildarchiv und dem Landesarchiv NRW organisiert und durchgeführt. Ich bedanke mich an dieser Stelle bei Dr. Ulrich Fischer für die konstruktive Zusammenarbeit.

² Die Präsentationen werden online zur Verfügung gestellt und die Beiträge in einem Tagungsband veröffentlicht.

³ <https://vfm-online.de/cms/processwire/de/tagung/fruhjahrstagung-2024/tagungsdokumentation-2024/> (aufgerufen am 26.08.2024).

⁴ <https://www.berlinerarchiv.de/7-berliner-archivtag-2024/> (aufgerufen am 22.11.2024).

⁵ S. <https://archive20.hypotheses.org/offene-archive-2-0-bis-2-2/konferenz-archivcamp-2022#programm> (aufgerufen am 13.12.2024).

⁶ Archivalia, archivfachlicher Blog, präsentiert von Hypotheses (<https://archivalia.hypotheses.org/?s=KI>, aufgerufen am 13.12.2024).

⁷ SaxArchiv-Blog, präsentiert von Hypotheses (<https://saxarchiv.hypotheses.org/>, aufgerufen am 13.12.2024).

⁸ <https://saxarchiv.hypotheses.org/17091> (aufgerufen am 26.08.2024).

⁹ https://www.saarland.de/landesarchiv/DE/aktuelles/aktuelle-meldungen/_functions/Pilotprojekt_Einsatz_KI_Handschriften.html (aufgerufen am 26.08.2024).

¹⁰ Manuela Lenzen: Künstliche Intelligenz. Fakten, Chancen und Risiken, 3. aktual. Aufl. 2024, S. 14.

¹¹ Katharina Zweig: Die KI war's! Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz, 4. Aufl. München 2023, S. 32.

¹² Lenzen 2024, S. 14.

¹³ Ebda., S. 10 f.

¹⁴ Isto Huvila: Letting AI lose in an archive: technology to manage or to manage with. In: ARCHIV. theorie & praxis, 78 (2025), H. 1, S. 12–15, hier S. 15.

¹⁵ Ebda., S. 14.

¹⁶ Miriam Meckel, Léa Steinacker: Alles überall auf einmal. Wie künstliche Intelligenz unsere Welt verändert und was wir dabei gewinnen können, Hamburg 2024, S. 43 f.

¹⁷ Ebda., S. 23 f.

Big Data“, so Meckel und Steinacker.¹⁸ Die beiden Autorinnen zeichnen nach, wie die chinesisch stämmige Wissenschaftlerin Fei-Fei Li in den USA „die damals größte Datenbank der KI-Forschung“ erschuf, um Anfang der 2000er Jahre die bislang ungekannten Datenmengen für die KI nutzbar zu machen: „Nach jahrelanger Arbeit hatte Lis Team [mit 50.000 Menschen v. a. aus dem globalen Süden; Anm. Joergens] 15 Millionen Bilder zusammen getragen, die nach 22.000 Objektklassen geordnet waren.“¹⁹ Im Jahr 2017 führte Google den „Mechanismus der ‚Self-Attention‘“ ein, der es ermöglicht, die Bedeutung verschiedener Wörter in einem Satz in Bezug zueinander abzuwägen. Dieser Mechanismus bildet die Grundlage der Transformer-Modellarchitektur, die eine Revolution im Bereich der Verarbeitung natürlicher Sprache (NLP) darstellte. Transformer-Modelle können durch diese Methode kontextuelle Zusammenhänge effizient erfassen und dadurch deutlich präzisere Ergebnisse bei Aufgaben wie maschineller Übersetzung oder Textgenerierung erzielen.²⁰ OpenAI veröffentlichte am 30. November 2022 ChatGPT, ein auf der Transformer-Architektur basierendes Modell. Mit der Einführung dieses Modells begann ein weitreichender Hype um ‚Künstliche Intelligenz‘, der sich trotz erheblicher Bedenken – auch von Seiten der Betreiber selbst – rasant verbreitete. ChatGPT machte die Möglichkeiten von KI erstmals für ein breites Publikum unmittelbar zugänglich und löste eine Welle von Innovationen und Diskussionen aus. Denn damit war KI nicht mehr nur ein Thema für Spezialistinnen, Nerds, Informatiker oder für Science Fiction, sondern plötzlich für alle. Es war – laut Meckel und Steinacker – das „iPhone-Moment“ der KI.²¹ Die KI reproduziert also im Grunde Vergangenes, also auch bestehende Diskriminierungen.²² Meckel und Steinacker sprechen geradezu vom „KI-Inzest“.²³ Was gesellschaftspolitisch kritisch ist, ist für Archive, die eigene Daten zugrunde legen, hilfreich.

Archive sind als Informationsstruktureinrichtungen, also als Halter von sehr großen Mengen an Informationen und Daten, prädestiniert für die Anwendung von KI-Technologien. Zum einen können Archive KI-Systeme einsetzen und sie auf der Basis großer Datenmengen trainieren, und sie so für verschiedene archivische Prozesse anwenden. Zum anderen werden Archive dadurch zu attraktiven Partnern bei Technologieentwicklungen.

Wie der vorliegende Themenschwerpunkt, aber auch zahlreiche andere Projektberichte zeigen, interessieren sich Archive zurzeit in erster Linie für den Einsatz von KI-Technologien zur inhaltlichen Erschließung von Archivalien im weiteren Sinne. Lange Zeit standen Texterkennungsprogramme im Vordergrund. Darüber hinaus werden zunehmend Tools zur Named-entity Recognition eingesetzt zur Erkennung und Klassifikation von bestimmten Begriffen, die Orte, Personen, Institutionen etc. bezeichnen. Mit großen Sprachmodellen (sog. Large Language Models) werden darüber hinaus Texte und Muster in unterschiedlichsten Quellen, erkannt und zugeordnet. Darüber hinaus werden in Archiven Bilderkennungstools eingesetzt, um insbesondere Fotobestände sinnvoller nutzbar zu machen. Seltener, aber z. T. angedacht, sind Tools, die die Bewertung großer Mengen an Unterlagen, i. d. R. born digitals, zu bewerten, oder Nutzenden in einem ersten Beratungsgespräch weiterzuhelfen. Beobachtbar ist,

dass derzeit in vielen Archiven erste Projekte starten, die sich z. T. durchaus ähneln.²⁴

Vielfach – so die Diskussion – entstand der Eindruck, dass zunächst gefragt wird, wie man nun „KI“ einsetzen kann, anstatt zunächst zu klären, welches Problem gelöst werden muss. Teilweise wird dabei kritisch diskutiert, dass etliche Erschließungsprojekte mit Unterstützung von KI-Technologien weit über den archivischen Auftrag hinausgehen anstatt etwa eine Basiserschließung anzubieten. Ob dieser Trend der „Tiefenerschließung“ technik- oder nutzergetrieben ist, wird man wahrscheinlich nicht abschließend feststellen können.

3) KI UND DER DATENSCHUTZ

Der Unterschied zwischen einem Computerprogramm des 20. Jahrhunderts, etwa einem Textverarbeitungsprogramm, und Technologien der Generativen KI ist, dass letztere keine bloßen „Tools“ sind, sondern gelernte Informationen speichernde Systeme. Die Folge ist, dass die Weitergabe oder Veröffentlichung von diesen Maschinen den Austausch von Daten nach sich ziehen kann. Dies erschwert gerade im archivischen Bereich die erforderlichen Kooperationen zwischen Archivinstitutionen, die etwa gleichartige Archivalien mit Large Language Models auswerten lassen wollen. Insofern sind Initiativen erforderlich, die eine Zusammenarbeit in sicheren Clouds, wie sie etwa für die Bundesverwaltung aufgebaut werden,²⁵ und die gemeinsame Nutzung von Technologien, etwa in einem „Tool-Pool“, ermöglicht. Aktuell muss jede Archivinstitution entscheiden, ob sie lizenzkostenfreie Software, on premise-Systeme oder Cloud-basierte Anwendungen nutzt, wie etwa Esther Lemmerz in ihrem Beitrag beim Deutschen Archivtag in Suhl 2024 über die Entwicklung einer KI-Strategie für das Bundesarchiv referierte. Während in on premise-Anwendungen in gesicherten und geschlossenen Netzwerken auch geschützte Daten verarbeitet werden können, sind sie gleichzeitig nicht so leistungsfähig wie Cloud- und Open Source-Tools, die auf einen wesentlich größeren Datenbestand zurückgreifen können. Jedenfalls müssen on premise-Anwendungen mit eigenen Daten trainiert werden und sind i. d. R. in der Entwicklung und Nutzung kostspieliger als Open Source-Technologien. Die Technologie-Entscheidung ist also zugleich eine strategische, rechtliche und finanzielle Abwägung.

4) MENSCH – MASCHINE/ ARCHIVAR*IN – MASCHINE

Es liegt auf der Hand, dass es bei dem Einsatz von KI genannten Technologien nur z. T. um Technik geht, sondern vielmehr um komplexe, neu zu gestaltende Soziotechnische Systeme aus Menschen und Maschinen²⁶.

Dies wird auch in der Geschichtswissenschaft verhandelt: Mit der Anwendung von KI-Tools und Big Data, also der Bereitstellung und Nutzung großer Mengen von Erschließungsdaten und Digitalisaten aus den Archiven kehrt die quantifizierende Herangehensweise an historische Phänomene wieder zurück. Gleichzeitig können – wie in dem SFB 1288 „Praktiken des Vergleichens“ – auf der Basis von großen Textmengen mithilfe von Machine Learning, Topic Model-



Ausschnitt eines mit Zeichnungen des Illustrators Peter Menne gestalteten Bauzaunes der Universität Bielefeld
(Quelle: „Unigewimmel: 50 Jahre Universität Bielefeld“, Abdruck mit freundlicher Genehmigung der Universität Bielefeld)

ling und Mustererkennung begriffshistorische, rechtshistorische, sprachhistorische oder kulturhistorische Vergleiche untersucht werden.²⁷ In der Sendung „KI in der Geschichtswissenschaft“ im Deutschlandfunk Kultur am 20.12.2023 von Carina Schroeder erläutert u. a. Matteo Valleriani, Forschungsgruppenleiter am MPI für Wissenschaftsgeschichte, wie mithilfe von KI etwa das Verständnis von Astronomie in der Geschichte erforscht werden kann und erklärt, dass KI dabei hilft, Quellen bzw. Bilder von Quellen miteinander zu verbinden, sie zeitlich und örtlich zu verbinden und dabei auch zu zeigen, wer von wem inspiriert wurde. Seiner Meinung nach könnten so viel mehr Archivalien ausgewertet werden, zumal – so Valleriani – bislang nur 10 % der Archivalien bekannt sind. In der selben Sendung bezeichnet Torsten Hiltmann (Professor für Digital History, FU Berlin) die „KI als Quantensprung“ für die Geschichtswissenschaft und warnt zugleich vor nicht leicht erkennbaren Quellen-Fälschungen.²⁸

So widmete sich auch eine Tagung der Historischen Kommission zu Berlin, deren Vorsitzende Prof. Ulrike Höroldt zugleich Direktorin des Geheimen Staatsarchivs Preussischer Kulturbesitz ist, im Juni diesen Jahres den „Chancen, Erwartungen und Herausforderungen“ von „KI in der Geschichtswissenschaft“.²⁹ Auch die Tagung „Digital History & Citizen Science: Call for all“ / „Digitale Methoden und neue Erkenntnisse zwischen digitaler Quellenerschließung, Forschung und Bürgerwissenschaften“ vom 19.–22.09.2024 in Halle a. d. Saale weist zahlreiche Schnittmengen mit Archiven und Archivar*innen auf und eröffnet ein weites Panorama an neuen Möglichkeiten der Datenauswertung mithilfe der KI für die historische und (wissenschaftliche) genealogische Forschung. Die Themen reichen von der Handschriftenerkennung über den Einsatz von Large Language Models für die Geschichtswissenschaften, die „Erfassung und Visualisierung von Daten des Gestapoterrors“ bis hin zu Fragen des Arbeitsmarktes für Historiker*innen und der Geschichtsvermittlung. Selbstverständlich bietet Künstliche Intelligenz neue bzw. schnellere Wege auch bei der individuellen Familienforschung, wie in zahlreichen Beiträgen auf der Seite

des Vereins für Computergenealogie e. V. nachzulesen ist.³⁰ Diese Aktivitäten auf Seiten der Kundschaft von Archiven kann durchaus als Aufforderung an Archive gelesen werden und verdeutlichen, dass die Grenzen zwischen Bereitstellung einerseits und Nutzung andererseits, also die klassische Rollenverteilung zwischen Archivinstitution und Archivnutzenden, durchaus ins Schwimmen geraten – mit Perspektive auf eine Art augmented customer relationship.³¹

Vielleicht kann man drei große Felder der gesellschafts-politischen und ethischen Implikationen von KI und der neuen Mensch-Maschine-Beziehung³² ausmachen. Schon alleine der Begriff „Künstliche Intelligenz“ legt nahe, dass Technik so intelligent sein könne wie Menschen. Die Vermenschlichung technischer Prozesse führt bisweilen zu einer problematischen Verlagerung von Verantwortung oder – wie Fabio Tollon schreibt – zu „responsibility gaps“³³. Fabio Tollon legte 2023 in Bielefeld seine Doktorarbeit in der Philosophie unter dem Titel „AI, Agency, and responsibility. From New

¹⁸ Ebda., S. 62.

¹⁹ Ebda., S. 63–65.

²⁰ Ebda., S. 78.

²¹ Meckel/Steinacker 2024, S. 17

²² Problem: KI reproduziert Vergangenes / Vorhandes, ist nicht kreativ, s. u. a. Meckel/Steinacker und ada Newsletter v. 21.7.2024 (https://www.join-ada.com/#ada_newsletter, aufgerufen am 13.12.2024).

²³ Meckel/Steinacker 2024, S. 150.

²⁴ Vgl. zum Einsatz von „Künstlicher Intelligenz“ in Archiven: Susanne Fröhlich: Standardisierte Metadatenhaltung als Basis für KI-Erschließung. In: *scrinium. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare*, 78 (2024), S. 17–26; Lise Jaillant (Hg.): *Archive, Access and Artificial Intelligence. Working with Born-digital and Digitized Archival Collection* (Digital Humanities Research 2), Bielefeld 2022; Tamara Kefer: „More product, less process“? Zukunftsweisende Erschließungsmethoden im Kommunalarchiv. In: *scrinium. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare*, 78 (2024), S. 50–66; Katharina Kinder-Kurlanda: Künstliche Intelligenz als Methode und Gegenstand der Forschung. Festvortrag auf dem 42. Österreichischen Archivtag in Klagenfurt. In: *scrinium. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare*, 78 (2024), S. 9–16; Günter Mühlberger: Transkribus für Archive oder wie Künstliche Intelligenz den Zugang zu historischen Dokumenten revolutioniert. In: *scrinium. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare*, 77 (2023), S. 111–127; Michael Schaffner: Digitale Zugänge schaffen. Die Aufbereitung zentraler Bestände und Serien im Staatsarchiv des Kantons Zürich am Beispiel des Weißen Registers. In: *scrinium. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare*, 78 (2024), S. 27–41.

²⁵ S. <https://www.itzbund.de/DE/itloesungen/egovernment/bundescloud/bundescloud.html> (aufgerufen am 22.11.2024).

²⁶ Katharina Zweig, Tobias D. Krafft, Anita Klingel, Enno Park: *Sozioinformatik: Ein neuer Blick auf Informatik und Gesellschaft*, München 2021 (978-3446452138).

²⁷ Anzeige von SFB 1288 INF: <https://biyecoll.uni-bielefeld.de/index.php/kwi/article/view/803/836> (aufgerufen am 26.08.2024).

²⁸ Carina Schroeder: KI in der Geschichtswissenschaft, Sendung von Deutschlandfunk Kultur, 20.12.2023 <https://www.mpiwg-berlin.mpg.de/de/podcast/ki-der-geschichtswissenschaft>, s. a. <https://www.deutschlandfunkkultur.de/ki-in-der-geschichtswissenschaft-dlf-kultur-07d4d267-100.html> (aufgerufen am 26.08.2024).

²⁹ <https://www.hsozkult.de/event/id/event-143504> (aufgerufen am 26.08.2024).

³⁰ <https://www.compgen.de/category/kuenstliche-intelligenz-ki/> (aufgerufen am 26.08.2024).

³¹ Vgl. Marlene Huber: HAI – „Historical Media date Artificial Intelligence“. Was geschieht, wenn Zeitgeschichte auf KI trifft. In: *scrinium. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare*, 78 (2024), S. 42–49.

³² Vgl. ada Newsletter v. 7.7.2024.

³³ <https://grk2073.org/projects/responsibility-maps/> (aufgerufen am 26.08.2024).

Technologies to Old Problems, and Back Again“ vor.³⁴ Er behandelt v. a. die Frage: „What are the relevant capacities that are required for an entity to be considered a moral agent, and can an artificial system be included in this set?“³⁵

Dies führt zu weiteren gesellschaftspolitischen und ethischen Fragen, die uns im Zusammenhang mit KI beschäftigen sollten: Transparenz, also die Überprüfbarkeit von Entscheidungen, Verfahren und Fehlern. Zum einen sind „KI-Systeme“, so Katharina Zweig, „aus verschiedenen Gründen heute oft sogenannte Blackbox-Modelle, also Systeme, in die man nicht so einfach hineinsehen kann. Das liegt teilweise daran, dass die Systeme von Firmen gebaut werden, die sich nicht gern in die Karten sehen lassen. (...) Ein anderer Grund liegt aber in der Natur der Technologie: Die meisten Methoden der Künstlichen Intelligenz, die heute verwendet werden, sind von sich aus nicht vollständig einsehbar.“³⁶ Und wie oben bereits angedeutet, erwarten wir im demokratischen Rechtsstaat, dass Entscheidungen der Verwaltung weiterhin überprüft und kontrolliert werden können. Auf den Seiten der Bundesregierung heißt es: „Die Bundesregierung setzt sich ... für einen risikobasierten Regulierungsansatz ein. Sie folgt damit auch den Empfehlungen der Datenethikkommission. Unter anderem bei den Abstimmungen zur KI-VO stehen dabei Kontrollinstrumente, Transparenzvorgaben und die Erklär- und Nachvollziehbarkeit von KI im Fokus.“³⁷ Darauf zielt auch das „KI-Transparenzregister für die öffentliche Verwaltung“.³⁸ Jenny Bunn vom National Archives, London, plädiert als Archivarin bzw. Records Managerin in ihrem Beitrag „Working in Contexts for which transparency is important“ für das Konzept der „XAI“, also der „explainable artificial intelligence (XAI)“ – als Antwort auf die „‘black box‘ systems“ insbesondere in „deep learning models“. Bunn geht dabei von dem Grundsatz der Transparenz und der Kontrollierbarkeit von Verwaltungshandeln aus und diskutiert „the increasing use of more opaque AI techniques“.³⁹

Damit ist lediglich angedeutet, dass der Einsatz von KI das Maschine-Mensch-Verhältnis herausfordert und i. S. ethischer Grundsätze und i. S. des demokratischen Rechtsstaats reguliert und organisiert werden muss. Das prominenteste Beispiel ist der AI-Act der EU⁴⁰, auf dessen Basis etwa Menschenrechtsverletzungen verhindert werden sollen.⁴¹ Vielfach werden Konzepte, Frameworks und Governance-Regeln aufgestellt, um die Anwendung von KI in Firmen, Behörden, Städten, Ländern etc. zu rahmen und in den demokratischen Diskurs zu stellen, auch wenn dies für die eine oder den anderen wie eine „Spaßbremse“ wirkt.⁴² Beispielsweise sollte i. S. der Redlichkeit gelten, dass KI generierte Texte und Konzepte als solche gekennzeichnet werden.⁴³ Dies muss auch für KI generierte Daten gelten, die Nutzenden und Archivmitarbeitenden angeboten werden. Übrigens ist dieser Text von Menschen gemacht, nämlich von Tobias Krafft und Bettina Joergens.

Eine KI-Compliance oder eine „Kritikalitätsmatrix“, wie von Katharina Zweig und Tobias Krafft entwickelt: Diese dient der Risikoabschätzung und zur Einschätzung des Schädigungspotentials von KI-Anwendungen. Dabei geht es um mögliche individuelle und gesellschaftliche Risiken.⁴⁴ Und ChatGPT wies darauf hin: „Ein wichtiges Thema ist die ethische Nutzung von KI im Archivwesen. KI-Systeme

können Verzerrungen (Bias) in den Daten oder bei der Interpretation der Daten aufweisen, was die Neutralität und Objektivität der Archivierung beeinträchtigen könnte. Es wird entscheidend sein, Richtlinien zu entwickeln, die den ethischen Einsatz von KI sicherstellen und die Transparenz der verwendeten Algorithmen gewährleisten.“⁴⁵ Gleichzeitig steht die Geschichtswissenschaft bzw. Forschende allgemein, die Archivinformationen nutzen, vor neuen Herausforderungen der Quellenkunde und -kritik, also der ‚Archival Literacy‘, wie Christian Götter fundiert ausführt.⁴⁶ Neben den zunächst abstrakt erscheinenden ethischen und gesellschaftspolitischen Aspekten des Einsatzes von KI werden ganz konkret archivische Arbeitsprozesse herausgefordert. Die Voraussetzung für die Anwendung von KI-Technologien sind Daten, im Archiv also Digitalisate. Umso mehr digitalisiertes Archivgut einer KI-basierten Auswertung zugrunde gelegt wird, desto besser werden die Ergebnisse sein. In den meisten Archiven erfolgt Digitalisierung nach dem Schritt der Erschließung, um die Digitalisate mit den vorhandenen Metadaten zu verbinden. Um jedoch den menschlichen Aufwand der Erschließung einzusparen oder stark zu reduzieren, empfiehlt es sich, Archivalien zunächst zu digitalisieren (und nur mit rudimentären Metadaten zu versehen) und dann erschließen zu lassen. Dies kehrt nicht nur bekannte Prozesse um, sondern erfordert auch ein revidiertes Datenmanagement.

Auch die Qualitätssicherung muss neu konzipiert werden: Dabei ist zunächst festzustellen, welche Fehler die KI-Anwendung macht und wie diese Fehler erkennbar sind. Vor einem Projekt sollte entschieden sein, wie hoch die Fehler-toleranz ist, ggf. nach Fehlerkategorien, und wie die Fehlerüberprüfung durchgeführt werden soll. Dieser Prozess sollte nicht nur mit allen Beteiligten abgestimmt sein, sondern auch den Nutzenden transparent gemacht werden. Mit Blick auf die Praxis wird deutlich, dass der Einsatz von KI-Technologien ein neues Arbeitsfeld auch für Archive ist. So sehr die IT Aufgaben vereinfacht, erfordert deren Einsatz auch zunächst zusätzliche Aufwände.

FAZIT

Als Open-AI am 30. November 2022 ChatGPT veröffentlichte, entfachte dies Begeisterung einerseits und Bedenken andererseits⁴⁷. Die gerade begonnenen gesellschaftspolitischen und ethischen Diskurse wurden von Praxisanwendungen praktisch überholt. Man mag im Archivwesen ähnliche Entwicklungen feststellen: Während die Diskussionen über Methoden und Strategien gerade mal ansetzten, starteten schon die ersten Praxisversuche. Und mit den ersten Projekten setzte ein schneller Lernprozess über KI-basierte Technologien und deren Einsatzmöglichkeiten im Archiv ein. Dies wiederum befähigt die Archivcommunity, kompetent Methoden und Strategien zu entwickeln, die sie braucht, um weiterhin Vertrauen und Nachvollziehbarkeit zu gewährleisten.

Dr. Bettina Joergens

Landesarchiv NRW
 Fachbereich Grundsätze
 Schifferstr. 30, 47059 Duisburg
 E-Mail: bettina.joergens@lav.nrw.de

Dr.-Ing. Tobias Krafft

Trusted AI GmbH
 Am Kumb 5, 67731 Otterbach
 E-Mail: krafft@trusted-ai.com

- ³⁴ Volltext: https://pub.uni-bielefeld.de/download/2987436/2987437/FINAL_05.02.24.pdf (aufgerufen am 26.08.2024).
- ³⁵ https://pub.uni-bielefeld.de/download/2987436/2987437/FINAL_05.02.24.pdf, S. 12 (aufgerufen am 26.08.2024).
- ³⁶ Zweig 2023, S. 11; vgl. S. 67.
- ³⁷ <https://www.ki-strategie-deutschland.de/> (aufgerufen am 26.08.2024).
- ³⁸ <https://algorithmwatch.org/de/transparenzregister-oeffentliche-verwaltung-2023/> (aufgerufen am 26.08.2024).
- ³⁹ Jenny Bunn: Working in contexts for which transparency is important. In: *Records Management Journal*, 30 (2020), Nr. 2, S. 143–153, hier S. 144.
- ⁴⁰ <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32024R1689> (aufgerufen am 26.08.2024).
- ⁴¹ https://www.bmj.de/SharedDocs/Pressemitteilungen/DE/2024/0202_KI-VO.html?cms_mtm_campaign=linksFromNewsletter. Vgl. <https://futurium.ec.europa.eu/en/european-ai-alliance/best-practices/implementing-ai-governance-framework-practice> ada Newsletter v. 09.06., bmj.newsletter v. 02.02.24.
- ⁴² Vgl. <https://www.baden-wuerttemberg.de/de/service/alle-meldungen/meldung/pid/kuenstliche-intelligenz-in-der-verwaltung>; Stadt-Wien-Leitfaden: <https://digitales.wien.gv.at/digitale-agenda/#kuenstliche-intelligenz>; KI in öff. Verwaltung. Konzept vom Fraunhofer (s. o.); https://algorithmwatch.org/de/wp-content/uploads/2023/03/Konzept_KI-Transparenzregister_AlgorithmWatch_2023.pdf, <https://www.itzbund.de/DE/digitalemission/trendstechnologien/kuenstlicheintelligenz/kuenstlicheintelligenz.html> und <https://netzpolitik.org/2023/ki-in-der-verwaltung-bundesregierung-folgt-dem-hype/#netzpolitik-pw>.
- ⁴³ Vgl. ada Newsletter v. 30.6.2024.
- ⁴⁴ <https://www.denkfabrik-bmas.de/projekte/ki-in-der-verwaltung/interview-mit-tobias-krafft> (aufgerufen am 26.08.2024).
- ⁴⁵ Abfrage am 14.08.24.
- ⁴⁶ Christian Götter: Künstliche Intelligenz und Geschichtswissenschaft. Mehr als ein neues Werkzeug für die Digital History? In: *Historische Zeitschrift*, 319 (2024), H. 2, S. 299–330.
- ⁴⁷ Meckel/Steinacker 2024, S. 79.

LETTING AI LOOSE IN AN ARCHIVE

TECHNOLOGY TO MANAGE OR TO MANAGE WITH

by *Isto Huvila*

INTRODUCTION

Artificial intelligence has been called both a misnomer,¹ and a revolution coming very soon but not quite being here already for decades.² During the past couple years there has been a growing consensus that it actually might be here – for good and bad.³ A question remains, however, what is it that we have and what does it mean. The genuinely spectacular capabilities of rapidly advancing, especially machine learning based natural language and image processing techniques have raised optimism on their capacity to carry out tasks beyond the capabilities of traditional computational methods. In archives and records management, they provide opportunities to process large quantities of records beyond what has been possible before computationally or relying on human labour. However, at the same time, the rapid development of diverse AI techniques makes it difficult to grasp what consequences, implications, benefits and downsides they have. A large number of projects and initiatives around the world have started to work on questions pertaining to AI and archives. The author is affiliated with two such initiatives. The international InterPARES TRUST AI project⁴ has worked from 2021 to inquire into how AI can be used to provide support in archives and records management tasks. An example of a smaller project is the “Artificial Intelligence as a risk and opportunity for the authenticity of archives”⁵ funded as a part of the large Swedish Wallenberg AI, Autonomous Systems and Software Program – Humanity and Society (WASP-HS) research programme.

The work in these two projects and numerous initiatives beyond them, much of which is done at archival institutions, provide an increasing number of pivotal insights into how AI can assist in a broad range of critical records management tasks. Automatic character and hand-written text recognition makes transcribing texts faster in an entirely different order of magnitude compared to manual work.⁶ Automatic classifiers and indexing can help to organise and document archives on the level of individual records making them directly searchable and accessible to users much faster, and with at least slightly thorough understanding of their administrative and organisational provenance.⁷ Large language

models have also already now proven their usefulness in summarising contents of large quantities of text. They hold even a lot of promise in supporting archival appraisal.⁸ NLP can help to classify records according to a set criteria both for keeping and weeding.⁹ Further, similar techniques can be used to identify sensitive information helping archivists to make decisions on whether they can be released and to whom.¹⁰ These are merely examples of emerging opportunities.

AI AS AN IMPERFECT SOLUTION

All this is naturally more complicated in practice. Applying AI techniques requires that all processed information is digitised or born-digital. Especially with older material, digitisation still takes a lot of time meaning that a lot of material is and will remain non-digital for the foreseeable future. Considering the required resources, it is also relevant to ask how much is it reasonable to digitise. Further, even if AI is capable of producing remarkable results, they are far from perfect and it is difficult to know when they are trustworthy and on what premises. Automatic classification can be largely accurate but may well miss sensitive information or suggest that important records should be disposed. Similarly, while AI-based indexing would be mostly accurate, it can miss a lot of relevant details and perspectives and misrepresent records and their contents. Accuracy can be measured but the measurements reveal inaccuracies only in relation to what is measured, risking to lead to comparable problems information retrieval research has been struggling with for decades with how to measure and ensure the relevance of search results in information systems.

The imperfections of AI techniques in records management tasks are not, however, what makes AI truly a conundrum. Also records creators, managers and archivists are imperfect, missing details and making debatable decisions. The enigma of AI lies rather in how it is accurate and how it makes mistakes. The underpinning premise of AI development is what Birhane describes as turning challenges to ‘problem/solutions formulations’ where a challenge is formalised as a mathematical problem and solved by developing an ap-

propriate solution to the problem.¹¹ There is an obvious risk that the challenge is only partly captured by the 'problem'. The premise of how machine learning approaches 'learning' is closely similar. At best quantitative learnability – how an AI can 'learn' to produce statistical solutions that have a low likelihood of being erroneous – can be a close approximation of real-life learnability – how a human-being can learn to make informed decisions – but seldom exactly the same.¹² However, instead of the approximations themselves being an issue, the more serious problem is how they are often portrayed as solutions to the challenge itself. The deliberate black-boxing of AI systems as automatic question-answering machines coupled with the inherent opacity of many popular AI techniques put them at odds with the conventional logic of archives as process-bound information. In contrast to simpler algorithms with all decisions and threshold values explicitly encoded, probabilistic techniques do not follow set procedures producing outputs based on pre-programmed sets of instructions and consequently, their outputs are not fully predictable. This means also that making AI to work requires more than digitisation and data cleaning. It is at least equally critical to understand and communicate the opportunities and limits of AI to diverse stakeholders of archives from records creating organisations to researchers and the public, and manage their expectations. Instead of insisting on giving a (false) impression of the completeness of search results generated by AI-powered search engine, their incompleteness and possible complementary search strategies can be highlighted through the design of search systems and how the results are presented. Automatically generated summaries can similarly be flagged and perhaps not being shown by default at all in order to indicate their unverified status. Similarly, even if it would be tempting to flag for how the use of automatic classification of records facilitates and speeds up the public release of records, perhaps the best strategy is to avoid overestimating its impact and downplaying the importance of human labour in the process.

RECORDING AI

A parallel question to how to manage AI in archives and records management tasks is how to get grips with it in the broader context of how records and recording intersect with diverse areas of social life. In parallel to how AI is increasingly used to manage records at archives, they are to an even larger extent used to generate and process information and records across the society. Currently perhaps the most debated examples relate to automated and AI-supported decision-making that are becoming increasingly popular in administrative processes in areas like processing welfare benefit requests and credit applications. AI is also gaining popularity in detecting potential cases of tax and benefits fraud and other criminal activities.¹³ In all of such cases, AI is used in decision-making that need to be appropriately documented for adequate evidence much similarly to how AI needs to be kept accountable when used in archival tasks. In some cases, decisions can be completely automated, which raises a question of who *de facto* created a record: developer of an AI model or system, its operating organisation, or perhaps the AI itself. Even when AI is used to augment rather

than automatise human decision-making, it is relevant to ask how and to what extent AI use should be documented. The borderline between using AI as a piece of office automation with at the most minor consequences to records creation processes and as their key constituent is far from being sharp and straight.

The increasingly apparent risks relating to hallucinating generative AI systems, biased and incorrect outputs of AI classifiers and general difficulty to understand how AI techniques work and what problems they solve have unsurprisingly led to mounting demands of transparency, explanations and information on how different AI techniques function,¹⁴ including calls for recordkeeping of AI development itself.¹⁵ One of the measures proposed to provide neces-

- ¹ Jaron Lanier: There Is No AI. In: The New Yorker 20 (2023), pp. 04–23.
- ² Thomas H. Davenport: The AI Advantage: How to Put the Artificial Intelligence Revolution to Work. Cambridge, MA 2018; Michael I. Jordan: Artificial Intelligence – The Revolution Hasn't Happened Yet". In: Harvard Data Science Review 1.1. (2019); Spyros Makridakis: The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms. In: Futures 90 (2017), pp. 46–60.
- ³ Sarah Griffiths: Move over, Indy! AI Is Here. In: Engineering & Technology 18 (2023) H6, pp. 28–31; Stuart A. Thompson: The Dark Side of A.I. Is Here. In: The New York Times (2024), B1–B1; Niels van Hove: Commentary: AI Is Here to Automate the Knowledge Worker. In: Foresight 61 (2021), pp. 36–38.
- ⁴ On the Internet www.interparestrustai.org (accessed 04.11.2024). See also the article by Luciana Duranti in this issue of Archiv theorie & praxis.
- ⁵ Artificial Intelligence as a risk and opportunity for the authenticity of archives. On the internet at <https://www.uu.se/en/departement/alm/research/research-projects/ongoing-projects/artificial-intelligence-as-a-risk-and-opportunity-for-the-authenticity-of-archives> (accessed on 04.11.2024).
- ⁶ E. g., Dick Kasperowski, Karl-Magnus Johansson and Olof Karsvall: "Temporalities and Values in an Epistemic Culture: Citizen Humanities, Local Knowledge, and AI-supported Transcription of Archives". In: Archives & Manuscripts 51 (2024) H2, pp. 3–22.
- ⁷ E. g., Ciaran B. Trace: Maintaining Records in Context? Disrupting the Theory and Practice of Archival Classification and Arrangement. In: The American Archivist 83 (2020) H2, pp. 322–372; Jason Franks: Text Classification for Records Management. In: Journal on Computing and Cultural Heritage 15 (2022). H3, 42:1–42:19.
- ⁸ E. g., Shitou Zhang et al.: Archives Meet GPT: A Pilot Study on Enhancing Archival Workflows with Large Language Models". In: iConference 2024 Proceedings 2024.
- ⁹ E. g., Naya Sucha-xaya: Documenting Events in Times of Crisis: Investigating Traditional and Contemporary Thai Archival Practices. In: Archivaria 93 (2023), pp. 42–71; Basma Makhoul Shabou et al.: Algorithmic Methods to Explore the Automation of the Appraisal of Structured and Unstructured Digital Data. In: Records Management Journal 30 (2020) H2, pp. 175–200.
- ¹⁰ E. g., Camille Moret: Police Casework Recordkeeping in England and Wales: An Archival Denial of Evidence. In: Archives and Records 45 (2024) H2, pp. 176–195; Lise Jaillant, ed. Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitised Archival Collections. Bielefeld 2022.
- ¹¹ Abeba Birhane: Algorithmic Injustice: A Relational Ethics Approach. In: Patterns 2 (2021) H2.
- ¹² Edward B. Kang: Ground Truth Tracings (GTT): On the Epistemic Limits of Machine Learning". In: Big Data & Society 10 (2023) H1, p. 20539517221146122.
- ¹³ E. g., Hilde Weerts et al.: Can Fairness Be Automated? Guidelines and Opportunities for Fairness-aware AutoML". In: Journal of Artificial Intelligence Research 79 (2023), pp. 639–677 cf. Alexandra James and Andrew Whelan: "Ethical" Artificial Intelligence in the Welfare State: Discourse and Discrepancy in Australian Social Services". In: Critical Social Policy 42 (2022) H1, pp. 22–42.
- ¹⁴ Michael Ridley: Explainable Artificial Intelligence (XAI). In: Information Technology and Libraries 41 (2022) H2; Christian Borch and Bo Hee Min: Toward a Sociology of Machine Learning Explainability: Human–Machine Interaction in Deep Neural Network-Based Automated Trading. In: Big Data & Society 9 (2022) H2, p. 2053951722111361.
- ¹⁵ Jenny Bunn: Working in contexts for which transparency is important: A recordkeeping view of explainable artificial intelligence (XAI). In: Records Management Journal, 30 (2020) H(2), pp. 143–153.

sary transparency is *paradata*, defined in the InterPARES Trust AI project at the present as information about the procedure(s) and tools (for example, AI models and systems) used to create and process information resources, or in the broader sense, as information on practices and processes.¹⁶ In the AI context, the research on paradata has inquired so far into the concept and how to frame it, and in a more practical sense, what would need to be documented relating to AI use.¹⁷ On-going studies inquire in detail, for example, in how it is currently done, and what opportunities and challenges have been identified.

An underlying assumption of many of the propositions for advancing human understanding of AI stemming from the technology industry rather than contexts where AI is used is a belief that AI techniques can be comprehensively explained much similarly to how solving issues related to (un)fairness and bias in AI are matters of developing a better mathematical model that is fair and unbiased.¹⁸ The idea that documenting especially probabilistic AI techniques is a comprehensible solvable problem has been, however, increasingly challenged.¹⁹ Rather, it has become increasingly apparent that documenting and keeping AI accountable requires thorough attention to the adequacy of explanations for specified contexts and purposes – with an obvious repercussion of a need keeping record of AI and AI use. The diversity of AI techniques and to what extent they are explainable and how means that there is no one size fits all solution to paradata, or in more general terms, to recording AI and AI use. Drawing parallels from the research on documenting complex processes,²⁰ the problem is not solved by adding a field to a documentation standard or proposing a complex framework for representing chains of inferences and actions with a prompt to document AI systems and their use in sufficient detail. Instead of holding to the current technology driven approach that regularly downplays recordkeeping beyond immediate informational needs, there is a critical need to decide what is required of a record of AI processing and of AI processing of records. With criteria founded on the requirements stemming from recordkeeping rather than specific characteristics of technologies, AI techniques can be implemented and applied to support human processes instead of trying to force non-technical concepts such as accountability, ideals of fairness and lack of bias, transparency and sensitivity to a new mathematical model. If an appropriate record of AI used for a particular high-risk purpose cannot be generated, it could rather be a signal that a particular AI should perhaps not be used for that purpose at all rather than of the inadequacy of how a record is conceptualised.

REIMAGINING AI

What is then an appropriate record of AI, or more specifically, what is an appropriate record of which AI, and what characterizes appropriate use of AI in archives and records management? The difficulty of getting to grips with AI – how it fits specific tasks and how to record its pertinent aspects – might not interestingly be that much a question of what AI is but rather what it is imagined to be. Sociologist Lucy Suchman refers to the problematic assumption of the “thing-

ness of AI”²¹ as a critical reason why it is so difficult to discuss AI in a fruitful manner. The term is used as if there was an AI, a perpetrator that makes things happen, while in reality it refers to a vast constellation of existing and imagined techniques, technologies and ideas that have very little in common. Even if conceptual open-endedness can be helpful for catalysing discussion, the present concept of AI can be rightfully criticised of having become so vague that it has all but lost its role as a meaningful point of reference. It is too open, vague and fluid to properly facilitate communication between multiple, both technical and non-technical social worlds as a useful ‘boundary object’, a popular concept used to describe physical and conceptual things that facilitate communication and collaboration between communities.²² Rather than thinking of AI as a thing, Suchman urges to turning it controversial through considering instead of what the specific technologies and techniques under the AI umbrella are and what they do in specific contexts.²³ She refers to decision taken at the Center on Privacy & Technology at Georgetown Law to stop using the term AI.²⁴ Instead, the director of the centre, Emily Tucker, urges to describe what specific techniques do, for example, how face recognition is about matching images of human faces or hand-written recognition about converting digitised hand-written script to machine-encoded text. The major benefit of such a move is to bring the discussion back to what really can be done with particular technologies, what are their strengths and weaknesses, what is best achieved using other technologies or techniques – and also how to record the pertinent aspects of specific technologies and their use for archival purposes. A parallel caveat to the purported thingness of AI is in its close association with the neo-liberal ideal of speed and increased efficiency through automation. Admitting that automation has many foreseeable advantages, it is worth considering in the spirit of slow movement²⁵ what are the effects of automation and speed, and why certain processes previously have been slow and manual – not only because of the lack of appropriate technologies and sophisticated processes. Automation has been advocated as a new archival paradigm,²⁶ but rather than accepting it as a thing similar to how Suchman describes AI imaginaries,²⁷ it is worth resisting and seeing as potentially equally controversial as AI. The same applies to speeding up processes that previously have been slow. In many cases, slowness is a ‘problem’ but it is worth considering whether solving that problem will solve its underpinning challenge,²⁸ or merely cause a lot of unanticipated and unwanted side-effects.

By stopping to talk about AI and resisting the temptation of overvaluing speed, it is possible to shift discussion from being technology-focused to becoming more people-centric as advocated by as a critical move towards understanding and implementing automation in the everyday practice instead of subjecting human pursuits to the condition of poorly understood technologies.²⁹ Such a move is crucial to avoid using AI as a system developed to provide unknown answers to questions we are not yet capable of formulating and asking.³⁰ In archival settings, people-centrism means focusing on archival practice but also turning attention to the implications of specific technologies and techniques for records and recording in the broader societal context.

Rather than starting with the question of what technologies there are to use, it is more relevant to start by identifying what archival problems there are to solve. If the major problem is poor searchability, there are a plenty of useful techniques available from text summarisation based on large language models to relevance-ranking of search results based on matching search strings and target documents and beyond. At the same time, solving that particular problem does not automatically solve the challenge of *finding* all relevant records. Identification of particular types of records for appraisal or release can be enhanced, for example, using models that help to predict if a particular record contains specific type of information. However, solving that problem does not mean that necessarily all pertinent documents would be identified and that the resulting set of records would for sure contain only expected types of records. Much similarly, the use of AI in diverse administrative processes across the society to make decisions and generate records would benefit of careful consideration of what is the problem to be solved and to what extent different techniques actually solve the underpinning challenge.

CONCLUSION

Even if it spontaneously might feel reasonable to think that the most significant challenges posed by AI for archives and records management, both directly and indirectly through how it is entangled in an increasing number of societal processes, are technical by their nature, they are to a large extent elsewhere. At the heart of what makes AI as a 'thingy' concept, and the plethora of technologies and techniques flagged as AI difficult is how they in different ways affect how and what we feel we should keep accountable. Even if AI is purported to be artificial *intelligence* – that is something commonly considered as human but artificial – at least for now, it is only mimicking – even if increasingly often so well that it is difficult to tell the difference without a considerable effort. At the same time neural network-based AI techniques are operating in ways that do not follow the innate logic of how we have learned to understand how technologies work either. This makes it difficult to envisage how a particular AI could be held accountable in terms of how we should judge, criticise, accept, or refuse its outputs. While understanding how AI works in context and how it can be relied is important to different degrees in all areas of life, it is an especially critical concern for archives in how the accountability of all the technologies used fiddle with records is in the very essence of what archives and records management is all about. Currently we are not letting AI loose in archives only in how it is used for archival applications but also through the processes that generate and impact records and archives throughout the entire continuum of recordkeeping. AI is a technology to manage with, but to a far greater extent a technology that requires energetic managing.

Rapidly advancing artificial intelligence (AI) hold great promise to support archival work. At the same time, its diversity complicates grasping its implications through its use in archives and records management tasks and even more so, through its use for generating and processing records in different fields of the society. Understanding AI requires abandoning the simplistic ideas of AI being intelligent like humans and a single monolithic thing.

Isto Huvila

Uppsala University

Department of ALM

Box 625, 751 26 UPPSALA, Schweden

E-Mail: isto.huvila@abm.uu.se

- ¹⁶ Scott Cameron and Babak Hamidzadeh: Preserving Paradata for Accountability of Semi-Autonomous AI Agents in Dynamic Environments: An Archival Perspective. In: Telematics and Informatics Reports 14 (2024), p. 100135.; Isto Huvila: Improving the Usefulness of Research Data with Better Paradata. In: Open Information Science 6 (2022) H1, pp. 28–48; Huvila, Improving the Usefulness of Research Data with Better Paradata.
- ¹⁷ E. g., Cameron and Hamidzadeh, Preserving Paradata for Accountability of Semi-Autonomous AI Agents in Dynamic Environments; Patricia C. Franks: The Crucial Role of Paradata in AI Governance". In: SCEaR Newsletter Special Issue: Artificial Intelligence and Documentary Heritage. Paris 2024, pp. 54–60.
- ¹⁸ Maarten Buyl and Tijl De Bie: Inherent Limitations of AI Fairness. In: Communications of the ACM 67 (2024) H2, pp. 48–55.
- ¹⁹ E. g., Hans de Bruijn, Martijn Warnier and Marijn Janssen: The Perils and Pitfalls of Explainable AI: Strategies for Explaining Algorithmic Decision-Making. In: Government Information Quarterly 39 (2022) H2, p. 101666.
- ²⁰ E. g., in Isto Huvila, Lisa Andersson and Olle Sköld: Perspectives on Paradata: Research and Practice of Documenting Data Processes. Cham 2024.
- ²¹ Lucy Suchman: The Uncontroversial 'Thingness' of AI". In: Big Data & Society 10 (2023) 2, p. 20539517231206794.
- ²² Cf. Susan Leigh Star: The Structure of Ill-Structured Solutions: Heterogeneous Problem-Solving, Boundary Objects and Distributed Artificial Intelligence. In: Proceedings of the 8th AAAI Workshop on Distributed Artificial Intelligence, Technical Report, Department of Computer Science, University of Southern California. Los Angeles 1988, pp. 37–54; Isto Huvila et al.: Boundary Objects in Information Science". In: JASIST 68 (2017) 8, pp. 1807–1822.
- ²³ Suchman, The Uncontroversial 'Thingness' of AI.
- ²⁴ Emily Tucker: Artifice and Intelligence 2022; Suchman, "The Uncontroversial 'Thingness' of AI".
- ²⁵ Vincanne Adams, Nancy J. Burke and Ian Whitmarsh: Slow Research: Thoughts for a Movement in Global Health". In: Medical Anthropology 33 (2014) 3, pp. 179–197.
- ²⁶ Siham Alaoui: Peut-on Parler de l'automatisation Comme Cinquième Paradigme Archivistique?" In: The Canadian Journal of Information and Library Science 47 (2024) H1, pp. 18–34.
- ²⁷ Suchman, The Uncontroversial 'Thingness' of AI.
- ²⁸ Birhane, Algorithmic Injustice.
- ²⁹ See e. g., Sarah Pink et al.: Everyday Automation: Experiencing and Anticipating Emerging Technologies. London 2022; Lanier, There Is No AI.
- ³⁰ Jesse Josua Benjamin: Machine Horizons: Post-Phenomenological AI Studies". Unpublished Doctoral Dissertation. Enschede 2023; Hans-Jörg Rheinberger and Peter McLaughlin: From Experimental Systems to Cultures of Experimentation". In: Concepts, Theories, and Rationality in the Biological Sciences. Ed. by Gereon Wolters and James G. Lennox. Pittsburgh 1995, pp. 107–122.

ALEXA TRIFFT ARCHIV

DER AKTUELLE STAND DES EINSATZES VON KI IN ARCHIVEN

von *Elena Williams*

Seit ChatGPT für die Öffentlichkeit zugänglich gemacht wurde, ist Künstliche Intelligenz (KI) in aller Munde. In vielen Bereichen der Wissenschaft, Wirtschaft und Gesellschaft wird KI schon verwendet, unter anderem auch in Archiven. Als angehende Archivarin hat mich dieses aktuelle Thema sehr interessiert und daher hat meine Bachelorarbeit zum Abschluss meines Studiums an der FH Potsdam im Frühjahr 2024 den Titel „Einsatz von Künstlicher Intelligenz in Archiven“. Mit einer selbst erstellten, anonymen Umfrage, die sich an alle Archivarinnen und Archivare in Deutschland und Österreich richtete, habe ich eigene Daten erhoben, um mir einen Überblick über den aktuellen Stand der Thematik zu verschaffen. Die Umfrage lief im Herbst 2023 und der Link zur Umfrage wurde über verschiedene Kanäle verteilt. Der Fragebogen umfasste 20 Fragen und wurde in 4 Abschnitte unterteilt: Allgemeine Fragen zum Archiv (wie Archivalienarten, Mitarbeiterzahl oder digitale Ausstattung), Wissen um KI (wie Einsatzgebiete von KI und KI-Programme), Einsatz von KI (also praktische Erfahrung) und Zukunftsaussichten. Insgesamt haben 241 Archivarinnen und Archivare teilgenommen. Die Definition von KI, die meiner Umfrage und damit diesem Beitrag zugrunde liegt, lautet: KI beschreibt Computersysteme, die menschliche Fähigkeiten wie logisches Denken, Lernen, Planen oder Kreativität imitieren.

VORAUSSETZUNGEN

Eine wichtige Frage zum möglichen Einsatz von KI in einem Archiv behandelt organisatorische Themen, also die Voraussetzungen, die ein Archiv mitbringen muss, um KI anzuwenden sowie die möglichen Problemfelder, die beim Einsatz einer KI auftauchen können. Die Voraussetzungen für einen Einsatz lassen sich in drei Kategorien einteilen: die, die Archive mitbringen müssen, das Wissen, das die Archivarinnen und Archivare haben sollten und die Hürden, die vor einem Einsatz zu bewältigen sind.

Welche Voraussetzungen sollte ein Archiv also mitbringen? Hierbei geht es vor allem darum, wie viele Bestände schon digitalisiert sind, welche technische Ausstattung das Archiv

hat, welche Archivalienarten vorhanden sind und um die Mitarbeiterzahl. Da KI nur eingesetzt werden kann, wenn die Archivalien auch digital vorliegen, stellt die Digitalisierung die wichtigste Voraussetzung für deren Einsatz dar. Die allermeisten Archive (83 %) haben immerhin schon 1 – 15 % ihrer Bestände digitalisiert. Diese Voraussetzung wird also größtenteils bereits erfüllt. Zudem wird eine gewisse Mindestausstattung benötigt, um eine KI anwenden zu können. Die Mehrheit der Archive hat Computer, Digitalisierungsgeräte, Archivmanagementsoftware und einen Internetanschluss. Die Archivalienarten sind wichtig, da diese bestimmen, welche Art von KI-Programm überhaupt eingesetzt werden kann. Bei Dokumenten wären das zum Beispiel Transkriptions-Programme oder Text-Mining-Verfahren, bei Fotos hingegen spezielle Bild-Erschließungsprogramme. Meine Umfrage ergab, dass Dokumente und Fotografien mit 95 bzw. 93 % die am weitesten verbreiteten Archivalienarten sind, aber auch Ton- und Filmaufnahmen haben mit 85,3 bzw. 83,9 % eine gute Verbreitung. Somit könnten fast alle abgefragten Archive Programme zur Text-, Bild-, Ton- und Filmanalyse anwenden. Ein weiteres Kriterium ist die Mitarbeiterzahl. Die meisten der Befragten haben angegeben, dass 1 – 5 Mitarbeiterinnen und Mitarbeiter in ihrem Archiv arbeiten. Der Mitarbeiterzahl kommt im arbeitsökonomischen Umfeld eine große Bedeutung zu. Das Training einer KI ist oftmals sehr zeitintensiv und hat ein Archiv nicht genug Personal mit freien Zeitressourcen, ist es schwieriger eine KI einzusetzen. Es gibt hier keine absolute Grenze, jedoch ist bei nur einem oder zwei Mitarbeitenden die Erfüllung der Aufgaben in Bezug auf das Training und den Einsatz einer KI eher schwierig. Natürlich besteht auch die Möglichkeit, die Digitalisierung und/oder den KI-Einsatz an eine Drittfirma auszulagern. Diese Möglichkeit erfordert allerdings einen höheren finanziellen Aufwand. Welches Wissen zum Einsatz von KI sollten nun die Archivarinnen und Archivare mitbringen? Da wären zuerst einmal die verschiedenen Einsatzbereiche für eine KI. Meine Umfrage ergab, dass der mit Abstand bekannteste Bereich die Transkription (86 %) ist. Auch die Bilderschließung (70 %) und die Inhaltser-

schließung von Texten (66,7 %) sind den Befragten geläufig, während die Filmerschlüsselung (42 %) am wenigsten bekannt ist. Und welches Wissen ist in Bezug auf KI-Programme, die sich für den Einsatz im Archiv eignen, vorhanden? Auch hier sind die Programme aus dem Bereich der Transkription am bekanntesten. Spitzenreiter ist mit 82 % das Programm Transkribus, die Transkriptions-Programme OCR4all, OCR-D und eScriptorium sind immerhin noch gut einem Drittel der Befragten ein Begriff. Das Programm DeepVA zur Bilderschlüsselung kennen hingegen nur noch gut 9 % der Befragten. Welche Hürden stehen nun dem Einsatz einer KI entgegen? Hier zeigte meine Umfrage als allererstes die fehlenden personellen Ressourcen auf (45,7 %). Weitere Hürden sind das fehlende Wissen über den Einsatz von KI (34,2 %), die fehlende Fortbildung dazu (32,7 %) und die noch geringe Verbreitung von KI in Archiven (30,7 %). Die technische Ausstattung (28,6 %), die finanziellen Ressourcen (28,1 %) und die Digitalisierung (24,6 %) stellen dagegen keine allzu großen Hürden dar. Dass der Nutzen einer KI nicht ausreichend belegt ist, ist die kleinste Hürde (16,6 %).

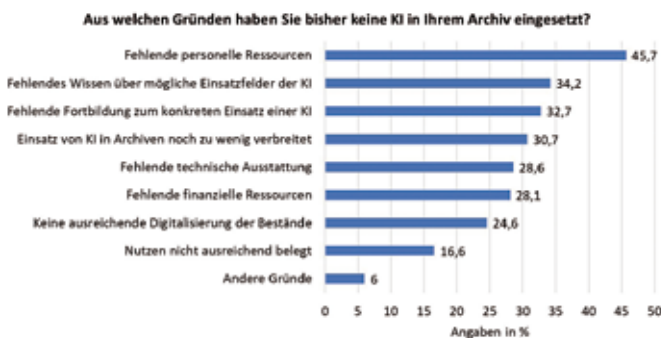


Abbildung 1: Ergebnisse der Frage 11 (Aus welchen Gründen haben Sie bisher keine KI in Ihrem Archiv eingesetzt?) aus der Umfrage (Eigene Darstellung, 2024)

EINSATZ EINER KI

Wie sieht es aber mit dem eigentlichen Einsatz von KI in Archiven aus? Die Mehrheit der Befragten hat noch keine KI in ihrem Archiv eingesetzt (71 %). Dort wo sie bereits eingesetzt wurde, sind die Bereiche Transkription und Inhaltserschließung von Texten am häufigsten vertreten. Da Dokumente die am häufigsten vorkommende Archivalienart ist, ist dies nicht überraschend. Bei der Bilderschlüsselung dagegen wird KI nicht so häufig eingesetzt und am wenigsten in der Filmerschlüsselung. Weitere Einsatzbereiche, die genannt wurden, sind Übersetzung, Genealogie und Datenanalyse und -verarbeitung.

Eine der wichtigsten strategischen Fragen, die sich einem Archiv stellt, ist die Frage, ob sich der Einsatz einer KI lohnt. Diese Frage konnte von den meisten Befragten noch nicht abschließend beurteilt werden. Von denen, die es beurteilen konnten, war der Großteil der Meinung, dass sich der Einsatz lohnt. Dies liegt aus Sicht der Befragten vor allem an der Arbeitserleichterung und an dem besseren Service für Archivnutzende. Auch die Zeitersparnis durch den besseren und einfacheren Zugang zu den Archivalien spielt eine

Rolle. Allerdings sollte hierzu das vorherige, häufig zeitaufwendige Training einer KI in Betracht gezogen werden, um zu beurteilen, ob eine ausgewogene Balance zwischen Arbeitszeiteinsatz und späterer Arbeitszeiterparnis erreicht werden kann. Aspekte, die für den Einsatz einer KI nicht so ausschlaggebend sind, sind die höhere Treffergenauigkeit bei Suchanfragen und der größere Kontext durch mehr Schlagwörter. Ein weiterer Aspekt, der genannt wurde, ist die Verarbeitung von Massenakten, Fotobeständen etc., die sonst nicht erschließbar wären.

In welcher Hinsicht hat sich der Einsatz der KI gelohnt?

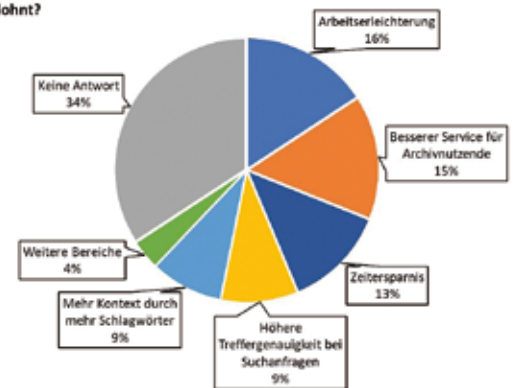


Abbildung 2: Ergebnisse der Frage 16 (In welcher Hinsicht hat sich der Einsatz der KI gelohnt?) aus der Umfrage (Eigene Darstellung, 2024)

Mit nur 2,6 % lohnt sich der Einsatz von KI für nur sehr wenige der Befragten nicht. Aber welches sind die Gründe, die gegen den Einsatz einer KI sprechen? Die wichtigsten der genannten Gründe hierfür sind der hohe Aufwand an Datenvor- und/oder -nachbereitung und die geringe Qualität der Ergebnisse der KI. Auch der große zeitliche Aufwand für – vor allem – das Training wurde als Grund genannt. Ein weiterer Abschnitt meiner Umfrage befasste sich mit den möglichen Problemen, die beim Einsatz einer KI auftreten können. Hierbei hielten sich die Aussagen, dass Probleme auftraten im Gegensatz zu einem problemlosen Einsatz einer KI in etwa die Waage. Die meisten Probleme gab es mit der Software selber, zum Beispiel, dass sie nicht für die gewünschte Aufgabe geeignet war. Ein Teil dieses Problems ist, dass die KI-Programme meistens nicht für den Gebrauch in einem Archiv entwickelt wurden. Gerade bei den vortrainierten KIs muss man die Software zusätzlich an die Daten im Archiv anpassen. Einige Probleme gab es auch mit den Daten selber, zum Beispiel, dass nicht genügend Datensätze für das Training der KI vorliegen. Auch im rechtlichen Bereich wurden Probleme gesehen, zum Beispiel mit dem Daten- und Personenschutz. Hierbei spielt es zum Beispiel eine Rolle, ob die KI-Programme auf einer internen Infrastruktur laufen können und damit die Anwendung der Richtlinien der Datenschutzgrundverordnung gewährleistet werden kann. Ein weiteres Problem waren ethische bzw. moralische Bedenken. Als Beispiel seien hier Akten des Reichskolonialamtes oder des „Dritten Reichs“ genannt, bei denen im virtuellen Raum die Kontextualisierungsmöglichkeiten rassistischer oder

antisemitischer Sprache fehlen.¹ Weitere Probleme, die die Befragten genannt haben, sind, dass die Software noch zu experimentell ist und dass es Probleme mit dem Import der erzeugten Metadaten in die Archivsoftware gibt. Weniger Probleme gab es mit der technischen Ausstattung und mit dem Umgang mit der Software auf der Anwenderseite.

BEISPIEL STADTARCHIV STOCKACH

Die Frage, ob sich der Einsatz einer KI lohnt und welche Vorteile er bietet bzw. mit welchen Herausforderungen eine Archivarin/ein Archivar konfrontiert werden können, habe ich in der Zwischenzeit als Archivarin der Stadt Stockach persönlich erlebt und möchte meine Erfahrungen gerne mit Ihnen teilen. In unserem Projekt wurden die Ratsprotokolle, Urbare, und einige der Lager- und Feuerversicherungsbücher der Stadt Stockach digitalisiert und mit Hilfe der KI-Software Transkribus transkribiert. Die Anwendung der KI fand nicht im Stadtarchiv selber statt, sondern wurde bei READ-COOP (den Entwicklern von Transkribus) durchgeführt. Transkribus hat für jede Handschrift, für Druckschriften und für jede Periode ein vortrainiertes Modell, jedoch kann man auch ein eigenes Modell trainieren. Im Falle des Stadtarchivs Stockach wurden drei vortrainierte Modelle verwendet: für Dokumente des 17.–18. Jahrhunderts das Transkribus Early Kurrent M1-Modell, für Dokumente des 19. Jahrhunderts das Transkribus German Kurrent-Modell und für die Druckschriften das The Text Titan I-Modell. Für die Handschriften des frühen 20. Jahrhunderts wurde für das Stadtarchiv ein eigenes Modell trainiert. Obwohl der Einsatz der KI in diesem Fall ausgelagert wurde, mussten dennoch Vor- und Nachbereitungsarbeiten vom Stadtarchiv Stockach geleistet werden. Vor der Anwendung der KI mussten die Unterlagen digitalisiert und die Digitalisate benannt werden. Nach dem Einsatz der KI müssen nun die Transkriptionen vom Archivpersonal persönlich mit den Originalarchivalien verglichen und auf Fehler hin durchgesehen werden. Dies stellt, auch wenn man es nur stichprobenartig macht, einen gewissen zeitlichen Aufwand dar. Bei der Durchsicht der Transkriptionen fällt auf, dass die KI – unabhängig vom verwendeten Modell – auch Fehler macht. Insofern gibt es zwar keine perfekten Transkriptionen, jedoch sind die Ergebnisse vollkommen ausreichend, um eine Vorstellung davon zu bekommen, was in dem jeweiligen Text steht. Sie stellen daher mit Sicherheit eine Arbeitserleichterung dar, da man den Text nun nicht mehr selber für die Nutzerinnen und Nutzer transkribieren muss. Neben dieser Zeitersparnis bedeutet der Einsatz der KI einen besseren Service für die Nutzenden, die nicht mehr persönlich ins Archiv kommen müssen, sondern von zu Hause aus online die Dokumente einsehen können und dabei direkt die zugehörige Transkription vorliegen haben. Im Falle unseres Projektes in Stockach würde ich sagen, dass die Vorteile überwiegen, jedoch muss das für jedes Projekt einzeln entschieden werden.

ZUKUNFTSAUSSICHTEN

Kommen wir nun zu den Zukunftsaussichten, d. h., in welchen Bereichen sehen die Befragten zukünftig Potential für den Einsatz einer KI und welche Herausforderungen können sie sich vorstellen. Mit Abstand die meisten Befragten kön-

nen sich vorstellen, eine KI in den Bereichen Transkription (80,9 %) und Erschließung (73,4 %) zu verwenden. Auch bei der digitalen Beratung (41,5 %), Genealogie (36,7 %) und Bewertung und Übernahme (33,0 %) können sich viele der Befragten vorstellen eine KI einzusetzen. Vergleichsweise wenige können sich die Verwendung einer KI in den Bereichen Bereitstellung (23,4 %), digitaler Auftritt (22,3 %), Lesesaal (21,8 %) und Magazin (21,3 %) vorstellen.



Abbildung 3: Ergebnisse der Frage 18 (In welchen Bereichen der Archivarbeit könnten Sie sich vorstellen, KI auszuprobieren?) aus der Umfrage (Eigene Darstellung, 2024)

Doch inwieweit könnte eine KI die Archivarbeit in bestimmten Bereichen übernehmen? Hierzu ergab meine Umfrage folgende Ergebnisse: Rund die Hälfte der Befragten (52 %) kann sich vorstellen, dass die KI die Arbeit im Bereich der Transkription in Teilen übernehmen können. 29 % sind der Meinung, dass die Arbeit in der Transkription sogar vollständig durch die KI ausgeführt werden könnte. 55 % der Befragten können sich im Bereich der Erschließung eine Übernahme der Arbeit in Teilen vorstellen, 28 % sehen die Rolle der KI dort eher unterstützend. Auch in der Digitalen Beratung sind die meisten der Befragten der Meinung, dass die KI die Arbeit in Teilen (41 %) oder unterstützend (35 %) übernehmen wird. Unterstützend oder in Teilen sehen jeweils gleich viele der Befragten die Rolle der KI in den Bereichen Digitaler Auftritt (36 %/36 %), Bereitstellung (27 %/26 %) und Genealogie (31 %/30 %). Vor allem unterstützend sehen die Befragten die Rolle der KI in den Bereichen Magazin (39 %), Lesesaal (37 %) und Bewertung und Übernahme (43 %).

Die Herausforderungen für die Zukunft sehen die Befragten wie folgt: die personelle Ausstattung könnte das größte Hindernis beim Einsatz von KI sein (71,8 %). Aber auch die finanziellen Ressourcen (65,2 %) könnten in Zukunft ein größeres Problem darstellen. Die fehlende Digitalisierung der Bestände (62,4 %), der zeitliche Aufwand zum Training der KI (61,3 %), fehlende Fortbildungen zum Thema (60,8 %) und die technische Infrastruktur (59,7 %) werden wohl auch relativ große Herausforderungen darstellen. Beim Daten-

und Personenschutz (51,9 %) sowie bei Berührungängsten in Bezug auf KI (50,3 %) sieht noch die Hälfte der Befragten eine Herausforderung. Weitere Probleme, die laut den Befragten in Zukunft auftreten könnten, sind zum Beispiel: die Akzeptanz der Verwendung der Ergebnisse in wissenschaftlichen Arbeiten, die Bewertung der Arbeit der KI, die Fehlerquote, die Möglichkeit der „Fütterung“ der KI mit Fehlern und ethische Bedenken.

FAZIT

Als Fazit aus meiner Umfrage und meinen persönlichen Erfahrungen lässt sich sagen, dass KI-Programme schon zahlreich in Archiven eingesetzt werden und auch für die Zukunft viel Potential haben. Wenn die Voraussetzungen für den Einsatz von KI in Archiven erfüllt sind, das heißt die nötige technische Ausstattung, die Digitalisierung der Bestände und das Wissen über KI vorhanden sind, dann kann sich der Einsatz durchaus lohnen. Allerdings gibt es auch einige Hindernisse und Probleme, die es zu beachten gilt. Am Ende müssen die Archivarinnen und Archivare individuell entscheiden, ob es sich für ihr Archiv lohnt eine KI einzusetzen. Sie müssen die Vorteile, wie Arbeitserleichterung oder besseren Service für Nutzende gegen die Nachteile, wie den hohen Aufwand an Datenvor- und -nachbereitung oder die potenziell etwas geringere Qualität der Ergebnisse, abwägen. Es wäre sicherlich wünschenswert, wenn zukünftig Software-Programme speziell für den Einsatz in Archiven entwickelt würden, um den Trainingsaufwand zu reduzieren und die Verwendung archivspezifischer Digitalisate zu erleichtern. Gleichzeitig werden die Ergebnisse der KI aber wohl in näherer Zukunft weiterhin von Menschen kontrolliert werden müssen, um Qualität und größtmögliche Fehlerfreiheit zu gewährleisten. Zusammenfassend sehe ich für Künstliche Intelligenz in Archiven viele Herausforderungen,

aber auch ein großes Potential, um die Arbeit der Archivarinnen und Archivare zu erleichtern und die Dienstleistungen der Archive zu erweitern.

ALEXA MEETS ARCHIVE – THE CURRENT STATUS OF THE USE OF AI IN ARCHIVES

This article presents the results of an anonymous survey carried out in the autumn of 2023 among archivists in Germany and Austria to discover the extent to which AI is or could be used in their work. The questionnaire's twenty questions covered archive type, size and equipment, knowledge of AI applications, practical experience and future prospects. It found that AI is already widely used and offers great potential for improving services and facilitating work. Challenges remain concerning costs, program training and quality assurance.

Elena Williams

Stadtarchiv Stockach
Adenauerstraße 4, 78333 Stockach
Tel. 07771 802-326
E-Mail: e.williams@stockach.de

¹ Andrea Hänger: Ohne Menschen können die Maschinen nicht lernen. Der Einsatz von KI im Bundesarchiv. In: Politik & Kultur. Zeitung des Deutschen Kulturrates. 29.03.2023. Im Internet unter: <https://politik-kultur.de/themen/kuenstliche-intelligenz/ohne-menschen-koennen-die-maschinen-nicht-lernen/> (aufgerufen am 08.11.2024).

I TRUST AI

THE FIFTH PHASE OF THE INTERPARES PROJECT

by *Luciana Duranti*

The goal of the InterPARES research project has been since 1998 to develop theory and methods for preserving the trustworthiness of digital records across technologies over the long term.¹ The fifth phase of InterPARES (2021–2026), *I Trust AI*, investigates how to design, develop, and leverage Artificial Intelligence to support the ongoing availability and accessibility of trustworthy public records. We aim to

- Identify specific AI technologies that can address critical records challenges;
- Determine the benefits and risks of using AI technologies on records; and
- Ensure that records concepts and principles inform the development of responsible AI.

Our short-term approach focuses on identifying high impact problems and limitations in records and archives functions, and applying AI to improve the situation. Our long-term approach aims to identify the tools that records and archives specialists will need in the future to flexibly address their ever-changing needs. This includes decision support and, once decisions are made, rapid implementation of AI-based solutions to those needs.

The *I Trust AI* project is a multinational interdisciplinary endeavour involving 36 countries in 4 continents and more than 200 researchers in several disciplines, from records and archives management to forensics, law, computer science, robotics, policy, and ethics. This means that our first effort had to be to understand each other, starting with the language we use. For example, archival professionals talk about records, while computer scientists and AI professionals talk about data. To archivists, data “are” the smallest meaningful unit of information in a record. To an AI specialist, data “is” organized information (possibly in a database), be it facts or not, regardless of size, nature and form. Thus, key to our work have been AI tutorials and workshops for non-AI researchers, and archival and diplomatics theory tutorials for non-archival researchers. These educational endeavours are supported by the Terminology Database that is developed in collaboration by a multidisciplinary team.²

The *I Trust AI* case studies, which are organized by archival function, from creation to preservation and access, are car-

ried out by teams, each of which includes researchers from different disciplines, type of organization, and countries, and are supported by labs where tools are designed, developed and tested, such as the Linguistics Lab at the University of British Columbia, directed by Dr. Prof. Muhammad Abdul Mageed, and the Vision Robotic Artificial Intelligence Lab at the University of Macerata, directed by Dr. Prof. Emanuele Frontoni. However, we also carry out general studies across functions, such as the study on “AI Competencies: from training programs to market needs.” The Linguistics lab has carried out studies and produced tools using Natural Language Understanding to facilitate named entity recognition (e. g. names of persons, places, and organizations within a text), sentiment analysis (e. g. to identify biases in a text), and speech tagging (e. g. determining the category of words). The lab has also used Natural Language Generation, enabling machines to create summaries of texts, to translate text from a language to another, and to provide answers to queries. Finally, the lab uses Natural Language Processing, such as Named Entity Recognition, to enable records classification, to enhance the engagement of users, and to support arrangement and description.

As we know that Large Language Models (LLMs) produce Hallucinations, that is, incorrect text, the UBC lab has also worked on producing a tool that avoids Hallucinations in Natural Language Processing. It uses Retrieval Augmented Generation (RAG), a technique that enhances LLMs by allowing them to extract relevant and timely information from various external databases as they work. This way, RAG enables the models to produce more accurate and reliable text that is supported by actual data. The researcher responsible for this project, El Moatez Billah Nagoudi, has applied this tool to archival literature and demonstrated that the model provides responses that are from verifiable sources and accurate. In addition, the lab has used RAG to create a tool for detoxifying content across various platforms while maintaining the original meaning (GreenLLaMA). Finally, the lab is working with Generative AI, at the core of which are LLMs trained on vast datasets.³

While the research carried out in the Labs aims to create tools that would support multiple studies, they often start with the intent to create a tool for a specific case and end up being useful to several cases and fulfilling different purposes. One example is the Perganet study. The purpose of the study was to develop a tool capable of identifying the key attributes of thousands of digitized parchments issued by city notaries in medieval times. The study, led by Emanuele Frontoni, used computer vision, a field of AI that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs and to take actions or make recommendations based on that information. The basic feature chosen to be identified for the research was the signum, a specific and personally drawn mark used by a single notary in the top part of the record and before his signature at the bottom of the record. Identifying the signum meant that every notary could be recognised and tracked in a virtually infinite series of documents. The AI made possible the creation of a register of signa, which would in fact be a registry of notaries, and the investigation of less visible features of the parchments.⁴ Once fully developed, this Deep Learning tool could be used in a wide range of applications, such as identification of the system of writing of individual authors, analysis of archival annotations on the back of the documents, recognition of recurring images or other features in large series of documents and of common patterns in manuscript maps or drawings, and others.

Another study taking advantage of the tools developed in the labs is led by Eng Sengsavang, the archivist of UNESCO, and aims to create metadata for a series of digitized interviews, speeches, press conferences, and various types of radio programmes (about 16,000 in total, of which only 800 have been described), which were created between the 1950s and the 1980s. This material is in more than 70 recognised languages, and 4% of the tapes is in multiple languages. Languages need to be identified and the talks transcribed, and the 57 metadata elements included in the UNESCO schema have to be extracted from such transcriptions. Difficulties related to speech-based records are the accent of the speaker, the way the recording is set up, and the short audio clips. Thus, much of the work focused initially on finding the right technology for identifying and transcribing the languages, especially those that are not commonly used, and then building a data set large enough, developing the model and testing it. This work was led by a doctoral student in the UBC Linguistics lab, Peter Sullivan, working with the Deep Learning & Natural Language Processing Group. For metadata identification, the team used diplomatics methodology to label the structural components of the audio recordings, as all the key identity metadata (e. g. names of persons, subject matter, date, etc.) are found in the same section of the same record genre. Diplomatics methodology is also used by other studies, such as the one on “Datafying Archives for Privacy Protection,” the purpose of which is to use labeling of the records intrinsic elements for sensitivity detection. Artificial Intelligence has been used in the past to block out names and dates, but we know that names and dates in most documents are not covered by privacy. This study aims to produce a tool capable of recognizing the types of records that are likely to contain sensitive information and the place in these records whe-

re such information might be. At the same time, Victoria Lemieux is conducting a study that uses existing Distant Reading technology, as well as a combination of such technology and Privacy Preserving Federated Machine Learning to protect personal private information in records.⁵

We know that AI tools, especially Generative AI, have had a strong impact on the trustworthiness of images, or lack thereof. Two *I Trust AI* studies are relevant to such matter, specifically the “Recordkeeping practices of creators using AI to generate images” and “Increasing access to photos, videos and social media records through AI-generated descriptive metadata”. Generative AI tools include image generators such as DALL-E, Midjourney and Adobe Firefly. Jessica Bushy leads those studies and has investigated the impact of the use of such tools in the fields of medicine, law enforcement, journalism and media communication. She found that the primary concerns in each of those fields were authenticity and its verifiability, manipulation and misinformation, bias and representation, attribution and intellectual property, transparency and explainability, and ethics.⁶ In the course of these investigations she has studied how image metadata can assist archivists in identifying images produced using AI and determining their provenance. She found that journalism and the technology sector are developing a framework that supports the capture of provenance data when taking pictures and videos, and creates a secure metadata profile for digital images (including those generated by AI) that cannot be altered and can be used to verify the authenticity of the images. Her research analyses the work of the Coalition for Content Provenance and Authenticity, C2PA, an open specification for attributing and asserting provenance for born-digital content that is disseminated online, and on emerging image metadata specific to AI-generated images, through the application of archival diplomatics.⁷

Images are also the object of *I Trust AI* studies based on data science, such as, for example, the study on Digital Twins. This study aims to understand how a digital twin is used and

¹ The first four projects can be accessed here: <http://www.interpares.org>.

² These materials are posted on the *I Trust AI* website at <http://www.interparestrustai.org>.

³ Much of this work has not yet reached the publication stage, as it is still being tested. A couple of articles can provide an idea of the work in the lab: Chen, W. R., Adebara, I., Doan, K. D., Liao, Q., & Abdul-Mageed, M. (2023). Fumbling in Babel: An Investigation into ChatGPT’s Language Identification Ability. In: Proceedings of the 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics. arXiv preprint arXiv:2311.09696 (2024). <https://arxiv.org/pdf/2311.09696.pdf> (Accessed on 09.10.2024); and Khondaker, M. T. I., Abdul-Mageed, M., & Lakshmanan, L. V. GreenLLaMA: A Framework for Detoxification with Explanations. In arXiv preprint arXiv:2402.15951 (2024). <https://arxiv.org/pdf/2402.15951.pdf> (Accessed on 09.10.2024).

⁴ See https://link.springer.com/chapter/10.1007/978-3-031-13324-4_25 (Accessed on 30.09.2024)

⁵ Victoria Lemieux, John Werner: Protecting Privacy in Digital Records: The Potential of Privacy-Enhancing Technologies. In: ACM Journal on Computing and Cultural Heritage (JOCCH), 16/4 (2024), pp. 1–18.

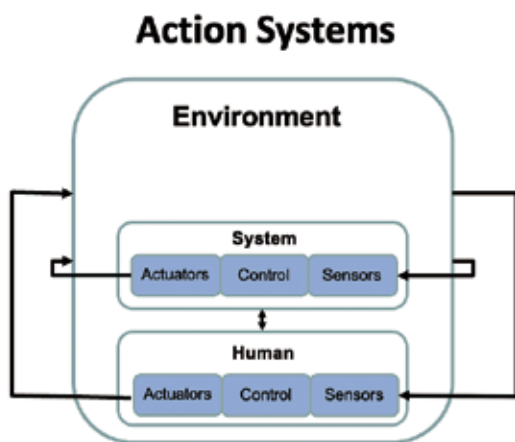
⁶ Jessica Bushy: AI-Generated Images as an Emergent Record Format. In: IEEE International Conference on Big Data (BigData). Sorrento, Italy (2023), pp. 2020–2031.

⁷ This work is in course of publication.

created so that it can be preserved over the long term. A digital twin is an ecosystem of multi-dimensional and interoperable subsystems made up of

- physical things in the real-world,
- digital versions of those real things,
- synchronized data connections between them and the people, organizations and institutions involved in creating, managing, and using these.⁸

As described by Cameron and Hamidzadeh, digital twin systems are “examples of cybernetic feedback-based systems. Common applications include traffic management, predictive maintenance systems, security, infrastructure, and manufacturing.



Examples of digital twin applications being developed or operationalized proliferate. For instance, Vancouver's YVR Airport uses a digital twin system to direct ground traffic, optimize pedestrian movement flows within the airport, and manage the building's maintenance and climate control systems.⁹

A team led by Tracey Lauriault studies the preservation of the Carleton University “digital twin prototype developed by the Carleton Immersive Media Studio (CIMS) for the real-time visualization of complex datasets to run simulations, increase efficiency of facility management and operations, and support University outreach activities such as virtual experiences for students.”¹⁰

Among the general studies that affect all archival functions, some are close to completion, such as the Comparative Assessment of Ethical Codes of the Archival/Records Management and Artificial Intelligence Communities,¹¹ the Educational Case Studies of Applications of AI to Archives and Records Management,¹² and the study of „Archival users and AI tools for reference and access”.¹³

The general study which will probably affect the most the archival institutions that wish to use AI tools to carry out traditional archival functions is the one on “Paradata.” The purpose of the study is to develop an approach for documenting the AI process that fulfills the need for archival accountability in the use of AI, thereby also supporting the authenticity of the outcome. The study is carried out by Pat Franks, Babak Hamidzadeh, Scott Cameron, Norman Moora-

dian, Alex Richmond, and Mario Beauchamp. Three articles have been published on its progress.¹⁴

It is important to understand that Accountable AI is different from Explainable AI as the latter focuses on why a given tool produced a given output from a given set of inputs. Building accountable AI must also consider the individuals, organization, and environment in which the AI tool operates, and paradata is necessary to explain why, how, by whom, and to what effect a given tool was used in a particular context. This short description of the study's current outcomes is based on the three articles mentioned earlier.

The I Trust AI definition of paradata is as follows: “Paradata is information about the procedure(s) and tools used to create and process information resources, along with information about the operation of the tools and the execution of the procedures, and about the persons carrying out those procedures and using the tools.”¹⁵ Paradata can be technical or organizational. The former include a description of the AI Model after being tested & selected, the metrics related to its evaluation and performance, the logs generated in doing so, the model training data set, the training parameters for the model, and, if it is a commercial tool, the vendor documentation, and the versioning information. The organizational paradata include the AI institutional policy, the plans for the design of the tool, the training received by the employees using the tool, the ethical considerations that were made when the tool was designed, the assessment of the impact of the use of the tool, the process of using the tool, and all the regulations put in place for the purpose of using the AI tool.

Among the types of tools that can be possibly used in an archives, the most useful would be an action system tool, one where system and human beings interact, as is the case with the digital twins described earlier. Because this kind of system is dynamic, the identification of the technical paradata may present issues related to the fact that the data must be recorded in real time, a particularly complex thing when humans' and machine's actions are concurrent rather than taken one at a time. Thus, with this kind of system, documentation and recording should be a mix of continuous, sampled, and event-based data captured at predefined trigger points. The documentation should follow a feedback cycle model, describing what the sensor recorded, what the control determined should be the consequent action, what the actuator action was, and what the sensors captured as consequence of the action of the actuator to document each of these phases. Association between what is sensed and how it is acted on, and control logic used must be documented. Also the temporal dimension has to be captured and documented. Some of the documentation will be automatic as part of the AI system or tool; some will be human-created prior to or after the creation and implementation of the AI system or tool (if the tool is commercial, the documentation will have to be provided by the vendor; if it is developed by the archives, it will have to be provided by its archives personnel). A substantial part of the human created documentation will be the Ethical Code of the archives, its AI policy, the qualifications and training of its personnel, regulatory requirements, and the process and procedures for either developing or choosing the tool, testing it, evaluating its performance, selecting the data

sets for training and those for evaluation, implementing the process, and assessing the outcome.

In conclusion, paradata is necessary to document the AI process and promote transparency and accountability, because the capture and preservation of paradata ensures that the AI process is documented in a way that preserves the authenticity of records but also supports their preservation.

The I Trust AI project began in 2021 and it has two more years of research time before completion. Technologically speaking, two years are equivalent to at least a couple of decades of the past century. We aim to complete the studies that are in course, but, if another mind boggling innovation were to hit us in the coming months, we will try to make sense of it in light of all the research completed to date and always in the respect of the fundamental archival concepts and principles underlining the continuing trustworthiness of records and their long term preservation.

This article outlines the goal, objectives, methodology, and progress to date of “I Trust AI”, the fifth phase of the InterPARES Project, a multinational, multidisciplinary collaborative research endeavour that began in 1998 with the purpose of ensuring the long term preservation of trustworthy records across technologies. It then presents some of the I Trust AI case studies and general studies, and directs the reader to key publications that have resulted from them.

Luciana Duranti

The University of British Columbia
Vancouver Campus | Musqueam Traditional Territory
#470 -1961 East Mall | Vancouver, British Columbia |
V6T1Z1 Canada
E-Mail: luciana.duranti@ubc.ca

- ⁸ See: <https://canadasdigitaltwin.ca/about-2-2/>. (Accessed 09.10.2024)
- ⁹ Scott Cameron and Babak Hamidzadeh: Preserving paradata for accountability of semi-autonomous AI agents in dynamic environments: An archival perspective. In: Elsevier, Telematics and Informatics Reports 14 (2024), pp.1–15, here p. 5.
- ¹⁰ See: <https://canadasdigitaltwin.ca/about-2-2/>. (Accessed 09.10.2024)
- ¹¹ See: <https://interparestrustai.org/assets/public/dissemination/MA01-ComparativeEthicsStudyPhase1ReportinclAppendix1-3-FINAL.pdf>. (Accessed 09.10.2024).
- ¹² See https://interparestrustai.org/trust/research_dissemination, Curriculum Materials. (Accessed 09.10.2024)
- ¹³ See Pierluigi Feliciati: Archival users and AI tools for reference and access: a study within the InterPARES Trust AI project. In: JLIIS.it Italian journal of Library Science, Archival Science and Information Science. 14, 3 (2023). <https://www.jlis.it/index.php/jlis/article/view/558> (Accessed 09.10.2024).
- ¹⁴ Pat Franks, Babak Hamidzadeh e Jeremy Davet: Archivist in the machine: paradata for AI-based automation in the archives. In: Archival Science, 23 (2023), pp. 275–295; Scott Cameron, Pat Franks, Babak Hamidzadeh: Positioning paradata: A conceptual frame for AI processual documentation in archives and recordkeeping contexts. In: ACM. Journal on Computing and Cultural Heritage (2023), see <https://interparestrustai.org/assets/public/dissemination/positioningparadata04-20preprintv21.pdf> (Accessed 09.10. 2024); and Scott Cameron and Babak Hamidzadeh: Preserving paradata for accountability. Cited.
- ¹⁵ Pat Franks, Babak Hamidzadeh, Jeremy Davet: Archivist in the machine: paradata for AI-based automation in the archives. Cited.

KÜNSTLICHE INTELLIGENZ IM ARCHIV

HERAUSFORDERUNGEN UND CHANCEN

von Martin Vogel

Die Entwicklung von Künstlicher Intelligenz (KI) hat in den letzten Jahren einen starken Aufschwung mit einer wachsenden Zunahme an KI-Modellen für verschiedene Bereiche erlebt. Dieser Artikel befasst sich mit den Herausforderungen, die Archive bei der Einführung und Benutzung von KI zu erwarten haben.

Es fließen Erkenntnisse aus der reinen Erprobung und Bereitstellung von KI-Modellen unter Verwendung von Docker und Jupyter Notebook im eigenen Haus sowie den fachlichen Diskussionen aus Workshops und Konferenzen mit ein, um die aktuellen Möglichkeiten und Chancen für die Benutzung von KI aufzuzeigen.

In diesem Zusammenhang werden erste Überlegungen angestellt, wie eine Ergebnissicherung und der Austausch mit anderen Archiven durchgeführt werden könnte.

Was versteht man unter KI und KI-Modellen?

KI bezieht sich auf die Entwicklung von Computern, die menschliche Intelligenz nachahmen können, indem sie Probleme lösen und Entscheidungen treffen. Im mathematischen Sinne kann KI als eine Methode betrachtet werden, um komplexe Funktionen zu modellieren oder zu approximieren, die dann in der Lage sind, bestimmte Aufgaben oder Probleme zu lösen.

KI-Modelle werden typischerweise durch maschinelles Lernen trainiert, einem Prozess, bei dem das Modell Daten aus einer bestimmten Quelle verwendet, um Muster und Beziehungen in diesen Daten zu erkennen. Der Trainingsprozess besteht darin, dass das KI-Modell iterative Anpassungen an seinen internen Parametern durchführt, um die Genauigkeit seiner Vorhersagen oder Entscheidungen zu verbessern. Das Ziel eines Sprachmodells wie z. B. bei einem Large Language Model (LLM) ist es, eine Wahrscheinlichkeitsfunktion zu lernen, die die Wahrscheinlichkeit einer bestimmten Sequenz von Wörtern oder Sätzen maximiert. Daher geht es bei KI und KI-Modellen, um die Maximierung von Wahrscheinlichkeiten und Probleme näherungsweise zu lösen.

1. HERAUSFORDERUNGEN FÜR ARCHIVE BEI DER EINFÜHRUNG VON KI

Die Einführung von KI in Archiven erfordert eine gründliche Vorbereitung.

Im Folgenden werden Aspekte aufgeführt, die im Vorfeld berücksichtigt werden sollten.

Heterogene Daten

In Archiven finden sich eine Vielzahl von Datentypen. Die Daten selbst sind ebenfalls sehr unterschiedlich (sowohl syntaktisch als auch semantisch). Die Herausforderung für Archive besteht darin, strukturierte wie unstrukturierte Daten zu verarbeiten. Einige Beispiele für unterschiedliche Datenformate sind Audiodateien, Bilddateien, Daten aus Fachverfahren (z. B. Personenstandregister, E-Akte, Geodaten), XML-Dateien und Datenbanken. Diese Vielfalt an Datenformaten macht es schwierig, vorherzusagen, welche Daten geeignet sind, um mit einer KI verarbeitet zu werden.

Datenschutz

Die Verarbeitung von Daten muss innerhalb des Archivs sichergestellt werden. Die Daten, die für das Archiv bereitgestellt werden, dürfen ausschließlich innerhalb der eigenen Räumlichkeiten oder einer geschützten Infrastruktur verarbeitet werden. Sie dürfen nicht mit externen Diensten wie z. B. ChatGPT, Microsoft CoPilot oder anderen Onlinediensten verwendet werden, da sie sensible, vertrauliche Informationen enthalten können. Diese Einschränkung wirkt sich auch auf die Nutzung von KI aus, da die technische Infrastruktur und die entsprechenden KI-Modelle nur innerhalb des Archivs betrieben werden können.

Kosten

Die Kosten für den Betrieb, gerade bei sehr großen KI-Modellen, können sehr hoch sein. Dazu tragen unter anderem die Beschaffung der notwendigen Infrastruktur bei, sowie deren eigenständiger Betrieb im Archiv.

Technische Infrastruktur

Die technische Infrastruktur beschränkt die Auswahl an möglichen KI-Modellen. Die Notwendigkeit, die technische Infrastruktur im eigenen Haus zu betreiben, setzt bestimmte Ressourcen voraus. Hierzu zählen insbesondere Personalressourcen, die über die notwendigen Kenntnisse und Fähigkeiten verfügen, um die eigene Infrastruktur aufzubauen, KI-Modelle einzurichten und deren Betrieb zu sichern.

Fehlende Anwendungsfälle

Die Vielfalt der Daten in einem Archiv macht es schwierig, vorab zu bestimmen, welche KI-Modelle für welche Daten geeignet sind. Da Angeboten und Abgaben sich stark voneinander unterscheiden können, ist es notwendig, Anwendungsfälle zu definieren, um die Daten mit KI nutzen zu können. Dabei ist auch zu beachten, dass die Ergebnisse und Resultate aus den entsprechenden Anwendungsfällen reproduzierbar sein sollten, um für weitere Angebote genutzt werden zu können.

Fehlende Erfahrungswerte

Die Welt der KI ist sehr groß, heterogen und unübersichtlich. Tagtäglich werden hunderte von neuen, trainierten KI-Modellen für spezielle Datentypen entwickelt und veröffentlicht. Hier besteht eine erhebliche Herausforderung darin einen Überblick zu bekommen, welche KI-Modelle sich für welche Anwendungsfälle und Daten eignen. Ebenso ist es schwierig, die Qualität einzelner KI-Modelle bei erstmaliger Verwendung im Voraus abzuschätzen. Dies liegt unter anderem daran, dass viele KI-Modelle für sehr spezielle Domänen und Datentypen entwickelt und trainiert wurden oder die deutsche Sprache nur rudimentär unterstützen. Daher ist es sehr schwierig, das richtige KI-Modell auszuwählen, dass den Anforderungen entspricht und zufriedenstellende Ergebnisse produziert.

Austausch von Erfahrungen zwischen Archiven

Da es aktuell noch wenig Erfahrungswerte bei der Verwendung von KI in Archiven gibt, ist es wichtig und notwendig, sich in dieser frühen Phase mit anderen Archiven auszutauschen. Jedes Archiv steht vor ähnlichen Herausforderungen. Daher sollte überlegt werden, eine einheitliche Beschreibung zur Ergebnissicherung, Erfahrungsaustausch und Nachnutzung (siehe Kap. 4.) von KI und KI-Modellen in Archiven zu etablieren.

2. ANFORDERUNGEN AN DIE TECHNISCHE INFRASTRUKTUR FÜR DEN BETRIEB VON KI

Im Folgenden werden die technologischen Aspekte erörtert, die für die Implementierung von KI relevant sind.

- Welches Budget steht zur Verfügung?
- Technisches Personal im Haus
- Schulung des Personals
- Single User/Multiuser
- Welche KI-Modelle will man einsetzen
- Folgekosten

Welches Budget steht zur Verfügung?

Die Kosten für die Infrastruktur, die benötigt wird, um KI im eigenen Haus zu betreiben, können sehr hoch sein. Um die Kosten gering zu halten, ist es ratsam, sich im Vorfeld mit den Anwendungsfällen auseinanderzusetzen, für die man die KI einsetzen möchte. Um erste Erfahrungen mit KI zu sammeln, reichen handelsübliche PCs oder Laptops (CPU ab I5, 16 GB RAM, GPU NVIDIA mit 8 GB VRAM) aus. Diese Ausstattung ermöglicht es, ohne Einsatz eines hohen Budgets erste Erfahrungen mit der KI zu sammeln. KI-Modelle benötigen für die Ausführung nicht zwingend Spezialhardware. Sie können auch auf PCs oder Servern ausgeführt werden, die über genügend RAM verfügen. Die Ausführungsgeschwindigkeit wird dadurch aber erheblich verringert.

Technisches Personal im Haus

Um KI im eigenen Haus zu betreiben benötigt man technisches Personal im Haus oder externe Experten, welche die Einrichtung und Betreuung der Systeme übernehmen. Dies muss bei der Budgetplanung mitberücksichtigt werden.

Schulung des Personals

Für die erfolgreiche Implementierung und Anwendung von KI benötigt man geschultes Fachpersonal. Um sicherzustellen, dass das Personal in der Lage ist, KI-Anwendungen erfolgreich einzusetzen, sollten erforderliche Schulungen (z. B. Programmieren in Python, Docker, LLM, KI in der Praxis, etc.) im Budget mit einkalkuliert werden.

Single User/Multiuser

Bei der Planung und Anschaffung der technischen Infrastruktur für KI muss entschieden werden, wie die Nutzung erfolgen soll. Da KI-Modelle oft große Speicherkapazitäten benötigen (RAM, VRAM), ist es wichtig, über die Art der Infrastruktur nachzudenken. Eine Option sind Einzelplatzrechner (Laptop, PC), die für eine Person ausgestattet werden und daher weniger Erweiterungsmöglichkeiten bieten. Eine andere Option sind Multi-User-Konfigurationen, bei denen ein Server zum Einsatz kommt, um mehreren Personen gleichzeitig Zugriff auf die KI-Anwendung zu ermöglichen. Bei gleichzeitiger Benutzung kann dies jedoch zu erhöhtem Speicherbedarf (RAM, VRAM) führen, was wiederum die Kosten für die Infrastruktur ansteigen lässt. Daher sollte im Vorfeld eine Prognose über den Bedarf und die Anzahl der nutzenden Archivar*innen erstellt werden.

Welche Modelle will man einsetzen

In der Welt der KI gibt es eine Vielzahl von unterschiedlichen KI-Modellen, die sich in ihren Anwendungen und Größen unterscheiden. Große Modelle benötigen oft mehr Speicher, liefern präzisere Ergebnisse, verursachen aber höhere Kosten für die Infrastruktur. Eine sorgfältige Auswahl der richtigen KI-Modelle ermöglicht es, Ressourcen effizient zu nutzen und Kosten zu sparen. Es ist daher entscheidend, sich mit den verschiedenen Modellen und Anwendungsbereichen (Computer Vision, Natural Language Processing, Audio) auf

z. B. (Hugging Face, 2024¹) vertraut zu machen. Nachfolgend betrachten wir beispielhaft zwei KI-Modelle von (LLaVA, 2024²) und (OpenAI, Whisper 2024³). Das KI-Modell (LLaVA, 2024) ist ein Multimodal Large Language Model (MLLM), das z. B. Bilder analysieren und beschreiben kann. Das KI-Modell liegt in unterschiedlichen Qualitätsstufen (98 Quantisierungen) vor. Je kleiner die Quantisierung (2bit statt 32bit) eines Modells ist, desto unpräziser werden die Ergebnisse. Das kleinste KI-Modell hat die Größe von 3,2 GB und das größte KI-Modell hat 69 GB. D. h. für den Betrieb des kleinsten KI-Modells benötigt man mindestens 3,3 GB RAM, VRAM und für das Größte mindestens 69 GB RAM, VRAM. Bei (OpenAI, Whisper 2024) zum Transkribieren von Audiodateien benötigt das kleinste KI-Modell (tiny) 1 GB RAM, VRAM und das Größte (large) benötigt 10 GB RAM, VRAM. Um verwertbare Ergebnisse zu bekommen, sollten die Modelle mit einer Quantisierung von mindestens 8 – 10 bit betrieben werden.

Folgekosten

Für den Betrieb von KI im eigenen Haus gehen Folgekosten einher, wie zum Beispiel:

- Kosten für weitere Hardware-Ausstattung (z. B. zusätzliche GPU Karten), um größere KI-Modelle nutzen zu können
- Stromkosten, die durch den Betrieb der Server entstehen
- Schulungskosten für Personal

3. ARCHIVFACHLICHE ÜBERLEGUNGEN FÜR DIE ANWENDUNG VON KI-MODELLEN

Im Folgenden werden Überlegungen aufgeführt, die archivfachlich im Vorfeld zu berücksichtigen sind.

Welche archivfachlichen Aufgaben sollen unterstützt werden?

Bevor man KI einsetzt, sollte überlegt werden, für welche Zwecke diese Technologie eingesetzt werden soll. Soll die KI zur Analyse, zum Generieren von Daten oder zur Unterstützung der alltäglichen Arbeit verwendet werden? Durch erste Überlegungen kann man bestimmen, welche archivfachlichen Aufgaben durch KI unterstützt werden sollen. Dazu gehören unter anderem z. B.:

- die Zusammenfassung von Dokumenten und Informationen,
- die Identifizierung von Personen in Bildern oder Videos,
- das Finden von Personen, Datumsangaben, Orten, etc. in Dokumenten,
- die Transkription von Audio- und Videoinhalten sowie die
- inhaltliche Beschreibung von Videodateien und Bildern.
- Angaben zu Dateiformaten (z. B. langzeitstabil)

Vorab-Analyse der Daten, die verwendet werden sollen

Durch eine sorgfältige Vorab-Analyse kann man feststellen:

- Welche Art von Daten vorhanden ist (z. B. Texte, Bilder, Videos, Audiodateien etc.),
- den Umfang der Daten, die verarbeitet werden sollen oder

- ob es sich um strukturierte/unstrukturierte Daten handelt. Diese Vorab-Analyse erleichtert die Auswahl der geeigneten KI-Modelle, um die erwarteten Ergebnisse zu erhalten.

Auswahl geeigneter KI-Modelle

Um sich für geeignete KI-Modelle zu entscheiden, muss eine Vorab-Auswahl der Daten erfolgen, die analysiert werden sollen. Als Anhaltspunkt kann z. B. die Kollaborationsplattform HuggingFace 2024 dienen. Hier sind frei verfügbare KI-Modelle entsprechend ihres Anwendungsgebiets (z. B. Computer Vision, Natural Language Processing, Multimodal, Audio) kategorisiert sowie die Speichergrößen der Modelle angegeben. Weiter können hier auch Programmbeispiele gefunden werden, wie die KI-Modelle auf der lokalen Hardware (z. B. mit Jupyter Notebook) in den eigenen Räumlichkeiten ausgeführt werden können.

Qualitätssicherung von Ergebnissen aus KI-Modellen

Wenn KI im archivischen Prozess verwendet wird, müssen die Ergebnisse einer Qualitätskontrolle unterzogen werden. Dies kann entweder durch Erfahrungswerte erfolgen, wenn man das KI-Modell bereits früher verwendet hat, oder es bedarf eines Prozesses (händisch/automatisiert) und klar definierter Kriterien für die Qualitätssicherung. Wenn ein neues KI-Modell zum ersten Mal verwendet wird, sollte man Daten wählen, bei denen man schon Ergebnisse hat. Durch den Vergleich der erwarteten Ergebnisse mit den tatsächlichen Ergebnissen kann beurteilt werden, wie zuverlässig das KI-Modell arbeitet und ob es eingesetzt werden sollte.

Verwendung der Ergebnisse aus der Analyse

Nachdem die Ergebnisse aus der Analyse vorliegen, muss festgelegt werden, wie mit ihnen umgegangen werden soll. Es gibt mehrere Möglichkeiten:

Unterstützung: Die Ergebnisse dienen lediglich der Unterstützung bei Aufgaben wie z. B. Finden von

- Personen
- Orten
- Datumsangaben

und ermöglichen dadurch einen gezielten Zugriff auf diese Daten in Dokumenten, Videos und Bildern ohne langes Suchen.

Erschließung: Die Ergebnisse sollen in die Erschließung mit einfließen. Wenn dies der Fall ist, muss überlegt werden, wie diese Ergebnisse als „KI-Produkte“ kenntlich gemacht werden sollen und wie sie in die Archivischen Fachinformationssysteme übernommen werden können.

Aufwand der Vorbereitung

Der Aufwand der Vorbereitung für die Anwendung von KI hängt stark von den spezifischen Anforderungen der Daten ab. Es ist wichtig, eine erste Analyse der Daten durchzuführen, um zu verstehen, ob sie überhaupt für die Anwendung von KI geeignet sind. Dabei sind folgende Aspekte zu berücksichtigen:

- Erwartetes Ergebnis: Welches Ergebnis ist mit der Anwendung von KI zu erwarten und inwieweit unterstützt es den Prozess?

- Qualitätssicherung: Wie wird das Ergebnis qualitätsgesichert? Welche Möglichkeiten gibt es, die Richtigkeit und Vollständigkeit des Ergebnisses zu überprüfen und Fehler zu erkennen. Wie präzise müssen die Ergebnisse sein?
- Umgang mit Ergebnissen: Wie werden die Ergebnisse verwendet und interpretiert?

Wenn der Aufwand der Vorbereitung sehr hoch ist, zum Beispiel bei Dateisammlungen, extrem unstrukturierten oder proprietären Daten, kann es durchaus sein, dass sich der Einsatz von KI nicht lohnt. In solchen Fällen ist es sinnvoller, alternative Lösungen zu suchen oder bewährte Methoden zur Datenanalyse anzuwenden.

4. ARCHIV-COMMUNITY UND ERGEBNISSICHERUNG

Im Folgenden werden erste Überlegungen vorgestellt, die für die Weitergabe von KI-Modellen innerhalb der Archiv-Community und zur Ergebnissicherung relevant sein können. Die detaillierte Beschreibung dieser Punkte ist nicht Gegenstand dieses Artikels und soll in 2025 auf Konferenzen und Workshops vorgestellt werden.

Überlegungen

- Welche Zielsetzung wurde verfolgt?
- Welche Daten wurden verwendet?
- Welche KI-Modelle wurden verwendet?
- Welche Fehlertoleranz wurde akzeptiert?
- Wie wurden Fehler festgestellt?
- Welche Parameter wurden für die KI-Modelle verwendet?
- Für welche Phase des archivischen Prozesses wurden die KI-Modelle eingesetzt (SIP, AIP, DIP, Magazin, Nacherschließung etc.)
- Hat das Ergebnis die erwarteten Anforderungen erfüllt?
- Wie hoch war der Aufwand (technisch, organisatorisch)
- Welche technische Infrastruktur wurde verwendet?
- Gibt es ein Code-Beispiel (z. B. Jupyter Notebook) zur Reproduktion/Nachnutzung?
- Welche Prompts wurden benutzt?

5. BEISPIEL AUS DER PRAXIS: DER EINSTIEG IN DIE KI

Um erste Erfahrungen mit KI-Modellen zu sammeln, wurde 2022 ein Workshop im Bereich Natural Language Processing besucht, um den technischen als auch fachlichen Hintergrund für die Verwendung von KI zu verstehen. Die ersten Tests mit KI-Modellen wurden dann auf lokal bestehender Infrastruktur durchgeführt. Hierfür wurde die Software Docker (Docker 2024¹) verwendet, um die entsprechenden KI-Modelle ausführen zu können. Für die Anwendung der KI-Modelle wurden zwei Webanwendungen genutzt, die in Docker ausgeführt werden können: Jupyter Notebook (Jupyter 2024²) mit Python für nicht generative KI-Modelle und Open WebUI (Open WebUI 2024³) für generative Large Language Models (LLM) und Multimodal Large Language Models (MLLM).

Die nachfolgenden KI-Modelle konnten erfolgreich auf den beiden Testumgebungen in den eigenen Räumlichkeiten ausgeführt werden. Für die ersten Tests wurden kleine

Verzeichnungen mit Bildern (Anzahl 38) (Niedersächsisches Landesarchiv 2017⁷) und Audiodateien (Anzahl 1, Länge 30 Minuten) (Niedersächsisches Landesarchiv, 2013⁸) ausgewählt.

Testumgebung 1 (Server):

- Intel Xenon (4 CPUs)
- 24 GB RAM
- Keine Grafikkarte (GPU Unterstützung)
- Langsame Ausführung der KI-Modelle

Testumgebung 2 (Laptop, handelsüblich, Preis unter 1.000,00 €):

- I5 (8 CPUs)
- 16 GB RAM
- NVIDIA RTX 4060 (8 GB VRAM, 3072 Cuda Kerne)
- Signifikante Erhöhung der Ausführungsgeschwindigkeit der KI-Modelle

Objekterkennung

- Input: Bilddokumentation der Altstadt während der Sanierung (Niedersächsisches Landesarchiv 2017)
- Modell: facebook/detr-resnet-50 (Hugging Face, AI at Meta 2024a⁹)
- Output: Liste mit erkannten Objekten

Segmentierung

- Input: Bilddokumentation der Altstadt während der Sanierung (Niedersächsisches Landesarchiv 2017)
- Modell: nvidia/segformer-b1-finetuned-cityscapes-1024-1024 (Hugging Face, NVIDIA 2024¹⁰)
- Output: Liste mit erkannten Objekten

Transkription

- Input: „Warum Umweltschutz“ von Prof. Dr. Adolf Brauns vom 13.11.1975 (Niedersächsisches Landesarchiv, 2013)
- Modell: OpenAI / Whisper (large, turbo) (OpenAI, Whisper 2024)
- Output: Textdatei mit der Transkription

- 1 Hugging Face (2024): The AI community building the future. <https://huggingface.co/models> (abgerufen 15.11.2024)
- 2 LLaVA (2024): <https://ollama.com/library/llava> (abgerufen: 15.11.2024)
- 3 OpenAI, Whisper (2024): Modell: OpenAI / Whisper (large) <https://github.com/openai/whisper> (abgerufen: 15.11.2024)
- 4 Docker (2024): Docker: Accelerated Container Application Development <https://www.docker.com/> (abgerufen: 15.11.2024)
- 5 Jupyter (2024): jupyter.org/ (abgerufen: 15.11.2024)
- 6 Open WebUI (2024): <https://github.com/open-webui/open-webui> (abgerufen: 15.11.2024)
- 7 Niedersächsisches Landesarchiv, NLA AU Rep. 227/24 acc. 2017/23 Nr. 58 (2017): Bilddokumentation der Altstadt während der Sanierung <https://www.arcinsys.niedersachsen.de/arcinsys/detailAction.action?detailid=v10623371> (abgerufen: 15.11.2024)
- 8 Niedersächsisches Landesarchiv, NLA WO 41 C Nds Zg. 2013/088 Nr. 380 (2013): Tonbandaufnahme „Warum Umweltschutz“ von Prof. Dr. Adolf Brauns vom 13.11.1975 <https://www.arcinsys.niedersachsen.de/arcinsys/detailAction.action?detailid=r11440416> (abgerufen: 15.11.2024)
- 9 Hugging Face, AI at Meta (2024a): Modell facebook/detr-resnet-50 <https://huggingface.co/facebook/detr-resnet-50> (abgerufen: 15.11.2024)
- 10 Hugging Face, NVIDIA (2024): Modell: nvidia/segformer-b1-finetuned-cityscapes-1024-1024 <https://huggingface.co/nvidia/segformer-b1-finetuned-cityscapes-1024-1024> (abgerufen: 15.11.2024)

Natural Language Processing

- Input: Ergebnis der Transkription (Textdatei)
- Modell: Spacy / de_core_news_lg (Hugging Face, spaCy 2024¹¹)
- Output: Liste mit erkannten Personen, Orten, Organisationen

Zusammenfassung

- Input: Ergebnis der Transkription (Textdatei), Verschiedene Textdokumente
- Modell: Falconsai/text_summarization (Hugging Face, Falcons.ai 2024¹²)
- Output: Textdatei mit einer Zusammenfassung der Transkription / der Textdokumente

Large Language Models

- Input: Dokumente, Audiodateien
- Modelle: llama3.2 (3b), llama3.1 (8b), mistral (7b) (Llama library 2024¹³)
- Output: Zusammenfassung von Texten, Transkription, Generative Texte

Multimodale Large Language Models

- Input: Bilddateien
- Modell: llava (7b) LLaVA (2024)
- Output: Inhaltliche Beschreibung der Bilder

Hauptziel war es, erste Erfahrungen mit KI-Modellen zu sammeln und herauszufinden, welche Daten wie verarbeitet werden können und welche minimale Infrastruktur notwendig ist, um KI-Modelle im eigenen Haus zu betreiben. Alle oben genannten KI-Modelle konnten erfolgreich auf beiden Testsystemen ausgeführt werden. Die Ausführungsgeschwindigkeit auf dem Testsystem 2 ist gegenüber dem Testsystem 1 signifikant schneller. Beispielsweise konnte das Transkribieren einer Audiodatei von ca. 30 Minuten Länge auf dem Testsystem 1 in etwa 2,5 Stunden und auf dem Testsystem 2 innerhalb von nur 8 Minuten abgeschlossen werden. Diese Ergebnisse zeigen, dass der Einstieg in die KI mit überschaubaren Kosten möglich ist.

6. FAZIT

Die Verwendung von KI stellt Archive vor viele Herausforderungen. Die wichtigsten Aspekte sind der Datenschutz (Verarbeiten der Daten nur im Archiv) sowie die Diversität der zu bearbeitenden Daten. Alle unter Kap 5. aufgeführten Beispiele konnten lokal auf den Testumgebungen ausgeführt werden. Da es sich bei den Daten um sehr gängige Dateitypen (Bild/Audio) handelte, gab es zudem gute KI-Modelle für die Analyse. Durch die Entwicklungen der letzten zwei

Jahre ist es mittlerweile möglich, mit sehr geringen Kosten für die Infrastruktur und der leichten Nutzung von KI-Modellen (auf Hugging Face 2024 gibt es zu fast jedem KI-Modell ein Beispiel, dass in (Jupyter 2024) lokal ausgeführt werden kann) erste Erfahrungen zu sammeln. KI-Modelle können lokal mit echten Daten ausprobiert werden. Dadurch wird es ermöglicht, Erfahrungen mit den unterschiedlichen KI-Modellen zu machen und neue Ideen für die Anwendung von KI im Archiv zu entwickeln. Gerade dieser sehr niedrigschwellige Zugang ermöglicht es, die Chancen, die KI für Archive bieten kann zu erkunden. KI hat durchaus die Chance, bestehendes Archivgut besser zugänglich zu machen z. B. durch nachträgliche Analysen oder Archivar*innen bei zukünftigem Archivgut zuzuarbeiten (Transkribieren, Finden von Personen/Datumsangaben in Dokumenten, etc.), um diese zu unterstützen. Da alle Archive vor ähnlichen Herausforderungen bei der Nutzung von KI stehen, sollte ein regelmäßiger Austausch zwischen ihnen erfolgen, um gemeinsam diese neue Technologie einzuführen und zu nutzen.

ARTIFICIAL INTELLIGENCE IN ARCHIVES: CHALLENGES AND OPPORTUNITIES

The development of artificial intelligence has experienced a strong upswing in the last two years. The possible use of this technology presents archives with new technical and organizational challenges. However, it is now possible to get into the topic at a low threshold using commercially available computers and many available examples and thus approach the challenges. The exchange of experiences between the archives is essential here.

Martin Vogel

Niedersächsisches Landesarchiv
Am Archiv 1, 30169 Hannover
Tel. +49 511 120-6666
E-Mail: martin.vogel@nla.niedersachsen.de

¹¹ Hugging Face, spaCy (2024): Modell: Spacy / de_core_news_lg https://huggingface.co/spacy/de_core_news_lg (abgerufen: 15.11.2024)

¹² Hugging Face, Falcons.ai (2024): Modell: Falconsai/text_summarization https://huggingface.co/Falconsai/text_summarization (abgerufen: 15.11.2024)

¹³ Llama library (2024): <https://ollama.com/library> (abgerufen: 15.11.2024)

EINFLUSS VON BILDFORMATEN AUF DIE KI-BASIERTE ANALYSE VON DIGITALISIERTEN DOKUMENTBESTÄNDEN

*von Oliver Tüselmann, Fabian Wolf, Tim Raven
und Gernot A. Fink*

1. EINLEITUNG

Die Bereitstellung von Archivalien verlagert sich zunehmend in den digitalen Raum. Neben nativen digitalen Unterlagen werden durch Digitalisierungsprojekte auch analoge Archivbestände in digitale Formate transformiert. Diese Entwicklung bietet klare Vorteile für die Online-Bereitstellung von Archivobjekten. Zusätzlich eröffnet die Digitalisierung neue Möglichkeiten für die automatisierte Analyse und Bereitstellung der Daten. Mit dem Fortschritt der Künstlichen Intelligenz (KI) werden die Möglichkeiten zur Analyse digitalisierter Archivbestände noch weiter zunehmen.

Bei der Digitalisierung und Archivierung sollte die zukünftige Nutzbarkeit von KI-Modellen bedacht werden. Trotz des geringen physischen Platzbedarfs können digitale Bestände hohe Speichieranforderungen mit sich bringen. Daher stellt sich beim Scanprozess die Frage, in welcher Qualität digitale Abbilder gespeichert werden sollen. Eine höhere Qualität und Auflösung bietet mehr Informationsgehalt, aber auch einen höheren Speicherbedarf. Neben der Bildauflösung spielen das Dateiformat und der eingesetzte Kompressionsalgorithmus eine wichtige Rolle. Dabei lässt sich durch Kompression der Speicherbedarf erheblich reduzieren. Dabei ist eine wichtige Fragestellung, welche Auswirkungen Kompressionsverluste auf die Analysefähigkeit durch (zukünftige) KI-Modelle haben. Im Folgenden werden zunächst die Grundlagen der Bildkompression und anschließend drei zentrale Aufgaben der Dokumentenanalyse untersucht, die üblicherweise durch KI-Techniken gelöst werden. Der Fokus liegt dabei auf der Frage, wie verschiedene Dateiformate und Kompressionsqualitäten die Leistungsfähigkeit dieser spezialisierten KI-basierten Modelle beeinflussen.

2. BILDFORMATE UND KOMPRESSION

Die am weitesten verbreiteten Bildformate sind das Tagged Image File Format (TIFF), Portable Network Graphics (PNG) und das Joint Photographic Experts Group Format (JPEG). Diese Formate unterscheiden sich vor allem in den Möglichkeiten zur Speicherung von Bildinformationen und den eingesetzten Kompressionsmethoden.

Das TIFF-Format wurde entwickelt, um detaillierte Scans zu speichern. Es nutzt meist die verlustfreie Lempel-Ziv-Welch-Kompression¹, bei der zunächst ein Wörterbuch der häufigsten Zeichenketten erstellt wird und auf dessen Basis anschließend die Zeichenketten im Bild durch die Codes ersetzt werden. So lässt sich Speicherbedarf sparen, ohne Bildinformationen zu verlieren.

Das PNG-Format wurde hingegen für die digitale Nutzung optimiert und bietet im Vergleich zu TIFF eingeschränkte Speichermöglichkeiten. Zum Beispiel wird der CMYK-Farbraum, der im Druck oft verwendet wird, nicht unterstützt. Die weniger komplexen Speichermöglichkeiten führen zu kleineren Dateigrößen. Auch PNG verwendet eine verlustfreie Kompression, basierend auf einer Erweiterung des Lempel-Ziv-Welch-Ansatzes.

JPEG hingegen verwendet üblicherweise eine verlustbehaftete Kompression. Dabei wird das Bild in 8x8-Pixel-Blöcke unterteilt und diese werden einzeln komprimiert. Die Kompression erfolgt mithilfe der diskreten Kosinustransformation, die Bilddetails durch sogenannte Kosinuskurven abbildet. Dabei entstehen 64 Basisformen, die kombiniert

¹ Tilo Stutz: Bilddatenkompression: Grundlagen, Codierung, Wavelets, JPEG, MPEG, H. 264, HEVC Springer Vieweg, 2017.



Abbildung 1: Visuelle Beispiele für Bildqualitäten der unterschiedlichen Kompressionsraten auf einem Ausschnitt eines Volkszählungsbogens aus dem Jahr 1950.

werden. Durch Reduktion der genutzten Basisformen kann der Speicherbedarf weiter verringert werden, was jedoch den Informationsverlust erhöht.

Die Speicheranforderungen der drei Formate variieren stark. Beispielhaft wird an dieser Stelle ein großformatiges Dokumentenbild der Volkszählung von 1950 betrachtet. Die Scans dieser Formulare besitzen eine Auflösung von 8121 x 5050 Pixeln und benötigen 89.5 MB im TIFF-Format und 67.2 MB im PNG-Format. Im JPEG-Format bei einer Kompressionsrate von 100 sinkt der Speicherbedarf weiter auf 32.8 MB. Bei dem JPEG-Format ist sogar eine weitere Reduktion des Speicherbedarfs möglich, indem die Kompression noch weiter erhöht wird. Dabei kann die Bildqualität und Dateigröße durch die sogenannte Kompressionsrate, einem numerischen Wert zwischen 100 (minimale Kompression, höchste Qualität) und 1 (maximale Kompression, niedrigste Qualität), gesteuert werden. Die Auswirkungen der Kompression auf den Speicherbedarf und die Bildqualität lassen sich in den Abbildungen zu den Kompressionsraten (Abbildung 2) und der visuellen Qualität (Abbildung 3) nachvollziehen.

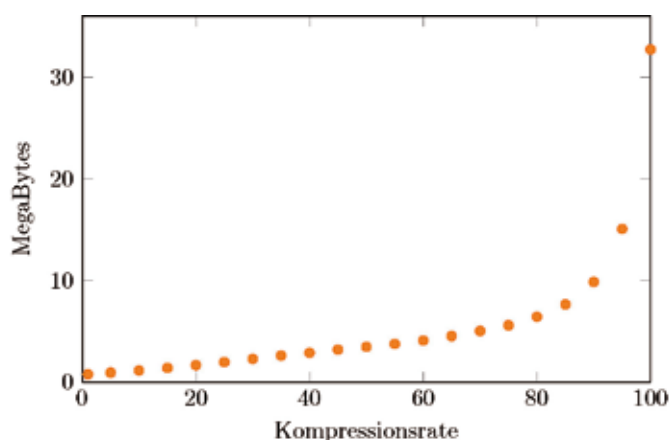


Abbildung 2: Einfluss der Kompressionsrate auf die Dateigrößen von JPEG-Bildern auf Basis eines im TIFF-Format digitalisierten Volkszählungsbogens aus dem Jahr 1950

3. HANDSCHRIFTERKENNUNG

Die Handschrifterkennung (engl.: Handwritten Text Recognition, HTR) ist eine Schlüsseltechnologie für die automatisierte Erschließung und Analyse von handschriftlich verfassten Dokumentenbildern. Mithilfe von HTR können handschriftliche Texte aus Digitalisaten in maschinenlesbare Formate umgewandelt werden. Diese Technologie ersetzt beziehungsweise erleichtert den manuellen Transkriptionsprozess. HTR leistet damit einen wesentlichen Beitrag zur effizienten Informationsbeschaffung und Nutzbarmachung digitalisierter Dokumente und eröffnet neue Möglichkeiten für umfassende Analysen und Erkenntnisse in Dokumentenbeständen mit handschriftlichen Inhalten. Aufgrund der hohen Praxisrelevanz und des vielfältigen Einsatzes in Archiven ist die Untersuchung der Auswirkungen von Bildkompressionen für KI-basierte HTR-Systeme von besonderer Bedeutung.

Methodik

Für die Untersuchung der Auswirkungen von Bildkompressionsqualitäten auf die Leistung von HTR-Modellen werden Experimente durchgeführt, die darauf abzielen, die Erkennungsgenauigkeit des Modells bei verschiedenen Kompressionsraten zu bewerten. Für die Experimente wird ein aktuelles und leistungsfähiges neuronales Netzwerk aus dem Forschungsbereich der Dokumentenbildanalyse verwendet, das von Lei Kang et al. (2018)² publiziert wurde und für die Transkription von handschriftlichen Wortbildern optimiert ist. Aufgrund der Verwendung von fortschrittlichen neuronalen Ansätzen bietet dieses Modell eine stabile und zukunftsichere Grundlage, um die Effekte verschiedener Bildkompressionsqualitäten auf die Erkennungsgenauigkeit zu untersuchen. Zur Bewertung der Modellleistung wird ein geeigneter Benchmark und eine Metrik verwendet. Für die Experimente wird die IAM-Datenbank³ genutzt, ein weithin anerkannter Benchmark-Datensatz zur Evaluierung von HTR-Modellen. Die IAM-Datenbank besteht aus Scans handschriftlich verfasster Dokumentenbilder im PNG-Format und enthält moderne Handschriften in englischer Sprache mit einer Vielzahl an Schreibern. Die Modellleistung wird anhand der Zeichenfehlerrate (engl.: Character Error Rate, CER) gemessen. Die CER berechnet den Anteil der falsch erkannten Zeichen an der Gesamtzahl der Zeichen

und liefert somit eine detaillierte quantitative Einschätzung der Erkennungsqualität. Da der Fokus auf den Effekten der Bildkompression liegt, werden die Dokumentbilder in das verlustbehaftete JPEG-Format konvertiert, wobei die Kompressionsrate zwischen 100 und 1 variiert.

In den Experimenten werden zwei Ansätze verfolgt: Im ersten Ansatz wird das Modell auf unkomprimierten Bildern trainiert und anschließend auf Bildern mit verschiedenen JPEG- Kompressionsraten getestet, um den Einfluss der Kompression auf ein bereits optimiertes Modell zu bewerten. Im zweiten Ansatz werden separate Modelle direkt auf verschiedenen JPEG- Kompressionsraten trainiert und evaluiert.

Ergebnisse und Diskussion

Im Folgenden werden die Ergebnisse der Experimente präsentiert und deren Bedeutung für die Erkennungsleistung des HTR-Modells unter verschiedenen Bildkompressionsqualitäten analysiert. Abbildung 3 zeigt die Zeichenfehlerraten des HTR-Modells bei unterschiedlichen JPEG-Kompressionsraten (1, 10, 50 und 100), sowohl ohne spezifische Anpassung als auch mit Anpassung an die jeweilige Kompressionsqualität.

Bei der höchsten Kompressionsrate (Qualität 100) erreicht das Modell eine CER von 6.2 %, was der Erkennungsgenauigkeit des verlustfreien PNG-Formats entspricht. Auch bei der Kompressionsrate 50 bleibt die Fehlerrate mit etwa 6.4 % nahezu unverändert, was darauf hinweist, dass moderate Kompressionen nur einen geringen Einfluss auf die Erkennungsgenauigkeit haben. Mit abnehmender Kompressionsqualität verschlechtert sich die Erkennungsleistung jedoch deutlich. Insbesondere bei der niedrigsten Kompressionsrate (Qualität 1) steigt die CER des Modells ohne spezifische Anpassung signifikant auf 16.3 %, was die negativen Auswirkungen extremer Verlustkompressionen auf die Erkennungsgenauigkeit verdeutlicht.

Die angepassten Modelle, die direkt auf JPEG-Bildern der jeweiligen Kompressionsqualitäten trainiert wurden, zeigen hingegen eine hohe Robustheit gegenüber Kompressionsverlusten. Selbst bei einer Kompressionsrate von 1 ist die Fehlerrate mit etwa 7 % nah an der Leistung des mit PNG-Bildern trainierten Modells. Bei den Kompressionsraten 10

bis 100 konvergiert die CER der angepassten Modelle sogar auf das Niveau der verlustfreien PNG-Erkennungsqualität, was darauf hindeutet, dass sich diese Modelle effektiv an die Eigenschaften komprimierter Bilder anpassen können. Zusammengefasst zeigen die Ergebnisse, dass verlustfreie Formate wie PNG zwar häufig die geringste Fehlerrate bieten, jedoch auch Modelle, die auf verlustbehafteten Formaten wie JPEG mit moderaten Kompressionsverlusten (ab Kompressionsrate 50) trainiert wurden, keine beziehungsweise kaum merkliche Einbußen bei der Erkennungsgenauigkeit aufweisen.

4. SCHREIBERANALYSE

Im vorherigen Abschnitt stand bei der Handschrifterkennung die Frage im Fokus, wie gut der Inhalt eines Textdokuments durch die verlustbehaftete Kompression erhalten wird. Im weiteren Verlauf wird nun die Frage erörtert, inwiefern der Stil der Handschrift erhalten bleibt.

Die Stilanalyse von Handschrift ist sowohl im Bereich der (digitalen) Geisteswissenschaften, als auch in der Forensik von großem Interesse. So erlaubt sie es Historiker(innen), eine (automatische) Zuordnung von Dokumenten zu einem möglicherweise bekannten Schreiber, einem bestimmten Zeitraum oder einem geografischen Ursprungsort. All diese Informationen können helfen, große, unsortierte Datenbestände in eine strukturierte Form zu überführen und so für weitere Forschung zugänglich zu machen.

Für die Untersuchung, inwieweit der Stil der Handschrift erhalten bleibt, wird nun die Aufgabe des Writer Retrievals betrachtet. Hierbei müssen für ein gegebenes Anfragedokument all jene Dokumente aus einem Datenbestand herausgesucht werden, welche vom Schreiber des Anfragedokuments geschrieben wurden. Ausschlaggebend ist hierbei, von wessen Hand das Dokument verfasst wurde.

Methodik

Für die Untersuchung in diesem Bericht wird die von Raven et al. (2024)⁴ veröffentlichte Methode verwendet. Hierbei wird zunächst ein Vision Transformer mit einem selbstüberwachten Lernverfahren trainiert und für die Extraktion lokaler Bildmerkmale eingesetzt. Anschließend werden die extrahierten Merkmale durch ein Kodierungsverfahren zu einem globalen Deskriptor einer festen Länge aggregiert. Die

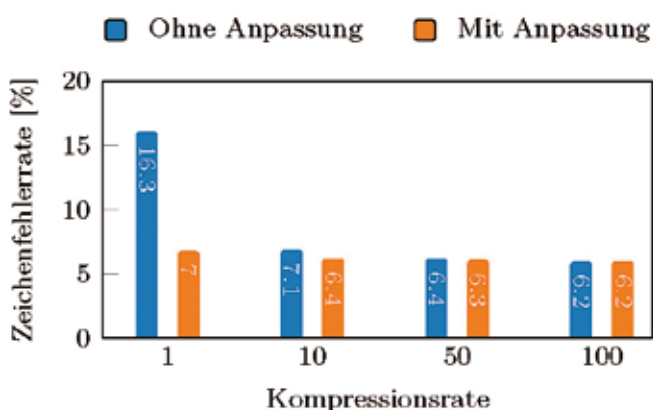


Abbildung 3: Zeichenfehlerraten des Handschrifterkennungsmodells auf der IAM-DB bei verschiedenen JPEG-Kompressionsraten. Mit Anpassung ist ein Modell gemeint, das auf die Kompression spezialisiert ist.

- Lei Kang, J. Ignacio Toledo, Pau Riba, Mauricio Villegas, Alicia Fornés, Marçal Rusiñol: Conville, Attend and Spell: An Attention-Based Sequence-to-Sequence Model for Handwritten Word Recognition. In: Proceedings of the German Conference on Pattern Recognition, 2018, S. 459–472.
- Urs-Viktor Marti, Horst Bunke: The IAM-Database: An English Sentence Database for Offline Handwriting Recognition. In: International Journal on Document Analysis and Recognition, Bd. 5, Nr. 1 (2002), S. 39–46.
- Tim Raven, Arthur Matei, Gernot A. Fink: Self-supervised Vision Transformers for Writer Retrieval. In: Proceedings of the International Conference on Document Analysis and Recognition, 2024, S. 380–396.
- Stefan Fiel, Florian Kleber, Markus Diem, Vincent Christlein, Georgios Louloudis, Nikos Stamatopoulos, Basilis Gatos: ICDAR2017 Competition on Historical Document Writer Identification (Historical-WI). In: Proceedings of the International Conference on Document Analysis and Recognition, 2017, S. 1377–1382.

Ähnlichkeit zwischen zwei Dokumenten kann dann mittels einer Distanzfunktion zwischen den zugehörigen globalen Deskriptoren bestimmt werden.

Als Benchmark für das Experiment wird der Historical-WI Datensatz⁵ verwendet, welcher einer der bedeutendsten Benchmark-Datensätze für das Writer Retrieval darstellt. Der Datensatz umfasst 4.782 Dokumentenbilder im PNG-Format, deren handschriftlich verfassten Texte zwischen dem 13. und 20. Jahrhundert von insgesamt 1.116 Schreibern in Deutsch, Französisch und Latein verfasst wurden. Die übliche Metrik zur Bewertung von Retrieval-Systemen ist die mean Average Precision (mAP). Diese misst für eine gegebene Menge von Dokumenten und einer gegebenen Suchanfrage, wie viele der vom Schreiber der Suchanfrage generierten Dokumente an der Spitze der vom System erzeugten Retrieval-Liste stehen.

Ergebnisse und Diskussion

Um den Einfluss von verlustbehafteter Kompression auf die Retrieval-Leistung zu bestimmen, wird analog zum HTR-Experiment zunächst ein Modell auf den unkomprimierten Daten trainiert und auf verschiedenen Kompressionsraten (1, 10, 50 und 100) evaluiert. Anschließend wird für jede Kompressionsrate ein eigenes Modell trainiert und auf der gleichen Kompressionsrate evaluiert.

Die verwendete Methode arbeitet auf binarisierten Bildern, wobei die Ausgangsbilder zunächst komprimiert und anschließend durch ein adaptives Schwellwertverfahren binarisiert werden.

Das auf unkomprimierten Daten trainierte und evaluierte Modell erreicht eine mAP von 78.2 %. Die Evaluation des Modells auf komprimierten Daten zeigt für die Kompressionsraten 100 (-0.0 % mAP), 50 (+0.1 % mAP) und 10 (-0.4 % mAP) keine praktisch relevanten Leistungseinbußen. Bei Kompressionsrate 1 hingegen werden deutliche Einbußen erkennbar (-13.6 % mAP). Die Anpassung des Modells auf den komprimierten Daten führt zu keiner Verbesserung der Systemleistung.

Zusammengefasst beeinträchtigt eine verlustbehaftete JPEG-Kompression die Stilanalyse handschriftlicher Dokumente erst bei sehr extremen Kompressionsverlusten. Selbst bei einer Kompressionsrate von 10 ist keine relevante Beeinträchtigung feststellbar. Eine Anpassung der verwendeten Methoden an die JPEG-Kompression (beispielsweise durch ein angepasstes Training) ist ebenfalls nicht erforderlich.

5. MULTIMODALE SPRACHMODELLE ZUR INFORMATIONSEXTRAKTION

Da die Speicherung komprimierter Bilder irreversibel ist, ist es für die vorliegende Analyse unerlässlich, die möglichen Auswirkungen der Kompression auch auf zukünftige KI-Modelle zu untersuchen. Obwohl diese Auswirkungen naturgemäß schwer abzuschätzen sind, werden multimodale Sprachmodelle (engl.: Multi-modal Large Language Models, MLLMs) als zukunftsweisende Technologie und damit als vielversprechende Möglichkeit für eine erste Abschätzung dieser Effekte betrachtet. MLLMs sind KI-Modelle, die auf einer neuronalen Modellarchitektur mit einer großen Para-

meteranzahl basieren und mit umfangreichen Datenmengen vortrainiert wurden. Diese Modelle verarbeiten verschiedene Eingabemodalitäten, wie Bild- und Textinformationen, und vereinen die visuelle und textuelle Analyse in einem einzigen Modell. Dadurch ermöglichen sie eine Vielzahl von Anwendungen im Bereich der Dokumentenbildanalyse, von der Extraktion spezifischer Informationen bis hin zur umfassenden inhaltlichen Erschließung großer Archivbestände. In dieser Arbeit wird speziell die Leistungsfähigkeit von MLLMs zur Informationsextraktion untersucht, um geeignete Kompressionsqualitäten zu identifizieren, die eine effiziente Langzeitspeicherung ermöglichen und gleichzeitig die Extraktionsgenauigkeit für zukünftige KI-basierte Analysen erhalten.

Methodik

Die Auswirkungen von Bildkompressionsqualitäten auf die Leistung von multimodalen Sprachmodellen werden in dieser Arbeit mit dem von Google veröffentlichten Paligemma-3B-Modell⁶ im Hinblick auf die Informationsextraktion aus handschriftlichen Dokumentenbildern bewertet. Das Modell basiert auf der Transformer-Architektur und wurde speziell für die Beantwortung natürlichsprachlicher Fragen auf Basis von Dokumentbildern angepasst.

Zur Evaluierung der Modellleistung wird ein selbst entwickelter Benchmark genutzt, der auf manuell annotierten Digitalisaten der CM/1-Umschläge⁷ basiert. Analysiert werden konkret die Deckblätter, die handschriftliche Informationen zu Vertriebenen aus dem Zweiten Weltkrieg und deren Begleitpersonen enthalten. Die zentrale Aufgabe besteht in der Extraktion des vollständigen Namens (Vor- und Nachname) der Hauptperson. Die Leistung des Modells wird anhand der Zeichenfehlerrate gemessen.

Da die CM/1-Bilder online bereits im JPEG-Format (ermittelte Kompressionsrate von 45) bereitgestellt werden, beziehen sich die verschiedenen Kompressionsqualitäten auf eine erneute Kompression der bereits verlustbehafteten Ausgangsdaten. Somit bleibt zu beachten, dass die Kompressionsraten nicht direkt mit den anderen Experimenten zu vergleichen sind und im Allgemeinen eine stärkere Kompression verwendet wird. Analog zu den vorangegangenen Anwendungsanalysen werden zwei Ansätze evaluiert: Zuerst wird der Einfluss der Kompressionsqualitäten auf ein Modell getestet, das auf den Ausgangsdaten trainiert wurde, und anschließend wird untersucht, wie sich ein Training auf speziellen Kompressionsqualitäten auf die Extraktionsleistung bei komprimierten Daten auswirkt.

Ergebnisse und Diskussion

Die Ergebnisse der Experimente in Abbildung 4 verdeutlichen die Auswirkungen von Kompressionsverlusten auf die Leistungsfähigkeit von multimodalen Sprachmodellen zur Informationsextraktion. Bei JPEG-Bildern mit einer Kompressionsrate im Bereich von 50 bis 100 zeigt das Modell eine nahezu identische Erkennungsleistung im Vergleich zur Verwendung der Ausgangsdaten. Bei stark komprimierten JPEG-Bildern ist hingegen eine deutliche Verschlechterung der Erkennungsleistung zu beobachten. Dies verdeutlicht, dass eine extreme verlustbehaftete Kompression die

Bildqualität und damit die Informationsextraktion erheblich beeinträchtigt. Bei einer Kompressionsrate von 10 sind die Ergebnisse zwar deutlich besser als bei Kompressionsrate 1, aber immer noch signifikant schlechter als bei Verwendung der Ausgangsdaten.

Im Gegensatz zu den HTR-Ergebnissen zeigt sich, dass auch nach gezielter Anpassung an stark komprimierte Bilder mit einer Kompressionsrate von 1 nur eine marginale Verbesserung erreicht wird. Die CER bleibt auf einem sehr niedrigen Niveau und auch bei einer Kompressionsrate von 10 bringt die Anpassung nur einen geringen Vorteil.

Insgesamt zeigt sich, dass die Leistungsfähigkeit der multimodalen Sprachmodelle bei der Informationsextraktion aus handgeschriebenen Dokumentbildern sowohl bei den Ausgangsdateien als auch bei moderat komprimierten JPEG-Bildern (ab Kompressionsrate 50) nahezu identische Erkennungsleistungen ermöglicht, hohe Kompressionsverluste jedoch problematisch sind.

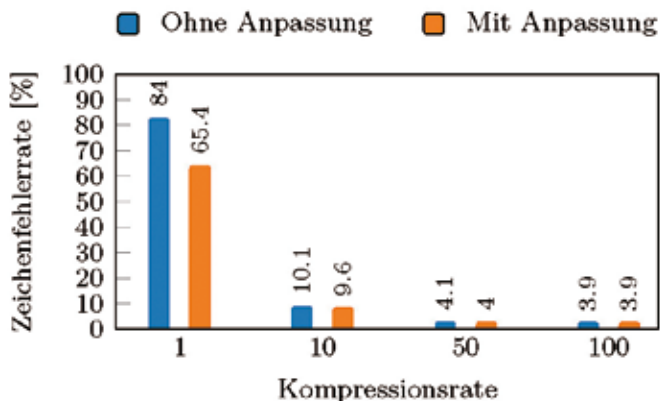


Abbildung 4: Zeichenfehlerraten des multimodalen Sprachmodells zur Informationsextraktion auf dem CM1-Datensatz bei verschiedenen JPEG-Kompressionsraten. Mit Anpassung ist ein Modell gemeint, das auf die Kompression spezialisiert ist.

6. DISKUSSION UND FAZIT

Die Untersuchungen in diesem Artikel haben gezeigt, dass im Allgemeinen die Leistungsfähigkeit von KI-Modellen robust gegenüber der Verwendung von verlustbehafteten Kompressionsalgorithmen ist. Insbesondere bei hohen Kompressionsqualitäten, welche einen kaum merklichen Einfluss auf die visuelle Erscheinung des Ausgangsbilds haben, nimmt die Leistung der betrachteten Modelle nur marginal ab. Motiviert durch die hohe Reduktion des Speicherbedarfs, lässt sich somit schlussfolgern, dass auch bei zukünftiger Verwendung von KI-gestützten Werkzeugen keine größeren Nachteile durch die Verwendung von komprimierten Bildformaten entstehen.

THE INFLUENCE OF IMAGE FORMATS ON THE AI-BASED ANALYSIS OF DIGITIZED DOCUMENT COLLECTIONS

This study examines the impact of image formats and compression on AI-based analysis of digitized documents. Focusing on handwritten text recognition, writer retrieval, and information extraction, it evaluates the effects of compression quality on model performance. Results show that high quality compression (e. g. JPEG 50+) has minimal impact on AI performance, while extreme compression degrades it. The findings highlight the feasibility of storage-efficient formats for long-term archival use without compromising future AI analyses.

Dr.-Ing. Oliver Tüselmann

Technische Universität Dortmund und Lamarr Institut für Maschinelles Lernen und Künstliche Intelligenz
Otto-Hahn-Str. 16, 44227 Dortmund
E-Mail: oliver.tueselmann@tu-dortmund.de

Dr.-Ing. Fabian Wolf

Technische Universität Dortmund und Lamarr Institut für Maschinelles Lernen und Künstliche Intelligenz
Otto-Hahn-Str. 16, 44227 Dortmund
E-Mail: fabian.wolf@cs.tu-dortmund.de

Tim Raven, M. Sc.

Technische Universität Dortmund
Otto-Hahn-Str. 16, 44227 Dortmund
E-Mail: tim.raven@tu-dortmund.de

Prof. Dr.-Ing. Gernot A. Fink

Technische Universität Dortmund und Lamarr Institut für Maschinelles Lernen und Künstliche Intelligenz
Otto-Hahn-Str. 16, 44221 Dortmund
E-Mail: Gernot.Fink@tu-dortmund.de

⁶ Lucas Beyer, Andreas Steiner, André Susano Pinto, Alexander Kolesnikov, Xiao Wang, Daniel Salz, Maxim Neumann, Ibrahim Alabdulmohsin, Michael Tschannen, Emanuele Bugliarello u. a.: PaliGemma: A versatile 3B VLM for transfer. In: arXiv preprint arXiv:2407.07726 (2024).

⁷ CM1 Akten (Mantelbögen), Arolsen Archives <https://collections.arolsen-archives.org/de/archive/3-2-1-1>

DIE ZÄHMUNG DES CHAOS

PERSPEKTIVEN ZUR KI-GESTÜTZTEN ERSCHLIESSUNG VON ENTSCHÄDIGUNGSAKTEN IM LANDESARCHIV BADEN-WÜRTTEMBERG¹

von *Andreas Neuburger*

TRANSFORMATION DER WIEDERGUTMACHTUNG

Der Petersberg bei Bonn am 1. Juni 2022: Vertreterinnen und Vertreter des Bundesministeriums der Finanzen (BMF) und der Archivverwaltungen des Bundes und der Länder unterzeichnen im Rahmen eines Festakts die Rahmenvereinbarung zum Aufbau des „Themenportals Wiedergutmachung“ im Archivportal-D. Hintergrund ist die Neuausrichtung der vom BMF bislang durch Zahlungen an NS-Opfer geprägten Bemühungen um Entschädigung von NS-Unrecht.² Seit 2017 macht sich das Ministerium intensiv Gedanken zur „Transformation der Wiedergutmachung“. Gemeint ist damit der Übergang von der materiellen Entschädigung zu einer zeitgemäßen, über die Bundesrepublik hinaus international vernetzten Erinnerungskultur und gezielten Impulsen für weitere Forschungen. In den Fokus rückten rasch die im Bundesarchiv und den Landesarchivverwaltungen umfangreich vorhandenen Aktenüberlieferungen und das Ziel, das „gesamte Dokumentenerbe der Akten aus der Wiedergutmachung (...) digital in einem Themenportal zugänglich“ zu machen³. Zusätzlich zu einer forcierten Erschließung der Unterlagen ging es von Beginn an auch um eine Komplettdigitalisierung und die rechtskonforme Onlinestellung der Akten. BMF stellt die zur Umsetzung dieses Großvorhabens über viele Jahre benötigten Finanzmittel zur Verfügung.⁴ Seit Mitte 2022 ist das Themenportal „Wiedergutmachung nationalsozialistischen Unrechts“ im Rahmen des Archivportals-D für die Öffentlichkeit zugänglich. Das Angebot wird Zug um Zug ausgebaut – neben bereits verfügbaren Erschließungsdaten und Beständeübersichten bietet die Präsentation heute bereits umfangreiche vertiefende Inhalte.⁵ Archivfachlich konzipiert und technisch betrieben wird das Themenportal in gemeinsamer Verantwortung durch das Bundesarchiv, das Landesarchiv Baden-Württemberg

sowie FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur.

AUSWIRKUNGEN AUF DIE ARCHIVE

Das Themenportal „Wiedergutmachung nationalsozialistischen Unrechts“ und die dort vorgesehenen Funktionalitäten stellen das Bundesarchiv und die Archivverwaltungen vor allem der alten Länder vor enorme quantitative und qualitative Herausforderungen. Bundesweit ist von circa 100 lfd.km einschlägigem Archivgut auszugehen.⁶ Allein in den Archivabteilungen des Landesarchivs Baden-Württemberg befinden sich insgesamt 3,5 Kilometer unmittelbar themenrelevantes Schriftgut, davon 1,8 lfd.km Einzelfallakten zu Verfahren nach dem Bundesentschädigungsgesetz. Sie stammen zum größeren Teil aus den ehemaligen Landesämtern für Wiedergutmachung – und werden bis heute vom Landesamt für Besoldung und Versorgung Baden-Württemberg angeboten und als komplett archivwürdig übernommen. Hinzu kommen die Überlieferungen der bei spezialisierten Amtsgerichten angesiedelten „Schlichter für Wiedergutmachung“, Unterlagen der Finanzverwaltung (v. a. der Oberfinanzdirektionen, aber auch von Finanzämtern), Landgerichtsunterlagen sowie Sachakten unterschiedlicher Provenienz sowohl in einzelnen Ministerien wie auch bei nachgeordneten Behörden.

Sehr unterschiedlich stellt sich die Erschließungssituation der Entschädigungs- sowie der Rückerstattungsakten in den Landesarchiven dar. Bereits 2009 hat eine Arbeitsgruppe der Archivreferentenkonferenz (ARK) Empfehlungen zur Erschließung dieser zentralen Überlieferung verabschiedet.⁷ Diese kamen seither jedoch nicht konsequent zur Anwendung. Hinzu kommt, dass das BMF technisch anspruchsvolle Benutzungsfunktionen für das Themenportal vorgesehen hat. 2023 hat sich das Ministerium mit der Konferenz der

Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA) zu diesem Zweck auf einen aus 17 Recherche- und Filtermöglichkeiten bestehenden Funktionskatalog verständigt. Jenseits auf die betroffene Person oder eine bestimmte Akte bezogene Angaben wie Namen, Lebensdaten, Orte und Aktenzeichen soll das Themenportal unter anderem die Recherche nach Verfolgungsgründen und Verfolgungsschäden, Haftstätten oder Zwangsmaßnahmen ermöglichen.

Die Abbildung dieser Anforderungen und die hierzu erforderliche strukturierte Erfassung zahlreicher Angaben setzt eine im deutschen Archivwesen bisher nie verlangte Qualität der Erschließungsdaten voraus. Zwischen 2021 und 2023 hat eine Arbeitsgruppe der KLA mit diesem Ziel einen gemeinsamen, aus 71 Feldern bestehenden Metadatenkatalog für die Entschädigungsakten der westdeutschen Bundesländer erarbeitet. Er ermöglicht die Umsetzung der mit dem BMF vereinbarten Anforderungen und dient als Standard zur Erschließung dieses Aktentyps.⁸ Verglichen mit der aus etwa zehn Feldern bestehenden Standardschließung einer personenbezogenen Akte ist bei konsequenter Anwendung des KLA-Katalogs mit einem hohen Arbeitsaufwand für die Erschließung und für die Anpassung der Archivischen Fachinformationssysteme zu rechnen.⁹

PERSPEKTIVEN DER KI-GESTÜTZTEN ERSCHLIESSUNG VON ENTSCHÄDIGUNGSAKTEN

Im Landesarchiv Baden-Württemberg sind die von den Archivabteilungen in Freiburg, Karlsruhe, Ludwigsburg und Sigmaringen übernommenen Entschädigungsakten der Wiedergutmachung bereits erschlossen und (soweit schutzfristenfrei) online zugänglich. Die Erschließung erfolgte flach und deckt zentrale Forschungsperspektiven ab, reicht jedoch bei weitem nicht für die für das Themenportal angestrebten Funktionalitäten aus.

Mit Blick auf die schmerzlichen Erschließungsrückstände an anderer Stelle und den sich verschärfenden Fachkräftemangel hat sich das Landesarchiv gleichwohl gegen die händisch-intellektuelle Nacherschließung der 1,8 lfd.km Entschädigungsakten entschieden. Stattdessen ist geplant, die Akten zu digitalisieren und möglichst viele Angaben des Metadatenkatalogs mithilfe Künstlicher Intelligenz erheben zu lassen. Und obgleich die technischen Fortschritte der letzten fünf Jahre äußerst bemerkenswert sind, steht die automatisierte Volltext- und Layoutmustererkennung von Verwaltungsunterlagen des 20. Jahrhunderts derzeit noch vor großen Herausforderungen. Dementsprechend bieten auch die Entschädigungsakten der Maschine ein zunächst undurchdringlich erscheinendes Chaos: Die Unterlagen bestehen aus einer Mischung von gedruckten, maschinen- und handschriftlichen Texten und weisen überdies nur bei wenigen Dokumenten ein einheitliches Layout auf. Mit der derzeit verfügbaren Software besteht für diese Ausgangslage keine ausreichende Erfolgsperspektive zur Volltexterkennung – schon allein der für Menge und Beschaffenheit der zu bearbeitenden Unterlagen notwendige Trainingsaufwand würde jedes Vorhaben zum Scheitern bringen.

Die ab diesem Jahr im Landesarchiv angestrebte Lösung setzt vor diesem Hintergrund nicht beim Volltext der gesamten Akte an, sondern wird sich auf einige wenige, stark standardisierte „Kerndokumente“ fokussieren, die (wenn gleich unterschiedlich vollständig) in jeder Entschädigungsakte vorhanden sind und zunächst maschinengestützt aus den Akten herausgefiltert werden. Neben dem Aktendeckel werden alle in einer Akte vorhandenen Anträge, der Bescheid sowie das sogenannte „Handblatt“ herangezogen. In diesen Dokumenten (in der Regel handelt es sich um einen Umfang von 12 Blatt je Akte) findet sich der größte Teil der zur Befüllung des KLA-Metadatenkatalogs erforderlichen Informationen. Zur Verbesserung der Ergebnisse wird der

- ¹ Der Beitrag geht zurück auf einen Vortrag im Rahmen des 1. Kölner Forums am 28. August 2024. Für Informationen und Hinweise danke ich herzlich meiner Kollegin Nastasja Pilz M. A.
- ² Vgl. Michael Unger: Feierliche Unterzeichnung der Rahmenvereinbarung über die Bereitstellung von digitalen Inhalten für ein „Themenportal Wiedergutmachung“. In: ARCHIVAR (2022) H 4, S. 354 f., online unter: <https://www.archive.nrw.de/sites/default/files/media/files/ArchivarHeft42022.pdf> (aufgerufen am 12.11.2024).
- ³ BMF-Monatsbericht 2021/1, S. 72, online unter: https://www.bundesfinanzministerium.de/Monatsberichte/2021/01/Inhalte/Kapitel-3-Analysen/3-7-archivierungsprojekt-wiedergutmachung-pdf.pdf?__blob=publicationFile&v=1 (aufgerufen am 12.11.2024).
- ⁴ Zum Projekt und seiner Genese vgl. Kai Wambach: Chancen an der Schnittstelle des Unmöglichen. Archive als Kommunikationsorte der Wiedergutmachung. In: An den Schnittstellen zwischen Archiv und Gesellschaft. Hg. v. Christian Keitel u. Gerald Maier, Ostfildern 2024, S. 59–72; Clemens Rehm: Die „Transformation der Wiedergutmachung“ und die Archive. Neue Perspektiven. In: Deuten und streiten, suchen und finden: neue Möglichkeiten der Kooperation zwischen Archiven und Geschichtswissenschaft beim Aufbau digitaler Infrastrukturen. Hg. v. Rainer Hering u. Gerald Maier, Ostfildern 2023, S. 59–70, online unter: <https://books.wlb-stuttgart.de/omp/index.php/regiopen/catalog/view/166/236/923>; Nils Meyer u. Mirjam Sprau: Neue Zugänge, neue Herausforderungen, neue Chancen – archivische Themenportale. In: Zeitschrift für Bibliothekswesen und Bibliographie 70 (2023) 1, S. 5–18, online unter: https://zs.thulb.uni-jena.de/receive/jportal.jparticle_01289880. Weitere Informationen zum Gesamtvorhaben im Webauftakt des BMF unter: https://www.bundesfinanzministerium.de/Web/DE/Themen/Oeffentliche_Finanzen/Vermögensrecht_und_Entschädigungen/vermoeogensrecht_entschaedigungen.html (alle Seiten aufgerufen am 12.11.2024).
- ⁵ Startseite des Themenportals unter: <https://www.archivportal-d.de/themenportale/wiedergutmachung> (aufgerufen am 12.11.2024).
- ⁶ Vgl. die Pressemitteilung der Bundesregierung zur Unterzeichnung der Rahmenvereinbarung unter: <https://www.bundesregierung.de/breg-de/suche/themenportal-wiedergutmachung-2047132>. Bei dem genannten Gesamtumfang handelt es sich derzeit noch um eine Schätzung, die auf einer 2010 bei den staatlichen Archiven durchgeführten Erhebung fußt (vgl. die Übersicht über die Überlieferung und Rechtsgrundlagen zur Wiedergutmachung nationalsozialistischen Unrechts in der Bundesrepublik Deutschland in den staatlichen Archiven, online unter: <https://www.bundesarchiv.de/assets/bundesarchiv/de/Downloads/Erklaerungen/wiedergutmachung-dokumentation.pdf> (aufgerufen am 12.11.2024)). Einbezogen sind auch bereits die für das Projekt relevanten kommunalen Unterlagen, zu deren Umfang bereits systematische Erhebungen laufen. Aktuell haben rund 80 Kommunalarchive Bestände gemeldet.
- ⁷ Vgl. den Abschlussbericht der ARK-Bund-Länder-Arbeitsgruppe „Wiedergutmachung“ unter: <https://www.bundesarchiv.de/assets/bundesarchiv/de/Downloads/Berichte/wiedergutmachung-abschlussbericht.pdf> (aufgerufen am 12.11.2024).
- ⁸ Online abrufbar unter: https://wiki.deutsche-digitale-bibliothek.de/download/attachments/126518213/Anlage_1_Metadatenkatalog_2023.05.12.pdf?version=1&modificationDate=1702050019616&api=v2 (aufgerufen am 12.11.2024). Die Anwendung des Metadatenkatalogs ist auch Voraussetzung einer Projektfinanzierung durch das BMF.
- ⁹ Auf Grundlage von Testerschließungen scheint für den KLA-Metadatenstandard eine Bearbeitungsdauer von 20 Minuten je Akte realistisch.

automatisierten Erfassung eine Wiedergutmachungs-Ontologie¹⁰ zugrunde gelegt, welche die semantische Verknüpfung der erkannten Entitäten unterstützt und beispielsweise einordnet, welche Art der Information (beispielsweise ein Name oder Geburtsdatum) wo genau in einem Dokument zu erwarten ist. Auf diese Weise lässt sich die entstehende Datenflut kanalisieren und die Qualität der Erkennungsergebnisse verbessern.

Ziel des Ansatzes ist es, eine Erfassungspipeline für Digitalisate zu entwickeln, an deren Ende ein möglichst vollständig automatisiert befüllter KLA-Metadatenkatalog steht. Automatisiert erfolgen das Auffinden und Auslesen von Informationen aus der Akte, die Umformung in ein strukturiertes, maschinenlesbares Datenformat sowie die Zuordnung zur entsprechenden Entitätengruppe. Zwar ist damit zu rechnen, dass insbesondere anfangs noch ein nennenswerter Teil der extrahierten Informationen von Archivpersonal geprüft und auch korrigiert werden muss,¹¹ ehe die Bereitstellung der Daten im eigenen Archivfachinformationssystem und anschließend dem Archivportal-D erfolgen kann. Gleichwohl ist von einer wesentlichen Zeitersparnis gegenüber einer rein händisch-intellektuellen Erschließung auszugehen. Hervorzuheben ist dabei, dass die angestrebte Erfassungspipeline vom Digitalisat über die strukturierten Informationen in maschinenlesbaren Formaten bis hin zu RDF-Tripeln nicht nur für Akten des Landesarchivs Baden-Württemberg, sondern für die Entschädigungsakten aller westdeutschen Bundesländer funktionieren soll. Um dies zu ermöglichen, soll die Software von Beginn an auf alle im Projekt bekannten Formulartypen der Entschädigungsakten trainiert und angewendet werden.¹² Die enorme Zeitersparnis bei der Erschließung wird im Erfolgsfall den Fortschritt des Gesamtprojekts enorm beschleunigen und zu der Situation führen, dass in absehbarer Zeit nicht mehr die Erschließung, sondern die Geschwindigkeit der Digitalisierung das Nadelöhr beim weiteren Ausbau des Themenportals „Wiedergutmachung nationalsozialistischen Unrechts“ bilden könnte.

ERSTE GRUNDLAGEN: TEXT- UND MUSTERERKENNUNG DER BUNDESZENTRAKARTEI

Im Landesarchiv Baden-Württemberg soll 2025 mit der automatisierten Erschließung von Entschädigungsakten begonnen werden. Hierbei kann auf vorhandene Modelle zur Informationsextraktion aus der Bundeszentalkartei (BZK) zurückgegriffen werden, die seit März 2023 bei FIZ Karlsruhe laufen. Zum Einsatz kommt eine Pipeline bestehend aus einer sprachmodellbasierten „End-to-End“-Informationsextraktion (IE), an die ein Post-Processing der Daten und eine RDF-Konversion angeschlossen werden.

Innerhalb des Wiedergutmachungs-Projekts kommt dem circa zwei Millionen Karteikarten umfassenden Korpus eine zentrale Rolle zu. Die BZK enthält neben personenbezogenen Daten der Antragstellerinnen und Antragsteller auch die für die Fallbearbeitung in den Landesämtern verwendete BZK-Nummer und somit den Schlüssel zu allen in der alten Bundesrepublik bearbeiteten Entschädigungsakten. Eine erfolgreiche Datenextraktion aus der BZK ist dabei aus zwei Gründen von enormer Bedeutung für das Gesamtvorhaben:

Zum einen wird über die Veröffentlichung der Daten im Themenportal je ein Basisdatensatz zu allen antragstellenden und verfolgten Personen verfügbar sein, über den alle zu dieser Person existierenden Akten über eine eindeutige ID verknüpft werden können. Zum anderen – und für das hier skizzierte Vorhaben unter Umständen noch wichtiger – bieten die aus der BZK extrahierten Daten zusammen mit den (etwa in Baden-Württemberg) schon vorhandenen Erschließungsinformationen einen Datenpool zur Qualitätsprüfung und Fehlerkontrolle für einen wichtigen Teil der später automatisiert aus den Entschädigungsakten zu erhebenden Erschließungsdaten. Mit dem Ziel einer Fehlerbereinigung und Datenkonsolidierung können perspektivisch auch Daten anderer Projektpartner herangezogen werden, etwa die umfangreichen Daten der Internationalen Holocaust Gedenkstätte Yad Vashem.

Im Lauf des vergangenen Jahres hat sich die von FIZ-Karlsruhe erreichte Erkennungsrate der BZK deutlich verbessert. Sie liegt bei Namen inzwischen bei einer Genauigkeit von über 90 %. In Verbindung mit dem raschen technischen Fortschritt der Texterkennungstechnologie berechtigt dies zu der Hoffnung, dass mit gezieltem Training bereits in Kürze auch die Erschließungsinformationen zu den Entschädigungsakten direkt aus den Digitalisaten extrahiert werden können.

AUSWIRKUNGEN DES KI-ANSATZES AUF DIE „AKTENDECKELORIENTIERTE ERSCHLIESSUNG“

Aus der Perspektive einer erfolgreichen Gewinnung der für den Metadatenkatalog der KLA vorgesehenen Erschließungsdaten über Texterkennung ergeben sich zwei unmittelbare Folgen. Erstens wird eine Umkehrung des bisherigen Workflows eintreten, indem die Digitalisierung anders als bisher der Erschließung vorgelagert sein wird.¹³ An die möglichst weitgehende Automatisierung der Metadatenextraktion schließen sich Nacharbeiten an. Zumindest in den nächsten Jahren ist noch mit größeren Aufwänden zur Qualitätssicherung und Fehlerbereinigung zu rechnen, vor allem bei weniger standardisierten Metadaten. Gleichwohl ist von einer sehr deutlichen zeitlichen Entlastung des Fachpersonals gegenüber der händisch-intellektuellen Erschließung nach dem KLA-Metadatenkatalog auszugehen.

Zweitens – und der eben skizzierten Arbeitsentlastung teilweise gegenläufig – ist damit zu rechnen, dass aus der semiautomatisierten Erschließung nicht selten Klärungsbedarf bezüglich der korrekten Schreibweise von Namen entsteht. Bislang hat sich archivische Erschließung stark auf den Aktendeckel fokussiert – schon allein, da die dort vorhandenen Informationen oftmals für die im Regelfall praktizierte flache Erschließung ausgereicht haben. Bei der semiautomatisierten Auswertung der oben genannten „Kerndokumente“ aus den Entschädigungsakten treten häufig abweichende Schreibweisen zutage, insbesondere bei Orts- und Personennamen. Während sich die Ortsnamen über normierte Thesauruslisten in vielen Fällen vergleichsweise einfach bereinigen lassen, liegen die Dinge bei den Personen schwieriger. Neben komplizierten und daher in den Akten fehleranfällig erfassten Namen aus dem osteuropäischen Raum finden sich insbesondere nach Israel und in die

USA geflohene Personen, die ihren Namen nach 1945 nicht selten bewusst geändert haben. Für ein und dieselbe Person ist dementsprechend mit Namensvarianten zu rechnen, die jeweils separat zu erfassen und der für das Themenportal eindeutig identifizierten Person zuzuordnen sind.¹⁴ Mit Blick auf mögliche Erfassungsfehler in der Akte und seither vorgekommene Übertragungsfehler ist dabei klar, dass nicht pauschal immer die derselben Fundstelle zu entnehmende Information den Ausschlag geben kann – sei es der von Archivarinnen und Archivaren bislang häufig favorisierte Aktendeckel oder der erste in einer Entschädigungsakte vorhandene Antrag einer Person.

AUSBLICK

Für die ab 2025 geplante Bearbeitung der Entschädigungsakten mithilfe automatisierter Pipelines zur Informationsextraktion laufen im Landesarchiv Baden-Württemberg derzeit die Vorbereitungen. Umfangreiche Trainingsdaten stehen aus den im Rahmen eines Pilotprojekts zwischen 2020 und 2023 im Staatsarchiv Ludwigsburg durchgeführten Testerschließungen und Digitalisierungsmaßnahmen bereits zur Verfügung.¹⁵ Auch mit Blick auf diese vor vergleichsweise kurzer Zeit erstellten Daten zeigt sich der rasante technische Fortschritt der Texterkennung. Im Rahmen des Pilotprojekts wurde noch eine separate Erfassung maschinen- und handschriftlicher Aktenteile für sinnvoll erachtet – inzwischen führt die gemeinsame Prozessierung der jeweiligen Passagen bereits zu besseren Ergebnissen.

Gleichwohl funktioniert die auf eine automatisierte Texterfassung gestützte Erschließung gegenwärtig bei weitem noch nicht mit der notwendigen Zuverlässigkeit, um die im Themenportal „Wiedergutmachung nationalsozialistischen Unrechts“ angestrebten Funktionalitäten zu ermöglichen. In Gestalt der Rückerstattungsakten zeichnet sich zudem bereits eine weitere Herausforderung ab. Auch für diese Archivaliengattung ist ein von der KLA-Arbeitsgruppe erstellter gemeinsamer Metadatenkatalog in Vorbereitung. Allerdings weisen die Rückerstattungsakten eine gegenüber den Entschädigungsakten weitaus weniger einheitliche Struktur auf – ein beträchtliches Hindernis für eine zielgerichtete automatisierte Texterkennung.

Es besteht aller Grund zur Zuversicht, dass sich das in den Akten der Wiedergutmachung aus Sicht einer Informationsextraktionssoftware vorhandene Chaos in den nächsten fünf bis sieben Jahren zähmen lassen wird. Und auch wenn der Weg der Daten nicht direkt aus der Volltexterkennung in die Erschließungsdatenbank und von dort ins Archivportal-D verlaufen wird, so scheint es doch unwahrscheinlich, dass zahlreiche Archivarinnen und Archivare überall in der

Republik die nächsten anderthalb Jahrzehnte (und vielleicht darüber hinaus) damit beschäftigt sein werden, mit den Methoden des 20. Jahrhunderts Entschädigungs- und Rückerstattungsakten zu erschließen, damit diese in einer innovativen Präsentations- und Nutzungsoberfläche des 21. Jahrhunderts zugänglich gemacht werden können.

TAMING THE CHAOS: PERSPECTIVES FOR THE AI-SUPPORTED INDEXING OF COMPENSATION FILES IN THE STATE ARCHIVES OF BADEN-WÜRTTEMBERG

State archives in Germany face a considerable challenge. Resulting from the project „Transformation der Wiedergutmachung“, around 100 km of analogue records need to be described in detail, digitized and shall be published (as far as legally possible) within the „Themenportal Wiedergutmachung“. The Landesarchiv Baden-Württemberg designs a workflow to extract metadata out of digitized documents with support of artificial intelligence. By this means, we hope to ease and significantly fasten the provision of metadata as base for various research functions in the Archivportal-D.

Dr. Andreas Neuburger

Landesarchiv Baden-Württemberg
Abteilung Archivischer Grundsatz
Urbanstraße 31A, 70182 Stuttgart
E-Mail: andreas.neuburger@la-bw.de

¹⁰ Basis sind vorhandene Ontologien, etwa die zu „Records in Contexts“ (online unter: <https://www.ica.org/resource/records-in-contexts-ontology/>, aufgerufen am 29.11.2024).

¹¹ Für die Korrektur stehen beispielsweise das auf dem KLA-Metadatenkatalog basierende EAD-Schema sowie Ansetzungsrichtlinien bei der WGM-Erschließung zur Verfügung.

¹² Gegenwärtig ist von einer überschaubaren Anzahl unterschiedlicher Formulare auszugehen.

¹³ Für die Verknüpfung der Digitalisate mit dem zugehörigen Erschließungsdatensatz wird im Vorfeld der Digitalisierung lediglich die Vergabe einer Signatur notwendig bleiben.

¹⁴ Innerhalb des Themenportals ist vorgesehen, eine eigene „Personenseite“ zu den in den Akten auftauchenden Personen zu erstellen. Dies setzt eine saubere Disambiguierung und Normreferenzierung der im Kontext der Akten in Erscheinung tretenden Personen voraus.

¹⁵ Informationen zum Projekt unter: <https://www.landesarchiv-bw.de/de/landesarchiv/projekte/transformation-der-wiedergutmachung/71002> (aufgerufen am 12.11.2024).

DAS DIGITALISAT ALS DATENQUELLE

EIN WORKFLOW FÜR DIE NUTZUNG BEREITS DIGITALISierter ARCHIVBESTÄNDE

von *Florian Probst und Thomas Reich*

NEUE WEGE?

Seit über zwei Dekaden wächst die Anzahl digitalisierter Akten erheblich – allein 2023 wuchs die Anzahl der Daten digitalisierten Archivguts am Landesarchiv NRW um 5,7 Millionen auf knapp 73 Millionen, dies entspricht über 16 Millionen Verzeichnungseinheiten.¹ Während die Digitalisierung von Archivgut mittlerweile ein gewöhnlicher Schritt im Zuge der Erschließung und nach bestandserhalterischen Maßnahmen ist, finden die Digitalisate selbst später hauptsächlich in der Beauskunftung Eingang in den Archivalltag. Unser Projekt zeigt, wie bereits digitalisierte Akten mit einfachen Mitteln tiefer erschlossen werden können. Dazu setzen wir auf leicht zugängliche Tools und bestehende Technologien, die ohne spezialisiertes Training von Modellen oder hohen technischen Aufwand anwendbar sind. Die vorgestellten Ansätze sind allesamt nicht-kommerzieller Natur, zudem

tragen wir rechtlichen Vorgaben und dem Datenschutz Rechnung, indem wir das Archivgut soweit wie möglich auf lokalen Geräten weiterverarbeiten.²

Unser Workflow umfasst drei Schritte: erstens die Text- bzw. Handschriftenerkennung, zweitens die Textverarbeitung mittels Natural Language Processings (NLP) und drittens die Auswertung durch eine generative Künstliche Intelligenz (KI). Als Grundlage dienen uns die im Landesarchiv Nordrhein-Westfalen, Abteilung Westfalen im Bestand C 001/Fürstabtei Corvey überlieferten und digitalisierten 156 Reichstagsberichte aus den Jahren 1681 bis 1803.³

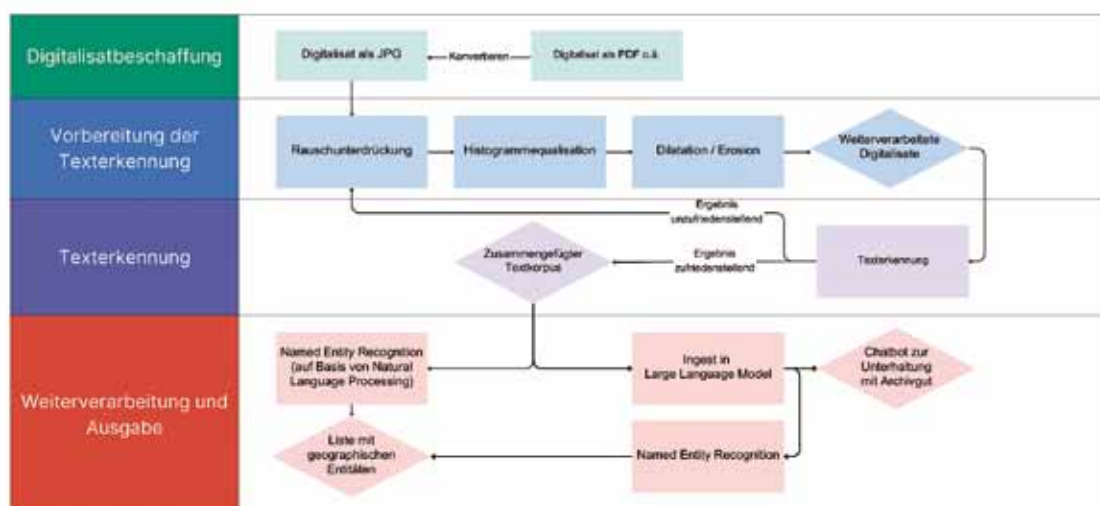


Abbildung 1: Arbeitsablauf zur Nutzung bereits digitalisierter Archivgüter zur tieferen Auswertung

DIE GRUNDLAGE SCHAFFEN: EXTRAKTION VON TEXT AUS DEN DIGITALISATEN

Im ersten Abschnitt, der Texterkennung, legen wir den Grundstein für die gesamte Auswertung. Diese beginnt zunächst mit der Vorbereitung der Digitalisate. Nachdem diese ggf. zunächst in Bilddateien umgewandelt wurden, werden die Bilder entrauscht. Dabei werden Pixelbereiche mit ähnlicher Struktur identifiziert und geglättet, um Störungen zu reduzieren. Anschließend wird die Bildkontrastierung durch Histogramm-Equalisation verbessert, um die Lesbarkeit des Textes zu erhöhen und ihn deutlicher vom Hintergrund abzuheben. Zuletzt wenden wir noch eine Methode aus dem Gebiet der morphologischen Operationen an und zwar Dilatation. Bei der Dilatation wird der Text visuell „verdickt“, d. h. Strukturen werden verbunden und kleine Löcher gefüllt.⁴ Der gezeigte Ablauf kann bei Bedarf auch übersprungen werden, einzelne Elemente werden von Programmen zur Texterkennung auch selbst durchgeführt. Die Vorbereitung der Digitalisate nach dem gezeigten Schema garantiert bessere Ergebnisse in der Texterkennung; das impliziert jedoch nicht, dass nicht optimierte Digitalisate ungeeignet sind für die weitere Verarbeitung. In unserem Test waren die unten angeführten Fehlerquoten ohne die gezeigte Optimierung etwa ein bis drei Prozentpunkte höher.

Nach dem Durchlaufen dieser Schritte sind die Digitalisate optimiert für die Text- und Handschriftenerkennung. Da die untersuchten Reichstagsprotokolle zu etwa der Hälfte aus handschriftlich verfassten Berichten und zur anderen Hälfte aus Drucksachen bestehen, wird ein Modell benötigt, das sowohl Druck- als auch Handschriften erkennen kann. Aus diesem Grund entfällt Tesseract als Standardlösung für Druckschriften ebenso wie Transkribus aufgrund seiner Kosten. Wie in der Einleitung beschrieben, lassen wir zugunsten der Umsetzbarkeit unseres Workflows auch alle Tools aus, die fortgeschrittene Programmierkenntnisse oder eine längere Phase der Einarbeitung benötigen. Die Wahl fiel daher auf die Verarbeitung der Reichstagsprotokolle mit PERO OCR, auch da wir damit bessere Ergebnisse als mit den kostenpflichtigen Angeboten von Transkribus erzielen konnten, siehe Tabelle 1.⁵ Durch die Texterkennung haben wir so aus etwa 36.000 Bilddateien der 156 Verzeichnungseinheiten rund 3.000.000 Zeilen erkannten Text extrahieren können.

TIEFERGEHENDE ERSCHLIESSUNG DURCH TEXTANALYSE

Nach der Texterkennung werden die erkannten Texte auf Ebene der Verzeichnungseinheiten weiterverarbeitet. Der Textkorpus, der aus den 156 Reichstagsberichten besteht, wird dabei so aufbereitet, dass gezielte Analysen und inhaltliche Erschließungen möglich werden. Unser Ziel ist es, mittels Natural Language Processing Informationen über die Inhalte der einzelnen Dokumente zu gewinnen, die über bloße Wortzählungen oder Angaben über den Umfang hinausgehen, sondern die Nutzung des Materials erleichtern. Eine Möglichkeit ist beispielsweise die Extraktion von Entitäten. Die Erkennung – in vorliegender Untersuchung geographischer – Entitäten bietet nicht nur einen wertvollen Einblick in die räumlichen Bezugspunkte der Dokumente, sondern kann auch gezielt für spätere Recherchen herangezogen werden. Zur Erkennung geographischer Entitäten setzen wir spaCy ein. Diese offene Software wurde speziell für die Verarbeitung natürlicher Sprache entwickelt und enthält unter anderem ein Werkzeug, welches Named Entity Recognition (NER), also das Erkennen von Entitäten wie Organisationen, Orts- oder Personennamen ermöglicht. Mit NER kann spaCy im texterkannten Archivgut nach erwähnten geographischen Entitäten suchen. Allerdings zeigen die

- ¹ Landesarchiv Nordrhein-Westfalen: Das Landesarchiv Nordrhein-Westfalen 2023. Duisburg 2024 (= Veröffentlichungen des Landesarchivs Nordrhein-Westfalen 96), S. 57 auch online unter https://www.archive.nrw.de/sites/default/files/media/files/LAV_Jahresbericht_2023_www.pdf (abgerufen am 14.11.2024).
- ² Der vorgestellte Workflow ist größtenteils in einer einzigen Python-Pipeline bzw. als Skript in R implementierbar. Zur (auch nur teilweisen) Anwendung im Arbeitsalltag und ohne größere Programmierkenntnisse empfehlen wir für die Texterkennung die Browseranwendung unter pero-ocr.fit.vutbr.cz, für die NER bietet sich die Ausführung online unter nopaque.uni-bielefeld.de an. Bei Online-Lösungen ist stets darauf zu achten, dass keine sensiblen Daten geteilt werden.
- ³ Landesarchiv NRW Abteilung Westfalen, C 001/Fürstabtei Corvey – Akten, Nr. 23–35, 40–160, 162–164 u. 166–182.
- ⁴ Falls dieser Schritt ein zweites Mal durchlaufen werden sollte, um die Fehlerrate der Texterkennung zu reduzieren, so empfiehlt sich statt reiner Dilatation eher Opening (bzw. Closing), also eine Kombination von Erosion gefolgt von Dilatation (bzw. andersherum). Erosion beschreibt die Entfernung aller Textstrukturen, die kleiner sind als das Strukturelement, Details werden entfernt, die Schrift verkleinert.
- ⁵ Dabei ist anzumerken, dass durch Training eines eigenen Modells bei Transkribus vermutlich eine niedrigere Fehlerquote bei ungleich höherem Arbeitsaufwand möglich wäre.

Tabelle 1: Fehlerquoten in der Texterkennung bei verschiedenen Modellen. Zeichenfehlerquote CER (Charakter Error Rate) berechnet sich aus der Levenshtein-Distanz / korrekter Vorschlag

Anbieter	Modell	CER Gedruckte Seite	CER Handschriftliche Seite
PERO OCR	Kurrent	3,10 %	11,60 %
PERO OCR	Universal	1,62 %	6,83 %
Tesseract	German 4.1.0	28,70 %	n/a
Tesseract	German Fraktur 4.1.0	13,90 %	n/a
Transkribus	The German Giant I	n/a	10,30 %
Transkribus	The Text Titan I	n/a	9,74 %
Transkribus	Transkribus Print M1	1,76 %	n/a

Ergebnisse in unserem Beispiel, dass die Reichstagsberichte aufgrund ihres thematischen Fokus vor allem bedeutsame Ortsnamen des Heiligen Römischen Reichs und wenige aus der Region um Corvey enthalten.

Um diese Problematik zu adressieren, wählten wir einen Ansatz, der zum einen den spezifischen Kontext der Dokumente berücksichtigt, zum anderen aber auch möglichst niederschwellig ist. Dazu erstellten wir auf Basis des Gemeindelexikons für die Provinz Westfalen von 1895 einen kleinen Datensatz und zogen aufgrund der geographischen Nähe Corveys zum Fürstentum Lippe zudem ergänzend die Ortsnamen aus dem Ortschaftsverzeichnis des Fürstentums Lippe von 1871 heran.⁶ Im Folgenden ließen wir den Text auf Basis der in unserem Datensatz auftretenden Ortsnamen neu durchsuchen.⁷

Für fortgeschrittene Benutzer*innen besteht an dieser Stelle zudem die Möglichkeit, die Ergebnisse um die Identifier aus der Gemeinsamen Normdatei (GND) zu erweitern. Dazu haben wir die tagesaktuellen geographischen Entitäten aus der GND-Datenbank heruntergeladen und die in den Berichten gefundenen Ortsnamen damit abgeglichen. Bei eindeutigen Treffern können auf Ebene der Verzeichnungseinheit nun nicht nur auftretende geographische Einheiten, sondern auch die dazugehörigen GND-Identifier ausgegeben werden. Die Verwendung der GND in archivischen Fachinformationssystemen (AFIS) erfordert geeignete Schnittstellen und Felder, welche bis dato nur wenige Hersteller anbieten. Die Erschließungssoftware „VERA Verzeichnung“ des Landesarchivs NRW sieht dies erst für künftige Versionen vor, so dass bislang in nur einzelnen Projekten Normdaten in externe Datenbanken eingepflegt werden können. Organisatorische Voraussetzungen wie eine GND-Agentur für Archive in NRW fehlen überdies, um die stark bibliothekarisch geprägten Normdaten aus archivischer Perspektive weiterzuentwickeln – auch als Vorbereitung für Linked-Data und den zu erwartenden einheitlichen Gesamtstandard „Records in Context“ (RiC). Unbedingt müssen also die AFIS für Normdaten und KI ertüchtigt werden, bevor moderne Verzeichnungsstandards zum Einsatz kommen können.

Dieser zweigleisige Ansatz – entweder die Verwendung vor-trainierter NER-Modelle oder das händische Erstellen und Durchsuchen eigener Register – beweist, wie flexibel und auf den jeweiligen Bestand abgestimmte Workflows im Archivalltag angewandt werden können.

EINSATZ VON GENERATIVER KI IM ARCHIVALLTAG

Der dritte Schritt unseres Workflows beschäftigt sich mit dem Einsatz generativer Künstlicher Intelligenz (KI). Während die vorgehenden Schritte – Texterkennung und Natural Language Processing – auf der Verarbeitung bereits vorhandener Texte basieren, eröffnet uns generative KI die Möglichkeit, neue Inhalte zu erstellen oder bestehende Informationen zu verknüpfen und zu interpretieren. Für das Archiv kann dies ein wertvolles Werkzeug sein, um die Erschließung und spätere Nutzung von Beständen zu unterstützen.

Zwar haben wir oben die Erkennung von Entitäten durch spezialisierte NLP-Tools wie spaCy übernehmen lassen,

jedoch zeigt generative KI ebenfalls Potential in diesem Bereich. Insbesondere durch ihre Fähigkeit, flexibel auf verschiedene Kontexte oder auch Fehler in der Texterkennung zu reagieren, können diese Modelle hier ergänzend eingesetzt werden.

Dabei ist zwischen zwei Zielen zu unterscheiden: Entweder ist ein hoher Recall anzustreben, d. h. es werden möglichst viele Entitäten erkannt oder der Fokus liegt auf einer hohen Precision, also dass möglichst viele der erkannten Entitäten korrekt identifiziert werden. Für die Unterstützung bei der Verzeichnung benötigen wir vorrangig eine hohe Precision, wofür Large Language Models (LLMs) nur bedingt geeignet sind. Sie verfügen zwar über ein breites Wissen, sind jedoch zum einen nicht für die strukturierte Extraktion von Entitäten optimiert und zum anderen basieren LLMs letztlich auf statistischen Wahrscheinlichkeiten, die Texte nach der Wahrscheinlichkeit der Passgenauigkeit erzeugen. Das heißt, während bei der Named Entity Recognition von spaCy, flairNLP u. a. ein speziell trainiertes Modell zur Erkennung von Entitäten hinterlegt ist, erraten LLMs die Entitäten nur einfach sehr zuverlässig. Anhand der untersuchten Reichstagsprotokolle erkennen wir, dass sich spezialisierte NER-Modelle für diesen Zweck zwar anbieten, generative KI dank ihrer Funktionsweise aber auch sehenswerte Ergebnisse hervorbringen: LLMs erkennen zwar eine etwas geringere Anzahl an Entitäten als die ordinäre NER, sie finden jedoch wiederum Orte, die spaCy nicht als geographische Entität identifizierte. Dies scheint vor allem in der Fähigkeit zu liegen, falsche Schreibweisen oder falsch erkannte Buchstaben ignorieren zu können.⁸

Ein wesentlicher Aspekt beim Einsatz generativer KI ist der Datenschutz. Diese Problematik entfällt in unserem Falle glücklicherweise, da wir in unserem Ansatz ausschließlich auf Archivgut des Alten Reiches zurückgreifen. Dieses bietet sich nicht nur wegen der grundsätzlichen Unbedenklichkeit hinsichtlich des Datenschutzes, sondern auch wegen der Fülle der Digitalisate zum Aufbau großer Textkorpora zum Ausprobieren neuer Methoden an. Dennoch gilt unverändert, dass viele der derzeit verfügbaren großen Modelle von Anbietern betrieben werden, die ihre Server nicht in Deutschland oder der Europäischen Union hosten. Es ist also trotz anderslautender Behauptungen unklar, wie OpenAI hinter ChatGPT oder Google hinter Gemini mit mitunter hochsensiblen Angaben aus dem Archivgut verfahren. Um diese Bedenken zu umgehen, haben wir uns bewusst für eine on premises-Lösung entschieden. Dafür setzen wir PrivateGPT ein, ein Framework, das LLMs lokal auf eigenen Geräten betreibt. Für unsere Zwecke setzen wir dabei auf das Modell Mixtral 8x22B, das eine gute Balance zwischen Leistung und Ressourcenverbrauch bietet. Diese Lösung hat zwei entscheidende Vorteile: Erstens bleiben sämtliche Daten auf dem lokalen Rechner, was Datenschutzbedenken minimiert. Zweitens ermöglicht dieser Weg nach einer einmaligen technischen Einrichtung die Nutzung auch ohne spezialisiertes IT-Wissen.

Unsere Zielsetzung sah vor, generative KI zur Unterstützung bei der Verzeichnung einzusetzen, also sowohl einen Aktentitel als auch einen Enthält-Vermerk vorschlagen zu lassen. An dieser Stelle punktet die Funktionsweise von LLMs, da sie auch aus mitunter fehlerhaft erkannten, unstrukturierten

rierten Textkorpora Informationen gewinnen kann. Der Arbeitsablauf gliederte sich für die 156 Reichstagsberichte auf folgende zwei Punkte: (1) Ingest der Verzeichnungseinheiten: Zunächst wurden die 156 Textdateien, die jeweils eine Verzeichnungseinheit enthalten, in das Modell eingelesen. Das LLM wurde darüber hinaus nicht weiter trainiert oder angepasst. Dieser Vorgang dauerte etwa eine Stunde. (2) Ausgabe von Aktentiteln und Vermerken: In einem zweiten Schritt ließen wir mittels eines präzise formulierten Prompts („Du bist eine Archivarin und Spezialistin auf dem Gebiet der Verzeichnung. Gib für den vorliegenden Text mit dem Namen Nr. [XXX] einen prägnanten Titel für unsere Findmittel aus. Schreibe zudem einen Vermerk, der stichwortartig die wichtigsten Themen, Personen oder Ereignisse abbildet. Beachte dabei, dass es sich innerhalb eines Textes um mehrere Schriftstücke oder Drucke handelt.“) uns für jede Akte einen Vorschlag für Titel und Enthält-Vermerk generieren.⁹ Die Verarbeitungsgeschwindigkeit ist dabei mit etwa einer Minute pro Abfrage vertretbar, sodass wir bei den exemplarisch vorgestellten Reichstagsberichten innerhalb weniger Stunden Ergebnisse für alle 156 Verzeichnungseinheiten ausgeben lassen können.



Abbildung 2: Bildschirmfoto des aufgesetzten und mit Reichstagsberichten angefüllten Chatbots während der Ausarbeitung von Verzeichnungsvorschlägen

Die Ergebnisse zeigen, dass der Einsatz generativer KI in der Verzeichnung vielversprechend ist. Zur Erweiterung bieten sich mehrere Ansätze an, die eine präzisere und anwendungsorientierte Nutzung im Archivalltag vereinfachen können. Ohne technischen Aufwand stellt Prompt Engineering einen zentralen Hebel dar. Durch die Optimierung der Eingaben können die generierten Ergebnisse besser auf die spezifischen Anforderungen im Archivkontext zugeschnitten werden. Etwas aufwändiger, aber bei wiederholtem oder größerem Einsatz empfehlenswert ist elaborierteres Retrieval-Augmented Generation, bei dem das LLM mit einer eigenen Wissensdatenbank verknüpft wird. Diese könnte beispielsweise einschlägige archivische und historische Literatur und Wissen enthalten, sodass die KI beim Generieren von Ergebnissen auf einen spezialisierten Kontext zugreifen kann. Für besonders anspruchsvolle Szenarien ist schließlich auch ein Fine-Tuning des Modells denkbar. Hierbei wird die generative KI mit spezifisch archivischen Daten und Texten trainiert, um ihre Fähigkeiten optimal an die Bedürfnisse

von Archiven anzupassen. Dieser Weg ist allerdings mit erheblichem Arbeits- und Zeitaufwand verbunden, der sich nur bei fortwährendem Einsatz einer „eigenen“ KI lohnt.

EFFIZIENZ UND NEUE MÖGLICHKEITEN

Die Ergebnisse des vorgestellten Projekts verdeutlichen, dass sich bereits mit vergleichsweise geringem Aufwand beachtliche Fortschritte bei der Verzeichnung und Nutzung von bereits digitalisiertem Archivgut erreichen lassen. Der vorgestellte Workflow aus Text- und Handschriftenerkennung, Natural Language Processing und dem Einsatz generativer KI ist flexibel, skalierbar und an die individuellen Bedürfnisse der Arbeit anpassbar. In kurzer Zeit konnten wir aus 156 Verzeichnungseinheiten gezielt Informationen gewinnen und Vorschläge zu einer erweiterten Verzeichnung generieren.

Die Texterkennung zeigt, dass auch ohne kostspielige Tools oder langwierige Vorbereitungen bereits eine solide Grundlage für die weitere Verarbeitung geschaffen werden kann. Für größere Bestände oder Anwendungsbereiche empfiehlt sich das Trainieren eigener Texterkennungsmodelle und das Verlagern der Erkennung auf eigene Programme, zum Beispiel auf eine Instanz von eScriptorium.¹⁰ Der Einsatz von Named Entity Recognition stellt durch eine automatische Extraktion von relevanten Begriffen, hier geographischen Entitäten und einer anschließenden Verknüpfung mit Normdaten die Qualität der Verzeichnung sicher. Zuletzt bietet die Integration generativer KI ein hohes Potential. Mit KI-generierten Vorschlägen für prägnante Aktentitel sowie korrekte, aussagekräftige und umfangreiche Enthält-Vermerke – mit Hilfe von KI dann oft überhaupt erstmals vergeben – lassen sich bislang nur oberflächlich erschlossene Verzeichnungseinheiten detaillierter beschreiben. Das ist nicht nur ein Gewinn für spätere Benutzer*innen, sondern auch für das Archiv selbst: Wissen, das bislang auf die Expertise einzelner Archivar*innen beschränkt war, kann so nachhaltig verfügbar gemacht werden.

- 6 Königlich Preussisches Statistisches Bureau: Gemeindelexikon für die Provinz Westfalen. Mit einem Anhang, betreffend die Fürstenthümer Waldeck und Pyrmont. Berlin 1897, Ortschafts-Verzeichnis des Fürstenthums Lippe. Detmold 1875.
- 7 Alternativ ist spaCy mit einem eigenen Corpus für die Erkennung geographischer Entitäten trainierbar.
- 8 Dieses Thema wird aktuell breit diskutiert, siehe u. a. Shuhe Wang et al.: GPT-NER. Named Entity Recognition via Large Language Models, 2023, online unter <https://doi.org/10.48550/arXiv.2304.10428> (abgerufen am 14.11.2024), sowie die Beiträge von Nina Rastinger: Named Entity Recognition mit LLMs. AGKI-AG, 2024, online unter <https://agki-dh.github.io/pages/page-7.html>, abgerufen am 14.11.2024 und Torsten Hiltmann et al.: NER, aber prompt! Named Entity Recognition mit Large Language Models für historische Texte. Digital History Berlin. 2024, online unter <https://doi.org/10.58079/11vni> (abgerufen am 14.11.2024).
- 9 Grundsätzlich ist auch die Angabe der Laufzeit möglich, aufgrund der fehleranfälligen Erkennung von Zahlen in Handschriften ist dies nur für Drucksachen empfehlenswert.
- 10 Gefördert vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg unterhält beispielsweise die Universitätsbibliothek Mannheim eine eigene eScriptorium-Instanz, siehe <https://ocr-bw.bib.uni-mannheim.de/> (abgerufen am 14.11.2024).

Ein langfristiger Ansatz könnte die Entwicklung einer eigenen AI, einer „Archivischen Intelligenz“ sein – einer KI, die durch kontinuierlichen Ingest weiterer Digitalisate und gezieltem Fine-Tuning sukzessive an die spezifischen Anforderungen eines Archivs angepasst wird. Diese AI in Form eines Chatbots im Frontend hätte nicht nur einen enormen Skaleneffekt, sondern würde auch die Basis für neue Nutzungsszenarien neben der Verzeichnung schaffen, etwa die Unterstützung bei der gezielten Beantwortung komplexer Benutzer*innenanfragen. Die vorgestellten Technologien sind bereits heute praktikabel und zeigen, dass der Weg zur Nutzung digitaler Methoden im Archivwesen unumgänglich ist. Die Herausforderung liegt nun darin, diesen Weg möglichst nachhaltig, effizient und rechtlich sicher einzuschlagen.

THE DIGITISED ARCHIVAL MATERIAL AS DATA SOURCE

*A workflow for the use of already digitised archival records
This paper presents a workflow for enhancing the accessibility of already digitised archival materials using accessible tools and technologies, consisting of 3 steps: text and handwriting recognition, text processing via Natural Language Processing (NLP), and evaluation with generative AI. It focuses on 156 Reichstag reports from 1681–1803, aiming to extract geographical entities and suggest archive metadata. The approach minimises technical complexity, ensures data privacy, and offers scalable solutions for improved archival documentation and usability.*

Florian Probst

Universität Münster
Historisches Seminar
Domplatz 20-22, 48143 Münster
E-Mail: florian.probst@uni-muenster.de

Dr. Thomas Reich

Landesarchiv NRW
Abteilung Westfalen
Bohlweg 2, 48147 Münster
E-Mail: thomas.reich@lav.nrw.de

DIE KI-BASIERTE ERSCHLIESSUNG VON ARCHIVBESTÄNDEN FÜR DIE PROVENIENZFORSCHUNG IN NRW

BEST PRACTICE – NEXT PRACTICE

von Ruth von dem Bussche und Jasmin Hartmann

Die Provenienzforschung hat einen nicht zu überschätzen- den Anteil an Restitutionen oder alternativen gütlichen Einigungen, sogenannter gerechter und fairer Lösungen im Sinne der sich im vergangenen Jahr zum 25. Mal jährenden *Washingtoner Prinzipien*. Verortet als Wissenschaft zwischen Recht und Moral, trägt sie entscheidend zur Identifizierung von NS-Raubgut und zur Eruierung der Fakten bei, um der historischen Wahrheit möglichst nah zu kommen und damit bestmöglich zur Lösungsfindung beizutragen. Ihr Erfolg hängt jedoch unmittelbar von dem Grad des Zugangs zu Archiven, Quellen und Informationen, sowie grundsätzlich von dem Wissen um die für die Fragestellung relevanten Archivbestände ab.¹ Bereits die *Washingtoner Prinzipien* (vgl. Prinzip Nr. 2),² als auch die neuen *Best Practices* betonen nochmals ausführlicher die essentielle Bedeutung von Archiven und Quellen (vgl. Punkte A, G, H, M).³ Eine Evaluation der Erfolge und Bedarfe sowie ein ganzheitliches Konzept zur Stärkung der Provenienzforschung hinsichtlich einer nachhaltigen und unabhängigen Forschungs- und Wissensinfrastruktur und der damit verbundenen Qualitäts- und Effizienzsteigerung von Recherchen steht jedoch bis heute aus. Dies unterstreicht auch Meike Hopp, Juniorprofessorin für Digitale Provenienzforschung der Technischen Universität Berlin und zugleich Vorsitzende des Arbeitskreises Provenienzforschung e. V., der einzigartigen internationalen Vereinigung von Wissenschaftlerinnen und Wissenschaftlern auf dem Gebiet der Provenienzforschung mit über 500 Mitgliedern.⁴ Sie konstatierte gar ein Leck an Lösungen hinsichtlich der Zugänglichmachung von Archiven und Quellen, von relevanten Forschungsdaten und Informationen.⁵

Die Provenienzforschung steht insbesondere vor der Herausforderung, umfangreiche und heterogene Archivbestände effizient auszuwerten, um die Herkunft und den Erwerbskontext von Kunst- und Kulturgütern aufklären zu können. Relevante Dokumente dazu finden sich in öffentlichen Archiven des Bundes, der Länder, der Kommunen oder Gemeinden, aber auch beispielsweise in den Registraturen von Museen und Verwaltungen oder im Privaten (Archive des Kunsthandels, des Auktionswesens, von Privatpersonen). Um Wissenslücken in Provenienzen schließen zu können, zielen Recherchen häufig auf sehr detaillierte Fakten: das

¹ Zu dem Verhältnis Provenienzforschung, Archive und Restitutionen, vgl. Susanne Hehenberger: Archive und Restitutionen. In: Restituiert. 25 Jahre Kunstrückgabegesetz in Österreich. Hg. v. Birgit Kirchmayr u. Pia Schölnberger (Bd. 9, Schriftenreihe der Kommission für Provenienzforschung), S. 89–103.

² Über die Washingtoner Prinzipien sowie weitere Maßgaben zum Umgang mit NS-verfolgungsbedingt entzogenem Kulturgut informiert das Deutsche Zentrum Kulturgutverluste auf deren Website, <https://kulturgutverluste.de/kontexte/ns-raubgut> (letzter Zugriff: 04.11.2024).

³ Im Rahmen einer von der World Jewish Restitution Organization (WJRO) und dem US-Außenministerium einberufenen Veranstaltung am 05.03.2024 wurden in 15 Punkten (A-O) Best Practices vorgestellt, welche die praktische Umsetzung der Washingtoner Prinzipien verdeutlichen und verbessern sollen. Sie wurden inzwischen von insgesamt 24 Staaten unterzeichnet, darunter auch Deutschland. Nähere Informationen zu den Best Practices, <https://wjro.org.il/best-practices-global-art-report/> (letzter Zugriff: 04.11.2024).

⁴ Vgl. <https://www.arbeitskreis-provenienzforschung.org> (letzter Zugriff: 04.11.2024).

⁵ Zur Veranstaltung Best Practices and the way forward am 05.03.2024 in Washington D.C., <https://www.youtube.com/watch?v=3djokRCc0rY> (letzter Zugriff: 15.11.2024).

Handeln einzelner Akteure im Rahmen von Transaktionen, Besitzwechsel (Ausstellungen, Leihverkehr, Transport) und Eigentumsübergänge (Kauf/Verkauf, Tausch, Schenkung), deren Handelsbeziehungen und die Translokation von Kulturgütern. In der Regel geht es um das Auffinden konkreter Nachweise zu einem Werk oder einem Konvolut.⁶

Für die maschinelle Verarbeitung möglicher relevanter Archivbestände und das Finden der benötigten Schlüsseldokumente zielt die notwendige Recherche auf die Ebene von Seiten und einzelnen in den Archivalien identifizierbaren Vorgängen. Die Suche der nötigen Puzzlesteine ist somit äußerst punktuell und kleinteilig.

Der Einsatz von Verfahren der sogenannten Künstlichen Intelligenz (KI) verspricht hier neue Möglichkeiten, relevante Dokumente schnell aufzuspüren und die Auswertung von Akten zu beschleunigen und zu vertiefen.

Im Rahmen eines vom NFDI4Memory Incubator Funds geförderten neunmonatigen Projekts,⁷ welches an der Koordinationsstelle für Provenienzforschung in NRW⁸ angesiedelt ist, hat das Projektteam untersucht, wie KI-basierte Ansätze für die Erschließung und Auswertung von Archivgut in der Provenienzforschung nutzbar gemacht werden können. Ziel war es, Wege für die Verarbeitung großer Mengen von unstrukturierten Daten zu erproben, die gleichzeitig den fachlichen Anforderungen in der Provenienzforschung gerecht werden muss.

Alle generierten Daten aus dem Projekt werden in einer Suchmaschine auf der Basis von Elasticsearch gebündelt, über welche die Dokumente zunächst einer Peer Group und zukünftig der Fachcommunity zur Verfügung gestellt werden. Der direkte Nutzen der generierten Daten für die Provenienzforschung ist das eine Ziel des Projekts. Gleichzeitig kann die Verarbeitung der Daten mittels KI grundsätzlich die Potenziale für die Archive ausloten helfen. Es wurde bewusst auch der Einsatz sog. großer Sprachmodelle (LLM) über Clouddienste erprobt, um verstehen zu lernen, welche Verarbeitungsqualität sich mit den aktuellen Modellen (sowohl mit kommerziellen Diensten als auch Open Source-Lösungen) derzeit erzielen lässt.

Dieser Beitrag stellt erste Ergebnisse des Projekts vor und diskutiert die Potenziale und Herausforderungen der KI-Anwendung für die Provenienzforschung auf Archivmaterial. Ein besonderer Fokus liegt dabei auf der notwendigen Verschränkung von kunsthistorischer und archivischer Expertise in Verbindung mit neuen technologischen Möglichkeiten. Abschließend werden Perspektiven für die Weiterentwicklung KI-gestützter Verfahren in der Archivarbeit exemplarisch skizziert.

1. THEORETISCHE GRUNDLAGEN

Der Begriff „Künstliche Intelligenz“ bezeichnet Computersysteme, die in der Lage sind, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern, wie Wahrnehmung, Erkennung, Entscheidungsfindung und Sprache.⁹ Im engeren Sinne ist mit KI heute meist maschinelles Lernen gemeint, also die Fähigkeit von Computerprogrammen, selbstständig Muster und Regeln aus Daten zu extrahieren, ohne dass diese explizit programmiert wurden.¹⁰ Im Archivkontext sind insbesondere Sprachmodelle rele-

vant, die auf Deep Learning-Architekturen wie neuronalen Netzen beruhen. Diese Modelle werden auf sehr großen Textkorpora trainiert und können so auf einer statistischen Grundlage Eigenschaften und Zusammenhänge auswerten und in natürliche Sprache ausgeben.¹¹ Bekannte Beispiele sind die GPT Modelle, sogenannte Decoder-Modelle, die auf die Generierung von Texten hin optimiert sind, oder BERT, ein sogenanntes Encoder-Modell, spezialisiert auf die Klassifikation von Texten.¹² Typische Aufgaben dieser Modelle sind, neben der erwähnten Aufgaben der Textklassifikation und Textgenerierung, Named Entity Recognition, Datenextraktion oder das Erstellen von Zusammenfassungen. Ein wesentliches Merkmal von Sprachmodellen ist, dass sie rein datengetrieben auf statistischen Grundlagen arbeiten und kein explizites Wissen über die Welt oder bestimmte Fachdomänen enthalten. Sie sind also keine Wissensmodelle, aber doch hervorragend in der Lage, mit großen Textmengen umzugehen.

Für den Archivbereich bedeutet dies, dass große Sprachmodelle insbesondere Potenzial im Bereich der Erschließung und Auswertung von Dokumentenbeständen bieten, gerade auch bei schwach strukturierten oder heterogenen Daten. Sie können beispielsweise eingesetzt werden, um relevante Informationen zu extrahieren, Dokumente zu klassifizieren oder zusammenzufassen.

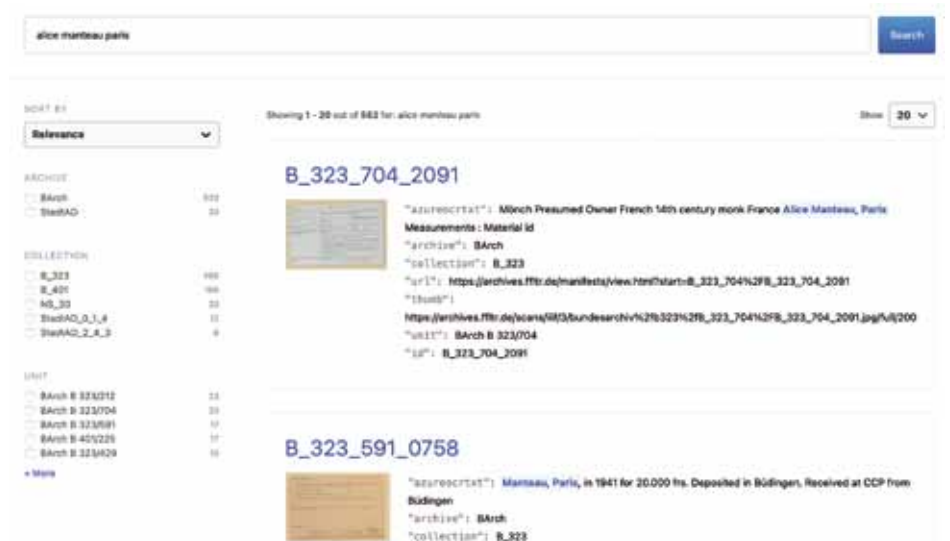
2. METHODISCHE UMSETZUNG

2.1. OCR und Suchmaschine

Unsere Arbeitspipeline im Projekt beginnt mit dem OCR. Zur Extraktion von Textinhalten aus den digitalisierten Dokumenten wurden zunächst verschiedene OCR-Verfahren evaluiert, u. a. ABBYY Finereader und Tesseract. Letztlich hat sich das Team für den Clouddienst von Azure entschieden, da bei diesem Dienst bereits intern KI-basierte Verfahren die Qualität des OCRs wesentlich beeinflusst haben und im Ergebnis eine deutlich bessere Qualität festgestellt werden konnte. Die Verarbeitung erfolgt über eine Cloud, über die sich große Dokumentenmengen in parallelen Prozessen gut skalieren lassen.¹³ Das zeilenweise OCR von Azure kann einen sehr großen Teil der Dokumente verarbeiten, was zum Teil auch handgeschriebene Ergänzungen beinhaltet. Azure Computer Vision API könnte ebenso für komplexe Layouts wie etwa Karteikarten zum Einsatz kommen. Nach diesem ersten Arbeitsschritt wird der OCR-Text seitenweise zusammen mit den Digitalisaten in einer Suchmaschine gespeichert.

Nutzend können damit auf kleinteiligen Einheiten (Einzelseiten) recherchieren. Das Handling der Digitalisate, über einen Bildserver (aktuell ist Cantaloupe im Einsatz) mit der Möglichkeit zu browsen und zu blättern, erlaubt es auch innerhalb einer Akte navigieren und sich damit Kontexte erschließen zu können.

Bereits mit dieser schnellen und „flachen“ Verarbeitung konnte ein für die Provenienzforschung sehr robustes und nützliches Recherchetool geschaffen werden, welches das Auffinden einzelner Textstellen beziehungsweise Dokumente teils überhaupt oder in jedem Fall sehr viel schneller und effizienter ermöglicht.



Suchmaschine auf der Basis von Elasticsearch

Die in jedem OCR vorhandenen qualitativen Probleme, konkret fehlerhafte Textwiedergaben, werden durch die eingesetzte Suchmaschine ausgeglichen, da diese eine Unschärfe (Fuzziness) von bis zu zwei Buchstaben bei der Suche über Elasticsearch bietet. So erscheinen auch verschriebene Namen oder Wörter bei der Suche noch in Ergebnislisten.

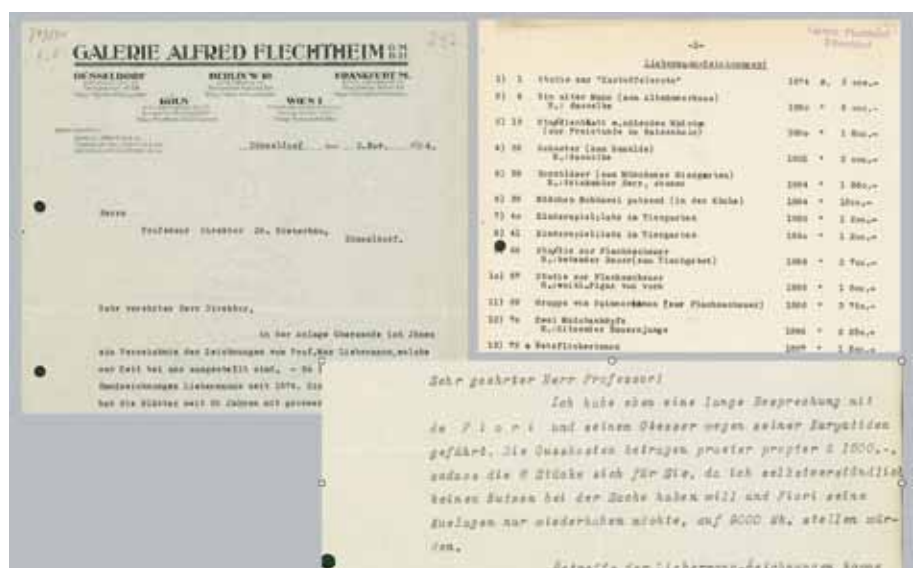
2.2. Text Summarization

Für die weitere Verbesserung der via OCR generierten Texte werden auf ausgewählten Beständen Zusammenfassungen (text summerarization) mittels Large Language Models generiert.

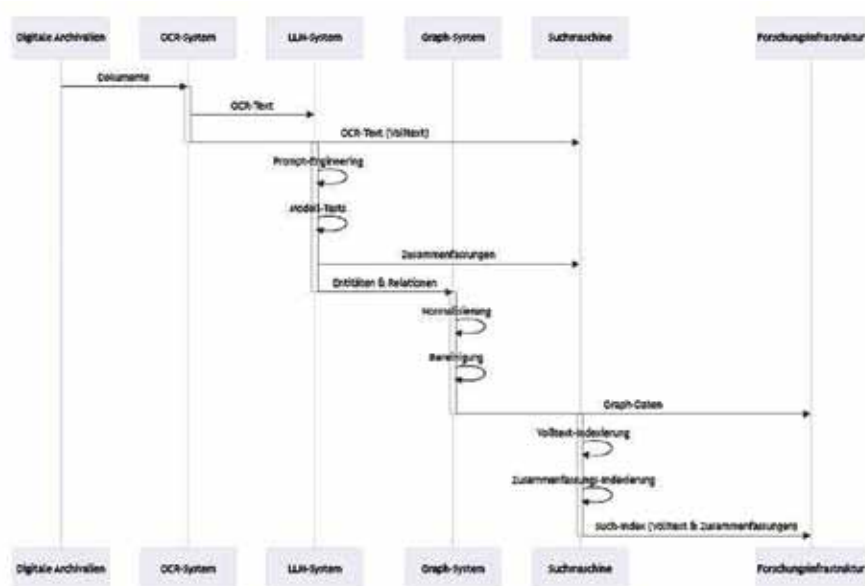
Grundlage für die Zusammenfassungen bildeten Archivbestände der ehemaligen Städtischen Kunstsammlungen

- 6 Meike Hopp, Ruth von dem Bussche: Provenienzforschung und ihre Quellenbestände. Aktuelle Nutzungsszenarien zwischen Open Access und Inaccessibility, DHd 2023, <https://zenodo.org/records/7715361> (letzter Zugriff: 20.11.2024).

- 7 Das Projekt „Maschinelles Sehen und Distant Reading auf Archivbeständen für neue methodische Standards in der Provenienzforschung“ ist bei der Koordinationsstelle für Provenienzforschung in NRW angesiedelt und läuft von April bis Dezember 2024 im Bereich TA2 Data Services bei NFDI4memory, <https://4memory.de/4memory-incubator-funds-projekte-2024/> (letzter Zugriff: 20.11.2024).
- 8 Die Koordinationsstelle für Provenienzforschung in NRW ist eine Serviceeinrichtung zur systematischen, flächendeckenden und nachhaltigen Etablierung der Provenienzforschung in allen Unrechtskontexten und für sämtliche Zielgruppen in NRW. Nähere Informationen vgl. www.kpf.nrw (letzter Zugriff: 04.11.2024).
- 9 Stuart Russell, Peter Norvig: Künstliche Intelligenz. Ein moderner Ansatz. Aktualisierte Auflage, München 2023.
- 10 Christopher M. Bishop: Pattern recognition and machine learning. New York, 2019.
- 11 Große Sprachmodelle – Grundlagen, Potenziale und Herausforderungen für die Forschung (Whitepaper), Deutsche Akademie der Technikwissenschaften, 23.05.2023, <https://www.acatech.de/publikation/grosse-sprachmodelle/> (letzter Zugriff: 14.11.2024); Palash Goyal, Sumit Pandey u. a.: Deep: Learning for Natural Language Processing. Bangalore 2018, [http://prof.khuisf.ac.ir/images/Uploaded_files/Deep_Learning_in_NLP_1\[4062289\].PDF](http://prof.khuisf.ac.ir/images/Uploaded_files/Deep_Learning_in_NLP_1[4062289].PDF) (letzter Zugriff: 20.11.2024).
- 12 Jacob Devlin, Ming-Wei Chang, u. a.: BERT: Pre-training of Deep Bi-directional Transformers for Language Understanding, 11.10.2018, <https://arxiv.org/abs/1810.04805> (letzter Zugriff: 14.11.2024).
- 13 Microsoft Azure OCR, <https://azure.microsoft.com> (letzter Zugriff: 04.11.2024).



Ausgewählte Einzeldokumente aus den Ankaufsakten der Städtischen Kunstsammlungen Düsseldorf im Stadtarchiv Düsseldorf



Datenverarbeitung und Daten Services im Projekt

Düsseldorf aus dem Zeitraum 1925 bis 1960 des Stadtarchivs Düsseldorf, insbesondere die Ankaufs- und Angebotsakten aus der Zeit des Nationalsozialismus. Die Akten dokumentieren unter anderem die Erwerbungsverfahren des Museums und enthalten wertvolle Informationen zur Provenienz der angekauften und angebotenen Kunstwerke, darunter Korrespondenzen, Rechnungen, Gutachten und Inventarlisten. Der Bestand wurde in einem vom Deutschen Zentrum Kulturgutverluste (kurz: Zentrum) finanzierten Projekt in Vorgängen (je Angebot, Ankauf, Tauschgeschäft, Schenkung etc.) erschlossen und in ACTApro detailliert nach den Bedarfen der Provenienzforschung dokumentiert (Adressat, Empfänger, Art des Vorgangs, Kurzbeschreibung, Hinweise auf Literatur, Provenienzen, Verlustzusammenhänge etc.). Insgesamt umfasst der tiefer erschlossene und zudem digitalisierte Bestand rund 65.000 maschinengeschriebene und teils handschriftlich annotierte Seiten.¹⁴

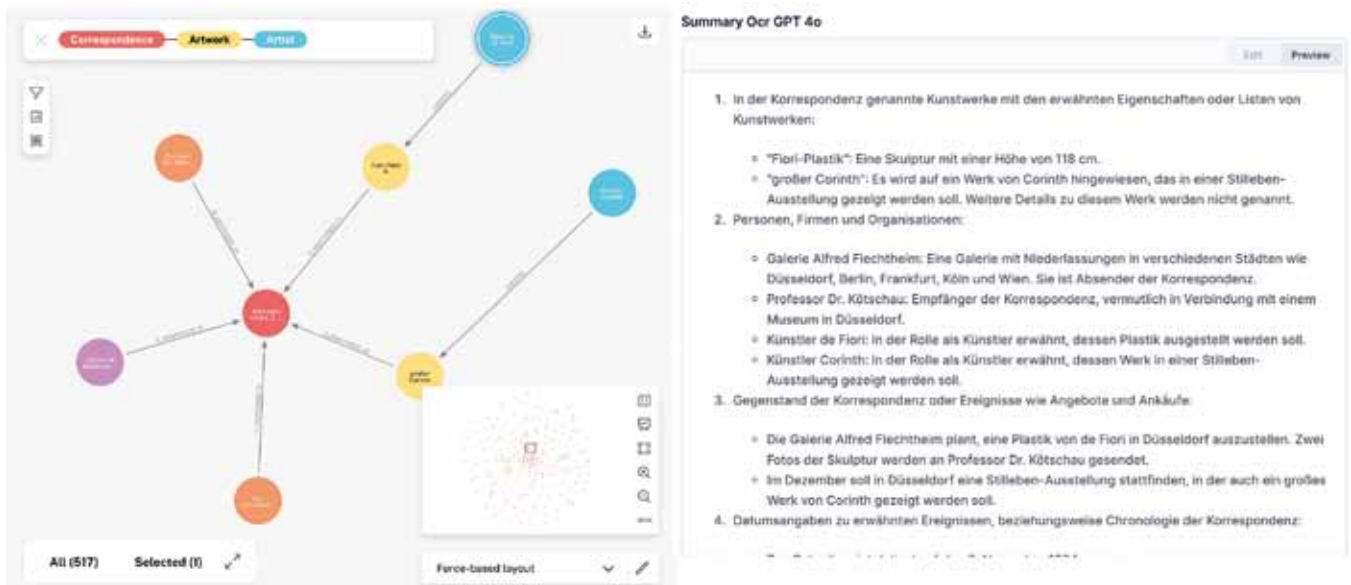
Die Bildung von Vorgängen in dem Düsseldorfer Projekt war im jetzt laufenden Projekt für den Projektschritt der *text summarization* von Vorteil, weil damit Texteinheiten in einer Größe handhabbar gemacht werden konnten, die ein Sprachmodell gut verarbeiten kann. Im Vergleich zu den Metadaten aus der Erschließung im vom Zentrum geförderten Projekt zeigt sich, dass die häufig in Archiven gebotene Kürze in der Beschreibung von Akten nun durch sehr viel ausführlichere Zusammenfassungen ergänzt werden kann.

Die Qualität der Zusammenfassungen einzelner ausgewählter Vorgänge wurde in Abgleich mit den Digitalisaten eingehend überprüft und über mehrere Iterationen hinweg verbessert. Erste Zusammenfassungen wurden in einer Peer Group, bestehend aus Mitgliedern des Arbeitskreises Provenienzforschung e. V., fachlich bewertet. Auf dieser Basis wiederum konnten die dazu verwendeten Texteingaben wie Fragen und Anweisungen (Prompts) weiter optimiert werden. Insbesondere war den Fachkolleginnen und -kollegen wichtig, dass alle vorhandenen Details zu den erwähnten Kunstwerken möglichst vollständig und genau wiedergegeben werden, inklusive Detailinformationen zu dokumentier-

ten Preisen oder Rollen und Funktion der in den Archivalien erwähnten Personen und Institutionen. Die Zusammenfassung selbst wurde dadurch immer weiter modifiziert. Dem Sprachmodell wurde zudem aufgetragen, die Daten der Korrespondenz selbst sowie die in den Schreiben erwähnten Ereignisse in einer chronologisch strukturierten Liste mit dem jeweiligen Datum wiederzugeben. Auch für den Fall, dass in einem Vorgang umfangreiche Werklisten eingefügt sind, wurde das Modell um entsprechende Hinweise gebeten, da diese in der Regel abgekürzt worden sind und in den Zusammenfassungen zu Anfang keinen Niederschlag fanden. Nach der Optimierung des Prompts wurden mit diesem verschiedene Modelle getestet, um die beste Qualität an Zusammenfassungen im Kosten- und Nutzenvergleich zu gewährleisten, da die Kosten sehr unterschiedlich ausfallen. Anschließend wurde der finale Prompt weiter ausgerollt, so dass weitere 100 Vorgänge aus den Akten über verschiedene Modelle verarbeitet worden sind. Die mit den verschiedenen Modellen generierten Zusammenfassungen wurden je Vorgang zu einer Datenansicht zusammengefasst, um die Ergebnisse der jeweiligen Modelle besser vergleichen und bewerten zu können.

In dieser Auswertung hat sich Claude 3.5 Sonnet von Anthropic als das zuverlässigste Modell erwiesen,¹⁵ mit dem die Zusammenfassungen für den Düsseldorfer Bestand nun produktiv umgesetzt werden. Die Zusammenfassungen ergaben einen gegenüber dem OCR stark verdichteten und bereinigten Text. Darüber hinaus könnten auf diesem Wege auch Zusammenfassungen in anderen Sprachen generiert werden. Das würde weitere Barrieren in der Nutzung der Akten senken.

Ein zentrales Ergebnis dieses Teilschritts war die Erkenntnis, dass für den gesamten Prozess – von den spezifischen Fragestellungen der Provenienzforschung an das Archivmaterial bis hin zu den erarbeiteten Prompts für die Datengenerierung – eine enge Zusammenarbeit zwischen inhaltlicher Fachexpertise (Provenienzforschung) und technischer Expertise (IT) entscheidend ist.



Extraktion von Entitäten und Relationen über LLMs sowie eine Graphdatenbank

2.3. Vom Text zum Knowledgegraph

Aktuell beschäftigt sich das Projektteam mit der Erstellung strukturierter Daten aus den Zusammenfassungen. Dabei wird in der Konversation mit einem LLM ein Datenmodell erstellt und über verschiedene Prompts sowie Testdaten weiter verfeinert. In der weiteren Verarbeitung werden Daten für einen Knowledgegraphen erstellt, die dann über weitere Verfahrensschritte bereinigt und vereinheitlicht werden.¹⁶

3. AUSBLICK

3.1. Problemstellungen und Hürden

KI-gestützte Verfahren sind immer statistische Verfahren. Das kann in den generierten Zusammenfassungen und Daten zu Unzulänglichkeiten führen, die bei manueller und gründlicher wissenschaftlicher Arbeit so nicht zu sehen wären. Diese Einschränkungen werden allerdings deutlich durch den Umstand aufgewogen, dass auf diesem Wege große Dokumentenbestände verarbeitet werden können und damit Dokumente überhaupt erst auffindbar werden und der Forschung zur Verfügung stehen.

Suchergebnisse aus KI-unterstützter Forschung müssen auf der Ebene von Einzeldokumenten verifiziert werden. Dafür wurde im Projekt für die Digitalisate und Zusammenfassungen mit der Suchmaschine eine eigene Umgebung geschaffen. Ziel des Projektes ist es, der Provenienzforschung ein Tool in die Hand zu geben, mit dem Einzeldokumente (rascher) auffindbar werden, die in der wissenschaftlichen Auswertung als Belege dienen. Letztendlich wird es notwendig sein, seitens der Archive stabile Links auf Einzelseiten einzurichten, die auch in publizierten Fachtexten referenziert werden können.

Ein zentrales Thema im Projekt war die Frage der Datenqualität.¹⁷ Bereits die Qualität von Scans bestimmt ganz wesentlich die des daraus generierten OCR und damit alle weiteren

Arbeitsschritte. Die Wahl des richtigen Werkzeugs für das OCR hat einen ebenso großen Einfluss. Hier stellen die heterogenen Dokument- und Schrifttypen, eine Ansammlung aus Typoskripten, mitunter auch gedruckten Texten sowie sehr unterschiedlich ausgeprägten Handschriften in den Archivalien die wesentliche Herausforderung dar.

Die Qualität des OCR bedingt wiederum die der Zusammenfassungen. Die Testung und Auswahl eines geeigneten LLMs für die Zusammenfassungen sind wichtige evaluierende Maßnahmen, um die entsprechende Datenqualität zu gewährleisten.

Diese qualitativen Abhängigkeiten zeigen sich deutlich in weiteren Verarbeitungsstufen, d. h. in den Daten des maschinell erstellten Knowledgegraphen. Auch auf dieser abstrahierten Metadatenebene sind deshalb nachgeschaltete Bereinigungen notwendig.

3.2. KI im Archiv - KI für Archivare

Innerhalb des Internen Archivtags des Landesarchivs NRW am 29. Oktober 2024 in Gelsenkirchen konnte das Projekt-

¹⁴ Nähere Informationen zum Projekt siehe <https://www.duesseldorf.de/kulturamt/provenienzforschung-der-landeshauptstadt-duesseldorf/aktuelles> (unter „Online-Findmittel“, letzter Zugriff: 11.11.2024), Bestandsübersicht siehe https://www.duesseldorf.de/fileadmin/Amt41-Zoll/kulturamt/pdf/Provenienzforschung/Aktuel.UEbersicht_der_tiefenverzeichneten_und_gesannten_Akten_mit_Links.pdf (letzter Zugriff: 11.11.2024). Die Digitalisierung der Akten wurde im Rahmen des Förderprogramms „WissensWandel. Digitalprogramm für Bibliotheken und Archive innerhalb von NEUSTART KULTUR“ durchgeführt, die Tiefenerschließung der Akten finanzierte das Deutsche Zentrum Kulturgutverluste, <https://www.proveana.de/de/link/pro10000308> (letzter Zugriff: 11.11.2024).

¹⁵ Anthropic Claude 3.5, <https://www.anthropic.com/news/claude-3-5-sonnet> (letzter Zugriff: 22.11.2024).

¹⁶ Zuletzt vorgestellt auf dem Workshop (Generative) KI für Kultur- und Textdaten, Weimar, 4.–5. November 2024, Klassik Stiftung Weimar, Goethe- und Schiller-Archiv, <https://agki-dh.github.io/pages/activities/page-2.html> (letzter Zugriff: 22.11.2024).

¹⁷ Vgl. Beitrag von Tüselmann, Fink, Wolf und Raven in diesem Heft.

team mit Archivarinnen und Archivaren diskutieren, wie sich Archive besser auf Prozesse mit KI-generierten Daten einstellen können.

Ein Ergebnis, um Archivbestände künftig mithilfe von KI effizient erschließen und auswerten zu können, war es, dass zukünftig Akten in kleinere Untereinheiten untergliedert werden sollten. Diese Einheiten können entweder inhaltlich oder technisch definiert sein. Nur so lassen sich die Untereinheiten sinnvoll und mit vertretbaren Kosten durch Sprachmodelle verarbeiten. Modelle, die für die Massenverarbeitung geeignet sind, haben typischerweise ein begrenztes Kontextfenster. Dieses „Sichtfenster“ beschreibt die Menge an (Text-)Daten, die ein Modell verarbeiten kann. Es umfasst bei größeren Modellen in der Regel etwa 128.000 Tokens (entspricht ca. 100.000 Wörtern für Prompt und Antwort). In unserem Anwendungsfall durfte die Menge an übergebenem OCR-Text inklusive der generierten Zusammenfassung 128.000 Tokens nicht übersteigen.

Bislang spielt die Bildung kleinerer Einheiten innerhalb einer Archiveinheit in den Digitalisierungsworkflows von Archiven noch keine Rolle. Einschübe wie Trennblätter würden auf praktischem Level bereits eine Abhilfe schaffen und größere Archiveinheiten zur Weiterverarbeitung unterteilen. Die Arbeit mit Sprachmodellen über Prompts hat den Vorteil, dass dies ein realistischer Zugang zur Datengenerierung auch und insbesondere für Nicht-Informatiker ist. Sprachmodelle können umso bessere Ergebnisse zurückliefern, umso genauer Kontexte und Forschungsfragen definiert werden. Hier sind Archivarinnen und Archivare sowie andere Fachcommunities gefragt, diese Fragestellungen zu definieren und in Prompts zu übersetzen.

Es bedarf mehr Praxis, Erfahrung und Expertise, wie sich Arbeitspipelines zur Generierung von Archivdaten aufbauen und steuern lassen. Auch hier sind Archivarinnen und Archivare für qualitative Bewertung der KI-generierten Inhalte sicherlich die besten Ansprechpartnerinnen und -partner. Unsere über verschiedene Sprachmodelle generierten Daten wollen wir den Archiven zur weiteren Analyse zur Verfügung stellen. Das könnte auch den Nutzen haben, neuere mittels KI generierten Daten durch qualitative Vergleiche besser verorten zu können.¹⁸

Der Einsatz von KI wird die Zugänglichkeit zu Archivmaterial in Zukunft stark verbessern. Gleichzeitig wird es nötig sein, für mehr Transparenz zu sorgen, auf welchem Wege und mit welchen Werkzeugen die zur Verfügung gestellten Daten erstellt wurden. Der Zugriff auf das Digitalisat (den Scan) ist im Forschungskontext der notwendige letzte Schritt, um das, was über generierte Daten und Abfragen greifbar ist, zu verifizieren. KI-generierte Inhalte sollten entsprechend mit einem Disclaimer gekennzeichnet sein:

„Dieser Inhalt wurde von einer Maschine erstellt. Bitte überprüfe ihn sorgfältig, indem Du die vorhandenen Digitalisate von den bzw. die Originaldokumente überprüfst“.

Der Verarbeitung mittels KI sind aktuell in der EU rechtliche Grenzen gesetzt. Im Projekt konnten diese Einschränkungen insofern umgangen werden, indem man sich auf Bestände

beschränkt hat, die keine schutzwürdigen Belange Dritter umfassten und deshalb bereits öffentlich online verfügbar waren. Letztendlich ist aber absehbar, dass ein umfangreicher Einsatz von KI in Archiven erst wesentliche Veränderungen bringen kann, wenn rechtliche Rahmen- und Nutzungsbedingungen definiert werden. Die besondere Aufklärung von Völkerrechtsverbrechen und Unrechtskontexten müssen gegenüber schutzwürdigen Belangen Dritter auf Dauer besser abgegrenzt und priorisiert werden. Provenienzforschung findet auf einer internationalen Ebene statt. Diesbezüglich sind die Recherchen transnational und müssen sich mit den (datenschutzbedingten wie auch technischen) Möglichkeiten in den USA messen lassen. Es ist an der Politik, diese heterogenen juristischen Rahmenbedingungen für eine Zugänglichmachung von Informationen der diversen Bestandshalter nicht nur mit gutem Willen zu überwinden, sondern dafür gesetzliche Grundlagen zu schaffen.

AI-BASED INDEXING OF ARCHIVE HOLDINGS FOR PROVENANCE RESEARCH IN NORTH RHINE-WESTPHALIA

Provenance research is essential for the restitution of Nazi-looted art, relying on efficient access to archives and sources. A project by KPF.NRW, funded by NFDI4Memory, explores the use of AI for analyzing archival materials, particularly through OCR and language models. Initial results demonstrate that diverse data can be generated using different AI pipelines. Full texts and summaries are utilized, among other applications, in a search engine. Challenges include data quality, legal frameworks, and close collaboration between domain experts and technology specialists.

Dr. Ruth von dem Bussche

Talstraße 116, 40217 Düsseldorf
E-Mail: rbussche@fotostoria.de

Jasmin Hartmann

Koordinationsstelle für Provenienzforschung
in Nordrhein-Westfalen
Bachstraße 5–9, 53115 Bonn
E-Mail: jasmin.hartmann@kpf.nrw

¹⁸ Die Korrespondenzen in den Düsseldorfer Ankaufsakten wurde über folgende Modelle vergleichend verarbeitet: Mistral Small, ChatGPT 4.0, Claude Sonnet 3.5, Llama 31 70b, Llama 31 8b, Llama 32 3b. Auf diese Weise konnte evaluiert werden, welche Stärken und Schwächen die unterschiedlichen Modelle mit sich bringen. Auf dieser Basis kann in Zukunft überprüft werden, ob weitere Verfahren eine vergleichbare (oder bessere) Datenqualität liefern.

BILD-ERSCHLIESSUNG MIT KI

SECHS VORSCHLÄGE FÜR DIE NUTZUNG KI-GENERierter METADATEN

von *Miriam Eberlein*

Eine der vielen Einsatzmöglichkeiten von KI-basierten Tools im Archiv ist die Erschließung von Bildquellen. Die Technologien zur Erkennung von pixelbasierten Mustern, zur Extraktion komplexer Merkmale sowie zur Bildverarbeitung in Hochgeschwindigkeit haben seit Mitte der 2010er Jahre einen solchen Reifegrad erreicht, dass sie in Produkten für eine breite Nutzerbasis bis hin zu internetbasierten Recherchertools Anwendung finden. In der Hoffnung, dass sie sich dann auch bei der archivischen Erschließung von Fotos nutzbringend einsetzen lassen sollten, hat das Stadtarchiv Heilbronn im Jahr 2020 eine solche KI-Anwendung eingeführt. Seither erfolgt die Identifizierung von Objekten und Gesichtern auf den im Archivischen Fachinformationssystem (künftig: AFIS) erfassten Fotos mit Unterstützung der Software „DeepVA“ der Firma Aiconix (ehemals TheChainless). Die Integration der per KI-Analyse erzielten Ergebnisse in die Erschließungsdatensätze im AFIS – das Stadtarchiv Heilbronn nutzt hierfür die Eigenentwicklung HEUSS – geschieht automatisiert über eine Schnittstelle zwischen beiden Systemen.¹

NEUE MÖGLICHKEITEN – NEUE FRAGEN

Für die Integration dieser KI-generierten Erschließungsdaten mussten zunächst die dafür vorgesehenen Datenfelder im AFIS festgelegt werden. Da unser Projektziel die namentliche Identifizierung von Personen und Gebäuden war, bot es sich an, die KI-generierten Metadaten als GND-Normdaten im jeweiligen zugehörigen Datensatz zu verankern. Die Zuweisung von Normdaten zu Bilddatensätzen im Heilbronner AFIS HEUSS erfolgt seither entweder automatisiert mit KI-Unterstützung, oder manuell, oder in einer Kombination aus beidem.

Doch bei der KI-Analyse erzeugt die eingesetzte Software DeepVA schon in ihrer Standardkonfiguration deutlich mehr Daten, als für das Projektziel verwendet wurden. Bei der Analyse von Gesichtern beispielsweise wird nicht nur

das Ergebnis einer dem verwendeten KI-Modell bekannten („antrainierten“) Person zurückgeliefert, sondern auch das Metadatum „Gesicht“ – und damit „Person“ – sowie, falls erkannt, das Geschlecht. Auch eine Klassifikation des ungefähren Alters der Person wäre möglich. Bei einem Gebäudefoto können die Analyseergebnisse außer der Identifizierung des Gebäudes beispielsweise auch eine Kategorisierung („Kirche“, „Hochhaus“, etc.), oder Informationen zu davor geparkten Autos, oder zu am Gebäude angebrachten Schriftzügen enthalten.

In unserem 2020 durchgeführten Projekt haben wir schon allein aus Zeitgründen strikt am ursprünglichen Konzept festgehalten, so dass diese zusätzlich von der KI-Software gelieferten Metadaten derzeit sämtlich nicht ins AFIS eingespielt werden. Doch mit etwas Abstand und ohne Zeitdruck können die damals getroffenen Entscheidungen kritisch hinterfragt werden: Warum sollten die unberücksichtigt gebliebenen Analysedaten nicht auch als Erschließungsinformationen verwendet werden?

Wenn man beginnt, über diese Fragen neu nachzudenken, folgen daraus schnell einige weitere: Welche der von der KI-Analyse gelieferten Daten sind ausreichend relevant, um Aufnahme ins AFIS zu finden? Um beim obigen Beispiel zu bleiben: Sollen die Informationen „Mensch“, „Mann“ oder „Frau“ im AFIS-Datensatz zu einem Bild abgespeichert werden? Oder sollten sie besser verworfen werden, weil eine

¹ Ausführlich zum Einführungsprojekt und Integration der KI in das Heilbronner AFIS siehe: Miriam Eberlein: Kollege KI: Künstliche Intelligenz unterstützt Erschließung von Fotos. In: *Archivar* 74 (2021) H. 4, S. 285–287; Dies.: Fotoerschließung mit Künstlicher Intelligenz. Ein Projektbericht aus dem Stadtarchiv Heilbronn. In: *Archivische Erschließung im Umbruch*. Vorträge des 80. Südwestdeutschen Archivtags am 17. und 18. Juni 2021. Hg. v. Franz-Josef Ziwes u. Peter Müller. Stuttgart 2022, S. 67–74.

Recherche nach diesen Begriffen in einem AFIS kaum als sinnvolle Suchstrategie vorstellbar ist? In welchen Datenfeldern sollen solche Informationen untergebracht werden – in bereits vorhandenen? Oder sollten KI-erzeugte Erschließungsdaten besser in separaten, neu hinzuprogrammierten Datenfeldern abgespeichert werden?

Zu bedenken gilt ferner: Mit der Menge von automatisiert ins AFIS übernommenen KI-Analyseergebnissen steigt der Aufwand für Qualitätssicherungs- und gegebenenfalls Korrekturmaßnahmen. Wäre es daher nicht sinnvoll, eine eher überschaubare Auswahl zu treffen? Oder werden solche Fragen vielleicht gar nicht mehr von so großer Bedeutung sein, weil KI-basierte Tools künftig auch Qualitätssicherungs- und Korrekturmaßnahmen übernehmen, oder die Relevanz von Metadaten abhängig vom jeweiligen Nutzerinteresse werden berücksichtigen können?

Anlass zum Nachdenken über einen angemessenen Umgang mit KI-erzeugten Bild-Erschließungsdaten bietet nicht nur ihre Menge, sondern auch ihre Art: Manche Informationen waren bislang schlicht nicht existent. So werden KI-Analyseergebnisse in der Regel in Form von Wahrscheinlichkeiten ausgegeben. Das Ergebnis einer Objektanalyse lautet beispielsweise nicht: „Das ist ein Auto“, sondern „Das ist mit einer Wahrscheinlichkeit von 75 Prozent ein Auto“. Sollen diese Wahrscheinlichkeiten im AFIS vorgehalten und den Nutzer*innen angezeigt werden? Welche Erkenntnisse aber gewinnen Nutzer*innen aus der Mitteilung, dass eine Person auf einem Foto mit einer 75-prozentigen Wahrscheinlichkeit ein ganz bestimmter Mensch ist? Sorgt die Angabe einer Wahrscheinlichkeit für das Zutreffen bzw. Nicht-Zutreffen einer Erschließungsinformation für Transparenz oder eher für Verwirrung bei den Nutzer*innen?

Und schließlich: Wie kann die Verlässlichkeit von Erschließungsinformationen gewahrt bzw. transparent dargestellt werden, wenn die Menge der KI-generierten Informationen den Ressourcenbedarf einer 100-prozentigen Nach-Kontrolle durch qualifiziertes Personal bei weitem überfordert? Die hier aufgeworfenen Fragen wurden von der Autorin beim 1. Kölner Forum am 28. August 2024 vorgestellt und im Teilnehmer*innenkreis diskutiert. Sie sind sicherlich nur ein Anfang im Nachdenken über den Umgang mit KI-generierten Erschließungsdaten und möchten eine weitere archivische Fachdiskussion dazu anregen. Denn die neuen Möglichkeiten, die KI-basierte Tools für die archivische Erschließung bieten, erfordern zwingend Entscheidungen über den Umgang mit diesen Möglichkeiten. Das bedeutet, dass für die Integration KI-generierter Erschließungsergebnisse im AFIS vorhandene Lösungen geprüft und verändert werden müssen: für die Speicherung dieser Daten, für ihre Darstellung, für die Prozesse zu ihrer Verarbeitung und gegebenenfalls für ihre Kennzeichnung als „KI-generiert“. Hierzu werden im Folgenden sechs Vorschläge gemacht, die sich aus der Praxis des Heilbronner Projekts ergeben haben sowie in der Diskussion beim 1. Kölner Forum am 28. August 2024 zum Thema „Jenseits des Hype – Anwendung von KI in Archiven“ wertvolle Anregungen erhielten.

VORSCHLAG 1: UMFÄNGLICHE NUTZUNG DER KI-ERGEBNISSE

Durch eine KI-Analyse stehen für die Bilderschließung in kürzester Zeit wesentlich mehr Informationen zur Verfügung, als durch manuelle Erfassung mit vertretbarem Personaleinsatz jemals erhoben werden können. Sachthematische Kategorisierungen oder Klassifikationen wie z. B. Tier/Hund/Dackel, Auto/Hersteller/Modell, Gebäude/Kirche/Name der Kirche, Person/Name der Person, OCR-erkannte Schriftzüge etc. können, bei entsprechender Konfiguration und entsprechendem Training der KI-Software, per Knopfdruck analysiert und ausgegeben werden. Diese per KI-Analyse gewonnenen Metadaten sollten möglichst umfänglich als Erschließungsinformationen genutzt werden, um möglichst viele unterschiedliche Recherchezugänge anzubieten. Der Sorge vor zu großen Treffermengen kann durch intelligente Recherche-Strategien oder durch ein Angebot an nachträglichen Filtermöglichkeiten von Suchergebnissen begegnet werden. Sicherlich ergeben sich aus einer umfänglichen Übernahme von KI-generierten Metadaten nicht wenige Herausforderungen an das Management dieser Daten und ihrer Integration in das AFIS. Beispielsweise muss entschieden werden, ob und welche dieser Daten als Normdaten im AFIS hinterlegt werden sollen, oder für welche Daten eine Speicherung als Textinformationen in einem Metadatenfeld genügt. Diese Herausforderungen sollten jedoch eher durch kluges Webdesign und dem Einsatz neuer Such- und Filtertechnologien gelöst werden als durch eine Auswahl und Einschränkung der Metadaten a priori.

VORSCHLAG 2: KENNZEICHNUNG KI-GENERierter ERSCHLIESSUNGSdaten

KI-generierte Erschließungsdaten müssen explizit als solche gekennzeichnet werden. Denn dies gibt Nutzenden einen relevanten Hinweis darauf, in welchem Maße diese Erschließungsinformationen als zuverlässig einzuschätzen sind. Eine 2024 präsentierte internationale Studie des Helmholtz-Zentrums für Informationssicherheit (CISPA) kommt zu dem Ergebnis, dass Menschen schon jetzt kaum noch in der Lage sind, KI-generierte Medieninhalte als solche zu erkennen – unabhängig von ihrem Bildungsgrad und ihrer Medienkompetenz.² Was für komplexe Medieninhalte gilt, dürfte erst recht auf Metadaten zutreffen: Ohne entsprechende Kennzeichnung sind KI-generierte von menschengemachten nicht zu unterscheiden.

Doch warum ist diese Unterscheidung überhaupt von Relevanz? Sowohl menschengemachte als auch KI-generierte Metadaten können fehlerbehaftet sein. Die Glaubwürdigkeit archivischer Erschließungsinformationen hängt davon ab, wie gut die Nutzer*innen deren Verlässlichkeit einschätzen können. Hierfür ist Transparenz notwendig: So wie Findbücher üblicherweise die Bearbeiter*innen benennen, so sollten KI-generierte Metadaten klar gekennzeichnet werden.

HEUSS
Heilbronner Fotoarchiv- und Such-System

Titel i
 Urheber i
 Anfangsdatum i
 Enddatum i
 Zeitangaben geschätzt
 Vorschau i

Altenfeier anlässlich der Harmonie-Einweihung
 Wendnagel
 1958-11
 1958-12
 ja

F-4624,66

Personen: Willy Brandt

KI-analysiert ja

Fehlerhafter KI-generierter Indexeintrag, gekennzeichnet als „KI-analysiert“. (Foto: Wendnagel, Heilbronn; Quelle: Stadtarchiv Heilbronn Foo1F-4624,66, AFIS-Datensatz)

VORSCHLAG 3: ANGABE DER KONFIDENZEN

Bereits erwähnt wurde, dass die Ausgabe von KI-Analyseergebnissen unter Angabe einer Wahrscheinlichkeit für das Zutreffen dieses Ergebnisses erfolgt – der „Konfidenz“. Die Integration der Konfidenz in die Präsentation von Erschließungsinformationen würde es den Nutzer*innen ermöglichen, deren Glaubwürdigkeit selbst einschätzen zu können: Trifft die Identifikation z. B. einer Person auf einem Foto tatsächlich zu? Oder bestehen erhebliche Unsicherheiten – die Archivar*innen bei der Erschließung mit Angaben wie

„vermutlich“, „vielleicht“, oder auch einem Fragezeichen kennzeichnen würden?

Allerdings hat ein Konfidenzwert für sich alleine nur eine begrenzte Aussagekraft. So zeigt die Erfahrung der Heilbron-

- ² Menschen erkennen KI-generierte Medien kaum. cva/dpa-Bericht vom 14.06.2024, im Internet unter <https://www.forschung-und-lehre.de/zeitfragen/menschen-erkennen-ki-generierte-medien-kaum-6430> (abgerufen am 31.10.2024). Publikation der vollständigen Studie: Joel Frank u. a.: A Representative Study on Human Detection of Artificially Generated Media Across Countries. 10.12.2023, im Internet unter: <https://arxiv.org/abs/2312.05976> (abgerufen am 31.10.2024).

Face Recognition

Status: Completed

Parameters:

Model: Heilbronn_v1

Filbinger, Hans, Ministerpräsident
 Similarity: 52%

Hoffmann, Hans, Oberbürgermeister
 Similarity: 73%

Unknown

Angabe der Konfidenz (hier: „Similarity“) bei der Gesichtserkennung (Fotograf: H. Kühlmann, Heilbronn; Quelle: Stadtarchiv Heilbronn Foo1N-28125, DeepVA-Datensatz)

ner KI-Analysen, dass selbst Ergebnisse mit einer Konfidenz von 50 % – ein auf den ersten Blick nicht sehr vertrauenserweckender Wert – in den allermeisten Fällen tatsächlich korrekt waren. Um die Aussage eines Konfidenzwertes einschätzen zu können, bedarf es daher ergänzender Informationen über die Ergebnisse und Prozesse der Qualitätsprüfung der KI-Analyseergebnisse.

VORSCHLAG 4: QUALITÄTSPRÜFUNGS- UND KORREKTURMASSNAHMEN TEILAUTOMATISIEREN

Mit der Qualitätsprüfung und dem Umgang mit Korrekturen ist ein weiteres Feld angesprochen, das es bei der Integration KI-generierter Erschließungsergebnisse zu bedenken gilt. Schon aufgrund der zu erwartenden Datenmenge sind Qualitätsprüfungs- und Korrekturmaßnahmen durch Fachpersonal nur stichprobenartig möglich. Daher sollte der Einsatz weiterer Instrumentarien erwogen werden. Eines davon könnten Feedback-Funktionen sein, die es Nutzer*innen ermöglichen, auf Datensatz-Ebene unmittelbar und niederschwellig Rückmeldung zu Erschließungsinformationen zu geben und somit – nach einer redaktionellen Prüfung durch Fachpersonal – zu einer Korrektur oder Ergänzung der Erschließungsinformationen beizutragen. Eine weitere Option wäre der Einsatz einer anderen KI-gestützten Anwendung bzw. die Kombination verschiedener Tools zur Plausibilisierung der Ergebnisse: So könnte beispielsweise geprüft werden, ob eine weitere Gesichtserkennungssoftware zum gleichen Resultat kommt. Lebensdaten identifizierter Personen könnten mit der Datierung eines Fotos oder auch anderer Kontextinformationen abgeglichen werden. Auch hier wäre wiederum zu entscheiden, wie mit den Ergebnissen der Qualitätsprüfung umgegangen werden soll: Widersprüchliche Ergebnisse könnten einem menschlichen Redakteur gemeldet und bis zu dessen Eingreifen im AFIS für den Nutzer*innenzugriff gesperrt bleiben. Bei übereinstimmenden Ergebnissen könnten die jeweiligen Konfidenzwerte separat hinterlegt, oder aber ein gemeinsamer Konfidenzwert errechnet werden.

VORSCHLAG 5: KI-ANALYSETOOL ALS „LERNENDES SYSTEM“ INTEGRIEREN

Neben der Etablierung von Prozessen zur Qualitätsprüfung und zur Korrektur von KI-generierten Erschließungsmetadaten sollten Wege gefunden werden, Korrekturen, aber auch korrekt erkannte Ergebnisse in die den KI-Analysen zugrundeliegenden Modelle zurückzuspielen. Für die Erkennung von bestimmten Gebäuden oder individuellen Gesichtern wird zunächst ein jeweils für diese Erkennungsart spezialisiertes Grundmodell mit Fotos dieser Gebäude bzw. Gesichter „trainiert“. Die Menge und die Güte der zur Verfügung stehenden Trainingsdaten beeinflusst die spätere Erkennungsgenauigkeit des jeweiligen Modells. Stehen durch erfolgreiche KI-Analysen oder durch Korrekturen weitere Daten zur Verfügung und werden in ein erneutes Training einbezogen, verbessert sich in der Regel die Erken-

nungsgenauigkeit des Modells. Auf diese Weise lässt sich ein Modell auch um ihm bis dato unbekannte Gebäude oder Gesichter ergänzen.

Um Korrekturen oder auch Bestätigungen korrekter Erkennungen wieder in das verwendete Modell einfließen zu lassen, müssten entsprechende Feedback-Funktionalitäten an der Stelle zur Verfügung gestellt werden, an der Qualitätsprüfungs- und Korrekturmaßnahmen durch Fachpersonal stattfinden. Zugleich sollte die Durchführung von KI-Analysen nicht als einmaliger Arbeitsschritt in den Prozess der Erschließung integriert werden, sondern stattdessen beliebig wiederholbar sein – mit entsprechender Kennzeichnung und Dokumentation des Analysezeitpunkts und der verwendeten Modellversion. Sowohl das Analysemodell, als auch der Erschließungsstand wären somit als Bestandteil eines lernenden Systems zu begreifen. Bezogen auf den einzelnen Erschließungsdatensatz ist ein abschließendes Ergebnis zwar vorstellbar – etwa, wenn sämtliche auf einem Foto abgebildeten Personen erkannt und die Korrektheit dieser Ergebnisse durch Fachpersonal bestätigt wurde – und im Datensatz sollte dies hinterlegt werden können (beispielsweise durch einen „geprüft“-Status). Doch für die Gesamtheit aller im AFIS oder auch nur in einem Bestand enthaltenen Datensätze dürfte eine vollständige Erkennung oder Indizierung in der Realität schwerlich erreichbar sein.

VORSCHLAG 6: NEUE ANFORDERUNGEN AN DAS PRODUKTDESIGN VON ONLINE-FINDMITTELN

Aus den bisherigen Überlegungen ergeben sich eine Reihe von Herausforderungen an die Präsentation von Erschließungsinformationen bzw. an das Produktdesign von Online-Findmitteln. Denn ein vertieftes Angebot an Erschließungsdaten sowie an präzisen Angaben über deren Herkunft und Konfidenz sollte bei den Nutzer*innen nicht zu einer Informationsüberflutung führen. Hier bieten sich Präsentationskonzepte mit abgestuften Informationsebenen an, wie sie in einigen Archivinformationssystemen praktiziert werden. Neuartige Informationen wie Herkunft („KI-generiert“) und Konfidenz müssen gut erkennbar dargestellt sein, ohne auf die Nutzer*innen verwirrend zu wirken. Feedback-Möglichkeiten für Nutzer*innen sollten niederschwellig und leicht bedienbar sein. Das gleiche gilt für Feedback-Funktionalitäten im Bearbeitungs- und Qualitätssicherungsprozess durch archivisches Fachpersonal, damit ein „lernendes System“ tatsächlich funktioniert. Die hier nur kurz genannten Beispiele mögen genügen, die bei der Integration von KI-generierten Erschließungsdaten ins AFIS ergebenden Herausforderungen zu skizzieren, ohne auch nur entfernt einen Anspruch auf Vollständigkeit erheben zu wollen.³

FAZIT

Die Nutzung von KI-gestützten Tools zur inhaltlichen Analyse und Identifizierung von Bildinhalten bietet enorme Chancen für die Erschließung. Es zeichnet sich ab, dass sich durch diese neuen Möglichkeiten sowohl Umfang und Inhalt der Erschließungsinformationen als auch die Art und Weise,

wie wir sie generieren und präsentieren, stark verändern werden. Wollen wir die Potentiale KI-gestützter Tools gut ausschöpfen und uns ihre Lernfähigkeit zu Nutze machen, müssen wir Erschließungsstände künftig wesentlich stärker als „Work in Progress“ verstehen und unseren Nutzer*innen auch so vermitteln. Zugleich müssen wir ihnen ermöglichen, die Zuverlässigkeit der Erschließungsinformationen einschätzen zu können. Aus alldem ergeben sich neue Herausforderungen für das Produktdesign von Online-Findmitteln. Dazu bedarf es einer weiteren Fachdiskussion, zu denen die sechs Vorschläge der Autorin ein Impuls sein möchten.

IMAGE INDEXING WITH AI: SIX SUGGESTIONS FOR THE USE OF AI-GENERATED METADATA

The study examines the use of AI-supported image analysis in the Heilbronn City Archives for photo cataloging. Using the tool „DeepVA“, faces and buildings are automatically identified. Six recommendations address key challenges: comprehensive data utilization, clear labeling of AI-generated metadata, inclusion of confidence scores, semi-automated quality control, adaptive systems, and new presentation concepts. The author emphasizes AI's potential but calls for further discussion on sustainable implementation in archival collections management systems.

- ³ Weiterführende Anregungen zu diesen Fragen enthält ein Aufsatz aus der bibliothekarischen Fachdiskussion von Michael Franke-Maier: Anforderungen an die Qualität der Inhaltserschließung im Spannungsfeld von intellektuell und automatisch erzeugten Metadaten. In: ABI Technik 38 (2018), S. 327–331 (<http://doi.org/10.1515/abitech-2018-4005>) (abgerufen am 31.10.2024).

Miriam Eberlein M.A.

Stadtarchiv Heilbronn

Otto Rettenmaier Haus/Haus der Stadtgeschichte

Eichgasse 1, 74072 Heilbronn

Tel. 01731 56-2290

E-Mail: miriam.eberlein@heilbronn.de

STANDARDBASIERTE AUSSONDERUNG ELEKTRONISCHER GERICHTSVERFAHRENSAKTEN ÜBER DAS DIGITALE ZWISCHENARCHIV DES BUNDES¹

Ausgewählte Verfahrensakte der obersten Gerichtshöfe des Bundes sind aufgrund ihres Ranges für die Rechtsprechung oder ihrer historischen Bedeutung von bleibendem Wert. Seit einiger Zeit führen alle obersten Bundesgerichte ihre Verfahrensakte elektronisch. Dabei sind die Anwendungen E-Akte als Service (eAS bzw. VIS-Justiz), ergonomischer elektronischer Arbeitsplatz (e2A) und Gerichtsoffene Software Anwendung (GOŠA) im Einsatz. Allen Anwendungen ist gemeinsam, dass sie Funktionen zur Aktenführung und gerichtliche Fachanwendungen (Justizfachsysteme) integrieren. Die Dokumente werden in elektronischen Verfahrensakten im PDF/A-Format geführt (sog. Repräsentat). Die Anwendungen sind zum Datenaustausch über den Standard XJustiz befähigt. Aktuell führen sämtliche obersten Bundesgerichte und der Generalbundesanwalt ihre elektronischen Akten vor Ort (on premises). Mittelfristig ist eine Konsolidierung der E-Akte-Führung bei einem zentralen Dienstleister geplant. Dies hat für die archivische Überlieferung aber keine Auswirkungen.

In der Justiz werden die Fachdaten – Daten zur Verfahrensorganisation wie Termine, Angaben zu den Verfahrensbeteiligten, Hinweise zu einschlägigen Rechtsvorschriften usw. – regelmäßig in gesonderten Fachverfahren (Justizfachsystemen) und Schriftgutobjekte im jeweiligen E-Akten-System vorgehalten. Für die Anbietung an die Archive, die archivische Bewertung und die Übernahme in das digitale Magazin des jeweiligen Archivs müssen die Daten zusammengeführt werden. Dabei sollen nach den Vorgaben der Bund-Länder-Kommission Arbeitsgruppe IT-Standards (BLK-AG IT-Standards) die E-Akten-Systeme den Aussonderungsprozess steuern. Für den technischen Ablauf bedeutet dies, dass die Aussonderung und Anbietung vom E-Akten-System initiiert wird. Der Datenexport muss den Vorgaben des Datenaustauschstandards für den elektronischen Rechtsverkehr XJustiz folgen.

Dies ist sehr vereinfacht der Rahmen, in dem es die Aussonderung der elektronischen Gerichtsakte der obersten Gerichtshöfe des Bundes sicherzustellen gilt. Die digitale Form schafft dabei neue Herausforderungen. Die schlichte Übertragung eines Gesamt-PDFs der Verfahrensakte reicht für die Übergabe an die Archive nicht. Vielmehr müssen mit den Dateien, den so genannten Primärdaten, auch die dazugehö-

rigen Metadaten mitübermittelt werden, um die Dokumente in ihrem aktenmäßigen Zusammenhang abzubilden und aufbewahren zu können. Anderenfalls droht ein Überlieferungsverlust. Die digitale Form der Verfahrensakte birgt gleichzeitig Chancen, den Prozess der Aussonderung zu vereinfachen. Mit E-Akten können auch die obersten Gerichtshöfe des Bundes am Zwischenarchivverfahren teilhaben (§ 8 Abs. 1 Satz 2 Bundesarchivgesetz (im Folgenden BArchG)). Der Bundesfinanzhof (im Folgenden BFH) ist das erste oberste Gericht des Bundes, das das Digitale Zwischenarchiv des Bundes (DZAB) nutzt. Gleichzeitig ist er das erste deutsche Gericht, das seine Verfahrensakte planmäßig und standardkonform auf XJustiz-Grundlage aussondert. Der Beitrag schildert diese Umsetzung. Dabei wird deutlich, wie bestimmte Vorgaben zur Standardisierung im Rahmen der digitalen Transformation ineinandergreifen.

RECHTSGRUNDLAGEN FÜR DIE ARCHIVIERUNG ELEKTRONISCHER GERICHTSVERFAHRENSAKTEN

Akten der Gerichte, die für das Verfahren nicht mehr erforderlich sind, dürfen grundsätzlich nach Beendigung des Verfahrens nur so lange aufbewahrt und gespeichert werden, wie schutzwürdige Interessen der Verfahrensbeteiligten oder sonstiger Personen oder öffentliche Interessen dies erfordern. Die Rechtsgrundlagen hierfür sind die Datenschutz-Grundverordnung (EU-DSGVO), ergänzt durch das Bundesdatenschutzgesetz (BDSG) sowie hinsichtlich der Anbietungspflicht der nicht mehr zur Erledigung der öffentlichen Aufgaben benötigten Unterlagen das BArchG. Das Justizaktenaufbewahrungsgesetz (JaktAG) sowie die Justizaktenaufbewahrungsverordnung (JaktAV) regeln die Aufbewahrung und Speicherung von Verfahrensakten in der Verantwortung des jeweils zuständigen obersten Bundesgerichts. In der JaktAV sind im Einzelnen die Aufbewahrungs- und Speicherfristen auf Ebene der Gerichtsakte und teilweise auf Ebene der Dokumente festgelegt. Diese variieren von fünf bis zu 70 Jahren und unterscheiden teilweise zwischen Akten- und Dokumentenebene. Um die digitale Archivierung möglichst einfach umsetzbar zu gestalten,

ten, haben einige Bundesgerichte ihre Regelungen zu den Aufbewahrungsfristen für Akten erheblich vereinfacht. So wurden teilweise dokumentbezogene Aufbewahrungsfristen aufgegeben, Aufbewahrungsfristen verkürzt und auf wenige Varianten beschränkt. Im Fall des BFH werden Verfahrensakten als Einheiten ausgesondert. In anderen Gerichtsbarkeiten vorhandene (z. T. sehr kleinteilige) Unterscheidungen hinsichtlich einzelner Dokumente oder Verfahrensarten sind im BFH daher nicht zu beachten. Das wirkt sich positiv auf die Strukturierung der Abgaben aus.

DAS DIGITALE ZWISCHENARCHIV DES BUNDES (DZAB)

Das Digitale Zwischenarchiv des Bundes (DZAB) ist ein zentraler Dienst im Rahmen der IT-Dienstekonsolidierung des Bundes, den das Bundesarchiv für alle Bundesbehörden und Bundesgerichte bereitstellt. Das DZAB wird durch das Bundesarchiv als Basisdienst unterhalten und von der Bundesagentur für Arbeit (BA) gehostet. Mit dem DZAB erbringt das Bundesarchiv einen Dienst, der das rechtssichere Aufbewahren elektronischer Unterlagen und damit eine mandantenfähige Langzeitspeicherung mitsamt Beweiswerterhaltung ermöglicht. Das DZAB gewährleistet die Langzeitspeicherung gemäß dem Standard BSI TR-03125 „Beweiswerterhaltung kryptographisch signierter Dokumente“. Das Bundesarchiv unterstützt die Bundesgerichte bei allen Fragen zum Betrieb des DZAB als solchem und der Netzanbindung einschließlich der damit zusammenhängenden technischen Fragen. Die Dienstleistung der BA umfasst dabei den Aufbau der Verbindung zwischen den Bundesgerichten und der DZAB-Infrastruktur sowie notwendige technische Maßnahmen wie die Einrichtung der einzelnen Mandanten für jedes Gericht, die Versiegelung der gespeicherten Daten durch digital signierte Zeitstempel sowie eine nachprüfbare Zugriffsprotokollierung.

Kosten entstehen den Bundesgerichten für die Nutzung der Speicherinfrastruktur und der dazugehörigen Services des IT-Dienstleisters. Die den Bundesgerichten vom IT-Dienstleister berechneten Kosten liegen einer BFH-internen Kalkulation zufolge aber erheblich (um ca. 90 %) unter den Kosten, die bei einer eigenen Aufbewahrung on premises einschließlich regelmäßig notwendiger Migrationen bis zum Ablauf der Aufbewahrungsfrist entstehen würden. Hier zeigen sich deutlich die Kosten- und Skaleneffekte einer zentralen Zwischenarchivierung gegenüber einer lokalen Aufbewahrung durch das einzelne Gericht.

Mit Speicherung der Verfahrensakte in einem Zwischenarchivgutmandanten werden die Unterlagen den Vorgaben nach § 5 Abs. 1 BArchG entsprechend dem Bundesarchiv angeboten. Dies entlastet die Gerichte von der bisher ihnen obliegenden Aufgabe der Aufbewahrung der Unterlagen bis zum Ablauf der Aufbewahrungsfrist. Zudem entfallen Arbeitsschritte im Zusammenhang mit der Anbietung. Das Bundesarchiv stellt dann auf Grundlage des § 3 Abs. 2 Satz 2 BArchG den bleibenden Wert fest.

ÜBERGABE MITTELS S.4-SCHNITTSTELLE

Nach den Vorgaben der BLK-AG IT-Standards ist für die elektronische Übermittlung eine Infrastruktur zu nutzen, die in der Lage ist, die zu übermittelnden Datenmengen vollständig zu transportieren. Dies kann entweder über die EGVP-Infrastruktur oder mittels einer gesonderten Schnittstelle erfolgen. Bei der Aussonderung besteht keine Verpflichtung zur Nutzung eines sicheren Übermittlungswegs im Sinne der Elektronischer-Rechtsverkehr-Verordnung – ERVV. Der BFH und das Bundesarchiv haben die Nutzung der in der technischen Richtlinie BSI TR-03125 beschriebenen S.4-Schnittstelle des DZAB vereinbart. Sie kommt auch bei der Übertragung der elektronischen Akten aus den anderen öffentlichen Stellen des Bundes in das DZAB zur Anwendung.

Die Verfahrensakte werden durch einen im Justizfachsystem GO\$A-EGA implementierten Scheduler-Job automatisch an das DZAB übertragen. Voraussetzung ist, die Verfahrensakte ist abschließend bearbeitet – in der gerichtlichen Terminologie „weggelegt“ – und eine gerichtliche Transferfrist abgelaufen. Diese ist notwendig, da ggf. noch Anfragen der Bundeskasse zum Verfahren erfolgen oder Dokumente nachgereicht werden, die zur Akte zu bringen sind. Das Hinzufügen von Dokumenten hat allerdings keine Auswirkung mehr auf die Aufbewahrungsfrist. Innerhalb der Transferfrist muss die Akte noch für einen Direktzugriff durch das Gericht zur Verfügung stehen. Sollte es zum Fall der Wiederaufnahme des Verfahrens kommen, kann die Verfügung des Weglegens sehr leicht wieder rückgängig gemacht werden. Nach Ablauf der Transferfrist erfolgt dann die Auslagerung der Akte in das DZAB. In dem vom BFH

¹ Da das wesentliche Material im Internet greifbar ist, weisen die Verfasserin und die Verfasser an dieser Stelle nur zusammenfassend auf weiterführende Fundstellen hin: Zur Überlieferung der obersten Gerichte des Bundes hat das Bundesarchiv 2023 ein Symposium veranstaltet, vgl. Thomas Henne: Tagungsbericht. Symposium des Bundesarchivs zur analogen und digitalen Überlieferung der Obersten Gerichte des Bundes – 24.–26. April 2023 in Koblenz, URL: <https://archivwelt.hypotheses.org/3562>. Auf der Homepage des Beauftragten der Bundesregierung für Informationstechnik sind die grundlegenden Dokumente zur IT-Dienstekonsolidierung des Bundes zugänglich, etwa die Architekturrichtlinie für die IT des Bundes und die Domänenarchitektur Elektronische Verwaltungsarbeit, in der das Digitale Zwischenarchiv des Bundes verortet ist, vgl. <https://www.cio.bund.de/Webs/CIO/DE/startseite/startseite-node.html>. Detailliertes Informationsmaterial zum Digitalen Zwischenarchiv des Bundes stellt das Bundesarchiv auf seiner Homepage bereit; gleiches gilt für ein Glossar zur Schriftgut- und Informationsverwaltung und für die 2023 gemeinsam mit dem Bundesbeauftragten für den Datenschutz und die Informationsfreiheit herausgegebene Broschüre „Archivrecht und Datenschutz“, vgl. <https://www.bundesarchiv.de/>. Die Spezifikationen der Standards für den XML-basierten Datenaustausch elektronischer Schriftgutobjekte, xdoma und Xjustiz, sind über das XRepository greifbar, vgl. <https://www.xrepository.de/>. Die Technische Richtlinie des Bundesamts für Sicherheit in der Informationstechnik BSI TR-03125 Beweiswerterhaltung kryptographisch signierter Dokumente trifft wichtige Vorgaben, die bei der digitalen Zwischenarchivierung zu berücksichtigen sind. Detaillierte Informationen und die Dokumente der Richtlinie stellt das BSI auf seiner Homepage bereit, vgl. https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/Technische-Richtlinien/TR-nach-Thema-sortiert/tro3125/TR-03125_node.html.

verwendeten Fachverfahren GOŠA erfolgt die Aktenführung in einem speziellen Modul für die elektronische Gerichtsakte (EGA). So fungiert GOŠA im Falle des BFH gegenüber dem DZAB als abgebendes System und holt aus dem EGA-Modul die verakteten Dokumente mit ihren Metadaten. Übergeben wird die gerichtseigene Verfahrensakte. Beiakten und Handakten, die in anderen Modulen gehalten werden, werden nicht mitübermittelt.

Für die Speicherung im DZAB werden die Gerichtsverfahrensakten vor der Übertragung in Datenpakete transformiert. Dabei gilt der Grundsatz der Aktenvollständigkeit. Das bedeutet, dass sämtliche Dateien eines Schriftgutobjekts, die veraktet wurden, einschließlich der Metadaten, diesem zuordenbar sein und mittels DZAB angeboten werden müssen. Hinsichtlich der Datenstruktur der in das DZAB auszusondernden Objekte setzt die vom DZAB verwendete Speicherlösung auf die in der Richtlinie BSI TR-03125 beschriebenen XML-basierten Archivinformationspakete (XAIP). Diese umfassen die prozesssteuernden, die archivfachlich relevanten und gerichtsspezifischen Metadaten, die Primärdaten sowie kryptographische Beweisdaten wie Zeitstempel oder Signaturen. Die Aufbewahrung im DZAB ist so konzipiert, dass die XAIP während der gesamten Aufbewahrungsfrist unverändert bleiben. Die Übertragung der XAIP-Pakete aus dem E-Akte-System zum DZAB erfolgt über die Netze des Bundes (NdB). Im DZAB werden anschließend eine XML-Schema-Validierung gegen das abgestimmte XAIP-Schema und eine Signaturprüfung durchgeführt. Im Fall einer positiven Validierung wird das XAIP gespeichert und mittels Zeitstempel versiegelt. An das E-Akte- bzw. das Justizfachsystem wird als Referenz eine eindeutige Archiv-Objekt-ID (AOID) zurückgeliefert. Verläuft die Übertragung erfolgreich, folgt auf eine positive Rückmeldung die Löschung von Primär- und Metadaten im EGA-Modul. Bestimmte Fachdaten und Angaben zu den Verfahrensakten bleiben bis zum Ende der Aufbewahrungsfrist zu Recherchezwecken und zur Abwicklung der Aussonderung im Justizfachsystem GOŠA erhalten.

XJUSTIZ ALS DATENEXPORTFORMAT FÜR DIE AUSSONDERUNG GERICHTLICHER VERFAHRENSAKTEN

Nach der gesetzlichen Regelung in § 5 Abs. 3 Satz 2 BArchG und den Abstimmungen der BLK-AG IT-Standards und des KLA-Ausschusses Records Management ist für die Abgabe von Verfahrensakten an die Archive das Datenaustauschformat XJustiz zu nutzen. Der XJustiz-Standard hält, neben vielen anderen Anwendungsfällen der Kommunikation zwischen Gerichten, Behörden, Kanzleien, juristischen und Privatpersonen, auch ein Fachmodul mit entsprechenden Nachrichtenschemata für die Aussonderung vor. Anders als beim Datenaustauschstandard xdomea-Standard gibt es in XJustiz kein gesondertes Schema für die Zwischenarchivierung. Im Modul Aussonderung des XJustiz-Standards sind in der zum Zeitpunkt der Produktivsetzung der DZAB-Anbindung des BFH gültigen Version 3.4.1 Nachrichtentypen für Anbieten, Bewertung und Aussonderung enthalten. Für die Abbildung der Metadaten der Verfahrensakten beim Datenexport in das DZAB wird das Schema der Ausson-

derungsnachricht, `nachricht.auss.aussonderung.1300005`, genutzt. Vor der Übermittlung einer Akte müssen aus dem Fachverfahren die Fachdaten angefordert werden. Nach Anlieferung der Fachdaten wird aus den Fachdaten und den Metadaten der Schriftgutobjekte die XJustiz-Nachricht erzeugt. Die Aussonderungsnachricht, wie auch die Primärdaten (die verakteten Dateien) werden in ein gemäß BSI TR-03125 gebildetes XAIP-Datenpaket eingebettet. Die Standards BSI TR-03125 und XJustiz bilden dabei zwei Datenschichten. Die gemeinsame technische Grundlage, das Datenformat XML, unterstützt die Verbindung. XAIPs unterliegen bei der Übertragung über die S.4-Schnittstelle Größenbeschränkungen. Deshalb wurde als spezielle Ausformung des Standards BSI TR-03125 für das DZAB eine Aufgliederung der XAIP vereinbart. Konkret wird je Verfahrensakte ein sogenanntes „Eltern“-XAIP erstellt. Es enthält die XJustiz-XML-Datei `nachricht.auss.aussonderung.1300005` zu einer Verfahrensakte und referenziert auf sogenannte „Kind“-XAIP. Diese enthalten die Primärdateien, die zu einer Verfahrensakte gehören. Das Eltern- und alle dazugehörenden Kind-XAIP bilden eine „Logische Einheit“. D. h., sie sind im Sinne der Aktenvollständigkeit, -wahrheit und -klarheit im Zusammenhang zu betrachten. Über die Schnittstelle werden die Datenpakete an das DZAB übertragen, wo sie unverändert und beweiswerterhaltend bis zum Ende der Aufbewahrungsfrist gespeichert werden. Die S.4-Schnittstelle deckt die Kommunikation der Empfangs- und Importbestätigung gegenüber dem abgebenden System ab, so dass der hierfür im XJustiz-Standard vorgesehene Nachrichtentyp `nachricht.auss.aussonderungImportBestaetigen.1300006` beim DZAB unmittelbar keine Verwendung findet.

RÜCKLEIHE AUS DEM DZAB

Innerhalb der Aufbewahrungsfrist kann der BFH Akten über eine im Justizfachsystem GOŠA implementierte Rückholfunktion (Rückleihe) reimportieren. Da beim BFH eine Gerichtsverfahrensakte mit dem förmlichen Verfügen des „Weglegens“ rechtlich abgeschlossen wird, besteht kein Bedarf sie weiterzubearbeiten oder zu ergänzen. Vielmehr erfolgt das Rückholen, um die Akte des abgeschlossenen Verfahrens als Beiakte zu noch anhängigen Parallelverfahren einsehen oder für andere Verfahren auswerten zu können. In GOŠA erfolgt die automatisierte Rückleihe über die Eingabe der oben erwähnten AOID oder des Aktenzeichens für eine zurückgeforderte Akte. Die aus dem DZAB rückübermittelte Akte wird direkt im Beiaktenmodul des E-Akten-Systems abgelegt. Die Beiakte enthält die aus den mitgegebenen XJustiz-Daten generierten Metadaten. Sobald die zurückgeholte Nutzungskopie im Gericht nicht mehr benötigt wird, wird sie gelöscht.

ÜBERNAHME IN DAS DIGITALE MAGAZIN

Das Bundesarchiv kann bereits während der Aufbewahrungsfrist auf der Grundlage von § 8 Abs. 2 Satz 3 BArchG die Bewertung der im DZAB liegenden Akten vornehmen. Es kann daher bereits vor Ablauf der Aufbewahrungsfrist mit-

tels Lesezugriffs auf zu bewertende elektronische Unterlagen (Primär- und Metadaten) im Zwischenarchivgutmandanten zugreifen. Dabei markiert es die Unterlagen nach archivfachlichen Kriterien als archivwürdig (A), noch zu bewerten (B) oder zu vernichten (V). Hierfür bedient sich das BArch einer eigenen Zugriffskomponente. Den Verfahrensakten des BFH wird bei Auslagerung in das DZAB ein Metadatum „Bewertungsvorschlag“ seitens des abgebenden Gerichts mitgegeben. So werden z. B. Akten als archivwürdig eingestuft, die im Jahresbericht des BFH enthalten sind, ein größeres Medienecho erfahren haben oder prominente Persönlichkeiten des öffentlichen Lebens betreffen. Diese Bewertungskriterien sind zwischen BFH und Bundesarchiv abgestimmt. Alle als archivwürdig bewerteten Verfahrensakten werden nach Ablauf der Aufbewahrungsfrist in einen bundesarchiveigenen Speichermantanten im DZAB übertragen. Von dort aus erfolgt die Übernahme in das Digitale Magazin. Zu vernichtende Akten werden nach dem Ende der Aufbewahrungsfrist physisch aus dem DZAB gelöscht. Über sein Justizfachsystem GO\$A stößt der BFH die Löschung der betreffenden Akten an. Die dabei entstehenden Protokollinformationen belegen die rechtskonforme Aussonderung.

ZUSAMMENFASSUNG

Im Rahmen der Abstimmungsgespräche zur Anbindung des BFH an das DZAB wurde für die Abgabe gerichtlicher Verfahrensakten Neuland betreten. Einige Aspekte mussten neu gedacht werden, etwa die Verwendung des Standards XJustiz oder die Anwendung des Zwischenarchivverfahrens. Die beim BFH gewonnenen Erkenntnisse können in die Abstimmungsgespräche mit anderen Obersten Gerichten des Bundes eingebracht werden. Im Laufe dieser Abstimmungen ist mit weiteren Herausforderungen zu rechnen, wie z. B. die Abbildung der Teilaussonderung mit elektronischen Mitteln. Am Beispiel BFH ist erkennbar, welches Potential an Automatisierung und Effizienzzuwachs unter bestimmten Voraussetzungen in der digitalen Zwischenarchivierung steckt.

*Lutz Bannert, Koblenz/Ines Müller, Berlin/
Nils Trossen, München*

DIE BERGUNGSERFASSUNG IM HISTORISCHEN ARCHIV MIT RHEINISCHEM BILDARCHIV KÖLN

ENTWICKLUNGEN SEIT DEM ARCHIVEINSTURZ IM JAHR 2009

Am 3. März 2009 stürzte das Kölner Archivgebäude an der Severinstraße in Folge von U-Bahn-Bauarbeiten ein. Bei diesem Unglück wurde nicht nur das Archivgut verunordnet und beschädigt, es verloren auch zwei Menschen ihr Leben. Somit stellt dies ein einschneidendes Ereignis in der Geschichte der Stadt Köln dar. Auch nach über 15 Jahren ist die Katastrophe des Einsturzes im Arbeitsalltag des Historischen Archivs mit Rheinischem Bildarchiv Köln (HASTK-RBA) weiterhin präsent.

Eine der Herausforderungen bestand zu Beginn darin, eine umfangreiche Dokumentation des Bergungsguts zu ermöglichen. Dabei musste zum einen der Fokus auf die inhaltliche Beschreibung gelegt werden, um das Ziel der Rückführung

des Archivguts in seinen ursprünglichen Zusammenhang erreichen zu können. Andererseits war auch die Schadensdokumentation sowohl aus rechtlichen als auch aus restauratorischen Gesichtspunkten notwendig. Dies wurde durch eigens für diesen Anwendungsfall programmierte Datenbanken sichergestellt, die die einzelnen Prozessschritte steuern, nachhalten und verknüpfen sollen. Im Folgenden soll ein Überblick über die Erfassungsarbeiten nach dem Einsturz bis hin zum gegenwärtigen Entwicklungsstand gegeben werden.

DER EINSTURZ UND SEINE UNMITTELBAREN FOLGEN

Der Einsturz stellte einen gravierenden Einschnitt in der Arbeit des Historischen Archivs der Stadt Köln dar. Durch diese Katastrophe veränderten sich Struktur und Arbeitsweise des Kommunalarchivs tiefgreifend und dauerhaft. Die Mitarbeiter*innen und Nutzer*innen verloren an diesem Tag sowohl ihren Arbeitsplatz als auch die Möglichkeit, Zugriff zu ihrem strukturierten Stadtgedächtnis zu erhalten. Somit stand plötzlich im Fokus der Tätigkeiten eine grundlegende Neuausrichtung und Aufarbeitung der durch den Einsturz verursachten Verunordnung der archivischen Bestände. Während des Einsturzes sackte das Gebäude ab, kippte nach vorne und zerbrach in der Mitte. Die dadurch verursachte vollständige Durchmischung des Archivguts mit unterschiedlichen Schadensbildern führte bei einem Großteil des Materials zum Verlust der Bestands- und Provenienzzusammenhänge. Dieser Notstand machte eine schnelle Entscheidungskompetenz und Konzeptionierung von ersten Bergungs- und Erfassungstätigkeiten notwendig. Natürlich war dies ohne die Unterstützung zahlloser Freiwilliger, die ihre Arbeitskraft zur Verfügung stellten, nicht möglich. Genau so entscheidend war die Bereitschaft von Kolleg*innen in vielen Archiven, Raum für die Lagerung von geborgenem Kölner Archivgut bereitzustellen. Die Erstversorgung des Archivgutes folgte dem Leitsatz „Wir lesen nichts, wir werfen nichts weg“ und war damit zunächst auf eine möglichst schnelle Bergung am Unfallort ausgelegt.¹ Die nächsten Jahre waren geprägt von der Aufarbeitung der Katastrophe sowie der Strukturierung der Bergungserfassung als Langzeitprojekt. Anders als man hätte vermuten können, gelang es dem Archiv schon sehr bald, Neuübernahmen zu ermöglichen und in die insbesondere seit Mitte des 19. Jahrhunderts angelegte Ordnung der Bestände zu integrieren. Somit wurden zwei Aufgabenstränge fokussiert und ausgearbeitet, wodurch es dem Archiv möglich wurde, sein Selbstverständnis als Bürger*innenarchiv wahrzunehmen.

MASSNAHMEN DES WIEDERAUFBAUS (BARCODE- UND DATENBANKSYSTEME)

Als Reaktion auf den Einsturz wurde im Historischen Archiv das Datenbanksystem ausgebaut und an die individuellen Bedarfe angepasst. Eine dieser Anforderungen war es, nun unter den Umständen des Einsturzes erstmals vollständig digital eine Ordnung und Recherchierbarkeit der Bestände zu ermöglichen.² Auf der anderen Seite wurden auch neue logistische Strukturen ausgearbeitet. Dies hatte zur Folge, dass ein neues Identifikationssystem eingeführt wurde, um sowohl alle Archivalien physisch auffindbar zu machen als auch ihre zugehörigen Informationen in den entsprechenden Datenbanken abzurufen. Dazu wird für jede Einheit ein eindeutiger aus neun Ziffern bestehender Zahlencode generiert, der sich über seinen entsprechenden Strichcode praktikabel einscannen lässt. Dieser sogenannte Objektbarcode ist auf der jeweiligen Mappe, in die die Einheit verpackt wurde, befestigt, um eine direkte Verbindung des Datensatzes zur physischen Einheit herzustellen. Durch das Einscannen

dieses Barcodes werden die zugehörigen Informationen unverzüglich im System aufgerufen und im Bedarfsfall ergänzt. Um zusätzlich alle vorhandenen physischen Zusammenhänge darzustellen, werden auch die Archivkartons, in die die Einheiten gebettet werden, mit einem eigenen neunstelligen Barcode versehen. Hierbei handelt es sich um die sogenannten Behältnisbarcodes. Zur direkten Unterscheidung von den Objektbarcodes wird den Behältnisbarcodes jedoch ein B vorangestellt. Auch die Regalplätze wurden seinerzeit in den Asylarchiven, ebenso wie heute im Neubau des Archivs, mit Barcodes versehen. Für diese logistische Ordnung wird inzwischen eine speziell für den Anwendungsfall in Köln entwickelte Software (das sogenannte „Restaurierungs- und Dokumentationsmodul“, kurz: RDM) verwendet. Sie macht es möglich, jeden Objektbarcode in ein Behältnis und jedes Behältnis auf einen Regalstandort zu buchen. Stets können damit in der Datenbank, die vom System zusätzlich unterscheidbaren Bergungseinheiten und Verzeichnungseinheiten in ihrem jeweiligen Behältnis gefunden werden. Ebenso ist auch jedes Behältnis mit den darin befindlichen Objekten auffindbar. Diese Systematik schaffte die Möglichkeit einer logistisch sowie konservatorisch bestmöglichen Lagerung für das jeweilige Archivale, da Rücksicht auf Größe und Klimaanforderung des Materials genommen werden kann. Zudem bestand besonders in der Vergangenheit für die eingestürzten Dokumente keine Möglichkeit, sie nach Beständezusammenhängen zu ordnen, da diese noch nicht final identifiziert worden waren. Genau diese Tatsache bot aber nun die Möglichkeit, durch eine dynamische Lagerung das Magazin ideal zu nutzen und die volle Kapazität auszuschöpfen. Zudem erlaubt es die Software, im Zuge der Arbeitssteuerung, Auswertungen zu individuellen Fragestellungen vorzunehmen und Übersichtslisten zu erstellen.

Die Unterscheidung von Bergungseinheit und Verzeichnungseinheit ist dabei besonders für die archivischen Identifikationstätigkeiten von Belang: Eine Bergungseinheit ist eine aus dem Einsturz hervorgegangene physische Einheit, die in ihrem Datensatz im System inhaltlich wie äußerlich beschrieben wird. Die Verzeichnungseinheit ist wiederum die inhaltlich zusammengehörende archivische Einheit, die in der Archivdatenbank hinterlegt ist. Die Verzeichnungseinheit kann somit sowohl ein eingestürztes Archivale, als auch ein neuübernommenes sein. Angestrebt wird immer die Zuordnung der Bergungseinheiten zu den bestehenden Verzeichnungseinheiten. Hierfür sind ebenfalls Sortierungsarbeiten, Zusammenführungen und Trennungen der Bergungseinheiten notwendig, um die ursprüngliche Verzeichnungseinheit wieder logisch und physisch herzustellen (s. Abb. 1). Eine einzelne Verzeichnungseinheit kann also beispielsweise aus fünf Bergungseinheiten bestehen, die jeweils lose Blattsammlungen enthalten. Die Identifizierung der ursprünglichen Verzeichnungseinheit kann dementsprechend sehr aufwändig sein und bedurfte auch stets einer umfangreichen und ausführlichen Erfassung des eingestürzten Materials, um die Teileinheiten verknüpfen zu können. Hierbei bildet neben der Bergungserfassung die Digitale Rekonstruktion Kölner Fragmente (DRKF) in Kooperation mit der MusterFabrik Berlin einen weiteren wichtigen Arbeitsstrang. Dieser Prozess konzentriert sich auf die Bearbeitung der durch den Einsturz entstandenen Puzzlefragmente mit

DIE ZUSAMMENFÜHRUNG UND NEUSTRUKTURIERUNG DER BERGUNGSERFASSUNG

Weiterhin wurde auch die räumliche Zusammenführung des Archivguts und der Kolleg*innen angestrebt. Im Zeitraum zwischen 2017 und 2021 wurden die Standorte weiter reduziert und schließlich ein Neubau am Eifelwall bezogen. In diesen Jahren wurde die Bergungserfassung ebenfalls immer weiter verfeinert und die Sortierung und Zuordnung des trocken geborgenen Archivguts als Großprojekt abgeschlossen (s. Abb. 4).



Abb. 4: Archivwagen mit Behältnissen

Nach einer Hochrechnung aus dem Jahr 2018 sind nach Abschluss der Erfassungsarbeiten knapp über 1,6 Millionen⁶ Bergungseinheiten zu erwarten. Zum jetzigen Zeitpunkt umfasst die Anzahl der Bergungseinheiten ca. 1,4 Millionen. Davon sind bereits über 900.000 in ihren ursprünglichen Bestandszusammenhang zurückgeführt worden. Aber nach wie vor sind noch nicht alle vom Einsturz betroffenen Verzeichnungseinheiten wiedergefunden oder erfasst worden. Im Zuge einer Neustrukturierung im Neubau am Eifelwall konnten die weit fortgeschrittenen Erfassungstätigkeiten den Status eines Sonderprojektes im Haus verlassen und in eine Langzeitaufgabe neben den üblichen archivischen Tätigkeiten überführt werden. Vor dieser Umstrukturierung war die Bergungserfassung in erster Linie an die Arbeit in den verschiedenen Asylarchiven geknüpft. Nach dem Umzug in den Neubau konnten diese Arbeiten zentralisiert werden, da nun das gesamte Archivgut und alle Mitarbeitenden an einem Ort versammelt waren. Die Bergungserfassung war lange Zeit ein Projekt, dessen Bearbeitung überwiegend durch Hilfskräfte in den Asylarchiven durchgeführt wurde. Somit bestand die Prozessstruktur

darin, nach der Erfassung des Archivguts zu den Standorten der Fachkräfte nach Köln zu transportieren. Hier wurden eine Sichtung durch die archivischen Fachkräfte zur weiteren Überprüfung (die sogenannte "Phase II"), die restauratorischen Prozesse sowie die abschließende Identifizierung durchgeführt. Dies bezweckte die endgültige Verknüpfung im AFiS und die Nutzbarmachung für den Lesesaal.⁷

Heute wird die Bergungserfassung in ihrer Hauptsache durch ein verkleinertes Team betreut. Hier werden Konzepte für einzelne Archivalientypen und -inhalte erstellt und entsprechende Übersichtslisten vorbereitet. Genauso wie die spätere Identifizierung, wird die Bergungserfassung als vorangestellter Prozessschritt in den Gesamtprozess der Wiedernutzbarmachung eingebettet. Da diese Tätigkeit – anders als früher – nur noch von archivischen Fachkräften durchgeführt wird, können Arbeitsschritte verdichtet werden. Die zusätzliche Kontrolle und Sichtung der ausgefüllten Erfassungsinformationen durch eine Fachkraft kann somit ausgespart und der Prozessesstrang effizient gestaltet werden. Die Integration der Erfassungsarbeiten in das Gesamtkonzept ist durch die Zentralisierung im Neubau sowie durch die Neuaufstellung der Prozesslandschaft erstmals möglich. Die Arbeiten konzentrieren sich heute auf eine möglichst genaue Erfassung durch eine Fachkraft, damit zeitnah die restauratorische Bearbeitung angeschlossen werden kann. Hierauf aufbauend werden in einem letzten Schritt alle Informationen in der Archivdatenbank ergänzt und die endgültige Verknüpfung erstellt, bevor das Archivgut wieder zur Benutzung freigegeben werden kann.

Seit dem Umzug Ende 2021 wurden die gesamten Strukturen des Historischen Archivs zusammengeführt und neu aufbereitet. Die gut vorgearbeitete Menge der übrigen Erfassungstätigkeiten konnte nun effizienter und zielorientierter gestaltet werden. Dennoch wird die weitere Abarbeitung voraussichtlich noch zwei oder drei Jahre in Anspruch nehmen. Im Anschluss wird sich denjenigen Einheiten zugewandt, die in der ersten Identifizierungsphase noch nicht zugeordnet werden konnten und somit bislang nicht wieder nutzbar sind. Zu erwarten ist somit weiterhin ein stetig ansteigender Datenpool von benutzbaren Einheiten.

Die Bergungserfassung erhielt eine neue Stellung in der Prozesslandschaft des Kölner Archivs und bildet somit einen der beiden essentiellen Stränge zur Nutzbarmachung des Archivguts. Die Erschließung der Neuübernahmen sowie die Bergungserfassung laufen parallel in ihren vorgesehenen Prozessen und werden vom Historischen Archiv mit Rheinischem Bildarchiv gleichwertig in der Bearbeitung fokussiert. Künftig wird der Schwerpunkt darauf liegen, die noch ausstehenden Sonderaufgaben der Trockenbergung zu bearbeiten. Von besonderer Bedeutung ist die Erfassung der sogenannten "Stichproben". Hierbei handelt es sich um nach dem Einsturz geborgenes Archivgut, das im Jahre 2014 zum Zwecke der Ermittlung einer Gesamtschadenssumme zusammengestellt wurde. Da der damalige Gutachter diese Behältnisse aus rechtlichen Gründen versiegelte, waren sie bis zum Abschluss des Regresses für die Bearbeitung gesperrt. Hierin findet sich erneut ein breiter Querschnitt des Kölner Archivguts, das bisher noch nicht für die Nutzung zur Verfügung stehen konnte. Ähnlich verhält es sich mit demjenigen Archivgut, das aufgrund seines Lagerortes im

Haus den Einsturz weitgehend unbeschädigt überstanden hat. Das daraus resultierende geringere Schadensbild ließ eine spätere Bearbeitung zu. Demgegenüber finden sich die aus dem Grundwasser geborgenen Archivalien, die aufgrund ihres überdurchschnittlich starken Schadensbildes bislang nicht bearbeitet werden konnten. Das Ausmaß des Schadens bei diesem kontaminierten Archivgut ist dabei so hoch, dass eine Identifizierung vor der aufwändigen Reinigung nicht möglich wäre. All diese Aufgaben erfordern eine besonders vorausschauende Planung mehrerer beteiligter Arbeitsbereiche des Hauses. Somit steht auch 15 Jahre später, nach einer intensiv erbrachten Erfassungs- und Identifizierungsleistung, immer noch eine Aufarbeitung der Katastrophe aus

archivischer Sicht im Vordergrund. Obwohl große Fortschritte gemacht wurden, sind die Arbeiten bei Weitem noch nicht abgeschlossen. Bis das Ziel der vollständigen Wiederherstellung des Kölner Stadtgedächtnisses erreicht ist, werden noch einige Jahre vergehen.

Selina Lurz/Sven Theemann, Köln

- ⁶ Vgl. Ulrich Fischer, Markus Späinghaus: Geschichte mit Zukunft. 10 Jahre Wiederaufbau des Kölner Stadtarchivs. Köln 2019, S. 118.
⁷ Vgl. ebd., S. 76.

„... DENN SIE WISSEN NICHT, WONACH SIE SUCHEN SOLLEN.“

EIN PLÄDOYER FÜR DIE ERWEITERUNG ARCHIVPÄDAGOGISCHER METHODEN ZUGUNSTEN ANGEHENDER HISTORIKER*INNEN

Im Vorlesungsverzeichnis des Historischen Seminars der Universität Heidelberg reihte sich im Sommersemester 2024 eine Veranstaltung mit Seltenheitswert ein. Die Übung „Mehr als nur Aktenlager! Archivwissenschaftliche Grundlagen für Historikerinnen und Historiker“ lief mit einem Programm auf, das über die sonst üblichen Quellenübungen hinausging. Die Beschreibung erklärte, man wolle dem deutschen Archivwesen auf den Grund gehen, um ein tiefgreifendes Verständnis der Strukturen zu entwickeln und es für das eigene Studium anwenden zu können. Konzipiert war die Übung in Kombination mit zwei Exkursionen für einen realitätsnahen Eindruck. Bereits eine halbe Stunde nach Öffnung des Anmeldefensters zählte die Anfrageliste über 60 Interessierte. Die Mehrheit der schließlich zugelassenen Studierenden gab bei einer Umfrage an, dem Begriff „Archiv“ zwar schon

einmal begegnet zu sein, aber abseits des stereotypen Kellergewölbes keine genaue Vorstellung davon zu haben, was er bedeutet. In der Vorstellungsrunde der Einführungs-sitzung wurde dann klar, dass die weitgreifenden Kontexte der Archivwelt auch für die Erfahreneren ein blinder Fleck waren. Ihre Motivation zur Teilnahme begründeten sie mit dem Bedürfnis, die Eindimensionalität ihres Archivbegriffs aufbrechen zu wollen. Die Evaluation nach Ablauf des Semesters offenbarte: Es war gelungen. Die Mischung aus Archivgeschichte zum historischen Verständnis der deutschen Archivlandschaft, die multiperspektivischen Zugriffe auf Bewahrungsstrategien und -entwürfe und die Fragen nach der Tiefenvernetzung von Geschichts- und Archivwissenschaft hatten das Archivbild der Studierenden verwandelt. Ein*e anonym*e Teilnehmer*in gab an, „[d]ass sich ein solches Format (Exkursion, Archivwesen, direkter Bezug

zum geschichtswiss. Studium) bitte etablieren sollte. Es ist so hilfreich!!!“

Ja warum, mag man fragen, hat es das nicht? Die Schnittstelle zwischen Geschichtswissenschaft und Archivwesen wird in Fachkreisen immer wieder zum Thema gemacht. Vor dem wissenschaftshistorischen Hintergrund mutet die Verzahnung an wie eine profane Selbstverständlichkeit. Auf Tagungen und Kongressen werden Vorschläge zum konstruktiveren Miteinander laut,¹ die Herausforderungen der digitalen Zukunft werden als Gemeinschaftsprojekt begriffen.² Doch „das Archiv“ scheint an modernen historischen Lehrstühlen und Instituten als Randerscheinung aufzutauchen. In den geschichts- und kulturwissenschaftlichen Vorlesungsverzeichnissen der Universitäten in Baden-Württemberg (Freiburg, Konstanz, Tübingen, Mannheim und Heidelberg) werden dem Themenkomplex Archiv in den letzten zehn Jahren zuweilen Gastauftritte eingeräumt. Meist verdinglichen die den Grund- bzw. Hilfswissenschaften verschriebenen Veranstaltungen die archivischen Zusätze aber als füllendes Beiwerk. Die Veranstaltungen, die archivarische Strukturen und Grundlagen als Schwerpunkt auserkoren haben, sind dabei klar in der Unterzahl.³ Die Einblicke in Arbeit und Räumlichkeiten der Gedächtnisinstitutionen passieren in der Folge durch Führungen. Diese erreichen zwar einen grundlegenden Eindruck – eine einstündige Führung durch einen langen Magazingang und der Beschuss mit statistischen Informationen – aber werden weder der archivarischen Mühe noch den Untiefen archivwissenschaftlicher (und damit verbunden auch geschichtswissenschaftlicher) Theorien gerecht.

So zeugt die Realität der Veranstaltungen im akademischen Kosmos nicht nur von einer geringen Dichte der Vernetzungsmöglichkeit zwischen Archivar*innen und Historiker*innen, an der das Gros der Studierendenschaft aufgrund der wenigen Plätze in den Lehrveranstaltungen keinen Anteil hat. Sie macht auch darauf aufmerksam, dass die wertvollen archivpädagogischen Trickkisten in ihrer Fülle die Bühne „Universität“ nicht erreichen.

VON QUALIFIKATIONEN UND ANTEILNAHME

Das Resultat bekommen Archivar*innen in den Lesesälen oder Dozent*innen in Seminaren zu spüren, wenn der akademische Nachwuchs aus Verzweiflung und Überforderung entweder den Weg zur Quelle nicht findet oder sie dann nicht entschlüsseln kann. Dabei lohnt sich ein Blick in die Ausrüstung der Kolleg*innen, die sich neben den Kerntätigkeiten mit der vielfältigen Vermittlung ihrer Arbeit auseinandersetzen: Einblicke für Schulklassen und Kindergartengruppen, wie sie beispielsweise das Landesarchiv Baden-Württemberg anbietet,⁴ wirken als Mitmach-Methoden wie Multiplikatoren und verstärken eine nachhaltige Erinnerung an „dieses große Gebäude, in dem ganz viel Geschichte liegt“. Nun handelt es sich bei Studierenden selbstverständlich um junge (oder ältere) Erwachsene, die sich bewusst und aus einer eigenständigen Handlung heraus für ihr Studienfach entschieden haben. Sie haben einen Bezug zur Geschichtswissenschaft oder ihren Nachbardisziplinen und sind im besten Fall bereit, sich auf die Aspekte und Herausforderun-

gen der akademischen Ausbildung einzulassen. Sie kommen im Laufe der Semester zumindest in Quellenübungen mit „Originalen“ in Kontakt – die ironischerweise meist in digitaler Form oder als Kopie vorliegen, und damit nicht eigentlich im Original. Wenn Student*in sich gegen eine entsprechende Exkursion oder das rare Lehrangebot entscheidet, bleiben Erfahrungen mit den „echten Originalen“ in Gänze aus. Die studentischen Eindrücke von Geschichtswissenschaft sind außerdem gefesselt an die Universitätspolitik der betreffenden Bildungsstätte und ihre Forschungsbereiche. Außerdem können das eigene Interesse und die daraus resultierende Schwerpunktsetzung die Kontaktmöglichkeiten mit Archiven ebenfalls dezimieren. In der Folge kommen Studierende der Geschichtswissenschaft häufig als unbeschriebene Blätter daher, die keine Idee von dem Umgang mit den alten beschriebenen Blättern gewinnen konnten.

Im Übrigen erschwert die Zugänglichkeit der Archive durch ihre Web-Präsenz das Ringen mit der fehlenden Archivkompetenz zusätzlich. Online-Findmittel verführen zu einer unbedarften und häufig unkoordinierten Suchstrategie, die dann nicht auf die gewünschten Ergebnisse stößt. Gleichwohl Handreichungen und Hilfestellungen dem Findmittel häufig angegliedert sind, können sie nur als Blaupause für ein umfassenderes Archivverständnis dienen. Ungereimtheiten und Kurzsichtigkeit scheinen die Studierendenschaft zu beherrschen; die Suchen werden – ohne Tendenz zur persönlichen Nachricht an die Archive selbst – kurzerhand abgebrochen, denn „im Katalog konnte ich nichts finden, also haben die wohl nichts.“⁵

Bei der Konzeption der Lehrveranstaltung im Sommersemester 2024 in Heidelberg stand also der arbeitspraktische Kontext im Vordergrund, in dem sich die theoretischen Diskussionen bewegen sollten. Zentral hierfür war die erste Exkursion in der zweiten Semesterwoche, die über zwei Tage im Staatsarchiv Würzburg abgehalten wurde. Die dort beteiligten Kolleg*innen erklärten anhand der eigenen Behörde ihre Arbeitsweise und Tektonik, gaben Einblicke in Quellengattungen und Archivgeschichte und vermittelten die Herausforderungen der archiv(ar)ischen Welt. Natürlich war hier eine Führung durch die Magazinräume in der Residenz inkludiert, doch ihre Inhalte und die Ausmaße der bewahrten Kilometer an Material wurden erst durch den Gesamtzusammenhang anschaulich und erkennbar. Auf dieser Basis gelang es den Teilnehmenden der Übung in den folgenden Wochen, die Theoreme mit praktischen Eindrücken zu unterfüttern. Sie fielen nicht dem Bild anheim, das nur künstlerische Einzelstücke und verstaubte Kellerräume zeigt. Ihr Verständnis umfasste das Archiv als Behörde, die Verantwortlichkeit der Bestandsbildung, die Herausforderungen der digitalen Fortentwicklung und ihre Bedeutung für Historiker*innen. Die Diskussionen und Überlegungen in den heimischen Universitätsräumen wurden in der Folge mehrdimensional, und die Facetten der deutschen Archivalandschaft konnten in vollem Ausmaß begriffen werden. Die Studierenden wussten mit Abschluss der Veranstaltung nicht nur, wie sie ein Findmittel benutzen sollten, sondern auch um Zuständigkeiten und Zugänglichmachung. Die abschließende Exkursion zum Ende des Semesters im MARCHIVUM in Mannheim führte die Vielartigkeit der deutschen Archive erneut vor Augen und verringerte die

Berührungsgänge mit dem E-Mailpostfach der Gedächtnisinstitutionen noch einmal mehr.

VON CHANCEN UND MÖGLICHKEITEN

Die Nachhaltigkeit dieser Veranstaltung wird sich – so die Hoffnung bei Konzeption – für die Studierenden dann zeigen, wenn sie auf das Archivwissen zurückgreifen, um sich historischen Fragestellungen zu nähern. Allein das Wissen um die Existenz und Funktionsweise der behördlichen und Freien Archive in Deutschland eröffnet ihnen einen breiten Horizont. Die Erfahrung der beteiligten Kolleg*innen sei an dieser Stelle zusätzlich erwähnt: Man habe sich daran erfreut, noch einmal in den theoretischen Lehrinhalten kramen zu können und zurückzudenken an das kleine Einmaleins der Archivschule. An dieser Stelle möchte ich noch einmal meinen herzlichsten Dank für Ihre Mühe und Zeit aussprechen – und für die Vermittlung nicht nur des Fachlichen, sondern auch für das Weitergeben Ihrer Begeisterung. Das Staatsarchiv Würzburg und das MARCHIVUM Mannheim erhielten im Rahmen dieser Übung den Status außeruniversitärer Lernorte. Ihre Räumlichkeiten gaben dem Gelernten zusätzliches Gewicht; die Kontaktaufnahme mit der Archivwelt verstärkte die gesammelten Eindrücke. Es war klar, dass diese Exkursionen keiner Berufsorientierung dienen sollten, sondern als Besuch von Angehörigen der Geschichtswissenschaft bei Angehörigen des Archivwesens gedacht waren. Diese Gewissheit veränderte den Umgang der Studierenden mit ihrer Lernsituation: Das Archiv avancierte zu einem Bezugspunkt geschichtswissenschaftlicher Selbstreflexion. Was macht Geschichte aus? Was bedeutet es, mit Quellen zu arbeiten? Welche Quellen stehen mir zur Verfügung und wie finde ich sie? Was bedeutet es, dass Archive nicht alle Unterlagen übernehmen? Diese und weitere Fragen hätten historiografische Grundlagen allein mit hoher Wahrscheinlichkeit nicht provoziert, ebenso wenig eine für sich erlebte Führung. Die punktuellen Veranstaltungen und Führungen als Schlaglichter vermitteln Studierenden zumeist einen Weg in Einrichtungen, die mit ihrem Fachbereich verwandt sind – seien es Archive, Museen oder Gedenkstätten. Doch das Verständnis dessen, was Archivar*innen täglich leisten, um historische Forschung möglich zu machen, bildet Anfangs- und für manche auch den Endpunkt ihrer akademischen Ausbildung. Nicht nur die Methodenkompetenz wird gefördert und das Fachwissen erweitert, für manche wird mit dem Archiv ein Berufsbild offenbar. Aber erst die Horizonterweiterung, die das Archiv nicht nur als Berufsoption begreift, sondern vor allem nach dem fragt, was vor der Beforschung geschieht, macht aus den Räumlichkeiten ein Wunderland der eigenen Entdeckung – und das völlig unabhängig von den eigentlichen Beständen. Das Geschichtsstudium braucht Archiv, um seine volle Breite erlernen zu können: Das Selbstverständnis von Historiker*innen ist in erheblichem Maße davon abhängig, was Archivar*innen zuvor bereitet haben. Um das nachhaltig zu vermitteln, lohnt sich ein tiefer Griff in die archivpädagogische Trickkiste. Denn die Annahme, Studierende seien aufgrund ihres Studienfaches Anteilnehmende am Erfahrungshorizont archivischer Grundlagen und bräuchten deshalb keine Vermittlung derselben, bestätigt sich nur in

den seltensten Fällen. Historiker*innen in spe fremdeln mit der Welt der Archivar*innen.

Die Geschichtswissenschaft ringt mit digitalen Entwicklungen; die Forschungsgebiete nehmen zu. An den Universitäten treffen viele Inhalte auf zu wenig Zeit – und häufig auf wenig Lehrpersonal. Dass das Archivwesen hier nicht standardisiert abgebildet werden kann, ist tragisch, aber eine logische Konsequenz. Zwar ist es beispielsweise in Heidelberg im Kanon der Hilfs- bzw. Grundwissenschaften mitverstanden,⁶ in der Lehre schlägt sich dies jedoch nicht nieder. Studierende sind in der Folge Adressat*innen eines Komplexes, der ihre akademischen Grenzen sprengen kann. Die Quellen selbst wirken bereits als Zündstoff für Faszination, die nicht nur durch Anschauen ohne Anfassen gelockt werden würde. Spricht etwas dagegen, Studierende auch bei Kurzbesuchen Anteil nehmen zu lassen an einer Archivverfahren, die gemacht statt gezeigt wird? Dass der wissenschaftliche Nachwuchs das Archiv erlebt, statt es nur zu bestaunen? Zum Beispiel eigens zu überlegen, was von den ihnen vorgelegten Akten kassiert würde, oder darüber in proaktiven Austausch zu gehen, welche Nachlässe warum von einem staatlichen Archiv übernommen werden könnten – die Trickkiste der Archivpädagogik muss keinen Boden haben.

Die Lehrveranstaltung in Heidelberg ermöglichte einerseits den Blick dafür, wie restriktiv die Anteilnahme an „dem Archiv“ ist, die Studierenden von Seiten ihrer akademischen Ausbildung eröffnet wird. Aber andererseits zeigte sie auch, dass es den Historiker*innen von morgen guttut, mit archivischer Magie zu spielen, statt nur bei der Vorstellung derselben anwesend zu sein. Das muss nicht bedeuten, fachliche Ansprüche und Qualität der Vermittlung zu kompromittieren. Das Niveau der Veranstaltung leidet nicht unter spielerischen Zugängen; Neugierde und Interesse sind keine Symptome für Wissenschaftlichkeitsverlust. Doch sie machen Zugänge einfacher, wecken Fantasie und führen zu einer konstruktiveren Vernetzung des Gelernten. Das hilft insbesondere beim Blick über den Tellerrand: Gerade die Komplexität dessen, was zwischen Archiv- und Geschichts-

¹ Vgl. Rainer Hudemann: Forschung und Archive. Überlegungen zur wissenschaftlichen Kooperation. In: Archive und Forschung. Referate des 73. Deutschen Archivtags 2002 in Trier. Hg. v. VdA Verband Deutscher Archivarinnen und Archivare. Siegburg 2003, S. 347–358 / Friederike Scholten: Nutzen und Nachteil der Archivarbeit für die historische Forschung. In: Archivpflege in Westfalen-Lippe 84 (2016), S. 6–10.

² Vgl. Rainer Hering: Archive und Geschichtswissenschaft im digitalen Zeitalter. Probleme und Herausforderungen aus der Sicht eines deutschen Archivars und Historikers. In: Mitteilungen des Instituts für Österreichische Geschichtsforschung 120 (2012), S. 116–138.

³ Vgl. Vorlesungsverzeichnis der Universität Heidelberg SoSe 2014 bis SoSe 2024, erreichbar über https://lsf.uni-heidelberg.de/qisserver/rd?s?state=change&type=5&moduleParameter=veranstaltungSearch&nextdir=change&next=search.vm&subdir=veranstaltung&_form=display&function=search&clean=y&category=veranstaltung.search&navigationsPosition=lectures%2CsearchLectures&breadcrum=searchLectures&topitem=lectures&subitem=searchLectures (aufgerufen am 20.08.2024).

⁴ Vgl. <https://www.landesarchiv-bw.de/de/themen/archivpaedagogik--angebote-fuer-schulen/45978> (aufgerufen am 20.08.2024).

⁵ Diese Aussage begegnete der Verfasserin mehrmals, in voneinander unabhängigen Lehrveranstaltungen.

⁶ Vgl. <https://www.uni-heidelberg.de/de/studium/alle-studienfaecher/historische-grundwissenschaften> (aufgerufen am 20.08.2024).

wissenschaft passiert, birgt Spannung und bietet zugleich den Mehrwert für beide Fachbereiche.

Die fachlich hoch angelegte Latte kann also als Türsturz verbaut werden, um den Eingang in die Welt der Bestandserhaltung, Tektoniken und der Kilometer an Überlieferung zu kennzeichnen. Es lohnt sich, die fachliche Vorbildung der Nachwuchswissenschaftler*innen und -lehrkräfte als

archivische Chance für Multiplikation zu begreifen, statt nur als selbstverständliche Kooperationsmöglichkeit. Jede durch diesen gebauten Eingang schlüpfende Person ist ein Gewinn für Archive – und erst recht, wenn es die Historiker*innen von morgen sind.

Nathalie Pfeuffer, Würzburg

DAS „WIESBADENER MODELL“ ZUR UNTERSUCHUNG HISTORISCH BELASTETER STRASSENAMEN

EIN ZWISCHENFAZIT AUS ARCHIVISCHER SICHT

Wie in vielen anderen deutschen Großstädten¹ begann in der hessischen Landeshauptstadt Wiesbaden die Kontroverse um die NS-Vergangenheit von Namenspatroninnen und -patronen im öffentlichen Raum in den späten 1990er Jahren.² Der erste prominente Fall betraf den aus Wiesbaden-Naurod stammenden Heimatdichter Rudolf Dietz (1863–1942), nach dem u. a. 1957 eine Grundschule benannt worden war. Der Volksschullehrer Dietz war ab 1933 NSDAP-Mitglied und hatte neben fast 2.000 weiteren auch rund 40 Mundartgedichte mit eindeutig antisemitischer und den NS-Staat glorifizierender Tendenz verfasst. Um die Frage, ob die Nauroder Grundschule weiterhin Dietz' Namen tragen sollte, entspann sich Anfang der 2000er Jahre eine langanhaltende geschichtspolitische Debatte, die erst 2019 mit der Umbenennung in „Wickerbach-Grundschule“ im Rahmen eines Neubaus der Schule endete.³

Parallel zu den Ausläufern der Diskussionswelle um Dietz geriet auch der Komponist, Dirigent und Autor Hans Erich Pfitzner (1869–1949), nach dem in Wiesbaden eine Straße benannt war, in den Fokus einer historischen Neubewertungsdiskussion. Pfitzner war Antisemit und hatte sich auch nach dem Ende der NS-Diktatur entsprechend geäußert. Auch über diesen Fall wurde in Wiesbaden lange und kontrovers diskutiert, obwohl andere Städte Straßenbenennungen nach

Pfitzner bereits zurückgenommen hatten. Im Zuge dessen überstimmte die Stadtverordnetenversammlung mehrheitlich auch den eigentlich für die Straßenbenennung zuständigen Ortsbeirat, was eine übergeordnete Debatte über die Rechte der Ortsbeiräte auslöste.⁴

Zum Ende dieser zweiten, erbittert geführten Debatte um einen Wiesbadener Namensgeber wurde durch die Stadtverordnetenversammlung ein politischer Beschluss gefasst, der eine generelle Untersuchung der NS-Belastung aller Personen vorsah, nach denen Straßen, sonstige Verkehrsflächen und Einrichtungen im gesamten Stadtgebiet benannt worden waren.⁵

VON DER EINZELFALLPRÜFUNG ZUM „WIESBADENER MODELL“

Im Juli 2020 konnte das Stadtarchiv Wiesbaden als fachlich zuständige Abteilung der Verwaltung die Ergebnisse seiner ersten Recherchen zur Quantifizierung des von der Politik definierten Untersuchungsbedarfs vorstellen. Von den 2.004 benannten Verkehrsflächen in der Stadt trugen 651 den Namen von Personen. Unter diesen waren wiederum 60 Namensgeberinnen und -geber, die in der NS-Zeit gelebt und gewirkt hatten.

Entgegen der ursprünglichen Überlegungen, die eine weitere Überprüfung der Namensgeberschaften mit Bordmitteln des Stadtarchivs vorgesehen hatte, empfahl das Archiv den politischen Entscheidungsträgern, auch auf Grundlage von Quantifizierungen des zeitlichen Arbeitsaufwandes, ein zweistufiges Verfahren zur historisch-wissenschaftlichen Überprüfung und politischen Bewertung der Namenspatronate. Die erste Stufe sollte eine historische Beurteilung durch eine ausschließlich mit Fachleuten besetzte Historische Fachkommission bilden, die zweite Stufe die politische Bewertung durch die für Straßenbenennungen zuständigen Ortsbeiräte. Der Kommission zuarbeiten sollte ein beim Stadtarchiv für das Projekt befristet beschäftigter wissenschaftlicher Mitarbeiter im Umfang eines Vollzeitäquivalents. Dieses Grundkonzept wurde in einer Sitzungsvorlage formuliert,⁶ die neben einer Geschäftsordnung auch eine Vorschlagsliste für die Mitglieder der Kommission sowie einen Kriterienkatalog für die Überprüfung der Namensgeberinnen und -geber enthielt.⁷

ORGANISATION UND ARBEITSWEISE DER HISTORISCHEN FACHKOMMISSION

Die Kommissionsmitglieder wurden im Februar 2021 auf Vorschlag des Stadtarchivs durch die Stadtverordnetenversammlung berufen. Am 1. März 2021 konnte sich die Kommission unter den besonderen Rahmenbedingungen der Corona-Pandemie in einer Videokonferenz konstituieren. Der Kommission gehörten Professor Dr. Christoph Cornelißen (Frankfurt a. M.), der zum Vorsitzenden gewählt wurde, Professorin Dr. Sabine Mecking (Marburg), gewählt zur stellvertretenden Vorsitzenden, sowie Privatdozentin Dr. Nadine Freund (Kassel/München), Professor Dr. Andreas Hedwig (Marburg) und Professor Dr. Eckart Conze (Marburg) als stimmberechtigte, wissenschaftliche Mitglieder an. Qua Amt nahmen Stadtverordnetenvorsteher Dr. Gerhard Obermayr (CDU), als Vertreter der städtischen Beschlussgremien, und Dr. Peter Quadflieg, Leiter des Stadtarchivs Wiesbaden, als Vertreter der Verwaltung an den Sitzungen ohne Stimmrecht teil. Zum 1. Juni 2021 konnte ein öffentliches Personalauswahlverfahren mit der Einstellung des Historikers Dr. Dirk Stolper als Fachreferent und Geschäftsführer der Fachkommission abgeschlossen werden. Damit nahm die Fachkommission ihre kontinuierliche Arbeit auf.

Die Kommission tagte zwischen Sommer 2021 und Ende 2023 insgesamt zehnmal. Sie bewertete dabei 71 Biografien – darunter 66 Männer und vier Frauen – im Sinne einer Einstufungsentscheidung, die in Artikel § 7 Abs. 2 ihrer Geschäftsordnung politisch vorgegeben war. Die Geschäftsordnung sah die Einstufung der untersuchten Personen in eine der folgenden drei Fallgruppen vor. Fallgruppe 1: „historisch schwer belastet und Umbenennung empfohlen“; Fallgruppe 2: „historisch umstritten und diskussionswürdig. Umbenennung wird jedoch nicht empfohlen“; Fallgruppe 3: „historisch nicht oder minderbelastet und Beibehaltung des Namens empfohlen“. Diese nicht unproblematische Schematisierung war keine Erfindung des Stadtarchivs, das die Geschäftsordnung entworfen hatte, oder der Fachkommission, sondern bereits im ursprünglichen politischen Auftrag durch die Stadtverordnetenversammlungen vorgegeben

worden.⁸ Als Vorbild für diesen politischen Beschluss hatte den Antragsverantwortlichen ein kurz zuvor in Düsseldorf abgeschlossenes Vergleichsprojekt gedient, das ebenfalls mit drei Kategorien („schwer belastet/nicht haltbar“, „teilweise belastet/diskussionswürdig“ und „unbelastet“) gearbeitet hatte und das sich seinerseits auf das Konzept eines Untersuchungsprojekts in Freiburg im Breisgau stützte, das 2016 abgeschlossen worden war.⁹

Zur Vorbereitung der Diskussion und Entscheidungsfindung in der Fachkommission wurden, auf Grundlage der Auswertung des Forschungsstandes und von Primärquellen aus Archiven, Kurzbiografien durch den wissenschaftlichen Referenten der Kommission erarbeitet, die teilweise mehrfach zur Vorlage kamen, wenn sich aus der Kommissionsberatung

¹ Eine umfassende Gesamtdarstellung des Diskurses um Straßennamen in Deutschland liegt noch nicht vor. Einführend und systematisch behandelt Rainer Poppinghe: *Wege des Erinnerns. Was Straßennamen über das deutsche Geschichtsbewusstsein aussagen*. Münster 2007 das Thema. Einen spezifischeren Einblick zu Einzelaspekten geben die einzelnen Beiträge des Sammelbandes *Fragwürdige Ehrungen!? Straßennamen als Instrument von Geschichtspolitik und Erinnerungskultur*. Hg. v. Matthias Freese. Münster 2012. Der Band *Geschichtspolitik im öffentlichen Raum. Zur Benennung und Umbenennung von Straßen im internationalen Vergleich*. Hg. v. Birgit Nemeš u. Florian Wenninger. Göttingen 2019 (= *Zeitgeschichte* 46, H. 1 (2019)) widmet sich dem Thema aus internationaler Perspektive. Spezifische Beispiele zu Umbenennungsdebatten im Kontext deutscher Ehrregime behandeln die Sammelbände: *Ehrregime. Akteure, Praktiken und Medien lokaler Ehrungen in der Moderne*. Hg. v. Dietmar von Reeken u. Malte Thießen. Göttingen 2016, und *Verhandelte Erinnerungen. Der Umgang mit Ehrungen, Denkmälern und Gedenkortern nach 1945*. Hg. v. Matthias Freese u. Marcus Weidner. Paderborn 2018.

² Vgl. dazu demnächst Quadflieg: „Eine Straße gehört niemals anders genannt, als mit der treffenden Bezeichnung.“ *Diskurse um Straßenumbenennungen in Wiesbaden im 20. und 21. Jahrhundert*. In: *Nassauische Annalen* [2025].

³ Zur Debatte um Dietz vgl. Philipp Kratz: *Eine Stadt und die Schuld. Wiesbaden und die NS-Vergangenheit seit 1945*. Göttingen 2019. S. 328–338.

⁴ Vgl. dazu demnächst Quadflieg (wie Anm. 2)

⁵ Dies nachdem 2015 bereits zu 40 Namensgebern öffentlicher Einrichtungen, nicht jedoch Straßen, Gutachten durch eine beauftragte Historikerin erstellt worden waren. Vgl. dazu demnächst ebd. Vgl. Beschluss Nr. 0060 Stadtverordnetenversammlung vom 13.02.2020 zu Vorlage Nr. 20-F-40-0001 „Umgang mit kritischen Namenspaten für Straßenbenennungen – Umbenennung der Pfitzer-Straße“. Diese und alle anderen im folgenden genannten Dokumente des Wiesbadener Stadtparlaments können mit Hilfe der Vorlagennummer im „Politischen Informationssystem Wiesbaden (PIWi)“ online abgerufen werden: <https://piwi.wiesbaden.de/>.

⁶ Vgl. Sitzungsvorlage 20-V-41-0015 „Einrichtung einer Historischen Fachkommission zur Überprüfung nach Personen benannter Verkehrsflächen, Gebäude und Einrichtungen der Landeshauptstadt Wiesbaden“, die am 10.12.2020 gegen die Stimmen von CDU, AfD und einem einzelnen Stadtverordneten angenommen wurde.

⁷ Vgl. Anlagen 1 bis 3 zur Sitzungsvorlage 20-V-41-0015.

⁸ Vgl. Beschluss Nr. 0060 Stadtverordnetenversammlung vom 13.02.2020 zu Vorlage Nr. 20-F-40-0001 „Umgang mit kritischen Namenspaten für Straßenbenennungen – Umbenennung der Pfitzer-Straße“.

⁹ Vgl. Abschlussbericht des Beirats zur Überprüfung Düsseldorf der Straßen- und Platzbenennungen, Düsseldorf 2020, online: https://www.duesseldorf.de/fileadmin/Amt13/presseanhang/2001/200123Abschlussbericht_Strassennamen.pdf (abgerufen am 13.08.2024). Zum Freiburger Projekt: Abschlussbericht der Kommission zur Überprüfung der Freiburger Straßennamen, Anlage 2 zur Drucksache G-16/212, Freiburg 2016, online: https://www.freiburg.de/pb/site/Freiburg/get/params_E-1906631749/1028363/Strassennamen_Abschlussbericht.pdf (abgerufen am 13.08.2024). In Freiburg hatte man die dritte Kategorie noch unterteilt in die beiden Unterkategorien „Straßenname würde heute nicht mehr so gewählt“ und „unbedenklich“.



Dirk Stölper
Namen im öffentlichen Raum
 Abschlussbericht
 der Historischen Fachkommission zur Überprüfung
 nach Personen benannter Verkehrsflächen, Gebäude
 und Einrichtungen der Landeshauptstadt Wiesbaden
 Schriften des Stadtarchivs Wiesbaden
 Band 17

Der Abschlussbericht der Historischen Fachkommission (auch online abrufbar unter www.wiesbaden.de/strassennamen)

weiterer Recherchebedarf ergeben hatte.¹⁰ Während über einige Fälle in der Kommission mehrfach und kontrovers diskutiert wurde, wurden letztendlich alle Entscheidungen einstimmig getroffen.¹¹

Die Sitzungen der Kommission fanden vertraulich statt, allerdings wurde bereits am 16. Juli 2021 die Öffentlichkeit über die Arbeitsaufnahme der Fachkommission durch eine Pressemitteilung unterrichtet. Am 2. Dezember 2021 fand ergänzend ein Pressegespräch statt. Die Fachkommission bot den für Benennungsfragen zuständigen Ortsbeiräten frühzeitig die Möglichkeit zur Information und zum Austausch an. Am 4. Oktober 2021 wurde ein Informationsgespräch für die Ortsbeiräte durchgeführt. Zudem legte die Fachkommission nach zwei Jahren Tätigkeit im Mai 2023 einen Zwischenbericht vor, der vom zuständigen Ausschuss und der Stadtverordnetenversammlung zur Kenntnis genommen wurde.¹²

ERGEBNISSE DER KOMMISSIONSARBEIT, ÖFFENTLICHE REAKTIONEN UND POLITISCHE BERATUNG

Im Rahmen einer Pressekonferenz stellte die Historische Fachkommission gemeinsam mit dem Kulturdezernenten am 8. November 2023 ihren Abschlussbericht öffentlich vor, nachdem sie ihn am Vortag an den Magistrat der Landeshauptstadt Wiesbaden übergeben hatte. Zum selben Zeitpunkt wurde der Abschlussbericht im Internet veröffentlicht.¹³

In 18 der 71 untersuchten Fälle, darunter 17 Männer und eine Frau, empfiehlt die Fachkommission eine Umbenennung.¹⁴ Insgesamt sind 21 benannte Verkehrsflächen und Objekte betroffen, darunter ein Freibad, eine Grünanlage, eine Schule, drei Schutzhütten, eine Mehrzweckhalle und eine Sportanlage sowie 13 Straßen, wobei in zwei Fällen (dem „Opelbad“ und der „Kronprinzenstraße“) statt einer Umbenennung auch die Möglichkeit einer Aberkennung des Namenspatronats durch die Kommission ins Spiel gebracht wurde. Mit 25,4 % der untersuchten Fälle liegt der Anteil der Umbenennungsempfehlungen damit höher als in Düsseldorf

(15,2 % der 79 untersuchten Fälle) und deutlich höher als in Freiburg im Breisgau, wo zwölf der 175 (6,9 %) kategorisierten Fälle zur Umbenennung vorgeschlagen wurden. Es muss allerdings beachtet werden, dass auf Grund des über die Zeit des Nationalsozialismus hinausreichenden Auftrags der Freiburger Kommission dort deutlich mehr Fälle untersucht wurden. In Wiesbaden wurden zudem, anders als in Düsseldorf, ganz offensichtlich unkritische Fälle, wie beispielsweise die Hauptprotagonisten des Attentats auf Hitler am 20. Juli 1944, nicht klassifiziert, sondern bereits im Vorfeld von der Untersuchungsliste gestrichen, wodurch die relative Anzahl der Umbenennungsempfehlungen stieg. Allerdings hat die Wiesbadener Fachkommission prominente Fälle wie den Literaten und Nobelpreisträger Gerhart Hauptmann, den Musiker und Musikfunktionär Richard Strauss oder den Arzt und Forschungsfunktionär Ferdinand Sauerbruch kritischer eingeordnet als andere Kommissionen in anderen Städten zuvor. In zwölf weiteren Fällen empfahl die Wiesbadener Kommission auf Grund der Einordnung in die zweite Fallgruppe eine Kontextualisierung der Namensgebung. Wie nicht anders zu erwarten, wurden die Ergebnisse der Fachkommission kontrovers aufgenommen und diskutiert. Erstaunlich war in diesem Zusammenhang, dass es nur vereinzelte direkte Rückmeldungen an die Geschäftsstelle der Fachkommission gab. Die lokale Presse verhandelte die Diskussion sachlich und lobte in einem Kommentar die gründliche Arbeit der Kommission. Die überlokale Presse griff vor allem eine relativ nüchterne DPA-Meldung auf. Wie auch in den Lokalteilen der Frankfurter Allgemeinen Zeitung und der Frankfurter Rundschau waren die meisten der veröffentlichten Leserbriefe hingegen kritisch.¹⁵

Die öffentliche Auseinandersetzung spielte sich allerdings vor allem in den Sozialen Medien ab, wo etwa in den Diskussionsbereichen der Facebook-Seite der Stadt sowie verschiedener Info-Facebookseiten zum Stadtleben mit redaktionellen Inhalten mehrere Hundert Wortmeldungen innerhalb weniger Tage zusammenkamen. Ablehnende Meinungsäußerungen machten den weitüberwiegenden Teil der Kommentare aus, wobei jedoch nicht auf den Inhalt des Abschlussberichts Bezug genommen wurde. Zentrale Kritikpunkte waren die mit rund 245.000 Euro als zu hoch empfundenen Kosten für den gesamten Untersuchungsprozess, der Einwand, es gäbe „wichtigere politische Probleme“ sowie der pauschale Vorwurf, der Prozess sei politisch von linker respektive von grüner Seite gesteuert gewesen. Hinzu kamen zahlreiche mehr oder weniger kreative Beschimpfungen und Verunglimpfungen, die zum Teil auch auf falschen Aussagen zum angeblichen Inhalt des Abschlussberichts fußten, wenn etwa behauptet wurde, die Kommission habe die „Umbenennung des Freibades Opelbad in Nerobergbad beschlossen“. Tatsächlich beschloss die Kommission nicht über Umbenennungen, machte keine Vorschläge für neue Namen und hatte zudem im Fall des Opelbades einen Entzug des Namenspatronats für den Unternehmer und SS-Förderer Wilhelm von Opel, nicht jedoch zwangsläufig die Umbenennung des Bades empfohlen.

Während diese emotionalen Reaktionen im Netz nach wenigen Tagen abebbten, erhielten das Stadtarchiv und andere Stellen der Stadtverwaltung sowie der Oberbürgermeister und der Kulturdezernent Anfragen und Zuschriften einzel-



Pressekonferenz zur Vorstellung des Abschlussberichts der historischen Fachkommission am 08.11.2023 (StadtA WI, Best. Foo6 Nr. 3; Foto: Jörg Halisch)

ner Bürger und Interessensverbände, sowohl aus Wiesbaden als auch darüber hinaus. In zwei Fällen intervenierten zudem Familienangehörige öffentlich gegen die Umbenennung.¹⁶ Die nächste Stufe der öffentlichen Debatte findet zurzeit in den Ortsbeiräten statt. Erwartungsgemäß und gewollt verlaufen die Debatten in den sozial und politisch sehr unterschiedlich strukturierten Stadtbezirken unterschiedlich schnell und uneinheitlich im Vorgehen. Bis zum Beginn der politischen Sommerpause 2024 und mithin zum Redaktionsschluss dieses Textes hatten von den neun betroffenen Stadtbezirken drei bereits abschließend über die Umbenennungen beschlossen. In allen Fällen folgten die Ortsbeiräte den Empfehlungen der Fachkommission. Dies betrifft die Umbenennung von fünf Verkehrsflächen sowie den Entzug des Namenspatronats im Falle der „Kronprinzenstraße“. Zudem benannte die städtische ESWE Versorgungs AG eine nach Christian Bücher benannte Mehrzweckhalle um. Die übrigen betroffenen Ortsbeiräte haben Arbeitsgruppen gebildet, um die Entscheidung über die übrigen Umbenennungen vorzubereiten. Mehrfach fanden für Anwohner Informationsveranstaltungen statt, zum Teil stehen diese noch aus. Die Stadt hat zudem über die Folgen im Fall einer Umbenennung informiert und versucht, die Sorge der Anwohner vor hohen Kosten zu nehmen.¹⁷ Das Projekt, die politischen Umbenennungsentscheidungen und die betroffenen historischen Personen werden zudem im digitalen Stadtlexikon Wiesbaden¹⁸ und in der Applikation „Geschichte Wiesbadener Straßennamen“¹⁹, die Teil des digitalen Geoportals der Stadt ist, dokumentiert. Über die physische Form der Kontextualisierung von Straßennamen ist hingegen noch nicht entschieden worden.

FAZIT: ERFAHRUNGEN MIT DEM „WIESBADENER MODELL“

Der Kommissionsvorsitzende Christoph Cornelißen sprach gegenüber der Presse von einem „Wiesbadener Modell“ bei der Untersuchung der Straßennamen.²⁰ Für dieses sind

folgende Eigenschaften des Wiesbadener Vorgehens konstituierend. Erstens die Zweistufigkeit, bei der die wissenschaftliche Fachuntersuchung und Debatte in der ersten Stufe von der öffentlichen und politischen in der zweiten Stufe getrennt ist. Statt eines gemischten Beirats, wie in anderen Städten, wurde ein reines Expertengremium berufen, dessen Ergebnisse erst im zweiten Schritt politisch bewertet werden. Zudem wurde das Expertengremium ausschließlich mit auswärtigen Historikerinnen und Historikern besetzt, ein Umstand, der der fehlenden Universität in der hessischen Landeshauptstadt geschuldet ist. Zugleich trug diese Entscheidung, auf die sonst nicht unübliche Beteiligung historisch interessierter lokaler Gruppen in der ersten Runde zu verzichten und diese erst in der zweiten, der öffentlich-politischen Debattenrunde einzubinden, die Versachlichung und die Vertraulichkeit der wissenschaftlichen Beratung in der ersten Projektstufe. Dies schloss natürlich keineswegs die Beteiligung etwa von Heimat- und Geschichtsvereinen bei der Recherche zu den einzelnen Fällen ein, vermied aber eine Diskussion darüber, welche lokalen Druckgruppen in einem „gemischten Beirat“ hätten beteiligt werden müssen. Weitere wichtige Eigenschaften des Wiesbadener Vorgehens waren die professionelle Aufstellung der Zuarbeit durch

¹⁰ Zum methodischen Vorgehen vgl. Dirk Stolper: Namen im öffentlichen Raum. Abschlussbericht der Historischen Fachkommission zur Überprüfung nach Personen benannter Verkehrsflächen, Gebäude und Einrichtungen der Landeshauptstadt Wiesbaden. Hg. v. Peter Quadflieg und Katherine Lukat (Schriften des Stadtarchivs Wiesbaden. Bd. 17), Wiesbaden 2023, S. 21 ff. Der Band ist über den Buchhandel beziehbar und digital kostenlos abrufbar über die Seite www.wiesbaden.de/strassennamen (abgerufen am 14.08.2024).

¹¹ Über die Ergebnisse wurden Protokolle gefertigt. Vgl. Stadtarchiv Wiesbaden WI/3-4107 Nr. 127.

¹² Vgl. Zwischenbericht der Historischen Fachkommission zur Überprüfung nach Personen benannter Verkehrsflächen, Gebäude und Einrichtungen der Landeshauptstadt Wiesbaden nach zwei Jahren kontinuierlicher Arbeit, Sitzungsvorlage 23-V-41-0005.

¹³ Vgl. Stolper (wie Anm. 10).

¹⁴ Betroffen waren folgende 18 Personen: Die Mediziner bzw. Naturwissenschaftler Gerhardt Katsch (1887–1961), Heinrich Pette (1887–1964), Ernst Ferdinand Sauerbruch (1875–1951) und Jonas Schmidt (1885–1958), die Unternehmer bzw. Wirtschaftsfunktionäre Adolf Todt (1886–1960), Friedrich Theodor Overbeck (1897–1972), Wilhelm von Opel (1871–1948), Adam Herbert (1887–1976), Christian Bücher (1878–1949) und Alexander von Engelberg (1894–1960), die Künstler Rudolf Dietz (1863–1942), Karl Elmendorff (1891–1962), Gerhart Hauptmann (1862–1946) und Richard Strauss (1864–1949), die Angehörigen der Hohenzollernndynastie Kronprinz Wilhelm (1882–1951) und Viktoria Luise von Preußen (1892–1980) sowie der Sportfunktionär Otto Schmelzeisen (1892–1983) und der ehemaligen NSDAP-Oberbürgermeister Alfred Schulte (1872–1957). Zur Einordnung der Personen vgl. demnächst Quadflieg (wie Anm. 2) und Stolper (wie Anm. 10).

¹⁵ Ein Pressespiegel findet sich in Stadtarchiv Wiesbaden, WI/3-4107 Nr. 128.

¹⁶ Reaktionen auf die Veröffentlichung des Abschlussberichts finden sich in Stadtarchiv Wiesbaden, WI/3-4107 Nr. 129. Es handelte sich um die Fälle Opel und Engelberg.

¹⁷ Vgl. Nils Lünser: Neuer Name: Was passiert, wenn eine Straße umbenannt wird. In: Wiesbadener Kurier, 15.02.2024, online: <https://www.wiesbadener-kurier.de/lokales/wiesbaden/stadt-wiesbaden/neuer-name-was-passiert-wenn-eine-strasse-umbenannt-wird-3286602> (abgerufen am 16.08.2024).

¹⁸ Vgl. www.wiesbaden.de/stadtlexikon (abgerufen am 16.08.2024).

¹⁹ Vgl. https://geoportal.wiesbaden.de/kartenwerk/application/Geschichte_Wiesbadener_Strassennamen (abgerufen am 16.08.2024).

²⁰ Vgl. beispielsweise Diana Unkart: Historikerkommission rät zu neuem Namen für Opelbad. In: Frankfurter Rundschau, 08.11.2023, online: <https://www.fr.de/rhein-main/wiesbaden/wiesbaden-historikerkommission-raet-zu-neuem-namen-fuer-opelbad-92663083.html> (abgerufen am 16.08.2024).

eine zusätzliche wissenschaftliche Kraft beim Stadtarchiv, die begleitende Öffentlichkeitsarbeit, die Vermittlung der Ergebnisse auf verschiedenen Kanälen sowie eine bereits zu Beginn der Kommissionsarbeit in Geschäftsordnung und Kriterienkatalog klar festgelegte Vorgehensweise. Zwei hessische Großstädte haben sich bereits über die Wiesbade-

ner Vorgehensweise erkundigt und angekündigt, zumindest Teile der öffentlich zugänglichen Organisationsprinzipien zu übernehmen. Für weitere Informationen steht das Stadtarchiv Wiesbaden auch anderen Kolleginnen und Kollegen gerne zur Verfügung.

Peter Quadflieg, Wiesbaden

EIN NEUES URTEIL ZUM VERHÄLTNIS VON WISSENSCHAFTSFREIHEIT VERSUS POSTMORTALEM PERSÖNLICHKEITSSCHUTZ¹

Das Archivrecht ist ständiger Begleiter der Archivarbeit geworden. Entsprechend gewachsen ist die archivrechtliche Literatur². Es ist vor allem eine Diskussion um die Rechtsnormen und deren Umsetzung in der Archivpraxis, während Gerichtsentscheidungen selten sind. Insofern verdient ein neueres Urteil auch über den aktuellen Streitfall Beachtung, und zwar gerade weil er nicht über die erste Instanz hinauskam, und somit die dabei vorgenommene Rechtsabwägung auch der unterlegenen Partei akzeptiert wurde.

1. STREITFALL

Es handelt sich um folgenden Fall. Im Oktober 2020 publizierte die katholische Theologin und Historikerin Dr. Alexandra v. Teuffenbach³, die auch den vatikanischen Archivkurs absolviert hat, ein Buch über das ehem. Mitglied der Schönstätter Marienschwestern, Georgia Wagner (1905–1989)⁴, in dem sie auf der Basis von Briefen von Sr. Wagner und anderen Schwestern dem Ordensgründer P. Josef Kentenich (1885–1968)⁵ geistlichen, emotional-psychischen und körperlich-sexuellen Machtmissbrauch vorwarf. Das Buch wurde als Wortmeldung in dem seit 1975 laufenden Seligsprechungsprozess für P. Kentenich verstanden und löste eine heftige Kontroverse aus, die vor allem in den katholischen Internetmedien geführt wurde. Zwar wurde in ersten Reaktionen von Schönstattseite eine transparente Aufarbeitung und die Öffnung aller Archive zugesichert, aber dann ließ das Säkularinstitut der Schön-

stätter Marienschwestern am 24. November 2020 der Herausgeberin eine Abmahnung zugehen und erhob Klage vor dem Landgericht Berlin auf Erlass einer einstweiligen Verfügung zur Zurückziehung des Buches und Leistung von Schadensersatz. Der Prozess dauerte mit mehr als einem Jahr für ein Eilverfahren recht lange, denn das Urteil wurde erst am 16. September 2021 gesprochen und erlangte am 29. Dezember 2021 Rechtskraft⁶. Das Urteil hatte auch Folgen, indem der für den Seligsprechungsprozess von P. Kentenich zuständige Bischof von Trier mit Dekret vom 3. Mai 2022 das Verfahren bis zu Klärung der Vorwürfe vorläufig aussetzte, was den kompletten Verzicht auf weitere Aktivitäten des Bistums in der Sache bedeutete.

2. DAS URTEIL⁷

Im Folgenden seien die hauptsächlichen Entscheidungsgründe referiert:

Nicht streitgegenständlich war die Beachtung der archivrechtlichen Schutzfristen. Eine solche Klage hätte sich allerdings gegen das Archiv, das die Dokumente vorgelegt hatte, richten müssen. Beklagte wäre also die Pallottinerprovinz gewesen. Freilich wäre dadurch das historisch angespannte und nur mühsam wiederhergestellte Verhältnis der Pallottinern zu den Schönstätttern erneut belastet worden. Außerdem ist unklar, inwieweit von Dritten vor staatlichen Gerichten gegen allfällige Verletzungen von innerkirchlichen (KAO) und zumal von ordensspezifischen Archivregelungen

(KAO-O) geklagt werden kann, denn Ordensarchive sind rechtlich Privatarhive und einklagbar ist nur die Verletzung staatlicher Rechtsnormen. Da aber in diesem Punkt keine Klage erhoben wurde, hatte das Gericht auch keine Veranlassung, darauf einzugehen.

Die Klage griff vielmehr die Publikation an, wo der postmortale Persönlichkeitsschutz unabhängig von den archivischen Spezialgesetzen gilt. Grundlage ist der grundgesetzliche Schutz der Persönlichkeit, der sich zunächst auf lebende Personen bezieht, aber auch über den Tod hinaus noch gelten kann, wenn auch in einer sich fortwährend abschwächenden Intensität. Feste Fristen gibt es dafür nicht, so dass Raum für Klagen im Einzelfall besteht. Hier setzte nun die Klage an und wurde im Urteil wie folgt entschieden:

1. Wichtigster Punkt der Klage war, dass der Kläger (also das Säkularinstitut) beanspruchte, als **Wahrnehmungsberechtigter** anerkannt zu werden und damit ein Klagerecht zur **Wahrung der postmortalen Persönlichkeitsrechte** von P. Kentenich zu haben.

Ein solches Klagerecht steht nicht den Erben, sondern nur den Angehörigen zu, die direkt von einer solchen Verletzung betroffen sind. Das sind in der Regel die unmittelbaren Nachkommen der ersten Generation, nicht dagegen juristische Personen wie eine Schwesterngemeinschaft. Ausnahmen davon sind möglich, müssen aber hohen Anforderungen genügen, wie z. B. einer formellen notariellen oder testamentarischen Beauftragung. Eine solche lag aber nicht vor. Dass sich die Marienschwestern selbst als „Familie“ bezeichnen, konnte schon wegen der großen Zahl der in Frage kommenden Personen nicht durchgreifen. Wenn der Kläger auf die „Vater-Stellung“ von Kentenich verwies, so war dies für das Gericht schon durch die Tatsache des Prozesses nicht unumstritten in der Gemeinschaft. Es fehlte nicht nur an einem Testament von Kentenich, sondern insbesondere an einer formellen Übertragung der Wahrnehmung seiner Persönlichkeitsrechte nach seinem Tod. Dies wäre aber erforderlich gewesen. Es ist also festzuhalten: Eine geistliche Gemeinschaft ist nicht automatisch zur Wahrnehmung der postmortalen Persönlichkeitsrechte ihres Gründers berechtigt, auch wenn sie sich dazu berufen fühlt. Vorliegend billigte die Kammer dem Kläger deshalb kein Klagerecht zu.

2. Insofern brauchte die Kammer die Frage, inwieweit durch das Buch der **postmortale Persönlichkeitsschutz** von Kentenich verletzt worden wäre, nicht abschließend zu erörtern. Sie gab aber eine Zusammenfassung des Standes der Rechtsprechung.

Der Postmortale Persönlichkeitsschutz kann nicht abgeleitet werden aus dem Grundrecht auf informationelle Selbstbestimmung nach Art. 2 GG (*Jeder hat das Recht auf die freie Entfaltung seiner Persönlichkeit*). Denn dies kann nur einer lebenden Person zustehen sowie mit Einschränkungen seinen unmittelbaren Angehörigen, sofern sie selbst davon betroffen sind.

Der Postmortale Persönlichkeitsschutz basiert vielmehr auf Art. 1 GG (*Die Würde des Menschen ist unantastbar*) und beinhaltet einen doppelten Schutz: a. Schutz gegen Herabwürdigungen in einer die Menschenwürde verletzenden Weise und b. Schutz gegen schwerwiegende Beeinträchtigungen des sittlichen, personalen und sozialen Geltungswertes, den eine Person durch ihre eigene Lebensleistung erworben hat.

Dabei gilt, dass eine solche Verletzung sehr schwerwiegend sein muss. So genügt z. B. nicht die bloße Infragestellung des erworbenen Geltungsanspruches.

Das aber sah die Kammer nicht als gegeben an. Dabei griff das Gericht den vom Kläger selbst vorgetragenen Punkt auf, dass der Vorwurf des körperlich-sexuellen Missbrauchs gegen P. Kentenich gar nicht neu sei, sondern in ähnlicher Form schon bei der päpstlichen Visitation der Schwesterngemeinschaft Anfang der 1950er Jahre erhoben worden war. Damit waren für das Gericht diese Vorwürfe aber bereits Teil des Lebensbildes von P. Kentenich, so dass sein persönlicher Geltungswert dadurch nicht mehr beeinträchtigt werden konnte. Außerdem liegt ein Abstand von mehr als 52 Jahren seit dem Tod von P. Kentenich jenseits aller staatlichen und kirchlichen Schutzfristen und gilt damit nur in allgemeiner Form als Schutz gegen Beleidigungen und bewusste Diffamierungen, was die Kammer aber nicht als gegeben ansah.

3. Der Persönlichkeitsschutz konkurriert mit der **Wissenschaftsfreiheit** nach Art. 5 Abs. 3 GG. Auch hier referierte die Kammer den Stand der Rechtsprechung. Grundlegend ist die Definition von Wissenschaft aus der ersten Hochschul-Entscheidung des Bundesverfassungsgerichtes von 1973. Danach ist wissenschaftliche Tätigkeit *„alles, was nach Inhalt und Form als ernsthaft planmäßiger Versuch zur Ermittlung der Wahrheit anzusehen ist“*. Das klingt etwas unbestimmt. Aber Wissenschaft ist gerade nicht auf eine bestimmte Methode festgelegt, sie ist nicht an bestimmte Qualifikationen gebunden. Sie muss nur die wissenschaftliche Diskussion zumindest fördern und kann dazu auch Mindermeinungen vertreten, Einseitiges vorstellen und auch Thesen vertreten, die sich als irrig oder fehlerhaft erweisen. Alles dies ist von der wissenschaftlichen Meinungsfreiheit gedeckt. Gegen ein Buch, das seine Quellen nicht nur nennt, sondern sogar publiziert und seine Vorwürfe sachlich begründet, bestand damit für das Gericht kein Unterlassungsanspruch. Damit waren die Hauptpunkte abgehandelt und es blieb nur noch übrig, die Liste der Anträge der Kläger abzuarbeiten. Interessant ist der Versuch des Klägers, mit Anleihen beim Presserecht und beim Strafrecht gegen das Buch vorzugehen.

¹ Der Aufsatz geht auf einen Vortrag zurück, den der Verfasser am 8. April 2024 auf der Jahrestagung der Arbeitsgemeinschaft der Ordensarchive in Fürstenried (München) gehalten hat. Für freundliche Hinweise danke ich Herrn Kollegen Dr. jur. Jost Hausmann, Vallendar, und Herrn Richter Joachim Hennig, Koblenz.

² Archivrecht für die Praxis. Ein Handbuch. Hg. v. Irmgard Christa Becker u. Clemens Rehm. München 2017. - Jost Hausmann: Archivrecht. Ein Leitfad. Frankfurt a. M., Berlin 2. Aufl. 2021.

³ Zu biobibliographischen Angaben über die Verfasserin vgl. den biobibliographischen Artikel in Wikipedia

⁴ „Vater darf das!“ Eine Archivdokumentation. Sr. M. Georgina Wagner und andere missbrauchte Schönstatter Marienschwestern. Zusammengefasst und eingeführt von Alexandra von Teuffenbach. Nordhausen (Bautz) 2020.

⁵ Biographischer Artikel in Wikipedia, der auch über den hier behandelten Streitfall berichtet.

⁶ Landgericht Berlin 27 O 443/20, publiziert: <https://www.Schoenstatt.org/wp-content/uploads/2022/01/Urteil.pdf>.

⁷ Das Urteil zu den Schönstadtswestern und Pater Kentenich wird erscheinen in: Entscheidungen in Kirchensachen 78 (2025). Hg. v. Manfred Baldus, Stefan Muckel u. Norbert Diel.



4. Die Klage rügte, dass das Buch gegen die Grundsätze **öffentlicher Verdachtsäußerungen** verstoße, die im **Presserecht** dem Betroffenen ein Recht zur Stellungnahme einräumen. Dies sei, so der Vortrag des Klägers, hier aber nicht geschehen. Das Buch würde nur subjektive Äußerungen ohne objektiven Wert verbreiten. Die aussagenden Schwestern wären alle psychisch krank und ihre Aussagen beträfen keine realen Vorgänge. Trotzdem sei ihnen die Herausgeberin einfach gefolgt, ohne dass man den Marienschwestern als Wahrnehmungsberechtigten Gelegenheit zu einer Stellungnahme gegeben hätte. Dieser Versuch musste aber scheitern, da es sich vorliegend nicht um eine journalistische, sondern um eine wissenschaftliche Publikation handelt. Außerdem ist die Einholung von Gegenmeinungen in der Wissenschaft nicht üblich. Andersdenkenden steht die Diskussion offen.

5. In ähnlicher Weise wurde der Gebrauch der Begriffe „Täter“ (für P. Kentenich) und „Opfer“ (für die Schwestern) vom Kläger angegriffen, da damit der **Vorwurf einer strafrechtlichen Tat** verbunden sei. Der Vorwurf würde von der Verfasserin aber nicht umfassend geklärt und Argumente, die für eine Unschuld von P. Kentenich sprächen, würden nicht in Erwägung gezogen. Das Gericht stellte aber fest, dass der Vorwurf einer strafrechtlichen Handlung von der Verfasserin gar nicht erhoben würde, diese vielmehr sorgfältig zwischen damaliger und heutiger Beurteilung solcher Missbrauchsfälle differenziere. Die Begriffe Täter und Opfer würden so immer in einem sachlichen Zusammenhang mit Rückbezug auf die Quellen verwendet. „Allein die Verwendung der Begriffe“

fe“, so das Gericht, „vermag daher weder Kentenichs allgemeinen Achtungsanspruch zu verletzen noch dessen Lebensbild schwerwiegend zu verzerren“.

6. Am hartnäckigsten wurde der Begriff des **körperlich-sexuellen Missbrauchs** angegriffen. Der Gebrauch des Begriffes wurde von dem Kläger zurückgewiesen, weil die Berichte ja nur subjektive Schilderungen enthielten und keine objektiven Taten nachwiesen. Das Gericht stellte dagegen fest, dass die Verfasserin in Anlehnung an die heutige wissenschaftliche Diskussion sorgfältig die Begriffe des emotional-psychischen, des geistlichen und des körperlich-sexuellen Missbrauchs umschrieben habe und auch auf die unterschiedliche zeitgenössische Bewertung in den Jahren um 1950 und heute eingegangen sei. Außerdem würde durch die Verbindung mit den abgedruckten Quellen immer angegeben, worauf sich die Begriffe bezögen. So sah das Gericht hier eine zulässige wissenschaftliche Meinungsäußerung. Dabei störte den Kläger besonders, dass zur moralischen Wertung der damaligen Handlungen heutige Standards herangezogen würden. Vielmehr wurde gefordert, dass eine Bewertung allenfalls an den Maßstäben der damaligen Zeit zulässig sei. Dazu erklärte das Gericht lapidarisch: „Ein solches Verständnis machte indes jegliche Aufarbeitung historischer Ereignisse schlicht unmöglich und ist mit den grundgesetzlich verankerten Freiheiten, insbesondere der hier relevanten Meinungs- und Wissenschaftsfreiheit des Grundgesetzes nicht zu vereinbaren“. Das Urteil lässt also nichts an Deutlichkeit zu wünschen und zeigt, dass das Grundrecht der wissenschaftlichen Meinungsfreiheit nicht nur als Norm existiert, sondern auch im Konfliktfall durchsetzbar ist. Es stärkt somit die Position der Wissenschaft gegenüber Traditionspflegern.

3. RECHTSSYSTEMATISCHE EINORDNUNG

Versuchen wir zusammenfassend noch eine rechtssystematische Einordnung des Urteils⁸. Der postmortale Persönlichkeitsschutz besteht aus zwei unterschiedlichen Rechtsinstituten.

Er folgt einmal aus dem **Grundrecht auf informationelle Selbstbestimmung** nach Art. 2,1 GG: „Jeder hat das Recht auf die freie Entfaltung seiner Persönlichkeit, soweit er nicht die Rechte anderer verletzt und nicht gegen die verfassungsmäßige Ordnung und das Sittengesetz verstößt.“

Das ist ein **individuelles Recht**, das jedem Staatsbürger aus eigenem Recht zukommt. Es umfasst die Kontrolle über die eigenen Daten und das eigene Privatleben, so dass diese Daten nicht gespeichert und personengebundene Akten nicht kommuniziert werden dürfen. Als Individualrecht bedarf es aber immer eines Trägers, und so erlischt das Recht grundsätzlich mit dem Tod des Trägers. Es gibt also keinen grundrechtlichen Schutz des Andenkens eines Verstorbenen. Nur dann kann daraus ein Schutz gegen die Offenlegung persönlicher Lebensumstände eines Verstorbenen abgeleitet werden, wenn dadurch zugleich überlebende direkte Angehörige in der Entfaltung ihrer eigenen Persönlichkeitsrechte verletzt werden. Dieser Rechtsschutz nimmt aber mit der Zeit ab. Öffentlich-rechtlich gilt heute in der Literatur auf der Basis des Bundesarchivgesetzes und den Archivgesetzen von 15 der 16 Bundesländer eine Schutzfrist von 10 Jahren.

Der postmortale Persönlichkeitsschutz folgt zum anderen aus dem Grundrecht der Menschenwürde nach Art. 1.1. GG: *Die Würde des Menschen ist unantastbar*. Dies ist ein objektiv-rechtliches Verfassungsprinzip. Es ist die Pflicht des Staates, die Menschenwürde zu achten und für ihre Wahrung zu sorgen. Damit ist es unabhängig von der Existenz eines Rechtssubjektes und gilt absolut. Es verbürgt Schutz vor Beleidigungen und bewussten Diffamierungen auch über den Tod hinaus. Das soll den Verstorbenen nun freilich nicht von aller Kritik freistellen. Vielmehr soll gerade eine Diskussion und auch eine kritische Diskussion ermöglicht werden. Der postmortale Schutz verhindert nicht das Bekanntwerden von Lebensumständen von Verstorbenen, die ihr Andenken beeinträchtigen können, er sorgt nur dafür, dass dies in sachlicher Form geschieht. Insofern plädiert die Literatur dafür, dass wahrheitsgemäße Angaben etwa über Straftaten, NSDAP-Mitgliedschaft, Stasi-Tätigkeit etc. kein Verstoß gegen das Achtungsgebot von Art. 1 GG darstellen. Als Konsequenz daraus ergibt sich, dass zu prüfen wäre, ob der postmortale Persönlichkeitsschutz in Archiven immer

grundgesetzkonform gehandhabt wird, wenn z. B. Gerichtsakten deswegen gesperrt werden, nur weil darin ein Zeuge mit Geburtsdatum oder anderen Bagatelldaten genannt wird⁸. Das könnte einer nicht grundgesetzkonformen Hagiographierung des Andenkens von Verstorbenen oder sogar einer Art Täterschutz Vorschub leisten. Demgegenüber steht die Freiheit der Wissenschaft, die das Grundgesetz nur an die Verfassung bindet, nicht aber an einen etwaigen Gesetzesvorbehalt. Insofern gilt uneingeschränkt: *„Kunst und Wissenschaft, Forschung und Lehre sind frei.“*

Wolfgang Hans Stein, Koblenz

⁸ Bartholomäus Manegold: Archivrecht. Die Archivierungspflicht öffentlicher Stellen und das Archivzugangsrecht des historischen Forschers im Lichte der Forschungsfreiheitsverbürgung des Art. 5 Abs. 3 GG, Berlin 2002, S. 108–129; Bundesarchivgesetz, Handkommentar. Hg. v. Christoph Partsch. Baden-Baden² 2021, S. 70–71, 217–242; Dreier Grundgesetzkommentar. Hg. v. Frauke Brosius-Hersdorf, Bd. 1. Tübingen 4. Aufl. 2023, S. 183–187, 812–817.

⁹ Manegold: Archivrecht, S. 113, 279.

KÖLNER FORUM „JENSEITS DES HYPE – ANWENDUNG VON KI IN ARCHIVEN“



Das Thema Künstliche Intelligenz wird zurzeit nicht nur in der Archivwelt vielfach diskutiert. Der dabei entstehenden Lücke zwischen Hype und Realität nahm sich am 28.08.2024 das erste Kölner Forum, organisiert von Dr. Ulrich Fischer (HASTK-RBA) und Dr. Bettina Joergens (LAV NRW), an und brachte Kolleg*innen aus der Archivwelt mit Expert*innen aus Wissenschaft und Wirtschaft zusammen, um offen über aktuelle Anwendungsfälle und Projekte, Chancen und Fehlschläge zu diskutieren. Entstanden ist das Format in der Nachfolge des Fachbeirats für den Wiederaufbau des Historischen Archivs der Stadt Köln. Dessen fruchtbare Zusammenarbeit sollte, wie Dr. Bettina Schmidt-Czaia (HASTK-RBA) in ihrer Begrüßung erläuterte, an anderer Stelle als Forum für Vernetzung und Diskussion von Expert*innen zu aktuellen, klar definierten Themen fortgesetzt werden.

Um Austausch und Diskussionen zu fördern, waren die einzelnen Beiträge in ihrer Länge bewusst begrenzt. Da auch dieser Bericht einer ähnlichen Limitierung unterliegt, werden Vorträge und Diskussionen im Folgenden zusammengefasst.

Die fachlichen und historischen Grundlagen zu KI im Allgemeinen und im Archivwesen im Besonderen erläuterten Dr. Bettina Joergens (LAV NRW), Tobias Krafft (Universität Kaiserslautern/PD GmbH/AI Trust) und Elena Williams (FH Potsdam/Stadtarchiv Stockach) in ihren jeweiligen Beiträgen. Die Tagung sollte, so Joergens, dazu beitragen, ein besseres Verständnis für das Phänomen KI in der Archivwelt zu entwickeln und den Lernprozess der Archive durch Einblick in verschiedene Projekte voranzubringen. Zunächst gab Krafft einen konzisen Einblick in grundlegende Begrifflichkeiten des Themenraumes. KI definierte er dabei als „Alle von einem Computer ausgeführten Tätigkeiten, die bisher von Menschen durchgeführt werden und mit menschlicher Intelligenz assoziiert werden.“ Der verkürzende Begriff wird hier Einfachheit halber verwendet. Anders als bei der einfachen Ausführung von Algorithmen bedeutet die Nutzung von KI auch, dass Systeme lernen und Fehler vorkommen. Maschinelles Lernen sollte von einem regulatorischen Ökosystem begleitet werden, um KI zielgerichtet und sinnvoll einsetzen zu können. Dass diese Nutzung im deutschsprachigen Archivwesen nur sehr punktuell Realität ist, zeigte Williams. Ihre Umfrage unter Fachkolleg*innen zeichnet ein sehr heterogenes Bild zwischen Potenzialen von KI im Archivwesen und großen realen und gefühlten Hürden wie Personalmangel, fehlender technischer Ausstattung und Wissenslücken. Dass insbesondere Erschließung als Einsatzfeld im Fokus stand, passte gut zu den auf dem Kölner Forum präsentierten Projekten.

Den Einsatz von KI bei der Erschließung von Texten und Formularen thematisieren Dr. Andreas Neuburger (LABW) mit Blick auf die Tiefenerschließung von Wiedergutmachungsunterlagen im Rahmen des Projekts „Transformation der Wiedergutmachung“, Professor Dr. Gernot Fink (TU Dortmund) und Dr. Astrid Küntzel (LAV NRW) hinsichtlich der Nutzbarmachung von Volkszählungsunterlagen von 1950 sowie Jasmin Hartmann und Dr. Ruth von dem Bussche (beide KPF NRW) bezüglich der Extraktion von relevanten Informationen für die Provenienzforschung aus unterschiedlichen Archivbeständen. Insbesondere der Einsatz von Large Language Models in Kombination mit OCR und Layoutanalyse sowie multimodaler Sprachmodelle haben, darin waren sich die Beiträge einig, das Potenzial, Erschließungstiefen und -formen möglich zu machen, die heute ob des manuellen Aufwands undenkbar sind. Im Bereich der Provenienzforschung zeigte sich bereits heute praktischer Nutzen, so lassen sich über die Werkzeuge OCR und Sprachmodell Werke identifizieren und in einer Suchmaschine finden, die sonst wahrscheinlich verborgen bleiben würden.

Mit der Nutzung freier KI-Modelle zur Erschließung von Bild und Text befasste sich zudem Martin Vogel (NLA), während sich Miriam Eberlein (StA Heilbronn) sowie Florian Leitner (Österr. BDA) jeweils ganz Projekten der KI-basierten Bilderkennung und -erschließung widmeten. Den Sprung zu bewegten Bildern vollzog dann Dr. Marco Chavarria (VITEC GmbH).

Insbesondere der Beitrag von Vogel zeigte auf, wie viel schon heute mit frei verfügbaren Modellen möglich ist. Dabei erscheint als wichtigste Lektion, dass die Anwender*innen ihre Daten genau kennen müssen, um entscheiden – und sinnvoll testen – zu können, welche Modelle an welcher Stelle des Prozesses hilfreich sind. Wie eng Nutzen und Fehlerhaftigkeit zusammenstehen und gedacht werden müssen, zeigte sich sowohl am produktiven Heilbronner Projekt als auch an den Arbeiten in Österreich. Beide KI-Anwendungen sind unzweifelhaft sehr nützlich, benötigen aber noch manuelle Nacharbeit und Kontrolle. Wie die Archivwelt damit umgehen möchte, bleibt eine offene Frage.

Eine Perspektive über die Erschließung hinaus demonstrieren dagegen Dr. Marc von den Linden (MusterFabrik Berlin) und Hanna Kannengießer (HASTK-RBA), die aufzeigten, welchen Anteil KI an der Wiederausammensetzung zerstörter Archivalien des Historischen Archivs der Stadt Köln haben kann. Während die Hauptarbeit beim Zusammenpuzzeln der Fragmente in Anwendungen klassischer mathematischer Methoden für Bildverarbeitung und Mustererkennung liegen, kann KI die Merkmalsextraktion unterstützen und

damit die Formierung in Folge des Einsturzes zerteilter Archivalien. Teilnehmer*innen konnten dabei zusehen, wie sich zerstört geglaubte Dokumente wieder zusammensetzen. Deutlich wurde, wie wichtig eine archivfachliche Diskussion über Bedarfe, Ziele, Fehlertoleranzen und Erfahrungen ist. Aus archivfachlicher Sicht hat KI Authentizitäts- und Qualitätsprobleme, aber auch große Potenziale. Hier braucht es offenen, kritischen Austausch und maximale Transparenz, welche Ergebnisse in welcher Weise von KI generiert oder unterstützt wurden. Darüber hinaus gilt es zu diskutieren, an welchen Stellen tatsächlich KI und an welchen andere Methoden sinnvoll anzuwenden sind, wie auch der Kölner Beitrag illustrierte.

Wie sensibel der Einsatz von KI insbesondere für Vertrauensinstitutionen sein kann, diskutierte schließlich u. a. die Abschlussrunde unter Leitung von Dr. Frank Bischoff (LAV NRW). Selbst bei Trefferquoten von 99 % können angesichts

der in Archiven vorliegenden Datenmassen schnell tausende Fehler ihren Weg in die archivischen Informationssysteme finden. Eine große Baustelle bleibt hier die Fähigkeit der Systeme, die eigene Fehlerhaftigkeit einzuschätzen. Für die Anwender*innen ist dagegen Datenkompetenz ein entscheidender Faktor. Diese sollte, so Professor Dr. Karin Schwarz (FH Potsdam), bereits in der Ausbildung großen Raum erhalten.

Bei allen Schwierigkeiten blieb unstrittig, dass KI-Anwendungen Potenzial haben, archivische Prozesse vielfältig zu unterstützen. In diesen Fragen wird noch sehr viel Entwicklung und Training nötig sein – und zwar nicht nur für die eingesetzten Modelle, sondern auch für die Archive. Die vorgestellten Projekte zeigten dabei deutlich, dass dieser Einsatz auf Dauer erheblich lohnen kann.

Hanna Kannengießer, Köln/Franziska Klein, Duisburg

ENTSÄUERUNG UND KALTLAGERUNG. BESTANDSAUFNAHME UND PERSPEKTIVEN DES ORIGINALERHALTS

WORKSHOP AM LANDESARCHIV SACHSEN- ANHALT VOM 18. BIS 19. JUNI 2024

Ist der Originalerhalt schriftlichen Kulturguts eine Fiktion, wie der Präsident des Bundesarchivs, Michael Hollmann, auf einer Berliner Tagung im Juni 2023 konstatierte? Die fachlichen Debatten um die Möglichkeiten und Grenzen des Originalerhalts und die Digitalisierung als Alternative des Informationserhalts haben sich in jüngster Zeit zugespitzt. Zugleich zwingen Staatsverschuldung und Haushaltskürzungen die Archive zur genauen Prüfung und Abwägung ihrer Optionen in der Bestandserhaltung – wie auch zu deren Vermittlung gegenüber Politik und Öffentlichkeit.

Vor diesem Hintergrund lud das Landesarchiv Sachsen-Anhalt Expertinnen und Experten aus Bibliotheken und Archiven und aus der Papierforschung und -technik sowie der Bauphysik zu einer Standortbestimmung nach Magdeburg ein. Der Hauptakzent des in Zusammenarbeit mit Johannes Kistenich-Zerfass (Hessisches Landesarchiv, Marburg) und Sven Kriese (Geheimes Staatsarchiv PK, Berlin) konzipierten und organisierten Workshops lag auf der Massensäuerung, zu der Ute Henniges (Staatliche Akademie der Bildenden Künste, Stuttgart) Ergebnisse retrospektiver Datenanalysen referierte und Manuela Fiedler (Papiertech-

nische Stiftung, Heidenau) Normen und Messmethoden erläuterte.

Dabei konnte Henniges konstatieren, dass bisherige Prüfungen nach unterschiedlichen Kriterien und unter den Bedingungen sowohl künstlicher wie auch natürlicher Alterung die Wirksamkeit von Entsäuerungsmaßnahmen zweifelsfrei nachweisen konnten. Zugleich stellte sie fest, dass die Ergebnisse fast durchweg anhand von Bibliotheksgut gewonnen wurden, sie deutliche Unterschiede zwischen den Entsäuerungsverfahren aufzeigen und dass die überprüften Verfahren teilweise obsolet sind. Fiedler betonte bei ihrer Darlegung der diversen verfügbaren Prüfmethode den Normierungsfortschritt, der 2018 mit der Verabschiedung der DIN 32701 erreicht wurde, indem nunmehr anhand definierter Testpapiere vergleich- und nachvollziehbar die alkalische Reserve und der pH-Wert im Kaltextrakt bestimmt werden. Daran schlossen sich Erfahrungsberichte aus der Schweizerischen Nationalbibliothek (Agnes Blüher) und der Sächsischen Landes- und Universitätsbibliothek (SLUB, Michael Vogel) sowie für das Hessische Landesarchiv (Johannes Kistenich-Zerfass), die staatlichen Archive Bayerns (Laura

Scherr), das Geheime Staatsarchiv PK (Sven Kriese, Ingrid Kohl) und für die Landesinitiative Substanzerhalt in NRW (LISE, Landesteil Westfalen-Lippe, Gabriele Rothkegel) an. Sie auf einen kurzen gemeinsamen Nenner bringen zu wollen, wäre vermessen, zeigte sich doch eine große Vielfalt der Erfahrungen mit unterschiedlichen Papiereigenschaften, Alterungszuständen und eingesetzten Verfahren. Da in Deutschland aktuell nur zwei Dienstleister nichtwässrige Flüssigverfahren anbieten, ist es, so die einhellige Einschätzung, um so wichtiger, Entsäuerungsergebnisse hinsichtlich der Wirksamkeit wie auch der Nebenwirkungen fortlaufend zu prüfen und mit den Dienstleistern zu erörtern, um die Behandlungsparameter an die konkreten Erfordernisse der eigenen Bestände bestmöglich anzupassen. Gleichwohl sprachen sich die anwesenden Vertreter*innen von Archiven und Bibliotheken, ungeachtet der sich teilweise abzeichnenden Haushaltskürzungen, für die Fortführung der Entsäuerungsprojekte aus.

Dabei spielte eine wichtige Rolle, dass die Bestandsaufnahme zur Kaltlagerung ernüchternd ausfiel. Vergleiche und Querbeziehungen zwischen den beiden Alternativen des Originalerhalts hatte Henrik Otto (Landesarchiv Sachsen-Anhalt, Dessau) einleitend in sehr anschaulicher Weise gezogen und dabei bekannte Faustformeln zur jeweiligen Wirksamkeit rekapituliert und hinterfragt. Zudem konnte er die mitunter postulierten Kostenvorteile der Digitalisierung als Alternative des Informationserhalts widerlegen. In einem zweiten Beitrag stellte er die Ergebnisse einer eigens durchgeführten internationalen Umfrage zu Kaltlagerungserfahrungen vor. Der inhaltliche Rücklauf fiel sehr überschaubar aus, einschlägig waren vor allem die Informationen aus der Dänischen Nationalbibliothek zu der dort praktizierten Kühltagerung in einem Temperaturkorridor von 8–16°C. Daran schloss sich nahtlos der Bericht von Sabine Preuß (Deutsche Nationalbibliothek, Leipzig) zu den klimatischen Anforderungen an den 5. Erweiterungsbau der DNB am Standort Leipzig an. Im Zuge der Planungen hatte sie die Gelegenheit, den hochmodernen Magazinbau der Library

and Archives Canada in Gatineau (Quebec) zu besichtigen, in dem das Biliotheks- und Archivgut größtenteils bei 10°C und 40 % r. F. archiviert wird. Das damit angestrebte Erhaltungsziel von mind. 500 Jahren liegt auch den DNB-Planungen zugrunde. Eine ebensolche Kaltlagerung für ähnlich große Mengen papiergebundenen Kulturguts ist jedoch in Deutschland bislang nirgends realisiert worden und, so die einhellige Einschätzung, angesichts der erheblichen Investitions- und Betriebskosten auf absehbare Zeit auch nicht realistisch. Diese bestätigte auch die Vorstellung der aktuellen Bauplanungen der SLUB Dresden, denn diesen liegt für das Gros des Bibliotheksguts ein Klimakernkorridor von 14–22°C bei 40–55% r. F. zugrunde. Dadurch wird eine erhebliche Kostenersparnis erreicht, die auch die Fortführung der ergänzend für erforderlich erachteten Mengenentsäuerung ermöglichen soll.

Abschließend erläuterte Lars Klemm (Fraunhofer Institut für Bauphysik / IKONYK GmbH, München) anhand zahlreicher Praxisbeispiele die bauphysikalischen Herausforderungen und energetischen Aufwände bei der Klimatisierung von Kulturgutdepots. Auch hier wurde deutlich, dass der Trend wegführt von Kaltlagerungskonzepten zu flexibleren Klimakorridoren, womit zwar Magazinbauten leichter und kostendämpfender realisierbar sind, aber keine alleinige Gewähr für eine hinreichende Eindämmung des endogenen Papierzerfalls geboten wird.

Die sehr engagierten Debatten um die grundsätzlichen Optionen zur Erhaltung des papiergebundenen schriftlichen Kulturguts verdeutlichten auch aktuell bestehende Forschungsbedarfe. Sie mündeten am Ende des Workshops in den Vorschlag, die in mehreren Einrichtungen über teils lange Jahre erhobenen Daten der Qualitätssicherung bei Entsäuerungsmaßnahmen zusammenzuführen, um sie für systematisch-vergleichende Auswertungen offenzulegen und somit Anreize für weiterführende Forschungsansätze zu geben oder Möglichkeiten zur Reflexion der je eigenen Preservation Policy zu bieten.

Ralf Lusiardi, Magdeburg

DAS INTERNATIONALE ARCHIVSYMPOSIUM 2024

Das Internationale Archivsymposium (IAS) wird jährlich mit begrenzter Teilnehmerschaft aus Belgien, Luxemburg, den Niederlanden und den staatlichen und kommunalen Archiven in Nordrhein-Westfalen und Rheinland-Pfalz durchgeführt. Am 16. und 17. Mai 2024 fand das IAS in Arnhem statt und stand mit dem Thema „Sind wir vorbereitet? Zukunftsfähigkeit der Archive“ im Zeichen der tiefgreifenden Transformationsprozesse, mit denen sich Archivar*innen in den vergangenen Jahren teils notgedrungen intensiver befassen mussten und müssen: Digitalität, Klimawandel, und auch Naturkatastrophen.

Dies spiegelte sich schon im Eröffnungsvortrag des ersten Tages wider. Valerie Trouet, wissenschaftliche Direktorin des belgischen Klimazentrums, sprach in ihrem Einführungsvortrag von den vielfältigen Möglichkeiten, die lokale, regionale und internationale Klimageschichte auf Grundlage unterschiedlichster, auch nicht klassisch archivischer Überlieferungen zu untersuchen. Neben klassischen Schriftquellen wie den irischen Annalen kam sie auf Datenbanken mit Daten zur Klimageschichte aus Holzringuntersuchungen, Wetteraufzeichnungen insbesondere zu Hurrikanen und der Quer-Referenzierung mit datierten Schiffswracks und u. a. auch auf Daten zur historischen Variabilität des Jetstreams zu sprechen, die in Bezug zu landwirtschaftlichen und gesellschaftlichen Entwicklungen gesetzt werden können. Im Ergebnis hielt Frau Trouet fest, dass angesichts der Klimaschwankungen und des Klimawandels die aufzufindenden Verbindungen der verschiedenen Datenquellen in den unterschiedlichen „Archiven“ wie z. B. den Baumringen und klassischen Archiven als Einrichtungen mit vornehmlich Schriftquellen einen wesentlichen Beitrag zur Gestaltung der gesellschaftlichen Zukunft leisten können.

Nach kurzer und intensiver Diskussion um diese Perspektive auf die Zukunftsfähigkeit von Archiven folgten die Beiträge von Jonas Feltes, Gilles Regener und Sophia van Hoek. Feltes beschäftigte sich mit der Frage, inwieweit das bundesdeutsche Modell des Kulturgutschutzes zukunfts- und in den kommenden Jahrzehnten noch tragfähig sei. Die Bundessicherungsverfilmung von Archivgut ist ein Projekt, das bereits in den 1960er Jahren startete, allerdings bis vor wenigen Jahren noch ausschließlich analoge Aufnahme- und Entwicklungstechniken zugrunde legte. Doch auch hier hat der digitale Wandel Einzug gehalten. Inzwischen wurde die digitale Ausbelichtung des digitalisierten Archivguts auf große Rollen des 35mm-Mikrofilms erfolgreich umgesetzt. Nichtsdestoweniger ist nach Feltes' Einschätzung nur ein erster Schritt in Richtung der Zukunftsfähigkeit des Bundesprojekts getan. Die Bundessicherungsverfilmung wurde ursprünglich zum Schutz des Kulturguts in bewaffneten Konflikten ins Leben gerufen, steht nun aber zunehmend vor der Herausforderung des Kulturgutschutzes im Falle

massiver Naturkatastrophen und der hybriden Kriegsführung. Insofern werden in den kommenden Jahren neue Wege betrachtet und auf ihre Tragfähigkeit hin erprobt werden müssen. Keramische Datenträger und das Arctic World Archive sind nur zwei der aus seiner Sicht zu betrachtenden Wege.

Regeners Vortrag betrachtete die Gebäudesituation und -planung des neuen Archivbaus für das Luxemburger Nationalarchiv in Esch-Belval. Neben der langwierigen Planungs- und Entwicklungsphase, die u. a. Pandemie- und marktbedingt einige Verzögerungen erfuhr, waren ebenso die Auswahl nachhaltiger und recycelter Materialien, Einhaltung von Green-IT-Vorgaben und Nachhaltigkeitsbedingungen des Luxemburger Staats Faktoren, die den Bau neben den gängigen Fachanforderungen maßgeblich beeinflussten. Am Ende steht ein Archivbau, der neben diesen Vorgaben auch die fachlichen Anforderungen erfüllt, erweiterbar ist, barrierefrei zu erreichen ist und im Ergebnis eine positive Energiebilanz vorweisen soll. Regener betonte abschließend noch einmal, wie komplex, doch nicht minder wichtig die nachhaltige Planung auch der Archive und der archivischen Tätigkeiten in Zukunft sein wird.

Green-IT-Fragen betrachtete vertiefend zum Schluss des ersten Tages Sophia van Hoek, die sich der Frage des ökologischen Fußabdrucks im Rahmen der archivischen Digitalisierung widmete. Nachhaltige digitale Archivierung, so ihre These, müsste noch einmal kritisch mit Blick auf die wechselseitigen Bezüge von Klimawandel und Energieverbrauch von für die digitale Archivierung kritischen IT-Infrastrukturen analysiert werden. Hierzu stellte sie verschiedene Ansätze vor, strukturiert nach internen und externen Einflüssen. Neben individuellen Verhaltensweisen der Archivar*innen stellte sie archivinterne Initiativen wie einen „Clean-Up-Day“ vor, an dem Mitarbeitende nicht mehr benötigte Daten löschten. Extern betonte sie, dass über übergreifende Verbände oder Institutionen auch künftig zunehmend solche Initiativen in die Fachwelt eingebracht werden müssten.

Am zweiten Tag wurden Themen aus dem Umfeld der nachhaltigen, digitalen Archivierung in den Fokus gerückt. Bieke Nouws stellte den Ansatz vor, den das belgische Staatsarchiv bei der digitalen Archivierung von Text- und Sofortnachrichten verfolgt, die Regierungsmitglieder und ihre Mitarbeitenden verfassen. Neben den bekannten sowohl technischen als auch rechtlichen Herausforderungen, vor die Archivar*innen in Europa und international gestellt sind, wenn es insbesondere um Social Media Archivierung oder eben auch die Archivierung von Daten aus Messenger Apps geht, schilderte sie auch die niederländische Rechtssituation: Chatnachrichten gelten als private Inhalte, für die das niederländische Archivgesetz keine, sondern nur für Daten der Verwaltung die Ermächtigungsgrundlage darstellt. Und so sind Beratung,

verschiedene technische Hilfsangebote, unterschiedliche technische Ansätze und insbesondere die Stück für Stück vertiefende Kooperation auf Grundlage beiderseitigen Vertrauens die Punkte, die sie in ihrem Vortrag betont hat. Den Blick auf Social Media warf dann im folgenden Vortrag Michel Cottin vom luxemburgischen Nationalarchiv. Und auch er hielt fest, dass sich neben Beratungsangeboten und technischen Informationsangeboten, die sich wechselseitig ergänzen müssten, der Aufbau einer auf Vertrauen fußenden Kooperation als der sicherste Weg zur erfolgreichen Archivierung von Social-Media-Inhalten wichtiger Vertreter aus Politik, Regierung und Verwaltung herauskristallisiert hat. Viele Personen des öffentlichen Interesses, wie Mitglieder der Regierung etwa, betreiben Social Media Accounts als private Accounts, weshalb für sie das luxemburgische Archivrecht nicht greift.

Von niederländischer Seite präsentierten schließlich Vincent Holt und Geert LeLoup den Ansatz des Staatsarchivs zur E-Mail-Archivierung. Hier gibt es eine entsprechende Verfügung, dass wichtige Vertreter aus Regierung und Verwaltung mit ihrem Ausscheiden verpflichtet sind, ihre E-Mail-Accounts anzubieten. Im Vorfeld der Übernahme der E-Mail-Daten erhalten die Mitarbeitenden zehn Wochen Zeit, persönliche Inhalte zu löschen. Im Anschluss werden die Maildaten für zehn Jahre geschützt aufbewahrt, bevor sie bewertet und übernommen oder kassiert werden. Grundlage der Bewertung, so hier der Ansatz, ist die Trennung von Person und Funktion, sodass Personen mit verschiedenen Accounts in Erscheinung treten können, wenn sie ihre Funktion wechseln. Insgesamt liegt der Fokus auf der Auswahl der E-Mail-Accounts von Personen, die zentrale Funktionen innehaben. Flankiert wird ihr Vorgehen, wie Holt und LeLoup schilderten, ebenso wie die Projekte in Belgien und den Niederlanden, von Handreichungen, Empfehlungen und persönlichen Beratungen.

Die in nicht wenigen Teilen ähnlichen Erfahrungen und Ansätze führten zur regen Diskussion und zum Konsens, dass

auch die intensive und gute Vernetzung von Archivar*innen mit den jeweiligen Archivträgern ebenso wie der Fachwelt (weiterhin) einen wesentlichen Faktor zur Zukunftsfähigkeit der Archive darstellen wird.

Den Abschluss bildete der Vortrag von Fred van Kan, der einen Blick auf die möglichen Entwicklungen des Archivwesens in den kommenden Jahrzehnten wagte. Im Zentrum der Ausführungen standen die Funktion der Archive in Regierung und Verwaltung, die Rolle der zukünftigen Archivar*innen als Informationsexpert*innen und -manager*innen, der Einsatz und die Rolle künstlicher Intelligenz gerade auch bei der Automatisierung von (Arbeits-)Prozessen, v. a. bei der Informationsaufbereitung für alle Bürger*innen, sowie die Entwicklung des Arbeitsplatzes insgesamt. Und so hielt er als eine mögliche Entwicklung bis etwa 2050 fest, dass jede Kommunikation zwischen Regierung, Verwaltung oder Archiven auf der einen und den Bürger*innen auf der anderen Seite bis dahin über ein und dieselbe technische Plattform erfolgen könnte. Auch Künstliche Intelligenz würde eine zentrale Rolle spielen und die Archivar*innen massiv in ihrer alltäglichen Aufgabenerledigung unterstützen, nicht zuletzt, um Inklusion und Zugang zu gewährleisten. Während also Archivar*innen unterstützt von KI künftig überwiegend von zuhause aus arbeiteten, würden einheitliche technische Lösungen und eine Ausweitung von Informationsaufbereitung und -verfügbarkeit die Diversität der Gesellschaft langfristig stützen.

Zu guter Letzt würdigte Frank Bischoff das große fachliche Engagement von Fred van Kan in der internationalen archivischen Zusammenarbeit. Nach 18 Jahren als Direktor des Gelders Archief tritt er nun seinen Ruhestand an. Die Vertreterinnen und Vertreter des Lenkungskreises des IAS verabschiedeten ihn gemeinsam und bedankten sich für eine sehr gute, ideenreiche und verlässliche Kooperation.

Sarah Rudolf, Brauweiler

ARCHIVARE ZWISCHEN KAISERREICH UND WEIMARER REPUBLIK. INSTITUTIONEN, SCHRIFTGUT, GESCHICHTSKULTUR

Hrsg. von Tom Toelle, Sarah Schmidt, Jessica von Seggern und Markus Friedrich. Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt Hamburg 25. Hamburg University Press, Hamburg 2023. 371 S., Druckausgabe: 44,80 €, ISBN 978-3-943423-80-8, Online-Ausgabe: DOI 10.15460/hup.262.2000, CC-BY.

Der vorliegende Sammelband vereint die Beiträge eines Workshops von 2020 am Staatsarchiv Hamburg. Die Einleitung nennt die thematisierte Gruppe zugespitzter „die Archivare der Weimarer Republik“. Gab es Weimar-spezifische Merkmale der deutschen Archivare und ihrer Berufspraxis? Der Band stellt diese zentrale Frage, ohne sie auch schon beantworten zu können. Er enthält neben der umfangreichen „Einleitung“, die eher den Versuch einer Synthese darstellt, zehn Einzelstudien zu Aspekten der Kriegsfolgenbewältigung, des Leitungspersonals, der Programmatik der Archivarbeit und ihrer kulturellen Einbettung. Anders als die Bezugnahme des Titels auf deutsche Zäsuren suggeriert, geht es dabei nicht nur um das deutsche Archivwesen. Zu den Kriegsfolgen gehörten massive Verluste und Abtretungen von Überlieferung. Sabine Herrmann zeichnet den Weg der Akten der deutschen Kolonialverwaltung in Kame-run vom Abzug des Personals 1916 über verschiedene Abwicklungsstellen bis ins Reichsarchiv nach, wohin sie 1938 als weitere Manifestation des Massenproblems der Kriegsakten gelangten. Herauszuheben ist der akteurszentrierte Ansatz, der den Bogen vom Kanzleihilfspersonal aus der kolonisierten Bevölkerung bis zu den Archivaren des Reichsarchivs (Meisner und Heinemeyer) schlägt. Diese Akten verblieben also den Kriegsverlierern, wenigstens bis zu ihrer weitgehenden Vernichtung durch die Bombardierung Potsdams 1945. Ansprüche der Sieger oder Nutznießer des Kriegsausgangs auf Archivalien im Zusammenhang mit abgetretenen Gebieten beschäftigten die Archivare aber Jahrzehnte lang. Zu den stärksten Beiträgen gehörte Sarah Schmidts Rekonstruktion der Archivverhandlungen um Nordschleswig, in dem die preußisch-deutsche Seite dänische Forderungen mit eigenen unerfüllten Ansprüchen aus dem Wiener Vertrag von 1864 zu kontern wusste, insgesamt aber unglücklich agierte und dem Provenienzprinzip als vorhandener Verhandlungsgrundlage eher nach Opportunität folgte. So kam der Archivalienaus-tausch, auf den es hinauslief, erst 1937 zustande. Die größten archivischen Kriegsfolgen schuf freilich die Auflösung des Habsburgerreiches. Die Nachfolgestaaten rangen um eine gemeinsame Linie gegenüber Rest-Österreich auf Archivkonferenzen, deren Verlauf Pavel Kahuda schildert und deren Protokolle er abdruckt. Das Provenienzprinzip und die fachliche Furcht vor der Zerstörung der alten Zentralarchi-ve prallte hier auf Pertinenzdenken und den Wunsch nach Verfügung über nationale Geschichtszeugnisse. Als Folge hielt archivisches Fachvokabular Einzug in den Frieden von Saint-Germain. Auf österreichischer Seite musste vor allem das Haus-, Hof- und Staatsarchiv mit diesen Forderungen umgehen, die Archivdelegierten der Nachfolgestaaten (oft alte Kollegen) aufnehmen, sich unter rigiden Sparzwängen

personell neu aufstellen, als Teil des Außenministeriums große Aktenmengen aufnehmen und die zeithistorische Auseinandersetzung um die Kriegsschuldfrage führen. Alle diese Aufgaben ruhten auf seinem Direktor Ludwig Bittner. Thomas Just hat sich bereits mit der Biografie dieses radikalen Nationalisten und Antisemiten beschäftigt. Hier stellt er nun Bittners fachliches Wirken in den Mittelpunkt, das imponierend zu nennen man nicht umhinkommt; seine bahnbrechende Studie zur Diplomatie neuzeitlicher Staatsverträge wäre noch hinzuzufügen.

Mit Paul Fridolin Kehr residierte ein anderer Titan zur selben Zeit in Berlin als Generaldirektor der preußischen Staatsarchive. Sven Kriese thematisiert seine beharrlichen, letztlich aber erfolglosen Bemühungen, aus der Archivverwaltung eine Einrichtung der außeruniversitären Großforschung neben den MGH, dem Preußischen Historischen Institut in Rom und anderen Akteuren, in die er bereits involviert war, zu machen. Sehr instruktiv wird an Beispielen nachvollzogen, wie Kehrs wissenschaftliche Beamte, unter ihnen Georg Winter, mehrheitlich aber die entgegengesetzte Wende von Historiker-Archivaren zum Berufsbild des Facharchivars vollzogen. Ihre Verwirklichung fand Kehrs 1915 verfasste Dankschrift nur in Teilen mit der Einrichtung des Instituts für Archivwissenschaft und im unheilvollen Engagement seines Nachfolgers Brackmann in der sogenannten Ostforschung. In der preußischen Provinz Westfalen wurde währenddessen mit Ludwig Schmitz-Kallenberg eine aus der Zeit gefallene Gestalt Leiter des Staatsarchivs Münster. Gestützt auf den politischen Katholizismus, kam damit ein akademisch gescheiterter Privatdozent von 57 Jahren in die Position eines landeshistorischen Strippenziehers alter Schule. Interessant und charakteristisch für eine neue Zeit war nur seine intensive Sorge um die Archivpflege, insbesondere bei den Adelsarchiven. Wilfried Reininghaus weist in seinem Beitrag aber auch nach, dass Schmitz-Kallenberg sich aus dem Verkauf eines Adelsarchivs sehr wahrscheinlich selbst bereichert hat.

Eine Stärke des Bandes liegt in unterschiedlichen Blickwinkeln auf wiederkehrende Themen. Ständige Finanz- und Personalnot in Münster wie in Wien gehörte dazu, aber auch die Archivpflege und die neue Wertschätzung nichtstaatlicher Überlieferung. Die Gegenseite zu Schmitz-Kallenburgs Wirken beleuchtet Tom Toelle anhand der Gründung der Vereinigten Westfälischen Adelsarchive und der Frage, warum dieses Modell selbstorganisierter Archivpflege im Rheinland und in Österreich trotz aufmerksamer Beobachtung keine Nachahmer fand. Während die Frage für Österreich nicht recht beantwortet wird, stießen im Rheinland die Kosten und die schillernde Persönlichkeit des Archivars der Vereinigung, Heinrich Glasmeier, ab. Eine Rolle spielte hier phasenweise auch die Debatte um gesetzlichen Kulturgutschutz als perzipierte Form von Enteignung. Toelle leitet die Ablehnungen konzeptionell aus einer „Verfügbarkeitsheuristik“ (S. 305) ab, was epistemologisch aber nicht über seine lockere Darstellung hinausweist. Mehr zur Einbettung archivischer Fachdiskurse in die Geschichtskultur der Zeit wirft Markus Friedrichs Studie über Armin Tilles ab, einen Schüler des Kulturhistorikers Karl Lamprecht mit genealogischen Interessen, die in die Richtung einer biologisierenden Sozialstatistik liefen. Als Publizist schrieb Tilles gegen Ver-

suche der Archive an, die um 1900 stark steigende Frequenz der privaten genealogischen Nutzung einzudämmen. Ins Amt des Direktors des Staatsarchivs Weimar und später aller thüringischen Staatsarchive gekommen, implementierte er Überlieferungsbildung im Sinne einer kulturhistorischen Histoire totale mit Fokus auf Archivpflege und Ergänzungsdokumentation. Schlüssig weist ihm Friedrich ein erweitertes, in Fachkreisen provozierendes Archivverständnis nach. Man hätte ergänzend noch den gemeinsam mit Woldemar Lippert aus Dresden versuchten Griff der Staatsarchive nach den Kirchenbüchern zur Bestätigung heranziehen können. Insgesamt wird aber auch deutlich, dass an Tilles' Wirken nichts spezifisch für Weimar ist. Vielmehr verweisen Tendenzen der Geschichtskultur des Kaiserreichs bereits auf das völkische Geschichtsverständnis des Nationalsozialismus', dem Tilles nahestand. Ebenso wenig lässt sich in der Archivtheorie eine Modernisierung erkennen. Dietmar Schenk setzt seine Beschäftigung mit dem Werk Adolf Brennekes fort und führt dessen Verständnis von Archivwissenschaft auf die Archivgeschichte zurück. Historisch individuelle Archiv-„Gestalten“ zu respektieren, war Kern von Brennekes Arbeitsverständnis, das um die Erschließung von Beständen aus der Zeit vor 1800 kreiste. Mehr als der Historismus des Kaiserreichs ist darin nicht zu erkennen.

Wie hätte man sich Archive im Geiste der Weimarer Republik vorstellen sollen? Etwa als Vernunftrepublikaner, die den Zugang liberalisieren, angefangen bei den Arbeiter- und Soldatenräten die Überlieferungsbildung ausweiten, die Erschließung im Interesse der Nutzung forcieren, auf Fachtagungen präsent sind und sich berufsständisch vernetzen? Und tatsächlich gab es diese Archive, wie Robert Kretzschmar eindrucksvoll und, das darf gesagt werden, sympathisch am Beispiel der aufeinander folgenden württembergischen Archivdirektoren Schneider, Müller und Winterlin zeigt.

Weitere Fallstudien wären wünschenswert. Wenn die Herausgeber als Fazit schreiben, „manchmal schuf sie [die Republik] sich ihrer Archive selbst“, meinen sie damit nicht-staatliche Akteure vor dem Hintergrund einer sich wandelnden Geschichtskultur. Der Band weist aber Lücken schon bei den direkten staatlichen Folgen der Revolution auf, insbesondere bei den evangelischen Kirchenarchiven und beim Politischen Archiv des Auswärtigen Amts, das als Werkzeug im Kampf gegen die sogenannte Kriegsschuldfrage vielleicht das Weimarer Archiv schlechthin war. Auch das Reichsarchiv wird man noch nicht als ausgeforscht betrachten können. Überhaupt zeichnet Tolles und Friedrichs Versuch einer Synthese mit breitem Pinsel ein stark auf kulturgeschichtliche Fluchtlinien zulaufendes Bild. Trotz einer breiten Literaturgrundlage werden archivwissenschaftliche Literatur und der allgemeine Forschungsdiskurs zur Weimarer Republik nur wenig rezipiert. Was unter „Weimar“ eigentlich zu verstehen ist, wird als Bedingung der Frage nach Weimarer Archiven nicht genügend bestimmt. Die weitreichenden Schlussfolgerungen beruhen in Teilen auf anekdotischer Evidenz: Eine Schädelvermessung auf dem Herrenabend der westfälischen Adelsarchive, die auffallend häufig als Beleg für generalisierende Aussagen dienen (S. 39), sagt noch nichts über Archive und Eugenik aus, und eine Bemerkung ausgerechnet von Ernst Müsebeck zum Volk als Staatsträger (S. 35) weist

sicherlich keine Ausrichtung der Archive am Parlamentarismus nach. Zum Massenproblem der Überlieferungsbildung aus Kriegsakten verweisen die Autoren auf das Bär'sche Prinzip, aber die Anfänge einer Bewertungslehre als Professionalisierungsschub werden nicht berücksichtigt. Aussagekräftiger ist die Analyse des Selbstbilds der Archive vor dem Hintergrund der Zäsurerfahrung von 1918, die freilich wieder auf die geschichtskulturelle Argumentation zugeschnitten ist. Dabei handelt es sich mehr um Hypothesen, die weiterer Untersuchung wert sind, als um Synthesen aus dem vorliegenden Band. In ihrer Divergenz erlauben dies die Beiträge auch nicht, stecken aber ein lohnenswertes Feld für weitere Forschungen ab.

Holger Berwinkel, Göttingen

VOR ORT UND VIRTUELL. ARCHIVE UND GESELLSCHAFT IM DIGITALEN ZEITALTER

Vorträge des 81. Südwestdeutschen Archivtags am 19. und 20. Mai 2022. Hrsg. von Ulrich Schludi (Werkhefte des Landesarchivs Baden-Württemberg, Heft 28). Jan Thorbecke Verlag, Ostfildern 2023. 107 Seiten, Broschur, 20,00 €. ISBN: 978-3-7995-2018-8

Die Pandemie ist vorbei. Aber ihre Auswirkungen beschäftigen uns weiterhin, auch im Archivwesen: Rückgang der Vor-Ort-Benutzung, Weiterentwicklung digitaler Nutzungs- und Vermittlungsangebote, Intensivierung virtueller interner und externer Kommunikation. Diese Entwicklungen wurden durch die Pandemie zumeist nicht ausgelöst, aber beeinflusst und beschleunigt. Und sie stellen uns vor ungeminde Herausforderungen bei der weiteren Ausrichtung der archivischen Angebote.

Daher lohnt die Lektüre der Beiträge einer Tagung, die sich im Mai 2022 genau diesen Veränderungen und Perspektiven widmete. Im Fokus des 81. Südwestdeutschen Archivtags in Reutlingen stand insbesondere auch die Positionierung der Archive in sich verändernden städtischen Kommunikationsräumen. Daher eröffnete das von Peter Schludi konzipierte Programm mit einem Blick von außen: Zunächst skizzierte die Reutlinger Stadtplanerin und Baudezernentin Angela Weiskopf die gegenwärtigen Veränderungen der Innenstädte, die unter der Konkurrenz von Einkaufszentren und Online-Handel leiden und zu Kultur- und Begegnungsräumen weiterzuentwickeln sind. Daran knüpfte der Beitrag von Johannes Milla direkt an, der ein Konzept seines Planungsbüros für das Staatsarchiv Ludwigsburg als Dritten Ort vorstellte und damit für einen auch baulichen Wandlungsprozess von einer archivischen „Trutzburg“ zu einem für alle Bürger*innen offenen Kommunikationsraum warb. Auf diesem Weg sind manche Archive schon weit fortgeschritten, wie Alain Dubois für das Staatsarchiv Wallis darlegen konnte. Als integraler Teil des 2016 eröffneten Kulturzentrums Les Arsenaux in Sitten profitiert es von der räumlichen Nähe zu anderen Kultureinrichtungen und sich

ergänzenden Informations- und Partizipationsangeboten, die sich auch auf den virtuellen Raum erstrecken.

Die digitalen Dimensionen der Vermittlung und Partizipation stehen im Fokus der von Joachim Kemper vorgestellten Kommunikationsstrategie des Stadt- und Stiftsarchivs Aschaffenburg. In enger Verbindung mit dem städtischen Digitalamt und im Rahmen eines städtischen Gesamtleitbildes als Dialog City wurden hier in kürzester Zeit ein „digitales Stadtlabor“ als Mitmachplattform und ein „Digitalladen“ als analoge, innerstädtische Anlaufstelle etabliert, die die Basis für weitere innovative Digitalprojekte bieten.

Dreh- und Angelpunkt der Öffentlichkeitsarbeit ist auch im Kreisarchiv Reutlingen die Verbesserung der Kommunikation durch die Verknüpfung digitaler und analoger Angebote. Im Mittelpunkt steht dabei eine neuentwickelte Kulturplattform, die alle Informationen bündelt, Partizipationsmöglichkeiten eröffnet und auch auf externe Module weiterleitet. Unter Abwägung von Aufwand und Nutzen kann Marco Birn eine deutliche Verbesserung der öffentlichen Wahrnehmung und Nutzung der archivischen Angebote konstatieren.

Die Chancen der in der Pandemie ausgeweiteten Kulturförderung hat das Stadtarchiv Münster entschlossen genutzt, um eine digitale Quellenkunde zu erarbeiten und ein umfassendes Angebot der Digitalisierung on demand zu etablieren. Zunächst zur Sicherstellung der Archivnutzung auch unter Pandemiebedingungen konzipiert, führte das Projekt nicht nur zu einem technologischen Fortschritt, sondern, wie Peter Worm hervorhebt, auch zu einer bewussten Neuausrichtung der Digitalisierungsstrategie: Sie führt weg von der Bearbeitung ganzer, vom Archiv ausgewählter Bestände hin zur nutzergesteuerten Auswahl kleinerer Einheiten, die über die Onlinestellung im NRW-Archivportal gleichwohl auch für alle anderen Interessierten zugänglich werden.

Im Geldersarchief in Arnheim wurde diese Entwicklung schon einige Jahre früher initiiert und in der Pandemie weiter forciert, so dass digitalisiertes bzw. digitalisierbares Archivgut im Lesesaal gar nicht mehr bereitgestellt wird. Dem Ziel eines vollständigen Online-Angebots und einer weitestgehenden Erübrigung von Lesesaalbenutzungen stehen ‚nur‘ noch der Datenschutz, das Urheberrecht und limitierte Scankapazitäten im Wege. Die Konsequenz, mit der dieses Ziel verfolgt wird, resultiert aus einem archivischen Selbstverständnis als Datenlieferant (auf Grundlage der FAIR Data Prinzipien) und könnte, so der Ausblick von Fred van Kan, sich einfügen in tiefgreifende Strukturreformen des niederländischen Archivwesens.

Nicht zuletzt vor diesem Hintergrund lässt sich bilanzieren, dass die Beiträge des Tagungsbandes nicht nur viele Anregungen für eine erfolgreiche archivische Öffentlichkeitsarbeit vermitteln. Sie lassen vielmehr immer wieder erkennen, dass nicht lediglich ein gelungener Mix von Aktivitäten „vor Ort und digital“ gefragt ist, sondern eine strategische Ausrichtung der Öffentlichkeitsarbeit – und damit eine Positionierung der Archive in einer sich rapide wandelnden Gesellschaft.

Ralf Lusiardi, Magdeburg

PHILIP HAAS UND MARTIN SCHÜRRER, WAS VON PREUSSEN BLIEB

Das Ringen um die Ausbildung und Organisation des archivischen Berufsstandes nach 1945. Quellen und Forschungen zur hessischen Geschichte, Bd. 183, Selbstverlag der Hessischen Historischen Kommission Darmstadt und der Historischen Kommission für Hessen, Darmstadt und Marburg 2020. 187 S., 27 s/w Abb., 24,00 €. ISBN 978-3-88443-338-6

Grundsätzlich ist es verdienstvoll, dass sich zwei Autoren mit archivarischem Hintergrund bereitgefunden haben, sich mit der Geschichte der archivischen Ausbildung nach dem Zweiten Weltkrieg zu befassen. Philip Haas und Martin Schürer vom Niedersächsischen Landesarchiv tun dies in Form einer knapp 200seitigen Publikation, die vom Hessischen Ministerium für Wissenschaft und Kunst gefördert wurde und die sich in fünf Hauptteile gliedert.

Die Einleitung (S. 9–14) setzt mit der Schilderung der Situation für die staatlichen Archive in Deutschland nach dem Ende der preußischen Archivverwaltung respektive des preußischen Staates (25. Februar 1947) nach Ende des Zweiten Weltkriegs ein. Die Autoren kritisieren das aus ihrer Sicht bis heute vorherrschende revisionsbedürftige Narrativ von der Alternativlosigkeit des „Weges nach Marburg“, das der Einrichtung der Archivschule an diesem Ort geradezu teleologischen Charakter beimesse.

Demgegenüber skizzieren Haas und Schürer im zweiten Kapitel („Ereignisse – Umkämpfte Archivpolitik zwischen 1945 und 1949“, S. 14–63) das bis zuletzt entscheidungs offene Ringen um die Einrichtung der Archivschule und deren Standort in Marburg. Erst Konzessionen gegenüber anderen Mitbewerbern, etwa aus dem neu konstituierten Land Nordrhein-Westfalen – hier ist namentlich Münster zu nennen –, hätten dann zum Erfolg der Marburger Bemühungen geführt. Auch die nach Kriegsende intensiv ventilierter Frage nach der Reorganisation des in enger Verbundenheit zum Preußischen Geheimen Staatsarchiv stehenden Instituts für Archivwissenschaft und geschichtswissenschaftliche Fortbildung (IfA), das von 1930 bis 1945 als zentrale Ausbildungseinrichtung für den höheren und mittleren Archivdienst in Deutschland fungiert hatte, wird aufgegriffen.

Das dritte Kapitel („Konzepte – Organisation, Ausrichtung und Anspruch der Ausbildungsformen“, S. 63–102) widmet sich den inhaltlich-konzeptionellen Überlegungen, die im Rahmen der Neuorganisationsdebatte rund um das Thema „archivarische Ausbildung“ von verschiedenen Protagonisten angestellt wurden. Der Reigen der Reflexionen, Erörterungen und durchaus auch kontroversen Auseinandersetzungen reichte von der Frage nach dem wissenschaftlichen oder eher praxisbezogenen Profil der Ausbildung über den zu bevorzugenden Namen der Ausbildungsstätte bis hin zur terminologischen Debatte hinsichtlich der archivischen – oder archivischen? – Berufssprache (S. 75, mit Hervorhebung der Unterscheidung zwischen „archivisch“, „archivarisch“ und „archivalisch“). Das Spannungsfeld zwischen „Archivkunde“ und „Archivwissenschaft“, das bis in die archivische Gegenwart hinein wahrnehmbare Vibrationen entfaltet, wird anschaulich umrissen.

Kapitel vier („Ein Ausblick: Die Folgen der Marburger Gründung“, S. 102–110) richtet den Fokus ein Stückweit in die Zukunft: Obgleich die Autoren die Errichtung der Archivschule in Marburg in institutioneller Hinsicht als Erfolg betrachten, äußern sie Kritik an der tatsächlichen qualitativen Beschaffenheit dieser Einrichtung und ihrem tatsächlichen Ausbildungserfolg. Insonderheit die von Haas und Schürer konstatierten „organisatorischen und konzeptionellen Eigentümlichkeiten“, die sich im Laufe der Zeit als „schwere Hypothek“ herausgestellt hätten (S. 102), sind in diesem Kontext zu nennen. So sei weder eine „Verwissenschaftlichung der Archivistik“ (S. 104) noch eine klare Profilierung zwischen den Paradigmen von praxisbezogener Archivkunde beziehungsweise Archivwissenschaft gelungen noch die Archivschule ihrer Aufgabe und ihrer Funktion als „Schule“ gerecht geworden, was den Autoren zufolge nicht zuletzt der mangelnden Bezugnahme des recht theoretischen Unterrichts auf den späteren Arbeitsalltag geschuldet gewesen sei (S. 107). Auch die Verflechtung von Archivschule und Staatsarchiv Marburg habe eher Probleme gezeitigt als Vorteile respektive nutzvolle Synergieeffekte generiert. Schließlich sei es eine große Herausforderung gewesen, qualitativ geeignete Lehrende für den Unterricht an der Archivschule zu gewinnen (S. 109 f.).

Eine zentrale Erkenntnis, die im Fazit (S. 110–116) vermittelt wird, ist die These, dass die Konflikte, die Auseinandersetzungen, ja Kämpfe innerhalb des archivarischen Berufsstandes – oder hier wohl besser: der staatlichen Archivverwaltungen und ihrer Akteure – des Deutschlands der Nachkriegszeit um die künftige Ausgestaltung der archivarischen Ausbildung bedingt durch eine „berufsständische[] Solidarität der Hüter der Quellen“ zugunsten von „geschichtsmächtigen Mythen“ umgeformt worden seien (S. 116).

Der im Verhältnis zum Fließtext recht umfangreiche Anhang umfasst „Zentrale Quellen zur Geschichte der Archivsulen“ (S. 117–169), ein Quellen- (S. 170 f.) und ein Literaturverzeichnis (S. 172–182) sowie ein Personenregister (S. 183–185) runden die Publikation ab.

Der Schwerpunkt der Untersuchung liegt wie bereits angedeutet auf der Archivschule Marburg, doch werden auch andere, letztlich konkurrierende Standorte und Institutionen nicht ausgespart, wenngleich – legitimer Weise – nachrangig gehandhabt. So wird etwa die bayerische Archivschule in München als denkbare Alternative zu „Marburg“ in einem eigenen Unterkapitel gewürdigt (S. 85–88).

Obzwar in einer Arbeit dieses recht überschaubaren Umfangs – was keineswegs abwertend zu verstehen ist – ein Mehr an Quellenstudium kaum einzufordern sein dürfte – immerhin wurden Bestände von zehn Archiven beziehungsweise regionalen Abteilungen ausgewertet –, stellt sich doch die Frage, ob ausschließlich in der Überlieferung staatlicher Archive unmittelbare und vor allem mittelbare Informationen zum Thema zu finden sind. Ließen sich nicht auch in nichtstaatlichen, unter Umständen sogar in nicht-öffentlichen Archiven Aussagen zur archivarischen Ausbildung nachweisen? Nahm man dort den Diskurs um die (Re-) Organisation der archivarischen Ausbildung wahr? Und (wie) positionierte man sich diesbezüglich? Dies sei lediglich als ein Aspekt genannt, der bei etwaigen weiteren Untersuchungen zum Thema Beachtung finden könnte, zumal die Zeit zumindest in technologischer Hinsicht – Stichworte: Digitalisierung, „Wundermittel“ Künstliche Intelligenz, „Keywordspotting“ etc. – für die Forschung zu arbeiten scheint.¹ So bleibt es im vorliegenden Falle erst einmal beim Universitätsarchiv Münster und den nichtdeutschen staatlichen National Archives and Records Administration (NARA), die sich im Quellenverzeichnis (S. 170 f.) zu den Beständen der deutschen staatlichen Archive (Bayerisches Hauptstaatsarchiv München, Geheimes Staatsarchiv Preußischer Kulturbesitz, Hessisches Hauptstaatsarchiv Wiesbaden, Hessisches Staatsarchiv Marburg, Landesarchiv Nordrhein-Westfalen mit Abteilungen Rheinland und Westfalen, Niedersächsisches Landesarchiv) gesellen.

Zu wünschen bleibt eine auf den von Haas und Schürer erarbeiteten Ergebnissen aufbauende, vertiefende und differenzierende Auseinandersetzung mit der Materie, gerne mit intensiverer Weitung in die Gegenwart sowie mit einem noch mehr vergleichenden Blick auf benachbarte Ausbildungsinstitutionen und deren Werdegang. Die vorgestellte Publikation bietet hierzu eine gute Ausgangsbasis. ■

Martin Schlemmer, Duisburg

¹ Vgl. hierzu jüngst beispielsweise *Scrinium* 78 (2024) mit mehreren Vorträgen des 42. Österreichischen Archivtags in Klagenfurt 24.–25. Oktober 2023 zum Thema der KI sowie das vorliegende Heft von *ARCHIV. theorie & praxis* mit diesem Themenschwerpunkt. Instrukтив zuletzt auch Christian Götter, *Künstliche Intelligenz und Geschichtswissenschaft. Mehr als ein neues Werkzeug für die Digital History?* In: *Historische Zeitschrift* 318, Heft 2 (2024), S. 299–330.

TRANSFORMATION DER WIEDERGUTMACHUNG

EIN NEUES MEHRJÄHRIGES ERSCHLIESSUNGS- UND DIGITALISIERUNGSPROJEKT IN DER ABTEILUNG RHEINLAND

Die Abteilung Rheinland des Landesarchivs NRW hat in den vergangenen Jahren von der zuletzt allein zuständigen Bezirksregierung Düsseldorf ca. 420.000 Entschädigungsakten nach dem Bundesentschädigungsgesetz (BEG) übernommen.¹ Diese seriellen Einzelfallakten überliefern Lebens- und Verfolgungsschicksale von hunderten Tausenden Antragstellenden, die in der Nachkriegszeit über Jahrzehnte die Möglichkeit hatten, Wiedergutmachungszahlungen für unter dem Nationalsozialismus erlittenes Unrecht zu beantragen. Zu diesem Zweck legten diese ansonsten größtenteils vergessenen Stimmen ihre Familiengeschichte und ihr Verfolgungsschicksal im Zuge eines gleichförmigen Verwaltungsverfahrens dar.² Es ist daher nicht verwunderlich, dass dieser einmalige Fundus zur Geschichte des 20. Jahrhunderts als ausnahmslos archivwürdig bewertet worden ist.

Angesichts ihrer enormen Menge stellt die Erschließung dieser Akten im laufenden Betrieb auch für große Archive einen Kraftakt dar. Sie ist aber angesichts der hohen Nachfrage nach diesen Unterlagen unumgänglich dringend.³ Alle Bestände in der Abteilung Rheinland (mit Ausnahme von BR 3001 und BR 3002) sind nun nach solchen mehrjährigen Kraftakten zumindest basisschlossen (Aufnahme von Antragsteller, Verfolgter sowie von deren Geburtsdaten). Stand heute sind ca. 32,04 % der Wiedergutmachungsakten zumindest in dieser Weise erschlossen und damit über unser AFIS recherchierbar. Als tiefer erschlossen dürfen dabei allerdings nur die Bestände BR 3003, BR 3004 und BR 3006 gelten⁴ – die hierzu notwendige Erschließungsarbeit wurde innerhalb des Fachdezernats R 3 nach und nach bewältigt. Die Bestände BR 3005, BR 3007 und BR 3008 hingegen wurden mithilfe glücklicherweise noch vorhandener behördlicher Findmittel (Namenskarteien), die zuvor in eine Excel-Tabelle ausgewertet worden waren, in Projekten abgeprüft und signiert, die mit der Duisburger Werkstatt für Menschen mit Behinderung durchgeführt wurden. Die geprüften Tabellen wurden im Anschluss in unser AFIS überspielt. Somit können die Akten zwar recherchiert und infolge auch genutzt werden, die personenbezogenen Schutzfristen aber müssen recht aufwendig individuell nachgeprüft wer-

den, da äußerst viele Akten auch schutzwürdige personenbezogene Daten jenseits von Antragstellenden und Verfolgten beinhalten.

Die beiden bislang noch gänzlich unerschlossenen Bestände stellen allerdings den größten Teil der Gesamtmenge dar (es handelt sich insgesamt um über 380.000 Einzelfall-Akten). Hier besteht noch immer ausschließlich die Möglichkeit einer umständlichen Recherche über die Bundeszentrale für politische Bildung (BZK).⁵ Dabei wird aber in jedem Falle das Geburtsdatum der antragstellenden Person (die häufig nicht mit den Verfolgten identisch ist) benötigt, denn erst mithilfe dieser Information kann in der nach Geburtsdaten geordneten BZK recherchiert

¹ Hinzu kommen über 27.000 Entschädigungsrentenakten sowie (von den rheinischen Landgerichten) über 52.000 Rückerstattungsakten. Außerdem werden laufend neue Akten durch die Bezirksregierung Düsseldorf angeboten – darunter werden noch Großakzessionen zu erwarten sein. Die Masse an Akten erklärt sich durch die überregionale Zuständigkeit zunächst der Bezirksregierung in Köln (u. a. für überregionale Verfolgengruppen, Staatenlose und Flüchtlinge im Sinne der Genfer Konvention), sodann der Bezirksregierung Düsseldorf, die noch immer im Dez. 15 Angelegenheiten nach dem BEG und dem Härtefonds NRW bearbeitet.

² Allg. zu den Wiedergutmachungsakten nach dem BEG, vgl. Cordula Rehr: Entschädigungsakten der Bezirksregierungen nach dem Bundesentschädigungsgesetz (BEG). In: Unbekannte Quellen. „Massenakten“ des 20. Jahrhunderts IV; Veröffentlichungen des Landesarchivs Nordrhein-Westfalen 75, S. 22–41.

³ Dies betrifft insbesondere Wiedergutmachungseinbürgerungen nach § 15 StAG. Täglich neu erreichen die Abteilung Rheinland Anfragen aus aller Welt, die um Ermittlung von Wiedergutmachungsakten bitten, da diese mittelbar als Voraussetzung für die Zuerkennung der deutschen Staatsangehörigkeit dienen. Außerdem sind die Akten in der Forschung sehr begehrt, beinhalten sie doch häufig als Einzige individuelle Ego-Dokumente. Darüber hinaus herrscht starke Nachfrage aus den Bereichen der Familien-, Provenienz- und NS-Forschung.

⁴ Dies sind zurzeit ca. 13,21 % aller Wiedergutmachungsakten in der Abteilung Rheinland. Tiefer erschlossen meint hier: im herkömmlichen Sinne. Zur Tiefererschließung nach Metadatenkatalog der KLA UAG, s. unten.

⁵ Die ca. 2,5 Millionen Karteikarten umfassende BZK wird ebenfalls vom Dezernat 15 der Bezirksregierung geführt und ist dort laufendes Registraturgut. Die analoge Version ist unter der Bestandsnummer BR 3015 Archivgut des Landes Nordrhein-Westfalen; die digitale Version wird im Jahr 2025 vom Landesarchiv ebenfalls als Archivgut übernommen werden.



Die Akten des Bestands BR 3007 wurden in der Bezirksregierung Aachen in Hängeregistraturen aufbewahrt (was zu Übermaß führt). Vor der Digitalisierung sind daher Maßnahmen zur Umbettung und Entmetallisierung unumgänglich.

werden. Dies führt dann zu einer ZK-Nummer, die über eine vorhandene Nummernblockübersicht einen Hinweis auf den Bestand gibt, in dem zu suchen ist (die beiden noch ausstehenden Bestände sind numerisch sortiert). Stand jetzt können die besonders wichtigen (da überregionalen) Wiedergutmachungsbestände BR 3001 und BR 3002 nur in dieser Weise erschlossen und genutzt werden. Diese aufwendige Methode stößt natürlich irgendwann an ihre Grenzen, da es sich bei den Wiedergutmachungsakten um eine immer stärker nachgefragte Aktengruppe handelt.

Die generelle Bedeutung der Bestände sowie das ausgesprochen hohe Nutzungsinteresse (insb. jenes zur Durchsetzung von Rechtsansprüchen) drängen also dazu, die Erschließung auch der beiden noch unerschlossenen Bestände prioritär anzugehen. Selbstredend wurde die Bedeutung dieser Thematik auch andernorts erkannt: Das Bundesarchiv begann 2021 mit der Konzeption eines eigenen Themenportals „Wiedergutmachung nationalsozialistischen Unrechts“ im Archivportal-D.⁶ Dort sollen sukzessive Wiedergutmachungsakten der Länder eingestellt und möglichst niedrigschwellig digital recherchierbar gemacht werden und zwar „angereichert mit historischen Hintergrundinformationen, der Vorstellung von Forschungsprojekten und Materialien für politische Bildungsarbeit“.⁷ Das Bundesministerium der Finanzen finanziert hierbei nicht nur den Aufbau dieses Portals, sondern auch einzelne Erschließungs- und Digitalisierungsprojekte der Länder im Rahmen der „Transformation der Wiedergutmachung“ mit.⁸ Im Rahmen dieser „Transformation“ hat nun auch das Landesarchiv NRW eine Finanzierung erhalten: Das dreiteilige Projekt ist auf 3,5 Jahre angelegt und setzt es sich zum Ziel, digitalisierte und tiefererschlossene Archivbestände zur Wiedergutmachung auf dem Themenportal in rechtskonformer Weise und unter Bereitstellung von Kontextinformationen zu veröffentlichen. Das erste Teilprojekt wird dabei den wissenschaftlichen und archivischen Rahmen für zukünftige Projekte bilden und die Basis für einen NRW-Auftritt im Themenportal Wiedergut-

machung legen. Hier ist, neben der wissenschaftlichen Aufarbeitung der spezifischen administrativen und rechtlichen Rahmenbedingungen der Wiedergutmachung in NRW, vor allem die landesweite Erfassung aller einschlägigen Unterlagen zur Wiedergutmachung in einem sachthematischen Inventar in Zusammenarbeit mit anderen Archiven, Institutionen und der Forschung geplant. Systematisch werden für NRW einschlägiges Registratur- und Archivgut, beteiligte Stellen und Behörden sowie Rechts- und Verwaltungsgrundlagen erfasst und beschrieben. Als Ergebnis steht am Ende der zweijährigen Projektphase ein sachthematisches Inventar, das als Ausgangspunkt für weitergehende Erschließungs- und Digitalisierungsprojekte dient und ebenfalls im Wiedergutmachungsportal zugänglich gemacht werden soll. Finanziert werden hierbei zwei EG 13-Stellen für zwei Jahre. Ziel des zweiten Teilprojektes ist die Digitalisierung und anschließende Tiefererschließung des Bestands BR 3007 (Wiedergutmachungsakten des ehemaligen Regierungsbezirks Aachen). Diese Akten umfassen rund 240 lfd. m. Archivgut und enthalten ca. 9.925 Entschädigungsverfahren von Verfolgten bzw. deren anspruchsberechtigten Angehörigen, die entweder am Tag der Antragstellung ihren Wohnsitz im Verwaltungsbereich der ehemaligen Bezirksregierung oder ihren letzten inländischen Wohnsitz während der NS-Zeit dort hatten. Sie sind nach obig beschriebenem Muster basierschlossen. Da die Akten sich in der Bezirksregierung Aachen in Hängeregistraturen befunden hatten, ist vor der Digitalisierung eine Umbettung aufgrund des Metallanteils und des Übermaßes zwingend erforderlich. Die Tiefererschließung wird dann anhand der Digitalisate gemäß dem von der KLA UAG vereinbarten Metadaten-set durchgeführt werden.⁹ Hierfür wird auch unsere AFIS-Anwendung V.E.R.A.-Verzeichnung ertüchtigt. Für die Umbettung, Digitalisierung und Tiefererschließung des Bestandes sowie die notwendigen Anpassungen an V.E.R.A. werden 3 Jahre und 6 Monate angesetzt. Finanziert werden zum Zweck der Tiefererschließung vier EG 9-Stellen für 1 Jahr und 6 Monate. Das dritte Teilprojekt schließlich beinhaltet die Erfassung der wichtigsten unerschlossenen Bestände zur Wiedergutmachung (BR 3001 und BR 3002) als vorbereitende Maßnahme für die spätere Digitalisierung und Tiefererschließung zwecks Präsentation im Themenportal. Es handelt sich hierbei um insgesamt über 387.000 Einzelfallakten. Die beiden behördlichen Findmittel, die uns zur Verfügung stehen (die Namenskarteien BR 3012 und BR 3014, insg. ca. 567.000 Karten), sollen hierbei zunächst gescannt werden. Im Anschluss werden die Digitalisate ausgewertet – so werden eine oder mehrere Excel-Listen gewonnen, die sich an die Ordnung des jeweiligen Bestandes einfach anpassen lassen, so dass anhand dieser Liste der Bestand kontrolliert und absigniert werden kann. So entstehen brauchbare Basis-Erschließungsdaten. Diese werden nach der Endkontrolle aus Excel exportiert und in das AFIS des Landesarchivs sowie in das Archivportal NRW und in das Themenportal Wiedergutmachung importiert. Die erstellten Scans der Namenskarteien werden darüber hinaus durch das Landesarchiv nachgehalten und können im Rahmen des Themenportals vorbehaltlich der personenbezogenen Schutzfristen des Archivgesetzes NRW online gestellt werden. Die so gewonnenen Erschließungsdaten bilden die Basis nicht bloß

für Recherche und Nutzung, sondern auch für die zukünftige Digitalisierung und Tiefenerschließung auch dieser Akten (sowie ihrer Bereitstellung im Themenportal). Mithilfe der Finanzierung durch den Bund sowie durch die technische Unterstützung der Projektpartner im Themenportal kommen wir so einen guten Schritt voran in unserem Bemühen, diese einzigartigen Quellen möglichst niedrigschwellig für den breitest möglichen Nutzendenkreis bereitzustellen.

David Hecken, Duisburg

- ⁶ Weitere Projektpartner stellen die Kolleginnen und Kollegen im Landesarchiv Baden-Württemberg, dem Portal Deutsche Digitale Bibliothek sowie dem Leibniz-Institut FIZ Karlsruhe dar.
- ⁷ <https://www.landesharchiv-bw.de/de/landesharchiv/projekte/transformation-der-wiedergutmachung/71002> (wie abgerufen am 24.10.2024).
- ⁸ https://www.bundesfinanzministerium.de/Monatsberichte/2021/01/Inhalte/Kapitel-3-Analysen/3-7-archivierungsprojekt-wiedergutmachung-pdf.pdf?__blob=publicationFile&v=1 (wie abgerufen am 24.10.2024).
- ⁹ https://wiki.deutsche-digitale-bibliothek.de/pages/viewpage.action?pageId=126518213&preview=/126518213/129007768/Anlage_1_Metadata-katalog_2023.05.12.pdf (wie abgerufen am 24.10.2024).

ZWISCHEN „INFOHAPPEN“ UND „BLOCKBUSTER“

NEUE WEGE DER ELEKTRONISCHEN BEHÖRDENBERATUNG IM LAV NRW¹

1. AUSGANGSSITUATION

Seit 2016 berät das Dezernat F 4 „Elektronische Unterlagen“ die Behörden, Betriebe und Einrichtungen der Landesverwaltung Nordrhein-Westfalen auf dem Weg von der analogen zur digitalen Schriftgutverwaltung, namentlich der elektronischen Aktenführung. „Gegründet“ wurde das Dezernat zunächst als Kompetenzteam im Rahmen eines anderen Dezernats (F 1 „Konzept- und Strategieentwicklung“), um nach kurzer Zeit ausgegliedert und als eigene Organisationseinheit im Fachbereich Grundsätze etabliert zu werden. Die derzeit 1,5 Stellen der Behördenberatung sind eingebettet in dieses Dezernat F 4, sodass die auf diesem Aufgabengebiet Tätigen auch die Entwicklungen rund um die Themen der Vorfeld- und Gremienarbeit sowie der digitalen Unterlagen respektive der digitalen Archivierung mitverfolgen können und in ihrer Beratungstätigkeit wiederum davon profitieren.

Im Zeitraum von 2016 bis 2023 wurden in 263 Veranstaltungen (2016/17: 30; 2018: 52; 2019: 33; 2020: 37; 2021: 39; 2022: 38; 2023: 34) für 109 Beratungspartner (Mehrfachberatungen eines Mandanten eingeschlossen), in aller Regel Einrichtungen der unmittelbaren und mittelbaren Landesverwaltung in Nordrhein-Westfalen, (2016/17: 10; 2018: 18; 2019: 16; 2020: 11; 2021: 15; 2022: 27; 2023: 12)² und mehr als 4.645 Beschäftigte (2016/17: keine Zahlen erfasst; 2018: ca. 2.000; 2019: 557;

2020: 573; 2021: 453; 2022: 658; 2023: 404) beraten, informiert und geschult.³

2. EVALUATION DER BERATUNG

Eine Evaluation der Beratung und ihres Erfolgs (wie immer dieser zu definieren wäre) gab es – zumindest im engeren Sinne des Wortes – lange Zeit nicht. Allerdings wurde bereits zu Beginn der Behördenberatung im Zusammenhang mit dem Umstieg auf die digitale Schriftgutverwaltung/elektronische Aktenführung eine Übersicht über die einzelnen konkreten Beratungsleistungen in Form einer Excel-Tabelle

- ¹ Geringfügig überarbeitete und ergänzte Fassung eines am 10.10.2024 im Rahmen der „Veranstaltung der Fachgruppe 1: Staatliche Archive“ auf dem 91. Deutschen Archivtag 2024 in Suhl gehaltenen Vortrags.
- ² Hinzu kommen noch hausübergreifende Veranstaltungen mit Teilnehmenden verschiedener Dienststellen, die hier nicht gesondert aufgeführt werden.
- ³ Bei Behörden und Beschäftigten liegen mitunter Doppelungen vor, da dieselben Einrichtungen und Personen durchaus mehrfach beschult worden sein können. Es handelt sich somit de facto um weniger beschulte Einzelbehörden und Einzelpersonen als es die rein rechnerische Summe auf den ersten Blick vielleicht vermuten lässt. Die Zahlen sind der seit 2016 geführten Excel-Tabelle sowie für die späteren Jahre den Jahresberichten des LAV NRW entnommen, wobei bei geringfügig voneinander abweichenden Werten den direkt der Tabelle entnommenen Zahlen der Vorzug gegeben wurde.

angelegt und bis heute geführt. Hier werden die beratenen Institutionen, die beratenden Beschäftigten des Landesarchivs sowie die Zahl der Teilnehmenden vermerkt, gegebenenfalls auch die Teilnehmenden am abschließenden Online- kahoot-Quiz, falls dieses Bestandteil der jeweiligen Veranstaltung ist. Seit einiger Zeit wird, wenn die Rahmenbedingungen und Absprachen dies ermöglichen, im Anschluss an eine Veranstaltung ein Feedback über das Online-Tool „mentimeter“ angeboten, von dem zwischen null und 50% der Teilnehmenden Gebrauch machen – in der Regel allerdings deutlich unter 50%, was jedoch immer noch ein gutes Resultat ist.⁴

Künftig möchte das Dezernat F 4 jedoch noch mehr „System“ in die Nachbereitung respektive Evaluation der einzelnen Beratungssituationen implementieren, wobei durchaus in die Richtung eines „Audience Development“ und „Relationship Marketing“ gedacht wird. Verwiesen sei in diesem Zusammenhang auf die jüngst an der Archivschule Marburg entstandene Transferarbeit „Government Audience Development. Wie steuern Archive ihre Verbindung zur Verwaltung?“ von Mareikje Mariak.⁵

Angedacht ist ein Nachsorge-Kontakt, eine Art „Follow-up-Kontakt“, der nach einer bestimmten Zeit, etwa nach sechs Monaten, im Nachgang zu einem Beratungstermin proaktiv erfolgen soll, um in Erfahrung zu bringen, wie sich die Situation bezüglich des Beratungsthemas zwischenzeitlich entwickelt hat. Erste Telefonate mit beratenen Einrichtungen wurden unlängst geführt und förderten erstaunlich positive Entwicklungen zutage, doch kann dies bei einem systematischen, breiter angelegten Nachfassen von Fall zu Fall schon wieder ganz anders aussehen. Um die Bereitschaft zu einer „Evaluation“ auf Behördenseite nicht zu schmälern, ist ein behutsames, eher in „qualitativen Interview-Situationen“ denkendes Vorgehen angezeigt, da statistisch operierende Fragebögen oder ähnliche auf Quantifizierbarkeit abzielende Verfahren definitiv prohibitiven Charakter tragen und auf den Mandanten allzu schnell abschreckend wirken können. Häufig möchte man sich ganz einfach eben „keine Blöße geben“, schon gar nicht eine messbare beziehungsweise quantifizierbare. Es gilt also auf „sanfte“ Art und Weise das Monitoring der Beratungsleistungen zu verbessern und eine mittel- bis langfristige Perspektive in die Beratungsplanung einzubringen.

3. BERATUNG ALS FORTBILDUNG

Eine weitere Neuerung ist das verstärkte Bestreben des Behördenberatungsteams in F 4, die Beratung zur elektronischen Aktenführung als zu perpetuierenden Bestandteil der hausinternen Fortbildung kenntlich zu machen und somit stärker als bisher die Fortbildungsstellen der Behörden und Einrichtungen des Landes zu adressieren. Bei der Staatskanzlei des Landes Nordrhein-Westfalen läuft die seit 2023 zweimal pro Jahr angebotene Beratung zur elektronischen Aktenführung über die Fortbildungsstelle (Referat I 4). Im September 2024 nahmen 26 Personen an der Viko-Schulung teil, die nächsten beiden Termine für das Jahr 2025, einmal in Präsenz und einmal als Videokonferenz, werden gerade im Moment abgestimmt.

Der Begriff des „Blockbusters“ für die Fortbildungsveranstaltung zur elektronischen Aktenführung („Basisschulung“) kam von der Staatskanzlei. Der zuständige Kollege der Staatskanzlei ließ die Anwesenden der Präsenzschiulung am 9. April 2024 wissen, dass diese Veranstaltung mit dem Landesarchiv inzwischen ein regelrechter „Blockbuster“ sei, „so begehrt [...] wie die Klopapierrolle während der Corona-Pandemie“.

Inzwischen hat sich die Schulungsveranstaltung zur elektronischen Aktenführung im Fortbildungsprogramm der Staatskanzlei verstetigt: Die Schulung hat dauerhaft Aufnahme gefunden mit zwei Schulungsterminen pro Jahr, einmal per Viko, einmal in Präsenz. Die Teilnehmendenzahlen schwanken zwischen 12 und 26 Personen, was jeweils bedeutet, dass die verfügbaren Plätze ausgebucht sind.

Wem das quantitativ zu wenig „Blockbuster“ ist, wird sich an der Zahl der Teilnehmenden an einer Viko-Kurz-Schulung für den landeseigenen Bau- und Liegenschafts-Betrieb (BLB) erfreuen, an der Beschäftigte verschiedener Niederlassungen bzw. Standorte teilnahmen. Normalerweise, so wurde ich im Vorfeld gebrieft, nähmen an diesem Info-Format ca. 550 Personen teil. Angesichts des vermeintlich recht „angestaubten“ Themas der elektronischen Aktenführung solle ich nicht enttäuscht sein, falls dieses Mal nur zwischen 400 und 500 Beschäftigte teilnahmen. Es wurden dann ca. 620 Teilnehmende...

Als wichtige Anknüpfungspunkte dienen bei der Andienung des Themas an die Fortbildungsstellen die zunehmende Fluktuation des Personals sowie der sich stetig verschärfende (Fach-)Kräftemangel, darüber hinaus die Tatsache, dass das Thema Schriftgutverwaltung/Aktenführung bei der Ausbildung keine Rolle spielt und die quereingestiegenen Beschäftigten ebenfalls kein entsprechendes Vorwissen mitbringen.

4. AKTENRELEVANZ ALS BERATUNGSTHEMA

Seit etwa anderthalb bis zwei Jahren wurden und werden wir immer häufiger auf die Frage nach Ablageorten für Informationen, namentlich nach der Aktenrelevanz, angesprochen. Die Verunsicherung ist groß, was man wann wo abzulegen hat – und was vielleicht auch nicht. Vor diesem Hintergrund legten wir nicht nur einen eigenen Workshop zur Aktenrelevanz als neues Beratungsformat auf, der binnen kurzer Zeit ausgebucht war (zwei Termine im September 2024) und recht erfolgreich verlief.

Wir erstellten vielmehr ein weiteres Erklärvideo zum Thema, das inzwischen online verfügbar ist.⁶ Gerade der Datenschutz und in diesem Zusammenhang speziell die DSGVO führen unter den Beschäftigten der Landesverwaltung regelmäßig zu großer Verunsicherung, ob denn personenbezogene Daten überhaupt veraktet werden dürfen. Der Datenschutz spielt infolgedessen auch in den übrigen Veranstaltungen eine größere Rolle als zu Beginn unserer Beratungstätigkeit. Wenn wir die E-Akte als bereichsspezifisches Instrument des Informations- und Wissensmanagements vorstellen, stößt dies nicht zuletzt in Kontexten der Frage nach der Veraktung in Verbindung mit den Aspekten der Informations- und Datensicherheit immer wieder auf große

Screenshot aus Erklärfilm zu Datenschutz und Löschgeboten



Aufmerksamkeit. Wir stoßen hier geradezu in eine bisherige „Blankstelle“ vor und nehmen vielen Beschäftigten Ängste und Sorgen, die regelrecht mit den Händen zu greifen sind und vor allem im bilateralen Gespräch artikuliert werden.

5. REELS ALS ERGÄNZENDES FORMAT IM BERATUNGSPORTFOLIO

Zuletzt haben wir uns an der Form der Reels oder besser „Kurzvideos“ versucht. Hier möchten wir die Hürden bezüglich der Befassung mit dem Thema der Aktenführung nach Möglichkeit abbauen oder doch zumindest absenken, indem in Filmen mit einer maximalen Dauer von 90 Sekunden Themen „ange teasert“ werden, die unserer Einschätzung nach von großer Relevanz sind und die Beschäftigten dazu bewegen sollen, sich nach dem Konsum des „Infohappens“ bei Bedarf an anderer Stelle vertiefend mit der Materie auseinanderzusetzen.

Die Reels stehen somit nicht in Konkurrenz zum Format der animierten Erklärfilme, sondern ergänzen dieses idealerweise. Neben der Erste-Hilfe-Information können diese Kurzvideos auch in Veranstaltungen als Einstieg genutzt werden, der die Teilnehmenden schnell in ein Thema einführt und zum unmittelbaren Austausch einlädt.

In wenigen Sätzen lässt sich ein solches Kurzvideo wie folgt definieren: Ein Reel im Kontext der Behördenberatung ist eine kurze Lerneinheit, die flexibel und wiederverwendbar ist. Sie vermittelt erste essentielle Inhalte oder auch Fragestellungen in komprimierter Form und kurzer Zeit und ist flexibel einsetzbar. Die Idee zu diesen Reels speist sich im Wesentlichen aus Erfahrungen mit dem Web 2.0 und aus dem Lernen mit mobilen Endgeräten. Sie gehört in den

Kontext des Mikrolernens. Dieses zeichnet sich durch kurze Lernschritte von wenigen Sekunden bis hin zu wenigen Minuten aus. Die Lerninhalte werden pragmatisch einfach dargestellt. Diese Reels zu Themen der elektronischen Aktenführung bieten im Idealfall einen Anlass zum Lernen und zur vertiefenden Beschäftigung mit der dargestellten Materie. In besonderer Weise sind sie geeignet, neue Lernbereiche kennenzulernen und attraktive Lern-Settings zu kreieren. Entscheidend ist die Tatsache, ob beziehungsweise dass diese Aktenführungs-Reels seitens der Lernenden als didaktisch sinnvoll und als im Wesentlichen relevant wahrgenommen werden. Die Kurzvideos sollen in vorderster Linie das Interesse der Lernenden am zu vermittelnden Lerninhalt wecken.

6. FAZIT

Auch im achten Jahr seines Bestehens ist das Kompetenzteam des Landesarchivs NRW zur Behördenberatung daran interessiert, das Beratungsportfolio in jeglicher Hinsicht an neue Entwicklungen und Gegebenheiten anzupassen, das Beratungsangebot weiter aufzufächern und somit das Kanalisieren und Bündeln von Beratungsbedarfen zum Vorteil der Mandanten zu ermöglichen. Die hier vorgestellten Schritte sind als Komponenten zur Verwirklichung dieses Ansinnens zu verstehen. An einem Austausch innerhalb der archivischen Gemeinschaft besteht nach wie vor großes Interesse. Die Einladung zu einem solchen Diskurs ist hiermit erneuert.

Martin Schlemmer, Duisburg

- 4 Martin Herzog, Freier Mitarbeiter des WDR, teilte dem Autor dieses Beitrags am 26.09.2024 mündlich mit, dass er bei seinen für den WDR produzierten „Zeitzeichen“ – bei über 200.000 Hörern – häufig kein einziges Feedback erhalte und bereits eine einzige eingegangene Rückmeldung als „gut“ zu betrachten sei. Bei allen Unterschieden, die es zwischen einem Rundfunk-Feature und einer Schulung zur Aktenführung – mit einem proaktiv kommunizierten niederschweligen Feedback-Angebot – zu konstatieren gilt, ist die gemeinsame Erfahrung hervorzuheben, dass es recht mühsam ist, von einer letztlich anonym bleibenden Zielgruppe ein freiwilliges Feedback einzuholen, selbst wenn dieses sehr niederschwellig über einen einzigen Link online möglich ist.
- 5 Vgl. Mareikje Mariak, Transferarbeit im Rahmen der Laufbahnprüfung für den höheren Archivdienst an der Archivschule Marburg (57. Wissenschaftlicher Lehrgang). Government Audience Development. Wie steuern Archive ihre Verbindung zur Verwaltung?, vorgelegt am 01.04.2024 (unveröffentlichtes Manuskript).
- 6 Vgl. Erklärfilm zum Thema „Datenschutz und Löschvorgaben“, <https://www.archive.nrw.de/landesarchiv-nrw/behoerdeninformationen/aktenfuehrung/aktenfuehrung-erklavideos> (Abruf vom 07.10.2024).

HEUTE, MORGEN, ÜBERMORGEN – ARCHIVE IM FOKUS

91. DEUTSCHER ARCHIVTAG 2024 IN SUHL

Tagungsbericht von Klara Deecke

Ausgerichtet vom VdA - Verband deutscher Archivarinnen und Archivare e. V. fand vom 8. bis 10. Oktober 2024 der 91. Deutsche Archivtag statt. Über 500 Teilnehmende aus Deutschland, Österreich, der Slowakei, aus Tschechien und aus Ungarn reisten zum fachlichen Austausch nach Suhl in Thüringen. Zum zweiten Mal nach 2019 wurde die größte deutsche Archivkonferenz dort veranstaltet.

Das Kongressthema „Heute, Morgen, Übermorgen – Archive im Fokus“ verstand sich als Einladung zur Standortbestimmung innerhalb sich rasch ändernder Rahmenbedingungen bzw. zum Austausch über Möglichkeiten der Weiterentwicklung. Inhaltlich breit gefasst bot das Rahmenthema ausreichend Spielraum, um im Fachprogramm ganz unterschiedliche Themen aufzugreifen. Wendepunkte im Archivwesen und aus archivischer Perspektive wurden an den drei Tagen intensiv diskutiert, beispielsweise die Auswirkungen der Digitalisierung, Künstlichen Intelligenz und des Klimawandels. Angesichts der jüngsten Landtagswahlergebnisse in mehreren Bundesländern, darunter Thüringen, und 35 Jahre nach der friedlichen Revolution in der DDR wurden auch Folgen der gesellschaftlichen und politischen Veränderungen der Demokratie mit Bezug zum Archivwesen in den Blick genommen.

ERÖFFNUNGSVERANSTALTUNG

Offiziell eröffnet wurde der 91. Deutsche Archivtag am Abend des 8. Oktober vom Vorsitzenden des VdA Ralf Jacob. Er verwies mit Nachdruck auf die Notwendigkeit von Archiven und ihre identitätsstiftende Funktion in der demokratischen Gesellschaft. Weiter skizzierte Jacob die rasante Veränderung des Arbeitsalltags und seiner Rahmenbedingungen mit nicht absehbaren Auswirkungen auf das archivische Berufsfeld, die auf dem Archivtag diskutiert werden sollten. Es folgten Grußworte des Oberbürgermeisters der Stadt Suhl André Knapp und der Präsidentin des Verbands Österreichischer Archivarinnen und Archivare Karin Sperl als Vertreterin der ausländischen Archivtagsteilnehmenden.



Markus Hilgert bei der Eröffnungsrede. Foto: VdA

Den Eröffnungsvortrag hielt Markus Hilgert, Generalsekretär der Kulturstiftung der Länder. Ausgehend von der aktuellen gesellschaftlichen Umbruchsituation der Zeitenwende identifizierte Hilgert Archive als wichtigen Faktor für die Widerstandsfähigkeit einer Gesellschaft und leitete daraus die Notwendigkeit ab, diese besonders zu stärken und zu fördern. Seine These „Eine resiliente Gesellschaft braucht starke Archive!“ konnte Hilgert als Altorientalist mit historischen Beispielen seit dem antiken Babylonien untermauern. Als Generalsekretär der Kulturstiftung der Länder betrachtete er sie aber auch aus kulturpolitischer und kulturfördernder Perspektive. Historizität, Authentizität, Integrität, Legitimität und Kontinuität seien fünf zentrale Güter, die Archive garantierten. Diese seien zugleich essenziell für freiheitlich-demokratische, rechtsstaatliche und plurale Gesellschaften. Als wesentliche kulturelle und gesellschaftliche Stabilitätsfaktoren brauche es für die Archive spezielle

Förderprogramme, die ihr Funktionieren sicherten und sie vor Naturkatastrophen, Sabotage und bewaffneten Konflikten schützten. Dafür sei es unabdingbar, bei den politisch Verantwortlichen und den Archivträgern ein Bewusstsein für die Relevanz von Archiven zu schaffen und dann zielgerichtete Förderprogramme zu entwickeln. Runde Tische und andere strategische Austauschformate mit den politischen Verantwortungsträgern könnten dazu beitragen. Hilgert konstatierte ein Missverhältnis zwischen der Bedeutung von Archiven und ihrer potenziellen Gefährdung einerseits und ihrer geringen finanziellen und politischen Förderung andererseits. Am Beispiel der Kulturstiftung der Länder zeigte er auf, dass deren Förderaktivitäten sich zwar auch an Archive richteten, aber ohne ein speziell auf Archive zugeschnittenes Programm. Bei der Analyse der Situation blieb Hilgert jedoch nicht stehen. Er richtete das konkrete Angebot an den VdA, finanzielle Unterstützung der Kulturstiftung der Länder für die Entwicklung und Umsetzung strategischer Austauschformate für die nachhaltige Stärkung von Archiven bereitzustellen. Außerdem stellte Hilgert die Entwicklung eines Förderprogramms der Kulturstiftung in Aussicht, das sich ausschließlich an Archive richtet – auch, um damit weiteren Förderbedarf illustrieren zu können.



Blick ins Auditorium. Foto: VdA

FACHPROGRAMM

Plenum 1: Archivierung audiovisueller Unterlagen

Das erste Fachforum des Archivtags befasste sich unter der Moderation von Undine Beier (Berlin) mit der Archivierung von AV-Medien. Die Notwendigkeit der Sicherung der teilweise fragilen Materialien, so Beier, setze die Archivarinnen und Archivare unter Handlungsdruck, sei dabei aber sehr kosten- und ressourcenintensiv. Die Vorträge führten dabei auch vor Augen, wie unterschiedlich der Arbeitsalltag mit AV-Medien abhängig von Größe und Ausstattung des Archivs sein kann. Ralf Springer (Münster) berichtete über die Einführung des Media Asset Management-Systems „Cube Collections“ im Filmarchiv des LWL-Medienzentrums für Westfalen. Die Spezialentwicklung ermöglicht das automatisierte Auslesen von technischen Metadaten sowie die digita-

le Speicherung, Rechteverwaltung, inhaltliche Erschließung, Verwaltung der analogen Datenträger und Recherche. Andrea Hänger (Koblenz) stellte den Digitalen Lesesaal Film des Bundesarchivs vor und schilderte den Ablauf seiner Einführung als Teilprojekt der Vereinheitlichung der Rechtersysteme für Filme, Fotos und Akten im Bundesarchiv. Johannes Beermann-Schön (Frankfurt am Main) vom Archiv des Fritz Bauer Instituts hingegen brachte die Perspektive kleiner Archive ein, die mit vergleichsweise geringer technischer und finanzieller Ausstattung bei der Archivierung von AV-Medien vorankommen wollen. Zwar seien die Archivierung von AV-Unterlagen und damit in Verbindung stehende Prozesse wie Rechtereklärungen sehr zeitaufwendig. Doch Ton und Bild seien sehr nachgefragt, so dass AV-Medien Aufmerksamkeit für die eigene Institution generieren könnten. Vor dem schier unüberwindlichen Berg an Arbeit dürfe man nicht zurückschrecken, sondern könne pragmatisch und mit langem Atem allmählich Fortschritte erzielen. Beermann-Schön empfahl, sich an bereits existierenden Lösungen, Leitfäden und Handreichungen zu orientieren.



Plenum 1: Andrea Hänger, Moderatorin Undine Beier, Johannes Beermann-Schön. Foto: VdA

Plenum 2: Digitale Unterlagen

Das zweite Plenum diskutierte aus organisatorischer Perspektive über Aspekte digitaler Archivierung und wurde moderiert von Martin Schlemmer (Duisburg). So stellte der Vortrag von Franziska Klein und Benedikt Nientied (beide Duisburg) die Begleitung der Verwaltungsdigitalisierung als integriertes Gesamtprojekt im Landesarchiv Nordrhein-Westfalen (NRW) vor. Nientied, der den Vortrag allein präsentierte, erläuterte mit der Schulung zentraler Personen, Sensibilisierung der Führungskräfte und einer Handreichung zur Aktenrelevanz verschiedene begleitende Maßnahmen und Angebote der Landesarchivverwaltung NRW zur E-Akten-Einführung, die auch die Sichtbarkeit und das Standing des Landesarchivs innerhalb der Landesverwaltung erhöht hätten. Annette Birkenholz und Arnold Otto (beide Nürnberg) sprachen über die Digitalstrategie für das Stadtarchiv Nürnberg und die Umsetzung des Strategiepapiers in einem viele Bereiche umfassenden Prozess ausgehend von den Arbeitsgrundlagen über Prozesse, Ressourcenplanung, Audience Development hin zu konservatorischen Maßnah-

men, Übernahme, Benutzung und Methoden. Silvia Josties und Mareile Mansky (beide Würzburg) stellten in ihrem Vortrag die Perspektive der abgebenden Stelle neben die des Archivs und schilderten jeweils aus ihrer Sicht Herausforderungen und Erfolgsfaktoren in der Verwaltungsdigitalisierung der Universität Würzburg. Sie gaben dem Publikum mit auf den Weg, den Faktor Mensch nicht zu unterschätzen, sondern zu versuchen, alle Beteiligten mitzunehmen und ein digitales Mindset zu fördern. Digitalisierung sei nicht einfach nur Einscannen, sondern Digitalisierung müsse ganzheitlich gedacht werden, wobei die Archivierung von Beginn an mitzuberücksichtigen sei. Wichtig sei es u. a., eine gemeinsame Sprache zu finden, Probleme und Bedenken ernst zu nehmen, aber auch den Nutzen der Digitalisierungsprozesse herauszuarbeiten, Erfolge zu vermarkten und mutig zu sein und zu bleiben.



Plenum 2. Foto: VdA

Plenum 3: Neue Arbeits- und Ausbildungsmodelle

Michael Häusler (Berlin) führte durch das dritte Plenum des Fachprogramms. Martin Schlemmer (Duisburg) analysierte die Homeoffice-Landschaft im öffentlichen Dienst vor dem Hintergrund der Coronapandemie und schilderte die Praxis im Landesarchiv Nordrhein-Westfalen, die sich seitdem herausgebildet habe und gerade in einer Dienstvereinbarung zur alternierenden Telearbeit geregelt worden sei. Die Flexibilisierung der Arbeitsbedingungen sei mitarbeiterfreundlich und mache die Arbeitsstelle attraktiver. Schlemmer wies aber auch auf das strukturelle (Gerechtigkeits-)Problem hin, dass untere Gehaltsgruppen aufgrund ihrer Tätigkeiten in der Regel weniger im Homeoffice arbeiten könnten als höhere, da die dort häufigeren Verwaltungsaufgaben sowie konzeptionelle, vorbereitende, redaktionelle und planerische Arbeiten weniger Anwesenheit vor Ort erforderten. Elke-Ursel Hammer (Koblenz) berichtete über eine Antwort des Bundesarchivs auf den Fachkräftemangel: den verkürzten Vorbereitungsdienst im gehobenen Dienst. Das Bundesarchiv habe einen jährlichen Bedarf von 20 Archivarinnen und Archivaren des gehobenen Dienstes, doch die Kapazitäten der Ausstellungstätten seien begrenzt. So würden an der

Archivschule Marburg für die entsprechende Laufbahn jährlich nur 25 Personen ausgebildet. Viele der Bewerberinnen und Bewerber auf eine Ausbildung im gehobenen Dienst verfügten bereits über ein abgeschlossenes Studium in Geschichte, die nun aufgrund eines Erlasses der Beauftragten der Bundesregierung für Kultur und Medien in einem eigenen Kurs einen verkürzten und inhaltlich auf das Bundesarchiv zugeschnittenen Vorbereitungsdienst absolvieren könnten – ein, so Hammers Fazit, zukunftsweisendes und auch für die Nachqualifizierung von Quereinsteigern geeignetes, wenn auch auf kleinere Einrichtungen nicht ohne Weiteres übertragbares Modell.

Plenum 4: Nachhaltigkeit als Auftrag. Archive in der Klimakrise

Das vierte Fachforum setzte sich mit der Bedeutung von Klimawandel und Nachhaltigkeit im Archivwesen auseinander. Es wurde von Johannes Rosenplänter (Kiel) moderiert, der den in diesem Plenum erfolgenden Blick auf Museen und Bibliotheken aus deren Vorreiterrolle im Umgang mit Klima- und Energiekrise und Nachhaltigkeit herleitete. So stellte Hanna Belz (Berlin) den Prozess vor, mit dem die Stiftung Preussischer Kulturbesitz ihre Vision von ökologischer, sozialer und ökonomischer Nachhaltigkeit entwickeln und umsetzen möchte. Um das Ziel der CO₂-Neutralität der Stiftung Preussischer Kulturbesitz vor 2035 zu erreichen, werde das Umweltmanagementsystem EMAS (Eco Management and Audit System) eingeführt. Bereits erfolgt sei eine „Grüne Inventur“, die Erstellung einer Klimabilanz. Eine erfolgreiche signifikante CO₂-Reduktion sei nur möglich, wenn bekannt sei, wo CO₂-Ausstoß erzeugt werde, woraus sich dann konkrete Maßnahmen ableiten ließen. Als Beispiele im Geheimen Staatsarchiv Preussischer Kulturbesitz nannte Belz die Verlängerung der Nachtabenkung der Heizung, die Reduzierung der Anzahl der Leuchten und die Entfernung von Dauerarbeitsplätzen aus dem Magazin.



Plenum 4: Johannes Rosenplänter, Hanna Belz, Laura Scherr. Foto: VdA

Laura Scherr (München) sprach über Nachhaltigkeit und Klimaneutralität im Archivbau. Ausgehend vom Auftrag der Archive, Kulturgut gegen schädliche Einflüsse zu schützen, rief sie die unmittelbaren Folgen des Klimawandels für den Erhalt des Kulturguts in einer immer feindlicheren Umwelt in Erinnerung. Notfallvorsorge und Stabilität auch bei Energieausfällen und im Klimawandel seien unabdingbar. Gleichzeitig müsse aktiv zum Klimaschutz beigetragen und den Verpflichtungen aus den Klimaschutzgesetzen nachgekommen werden. So müsse die Bayerische Staatsverwaltung Klimaneutralität bis zum Jahr 2028 erreichen. Mit dem Fokus auf Schadensprävention oder den langen Gebäudenutzungen seien Archive bereits recht nachhaltig aufgestellt, wobei dem Thema Archivbau eine Schlüsselrolle zukomme und man sich dort möglichst nah an den Passivhausstandard annähern müsse. Auch wenn es oftmals ohne Klimaanlage nicht gehe, könne diese auch nur zum bedarfsweisen Gegensteuern eingesetzt und möglichst viel über passive bauliche Maßnahmen erreicht werden. Diese sollten durch organisatorische Schritte wie die Beschränkung von Magazinführungen auf Kleingruppen und Magazinreinigung mit möglichst wenig Feuchtigkeit ergänzt werden. An verschiedenen aktuellen Beispielen wie dem Neubau der Staatsarchive Landshut und Würzburg und dem Magazinneubau für das Bayerische Hauptstaatsarchiv in der Leonrodstraße in München zeigte Scherr auf, wie ausgehend von den jeweiligen Gegebenheiten und aufgrund genauer bauphysikalischer Analyse Potenziale durch Photovoltaik, eine Holzpelletsanlage, eine Grundwasserwärmepumpe, gute Dämmung oder Klimaschleusen möglichst genutzt werden könnten. Sie verschwieg dabei aber auch nicht, dass gerade im Bereich der Altbausanierung schnell Grenzen erreicht würden und auch Zielkonflikte etwa mit dem Denkmalschutz aufgetreten seien. Abschließend übernahm Rosenplänter die auszugswise Präsentation des kurzfristig verhinderten Tim Schumann (Berlin) zum Erstellen von Klimabilanzen in Kultureinrichtungen. Schumann hatte zusammen mit einem Projektteam für eine Berliner Stadtteilbibliothek eine CO₂-Bilanz erstellt, auch hier nach dem Credo „What you measure is what you manage“. Heizung, Strom und mit Abstrichen auch Arbeitswege konnten als die wichtigsten Faktoren identifiziert werden. Gleichzeitig sei mit dem Projekt auch für das Thema Energiesparen und Klimaschutz sensibilisiert worden und Befürchtungen bei den Mitarbeitenden hätten besprochen werden können. In der anschließenden lebhaften Diskussion wurde deutlich, dass gerade im Kulturbereich nachteilige Folgen von CO₂-Bilanzierung befürchtet werden, da Besuchermobilität hohe CO₂-Emissionen verursacht. Manche Einrichtungen wählten daher statt oder zusätzlich zur Ermittlung des CO₂-Fußabdrucks der erzeugten Emissionen das Konzept des Handabdrucks, bei dem Einsparungen und Fortschritte aufgelistet werden. Zugleich zeigte sich in der Diskussion, dass bereits einige Archive CO₂-Bilanzen erstellt haben und dies als positiv bewerteten, so das Hessische Landesarchiv mithilfe des online frei verfügbaren CO₂-Kulturrechners auf Basis des CO₂-Kulturstandards, den die Kultusministerkonferenz zur Anwendung im Kulturbereich, auch für Archive, empfiehlt.



Plenum 4. Foto: VdA

Plenum 5: KI, Machine Learning und Algorithmen

Die letzte Plenarveranstaltung unter der Moderation von Bettina Joergens (Duisburg) war eine Veranstaltung des Ausschusses „Archivische Fachinformationssysteme“ der Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA); sie bildete zugleich den Abschluss des Archivtags. Das Panel fragte nach der Definition und den Einsatzmöglichkeiten von KI im Archivwesen, nach ersten positiven wie negativen Erfahrungen, Erfolgen und Herausforderungen, aber auch nach den strategischen Überlegungen mit Blick auf rechtliche, organisatorische und finanzielle Rahmenbedingungen, die dem Einsatz von KI vorausgehen müssen, d. h. nach der KI-Strategie von Archiven. Im ersten Vortrag sprach Maria Schmalenbach (Düsseldorf) von IT.NRW über Einsatzszenarien von KI in der öffentlichen Verwaltung und wies mit dem Zitat „AI or bye bye“ (Carsten Maschmeyer) darauf hin, die Bedeutung einer proaktiven Auseinandersetzung mit und Einbeziehung von KI nicht zu unterschätzen. Sie stellte das Beispiel des digitalen Verwaltungsassistenten NRW.Genius vor, plädierte aber auch dafür, sowohl verwaltungsintern wie in der Kommunikation nach außen Chatbots stärker zu nutzen. Software-Roboter, welche automatisiert Arbeitsschritte erledigen könnten, die wegen fehlender Schnittstellen zwischen unterschiedlichen genutzten Tools bisher ein Sachbearbeiter übernehmen musste, seien ebenfalls wichtige Helfer angesichts begrenzter Personalressourcen. Benjamin Rosemann (Stuttgart) referierte über KI im praktischen Einsatz bei der Aufbereitung und Nutzung archivischer Daten im Landesarchiv Baden-Württemberg. Er führte aus, dass es relativ einfach sei, kleine Testprojekte mit KI zu realisieren. Die flächendeckende Einführung im Arbeitsalltag sei hingegen mit ungleich größeren Schwierigkeiten wie fehlender Hardware, Software und fehlendem technischen Personal verbunden. Schwache, d. h. kleine, nur für eine spezielle Aufgabe nutzbare, aber dadurch handliche und energie- sowie hardwaresparsame KI-Modelle könnten für den Archibereich eine gute Lösung sein. Esther Lemmerz (Koblenz) stellte erste Erfahrungen aus dem KI-Programm des Bundesarchivs vor, in dem konkrete Einzelprojekte zu verschiedenen archivischen Ein-

satzmöglichkeiten von KI umgesetzt würden, auch mit dem Ziel, die dabei gewonnenen Erfahrungen in eine KI-Strategie einfließen zu lassen.



Plenum 5: Bettina Joergens, Maria Schmalenbach, Esther Lemmerz und Benjamin Rosemann. Foto: VdA

SITZUNGEN DER ARBEITSKREISE UND FACHGRUPPEN

Sitzungen der Arbeitskreise

Bereits vor der abendlichen Eröffnungsveranstaltung des Deutschen Archivtags veranstalteten die Arbeitskreise des VdA ihre Sitzungen. Der Arbeitskreis Offene Archive beschäftigte sich in einem von Antje Diener-Staeckling (Münster), Joachim Kemper (Aschaffenburg) und Johannes Waldschütz (Rottweil) moderierten Archivcamp u. a. mit den Themen Digitalisierung, Erschließung durch Hilfskräfte oder Mentoring.



Archivcamp des Arbeitskreises Offene Archive. Foto: VdA

Katharina Tiemann (Münster) moderierte den Workshop des Arbeitskreises Ausbildung und Berufsbild, der sich mit Ergebnissen aktueller Abschluss- und Projektarbeiten befasste. Bettina Nohr (Werne) gewährte einen Blick hinter die Kulissen der FaMI-Con 2023, einer Tagung von Fachangestellten für Medien- und Informationsdienste, während Franziska Ernst (Potsdam) ihre Bachelorarbeit über Gedenkstättenarchive vorstellte. Sarah Masiak (Detmold) hatte in ihrer Transferarbeit eine Orientierungshilfe für den Umgang mit dem postmortalen Persönlichkeitsschutz bei Tatortfotos entwickelt, die sie im letzten Referat des Workshops präsentierte. Der Arbeitskreis Archivpädagogik und Historische Bildungsarbeit verfolgte unter der Moderation von Annekatrin Schaller (Neuss) das Ziel, Akteure der historisch-politischen Bildungsarbeit aus Thüringen zu Wort kommen zu lassen und danach zu fragen, was historische Bildungsarbeit für die Stärkung des demokratischen Gemeinwesens leisten kann. Judith Mayer und Katja Mann (beide Erfurt) gaben Einblick in die Aktivitäten der Gedenk- und Bildungsstätte Andreasstraße Erfurt. Zu den Themen Haft, Diktatur und friedliche Revolution werden dort Objekte zur DDR-Repressionsgeschichte gezeigt. 8.000 Schülerinnen und Schüler jährlich besuchen die Gedenkstätte. Im zweiten Vortrag zeigte Laura Wesseler (Hamburg) aus Perspektive der Körber-Stiftung, welche Bedeutung Archive und Archivmitarbeitende als Partner für den Geschichtswettbewerb des Bundespräsidenten und damit eine demokratische Geschichtskultur haben können. Sie hob hervor, dass Archive die außerschulischen Lernorte mit der größten Bedeutung für den Geschichtswettbewerb seien, die authentische Quellen bieten und Multiperspektivität und die Grundlagen historischen Lernens vermitteln könnten. Im abschließenden Vortrag stellte Katharina Kempken (Jena) das Thüringer Archiv für Zeitgeschichte „Matthias Domaschk“ vor, ein Spezialarchiv für Oppositionsbewegungen im Thüringer Raum. Kempken schilderte, wie die Einrichtung durch ihre historische Bildungsarbeit das Ziel verfolge, junge Menschen darin zu unterstützen, selbständig mit verschiedenen Deutungs- und Identifikationsangeboten der DDR-Geschichtskultur umzugehen.

Sitzungen der Fachgruppen

Zeitgleich fanden am Vormittag des 10. Oktober die Fachgruppensitzungen statt, weshalb im Folgenden nur die Sitzung der Fachgruppe 2 Kommunalarchive ausführlich behandelt werden kann. Unter der Leitung der Fachgruppenvorsitzenden Katharina Tiemann (Münster) stand die Sitzung unter dem Leitthema „Die Rolle der Archive in der Demokratie – Garanten für Rechtsstaatlichkeit und Transparenz?“ Nachdem Marcus Stumpf (Münster) Neuigkeiten aus der Bundeskonferenz der Kommunalarchive (BKK) weitergegeben hatte, widmeten sich vier Fachvorträge dem Verhältnis von Archiven und Demokratie aus unterschiedlichen Perspektiven. Als Kommunalarchivarin mit 48 Jahren Berufserfahrung im Archiv sowie aus persönlicher Sicht zeichnete Sigrid Unger (Oelsnitz/Vogtland) den Wandel des archivischen Selbstverständnisses im Umbruch von der DDR zur BRD nach; fast auf den Tag genau, als im Zuge der friedlichen Revolution vor 35 Jahren auch in Plauen, der da-

maligen Arbeitsstätte Ungers, eine Demonstration durch die Stadt zog. In der DDR sei das Archivwesen zwar ideologisch in die Pflicht genommen worden, etwa mit den Aufgaben, die marxistisch-leninistische Gesellschaftswissenschaft zu unterstützen und die marxistisch-leninistische Archivwissenschaft weiterzuentwickeln, aber es habe auch Spielräume gegeben. Gerade im Kommunalarchiv sei es auch darum gegangen, den authentischen Alltag und die Lebensrealität zu dokumentieren. 1989/90 sei Unger klar gewesen, dass DDR-Quellen durch das Archiv zu sichern waren; es gab in dieser Zeit nicht nur verstärkt Anbietungen, sondern auch Aktenvernichtungen. Am Beispiel von Zwangsaussiedlungsakten machte Unger deutlich, wie wichtig die Überlieferung von DDR-Dokumenten gewesen sei. Zugleich sei der Systemwechsel von der DDR zur BRD für die im Archivwesen Tätigen eine Zeit der beruflichen Unsicherheit gewesen. Unger wies darauf hin, dass der DDR-Abschluss als Archivassistent bis heute nicht anerkannt sei. Sie dankte für die kollegiale Unterstützung der altbundesdeutschen Archivarinnen und Archivare, die in der Wendezeit eine wichtige Hilfe und Bestätigung gewesen sei. Das archivische Selbstverständnis habe sich im Vergleich von DDR und BRD in der Hinsicht gewandelt, dass die marxistisch-leninistische Doktrin nicht mehr gelte, in anderer Hinsicht aber nicht, vor allem das handwerkliche Können und die praktische Archivarbeit betreffend.



Fachgruppe 2: Katharina Tiemann, Sigrid Unger, Harald Stockert, Michael Scholz und Susanne Rappe-Weber. Foto: VdA

Harald Stockert (Mannheim) näherte sich dem Verhältnis von Archiv und demokratischem Rechtsstaat über die Themen E-Government, Datenschutz und Compliance. Anhand der Praxis im Marchivum zeigte er Möglichkeiten und Grenzen der archivischen Unterstützung einer funktionierenden rechtskonformen Schriftgutverwaltung auf. Stockert prognostizierte für die Zukunft eine Vereinfachung der Schriftgutverwaltung selbst, aber kurzlebigere und anspruchsvollere technische und rechtliche Rahmenbedingungen und damit neue und zeitintensive Herausforderungen für E-Akten-Verantwortliche im Archiv. Michael Scholz (Potsdam) rief in seinem archivrechtlichen Vortrag ins Bewusstsein, dass viele rechtliche Benutzungsfragen des Archivalltags Abwägungen zwischen Informati-

onszugangsfreiheit, Forschungsfreiheit und Datenschutz und damit letztlich Abwägungen zwischen verfassungsmäßig verankerten gleichrangigen Grundrechten seien. Bei der Kollision von Grundrechten seien Abwägungen unabdingbar, wobei kein Grundrecht im Zuge der Abwägung weiche, sondern immer nur im unvermeidlichen Umfang begrenzt werden dürfe. Mit den Schutzfristregelungen habe der Gesetzgeber eine Grundrechtsabwägung für den Normalfall vorgenommen. Trotzdem müssten Archivarinnen und Archivare bei Bedarf eigenständig abwägen, beispielsweise bei personenbezogenen Daten in Sachakten oder Schutzfristverkürzungen. „So viel Benutzung wie möglich, so viel Persönlichkeitsschutz wie nötig“, nannte Scholz dafür als Faustformel.

Abschließend widmete sich Susanne Rappe-Weber (Burg Ludwigstein, Witzenhausen) der Frage der Wahrnehmung von Archiven in Unrecht-Aufarbeitungsprozessen, insbesondere bei Betroffenen von sexualisierter Gewalt in Institutionen. Das Handeln nach Recht und Gesetz und die Neutralität von Archiven in Verbindung mit ausstehender Erschließung oder fehlender Anbietung jüngerer einschlägiger Unterlagen könne in der subjektiven Wahrnehmung von Betroffenen nachvollziehbarerweise den Eindruck von Behinderung der Aufklärung erwecken. Nach Rappe-Weber lägen aber in der gesteigerten Aufmerksamkeit für Archive infolge der Aufarbeitung von sexualisierter Gewalt an Kindern und Jugendlichen mehr Chancen als Nachteile für die Archive. Jeder Fall, der aufgeklärt werde, jede Akte, die eine Tatsache belege, jeder Nutzende, der bei der Recherche unterstützt werde, mache für die betroffenen Personen einen Unterschied, in juristischer, finanzieller oder psychologischer Hinsicht. Sich an die Seite der Betroffenen zu stellen und sie bestmöglich zu unterstützen sei daher wichtig, damit Archive für Betroffene nicht häufiger ein Problem, sondern eine Hilfe darstellten. In der alle vier Vorträge umfassenden Abschlussdiskussion wurde resümiert, dass Archive den Anspruch, Rechtsstaatlichkeit und Transparenz zu garantieren, alleine nicht einlösen, sondern höchstens einen Beitrag dazu leisten können. Auch die anderen Fachgruppen diskutierten für sie einschlägige aktuelle Themen: Fachgruppe 1 – Staatliche Archive tat dies mit Vorträgen zu Records in Context, einem archivrechtlichen Thema sowie zur archivischen Vorfeldarbeit, Fachgruppe 3 – Kirchliche Archive mit Referaten zur Digitalisierung und Onlinestellung von Archivgut sowie zur Webseitenarchivierung. Fachgruppe 4 – Herrschafts- und Familienarchive tagte zusammen mit Fachgruppe 5 – Wirtschaftsarchive mit Vorträgen zur Aufarbeitung der NS-Zeit in Adels- und Privatarchiven sowie zu Künstlicher Intelligenz im Archiv. Fachgruppe 6 – Archive der Parlamente, der politischen Parteien, Stiftungen und Verbände befasste sich vorwiegend mit archivbaulichen Themen, während Fachgruppe 7 – Medienarchive und Fachgruppe 8 – Archive der Hochschulen sowie wissenschaftlicher Institutionen gemeinsam eine Fishbowl-Diskussion zum Umgang mit emotional belastenden Beständen veranstalteten. Über die letzte Veranstaltung berichtet Christian George im nachfolgenden Artikel.

WEITERE PROGRAMMPUNKTE



Internationaler Austausch. Foto: VdA

Das Newcomer-Forum für neue Mitglieder im VdA oder der Internationale Austausch mit Tagungsteilnehmenden aus dem Ausland waren bewährte Formate, die auch in Suhl nicht fehlen durften, wie die Aktuelle Stunde. Moderiert von Ewald Grothe (Gummersbach) informierten dort Johannes Haslauer (Bamberg) und Nathalie Huck (Berlin) über die Ergebnisse der Tagung des Ausschusses „Archivische Fachinformationssysteme“ der KLA zu Records in Contexts, während Andreas Neuburger (Stuttgart) über die Arbeit von NFDI4Memory berichtete. Anne Thurmman-Jajes (Bremen) stellte die Kulturarchive vor, die demnächst eine eigene Fachgruppe im VdA bilden sollen. Im Workshop für Fachangestellte für Medien- und Informationsdienste, der von Evi Milz und Christiane Bruns (beide Berlin) moderiert wurde, berichteten Britta Hemme (Bergen), Roman Edner (Berlin) und Christian Rieck (Schwerin) von Erfahrungen im Berufsleben von FaMIs mit Faktoren, die zur Gewinnung und Bindung von Fachkräften beitragen wie die Vereinbarkeit von Familie und Beruf, Homeoffice, Work-Life-Balance oder Sabbaticals.



Workshop Fachangestellte für Medien- und Informationsdienste. Foto: VdA

Als Sonderveranstaltung wurde, moderiert von Kristina Starkloff (Berlin), das Projekt EMILIA zur Entwicklung einer teilautomatisierten Software für die Archivierung von E-Mails präsentiert. Nico Beyer und Felix Gericke (beide Berlin) erläuterten u. a. den Projektrahmen und die Projektziele und zeigten Einsatzmöglichkeiten des Produkts auf. Auf besonderes Interesse stieß der in Suhl aufgebaute Notfallanhänger „Ich rette (F)Akten. Mobile Notfallhilfe!“ des Bundesarchivs, mit dem unkompliziert Ausrüstung und Notfallmaterialien an die Magazinstandorte des Bundesarchivs transportiert werden können. An jedem Veranstaltungstag der Konferenz fand dazu eine halbstündige Präsentation statt, in der Jens Niederhut, Thilo Backhaus und Nina Eckardt (alle Berlin) sowohl die Materialien im Notfallanhänger als auch organisatorische Aspekte anschaulich erläuterten.



Der Notfallanhänger des Bundesarchivs. Foto: VdA

FACHMESSE ARCHIVISTICA 2024 SUHL

Mit dem Deutschen Archivtag verbunden war die Archivistica als größte europäische Fachmesse für das Archivwesen. Über 30 Aussteller waren mit ihren Ständen im Kongresszentrum in Suhl präsent und informierten und berieten die Besucherinnen und Besucher des Archivtags über ihre Produkte und Dienstleistungen oder führten technische Neuentwicklungen vor. Vertreten waren beispielsweise Hersteller von Archivverpackungsmaterial, Softwarehersteller und Archivierungsdienstleister, Scandienstleister und Anbieter von Digitalisierungstechnik, Anbieter von Archivrecherchen, Rechercheportalen und anderen archivischen oder archivnahen Dienstleistungen, Bestandserhaltungsfirmen, Regalanbieter, aber auch archivische Ausbildungseinrichtungen. Zwei moderierte Ausstellerforen auf dem Deutschen Archivtag gaben Herstellern und Dienstleistern zudem die Möglichkeit, in halbstündigen Präsentationen ihre Angebote vorzustellen und Fragen potenzieller Kundinnen und Kunden zu beantworten.



Fachmesse Archivistica. Foto: VdA

RAHMENPROGRAMM

Wer den Deutschen Archivtag besucht, erwartet in der Regel nicht nur inhaltliche Impulse aus dem Fachprogramm, sondern schätzt auch den persönlichen Austausch, die Gelegenheit, Kontakte zu knüpfen und zu pflegen, sowie die Möglichkeit, die Kongressstadt näher kennenzulernen. Das Rahmenprogramm und Suhl, die zweitkleinste kreisfreie Stadt Deutschlands – ein Aspekt, den Oberbürgermeister André Knapp in seiner Begrüßung bei der Eröffnungsveranstaltung humorvoll hervorhob –, bildeten dafür die passende Basis. So wurden verschiedene Führungen angeboten, unter anderem durch das Stadtarchiv Suhl mit vielen spannenden Informationen zur Stadtgeschichte und den Archivbeständen. Dabei gab es selbst für Teilnehmende, die bereits den Archivtag 2019 in Suhl besucht hatten, viel Neues zu erleben, war das Stadtarchiv doch zwischenzeitlich in das Haus der Geschichte umgezogen. Ein weiteres Highlight war der Empfang im Anschluss an die Eröffnungsveranstaltung. Hier sorgte der gekrönte Thüringer Bratwurstkönig mit einem launigen Vortrag für Unterhaltung und präsentierte die Thüringer Bratwurst als echtes Kulturgut. Passend dazu war die Versorgung mit den berühmten Bratwürsten reichlich, während vegetarische Alternativen eher eine Nebenrolle spielten.

FAZIT

Der Deutsche Archivtag 2024 in Suhl überzeugte mit seinem bewusst offen formulierten Rahmenthema „Heute, Morgen, Übermorgen – Archive im Fokus“. Verschiedene die Zukunft des Archivwesens bestimmende Themen konnten damit auf einer Konferenz zusammengeführt werden: Digitalisierung, Umgang mit dem Fachkräftemangel, Künstliche Intelligenz, Nachhaltigkeit und Klimawandel. Hinzu kam ein Eröffnungsvortrag, der weit mehr war als ein motivierender Impuls für das kommende Tagungsprogramm oder gar nur Seelenbalsam für den Berufsstand. Markus Hilgert skizzierte eine klare Strategie für die kulturpolitische Stärkung der Archive und stellte diese unter das prägnante Leitmotiv: „Eine resiliente Gesellschaft braucht starke Archive!“ Dabei verband er seine Botschaft mit konkreten Angeboten – bis hin zur Aussicht auf finanzielle Förderung –, um die politi-

sche und gesellschaftliche Wahrnehmung von Archiven zu schärfen und ihre Bedeutung für die Demokratie zu betonen. Auch die anderen Formate der Tagung, insbesondere die Veranstaltungen der Arbeitskreise und Fachgruppen griffen



Das Tagungsteam: Ida Hofmeister, Kai-Uwe Schellenberg, Isabel Chalamel, Tim Czemper, VdA-Geschäftsführer Thilo Bauer, VdA-Vorsitzender Ralf Jacob, Barbara Sauer, Hermann-Josef Klüber, Manuela Meyer-Grandpierre. Foto: VdA

hochrelevante Themen auf. Ein Schwerpunkt lag dabei auf der Rolle von Archiven, ihren Aufgaben und Aktivitäten in einer freiheitlich-demokratischen Gesellschaft. Gerade im Jubiläumsjahr der friedlichen Revolution und angesichts aktueller politischer Herausforderungen waren diese Diskussionen wichtig. Mit der Konferenz in Suhl, an der fast 80 Vortragende bzw. Moderatoren/Moderatorinnen aktiv mitwirkten, löste der Deutsche Archivtag seinen Anspruch ein, die wichtigste Konferenz des deutschen Archivwesens zu sein. Und dies nicht nur in inhaltlicher Hinsicht: Die Tagung war sehr gut organisiert, die Mitarbeitenden der VdA-Geschäftsstelle und das Tagungsteam gewährleisteten einen reibungslosen Ablauf, die angeschlossene Fachmesse Archivistica bot den Teilnehmenden zusätzliche Informationen und Angebote. Suhl als Tagungsort mag zwar für den einen oder die andere nicht ganz leicht zu erreichen gewesen sein, konnte aber u. a. durch kurze Wege, die idyllische Lage im Thüringer Wald, der sich in leuchtenden Herbstfarben zeigte, und ein zentral gelegenes, optimal dimensioniertes Kongresszentrum überzeugen. Als weiteres Plus erwies sich erneut die bereits 2023 erfolgte Abkehr vom früheren Format mit parallel stattfindenden Sektionssitzungen zugunsten von Plenarveranstaltungen. Die lineare Abfolge der Vorträge bot den Teilnehmenden eine entspanntere Tagungserfahrung ohne hektische Raumwechsel oder das Gefühl, spannende Beiträge zu verpassen. Wie bereits 2023 eingeführt, werden die in Suhl gehaltenen Vorträge in einem Tagungsband publiziert, der den VdA-Mitgliedern online im Mitgliederbereich zur Verfügung gestellt wird. Mit Vorfreude darf man also der in Suhl vom Oberbürgermeister der Stadt Fulda ausgesprochenen Einladung zum 92. Deutschen Archivtag in Fulda folgen: Der 92. Deutsche Archivtag findet dort vom 7. bis 9. Oktober 2025 unter dem Rahmenthema „Digitalisate im Archiv. Positionen – Prozesse – Perspektiven“ statt.

FACHGRUPPEN 7 UND 8

UMGANG MIT EMOTIONAL BELASTENDEN BESTÄNDEN. GEMEINSAME VERANSTALTUNG DER FACHGRUPPEN 7 UND 8 AUF DEM DEUTSCHEN ARCHIVTAG 2024

An der Veranstaltung im Bankettsaal des Congresszentrums Suhl nahmen am 10. Oktober 2024 rund 60 Personen teil. Die hohe Zahl der Teilnehmenden, darunter auch viele Mitglieder anderer Fachgruppen, weist auf ein großes Interesse an einem in der deutschen Archivlandschaft bislang noch wenig behandelten Thema hin. Für die Veranstaltung wurde das Format der Fishbowl-Diskussion gewählt, um in strukturierter Form möglichst vielen Teilnehmenden die Möglichkeit zu geben, ihre Meinung einzubringen. Die Diskussion wurde von Moderator Torsten Musial (Akademie der Künste, Berlin) in zwei thematische Blöcke gegliedert, die jeweils mit kurzen Impulsreferaten mit Blick auf die Perspektive der Archivmitarbeitenden und der Archivnutzenden eingeleitet wurden. Der ursprünglich vorgesehene – und im Tagungsprogramm auch so angekündigte – Themenkomplex diskriminierende Sprache wurde nicht berücksichtigt, um die Veranstaltung nicht zu überfrachten.



Fachgruppen 7 und 8: Johannes Beermann-Schön. Foto: VdA

Zum Auftakt gab Johannes Beermann-Schön (Archiv des Fritz Bauer Instituts, Frankfurt am Main) eine kurze Einführung in die Thematik, bei der er sich auf die Ergebnisse seiner an der FH Potsdam erstellten Masterarbeit stützte, für die er Interviews mit betroffenen Archivkolleginnen und -kollegen geführt hatte. Dabei konstatierte er zunächst einen lückenhaften Forschungsstand in Deutschland im Gegensatz zu den Vorreitern Kanada und Australien. Dieser Befund erscheint insbesondere deshalb erstaunlich, da laut einer ICA-Umfrage aus dem Jahr 2023, an der weltweit 1.138 Kolleginnen und Kollegen teilgenommen haben, 68 % der Befragten – darunter ein Großteil aus Deutschland – angaben, schon einmal mit belastenden Unterlagen in Kontakt gekommen zu

sein. Akten in Archiven dokumentierten die Welt mit ihren allgegenwärtigen menschlichen Traumata. Gewalt, Missbrauch, Entrechtung oder Handlungen eines repressiven Staates fänden daher Eingang in Archivalien, die dadurch als belastet empfunden werden könnten. Triggerpunkte und Belastungsgrenzen seien dabei individuell unterschiedlich, es sei aber zu beobachten, dass Stress – etwa Zeitdruck – die Belastung verstärke und häufiger Kontakt mit entsprechenden Beständen das Risiko einer negativen emotionalen Reaktion erhöhe. Diese können vom Gedankenkarussell über Müdigkeit und Schlafstörungen bis hin zu sozialer Isolation oder Burnout reichen. Entscheidend sei dabei nicht die Mediengattung – mitunter würden Texte als belastender empfunden als Bilder, da sie mehr Raum für eigene Vorstellungen ließen. Der Kontakt mit belastenden Beständen sei im Archiv in allen Arbeitsbereichen von der Übernahme und Erschließung bis hin zur Nutzerberatung und Vermittlung möglich. Viele Kolleginnen und Kollegen müssen dabei bislang meist eigene persönliche Bewältigungsstrategien finden, da eine Unterstützung durch den Arbeitgeber in der Regel nicht erfolgt.

Die Diskussion wurde mit der provokanten Frage eröffnet, ob das Thema ein Generationenthema sei, also jüngere Kolleginnen und Kollegen mehr betroffen seien als ältere. Schnell wurde jedoch klar, dass das Thema generationenübergreifend ist, sich aber der gesamtgesellschaftliche Umgang geändert habe und eine verstärkte Sensibilität für emotionale Belastungen erkennbar ist. Am Beispiel der Recherchen für die Missbrauchsstudie in Kirchenarchiven stellte Sabrina Heeren-Simon (Ev. Landeskirchliches Archiv in Berlin) Stress durch Zeitdruck als belastenden Faktor heraus und bestätigte, dass traumatisierende Unterlagen in allen Akten möglich sind. Im geschilderten Fall wurde Supervision und seelsorgerliche Unterstützung nach Intervention der Projektbeteiligten angeboten. Mehrfach wurde in der



Fachgruppen 7 und 8: Fishbowl-Diskussion. Foto: VdA

Diskussion auf fehlende Unterstützungsinfrastrukturen hingewiesen. Susanne Freund (FH Potsdam) machte zudem auf die Fürsorgepflicht der Führungskräfte in diesem Zusammenhang aufmerksam. Karin Schwarz (FH Potsdam) und Johannes Kistenich-Zerfass (LA Hessen) sprachen sich darüber hinaus dafür aus, präventive Maßnahmen zum Schutz der Mitarbeitenden zu ergreifen.

Der zweite Teil der Veranstaltung wurde mit einem Impuls von Corinna Keunecke und Nora Wohlfarth (LA Baden-Württemberg) eröffnet, die aus ihren Erfahrungen mit Betroffenen berichteten, die sie im Rahmen der drei seit 2012 beim LA Baden-Württemberg angesiedelten Drittmittelprojekte zu „Heimkindheiten“ beraten und betreut hatten. In diesen Projekten, deren letztes im Oktober 2024 ausgelaufen ist, wurden über 2.000 Recherchen zu individuellen Heimgeschicksalen erstellt. Die Verbindung zwischen den archivunerfahrenen Betroffenen, deren Verhältnis zu Behörden durch vielfache negative Erfahrungen belastet war einerseits, und ihren Hoffnungen, Antworten auf grundlegende Lebensfragen aus den Akten zu erhalten, andererseits, führte vielfach zu einer großen emotionalen Belastung der Betroffenen. Die Projektbeteiligten sahen sich dabei oft in die Rolle eines Sozialarbeiters oder einer Psychologin gedrängt. Für die Mitarbeitenden entstand die Schwierigkeit, sich emotional abzugrenzen und die Grenzen dessen aufzuzeigen, was im Rahmen des Projekts an Betreuungsarbeit zu leisten war. Diese Frage wurde in der anschließenden Diskussion vertieft. Als mögliche Lösung wurde die Schaffung von Schutz-

konzepten oder psychologischen Gefährdungsbeurteilungen für Mitarbeitende vorgeschlagen. Auch Supervision, eventuell in Zusammenarbeit mit übergeordneten Stellen, etwa in der Landesverwaltung, oder die Ermöglichung einer kollegialen Beratung könnten Wege aus dem Dilemma aufzeigen, den betroffenen Nutzern weitgehende Hilfestellung zu leisten, ohne dabei die eigenen Grenzen zu überschreiten. Susanne Freund (FH Potsdam) betonte die Wichtigkeit, schon in der Ausbildung für diese Fragen zu sensibilisieren. Francesco Gelati (Universitätsarchiv Hamburg) verwies auf die einschlägigen Hilfseinrichtungen des Arbeitsschutzes, die an Universitäten bereits bestehen. An einigen archivimmanenten Barrieren wie etwa der verstreuten Überlieferung von Unterlagen zu einzelnen Betroffenen in verschiedenen Archiven wird sich aber auch in Zukunft wenig ändern lassen.

Die rege Diskussion machte deutlich, wie wichtig es war, dem Thema emotional belastende Bestände Raum zu geben. In ihren Schlussworten verliehen die vier Hauptdiskutanten ihrer Dankbarkeit dafür Ausdruck, dass das Thema endlich in der deutschen Archivlandschaft angekommen sei. Es bleibt zu wünschen, dass der Umgang mit emotional belastenden Beständen künftig mehr Aufmerksamkeit erhält und durch Beiträge in Publikationen oder auf Veranstaltungen der Archivcommunity und durch Verankerung in den Ausbildungscurricula eine deutlich höhere Sensibilität für diese Thematik im deutschen Archivwesen geschaffen wird.

Christian George, Mainz

BERICHTE AUS DEM VERBAND

LANDESV ERBAND BERLIN

GROSSES INTERESSE AM 7. BERLINER ARCHIVTAG: KI IM ARCHIV

Was für ein Archivtag! Obwohl der Einsatz Künstlicher Intelligenz 2024 bereits auf anderen Archivkonferenzen diskutiert worden war, zog das Tagungsthema „KI im Archiv: Chance oder Risiko?“ so viele Interessierte an (weit über 300!), dass erstmals bei einem Berliner Archivtag die Anmeldung vorzeitig geschlossen werden musste. Denn die Platzkapazität des Tagungsortes, in diesem Fall der Münzenbergsaal am Franz-Mehring-Platz in Berlin-Mitte, war schlicht erschöpft. Um dennoch so vielen Interessierten wie möglich eine unmittelbare Teilnahme zu ermöglichen, wurde kurzfristig eine Videoübertragung der Konferenz in einen direkt neben dem Tagungsaal gelegenen Raum eingerichtet. So konnten schließlich am 20. November 2024 rund 260 Menschen am Berliner Archivtag teilnehmen.

Nach der Begrüßung durch den Vorsitzenden des VdA-Landesverbands Berlin Torsten Musial sprach der Vorsitzende

des Kulturausschusses des Berliner Abgeordnetenhauses Peer Mock-Stürmer ein Grußwort. Darin betonte er die zentrale Rolle der Archive in der digitalen Gesellschaft und hob die Vielfalt der Berliner Archivlandschaft hervor, die so auch ein relevantes Spiegelbild der Stadtgesellschaft sei.

In seinem Eröffnungsvortrag gab Matteo Valleriani, Honorarprofessor an der TU Berlin und Forschungsgruppenleiter am Max-Planck-Institut für Wissenschaftsgeschichte in Berlin, einen aufschlussreichen Laborbericht über die praktische Anwendung von KI in der Arbeit seiner Forschungsgruppe. Dabei unterstützt die Künstliche Intelligenz die Erschließung und Verknüpfung frühneuzeitlicher Quellen zur Wissenschaftsgeschichte. Mithilfe selbstentwickelter Modelle werden Quellen miteinander verglichen und auf Ähnlichkeiten überprüft. So ergeben sich neue Erkenntnisse, die ohne Einsatz von KI einen unvergleichbar hohen Zeitauf-



wand erfordern würden oder schlicht nicht leistbar wären. Ein klares Plädoyer für den Einsatz von KI.

Im ersten Panel wurden die rechtlichen und ethischen Fragestellungen rund um den Einsatz von Künstlicher Intelligenz in Archiven umfassend beleuchtet.

Zunächst ging Paul Klimpel (iRIGHTS, Berlin) detailliert auf die rechtlichen Aspekte der Anwendung von KI ein, insbesondere auf urheberrechtliche Fragen und datenschutzrechtliche Gesichtspunkte. Besonders thematisierte er Urheberrechtsfragen beim Training von KI-Modellen sowie datenschutzrechtliche Herausforderungen. Deutlich wurde, wie vielfältig – und auch noch nicht für alle Belange fixiert – die rechtlichen Rahmenbedingungen für den Einsatz von KI in Archiven sind.

Jessica Heesen (Universität Tübingen) widmete sich den Fragen einer Ethik der Künstlichen Intelligenz. Sie verwies Superintelligenz, also die autonome oder gar omnipotente Herrschaft Künstlicher Intelligenz über den Menschen, in das Reich von Science-Fiction. Dagegen machte sie auf das Problem aufmerksam, wenn Produkte Künstlicher Intelligenz nicht als solche gekennzeichnet werden. Und sie sieht zwei wichtige Herausforderungen: Die Reproduktion oder sogar Verstärkung unerwünschter menschlicher Muster durch KI wie beispielsweise Vorurteile oder rassistische Stereotype. Und die scheinbar menschliche Anmutung von Künstlicher Intelligenz, wenn derartige KI-Produkte nicht als solche gekennzeichnet werden. In ihrem kurzweiligen Vortrag zeigte sie viele aktuelle Beispiele und Möglichkeiten der Nutzung von KI wie beispielsweise die „Wiederauferstehung“ Verstorbenen als KI-Avatare. Insgesamt befindet sich aus ihrer Sicht die Diskussion ethischer Fragen jedoch erst am Anfang.

Aktuelles

Jana Bertels vom Berliner Beauftragten zur Aufarbeitung der SED-Diktatur stellte die Ergebnisse der aktuellen zweiten Teilstudie zur Archivierung und Dokumentation von Beständen mit SED/DDR-Provenienz im Land Berlin vor, die in der „Berliner Archivrundschau“ (Heft 1/2024) und im Blog „Berliner Archive“ (24.4.2024) bereits besprochen wurden. Neben der Feststellung, dass noch viel Handlungsbedarf bei der Erschließung und Bestandserhaltung besteht, waren wichtige Schlussfolgerungen der Studie: Vollständige Erschließung und verstärkte Digitalisierung der SBZ/DDR-Bestände, Prüfung der Möglichkeit des Aufbaus eines Berliner Archivportals, Einrichtung einer Archivberatung in Berlin sowie die Stärkung des fachlichen Austauschs auf Landesebene.

Nico Beyer und Felix Gericke (FU Berlin) präsentierten die von ihnen entwickelte Software EMILiA zur Archivierung von E-Mails, die die dauerhafte Erhaltung dieser Dokumentationsform sichern und die archivische Bewertung unterstützen soll.

Neuwahl des Vorstands des Landesverbands Berlin

Die anstehende Neuwahl des Vorstands des Landesverbands Berlin wurde diesmal nicht als geschlossene Mitgliederversammlung, sondern als eine allen Archivtagsteilnehmenden zugängliche Veranstaltung durchgeführt. Das sorgte für gro-

ße Transparenz und ermöglichte zugleich auch den (noch) nicht dem VdA angehörenden Fachkolleginnen und -kollegen einen Einblick in die Aufgaben und die Arbeit des Vorstands. Zunächst wurde jedoch dem bereits im Frühjahr 2024 aus persönlichen Gründen aus dem Vorstand ausgeschiedenen Yves A. Pillep (Domarchiv Berlin) gedankt. Pillep war Gründungsmitglied des Landesverbands Berlin und seit 2016 dessen stellvertretender Vorsitzender. Er wird sich jedoch auch weiterhin für den Landesverband engagieren, u. a. im Blog „Berliner Archive“ und bei der Herausgabe der Tagungsbände zum Berliner Archivtag.

Über die zurückliegenden Aktivitäten, gegenwärtigen Herausforderungen und neuen Pläne des Vorstands hatte der bisherige Vorsitzende Torsten Musial schon in seiner Begrüßung zu Beginn des Archivtags berichtet. Nun stellten sich die Kandidatinnen und Kandidaten für die Wahl zum neuen Vorstand des VdA-Landesverbands in zwangloser Runde den Fragen Dirk Ullmanns (Archiv für Diakonie und Entwicklung), der die Wahl moderierte. Ihn interessierte vor allem, warum diese sich – überwiegend ehrenamtlich – für den Landesverband engagieren oder dieses tun wollen. Als Beweggründe wurden vor allem genannt: Weitung des Blicks über das eigene Archiv hinaus, Netzwerken sowie die Chance, etwas bewegen zu können. Deutlich wurde, dass der Landesverband als gemeinsame Plattform des fachlichen Austauschs nur lebendig bleiben kann, wenn sich möglichst viele Kolleginnen und Kollegen engagieren und ihre Interessen einbringen.

Als Ergebnis der geheimen Wahl gehören dem neuen Landesvorstand an: Matthias Buchholz (Bundesstiftung Aufarbeitung), Philip Gorki (Berlinische Galerie), Rebecca Hernandez-Garcia (Archiv der DDR-Opposition), Torsten Musial (Akademie der Künste), Ines Oberling (Bundesarchiv - Stasiunterlagenarchiv), Karl Sand (Landesarchiv), Irina Schwab (Archiv der Technischen Universität) und Bianca Welzing-Bräutigam (Landesarchiv).

Inzwischen hat sich der neue Vorstand konstituiert und Rebecca Hernandez Garcia als Vorsitzende, Torsten Musial als deren Stellvertreter, Karl Sand als Schatzmeister sowie Ines Oberling als Schriftführerin bestimmt.

Praktische Anwendungen von KI

Das Nachmittagspanel widmete sich praktischen Anwendungen von KI. Dominic Eickhoff vom LWL-Archivamt für Westfalen bot einen fundierten und kurzweiligen Überblick über verschiedene Anwendungen, die sich nicht nur in der Darstellung von ChatGPT erschöpften, und führte in die technischen Grundlagen ein. Er ließ keinen Zweifel an den riesigen Chancen, die der Einsatz von KI gerade im Archiv bietet.

Karen Tieth und Stefan Geiser stellten die praktischen Anwendungen in der Bildagentur Preußischen Kulturbesitz, deren Nutzen, Aufwand und Grenzen dar. Sie zeigten, dass KI in einigen Einsatzbereichen bereits jetzt eine deutliche Hilfe ist und warben dafür, auch kleinere KI-Tools einfach auszuprobieren.

Esther Lemmerz schilderte die Ergebnisse KI-gestützter Erschließung von Karteikarten im Bundesarchiv. Am Beispiel der großen Masse an Findmitteln zu den Stasiunterlagen

wurde der hohe Entwicklungsaufwand im Vorfeld, aber auch deren effektiver Nutzen im Einsatz deutlich.

Alexander Wolff und Joanna Bars berichteten vom Einsatz der KI im Deutschen Rundfunkarchiv, die gerade die Erschließung von Audiodateien unterstützt und beschleunigt, insbesondere bei Voice Activity Detection, automatisierter Transkription und Metadatenextraktion. Hier sei man schon sehr weit. Bei der Anwendung von KI bei den Videobeständen bestünden jedoch noch einige Herausforderungen.

Podiumsdiskussion

In der abschließenden, von Matthias Buchholz und Rebecca Hernandez Garcia moderierten, Podiumsdiskussion hatten die Archivtagsteilnehmenden noch einmal die Möglichkeit, ihre Fragen zu Chancen und Risiken der KI an Paul Klimpel und Dominic Eickhoff zu richten.

Die Diskussion betonte noch einmal das große Potenzial von KI bei der Archivarbeit und machte Mut, auch kleine KI-Tools einfach auszuprobieren, zeigte zugleich aber auch die Unsicherheit einiger Teilnehmender, was die Anwendung von KI und die damit verbundenen Aufwände betrifft. Deut-

lich wurde aber auch, dass für die Anwendung von KI digitale Daten vorliegen müssen und daher für viele Archive derzeit zunächst noch die Digitalisierung im Vordergrund steht.

Fazit

KI im Archiv ist ein hochaktuelles Thema, das unbedingt weiterer Beachtung bedarf. Denn nicht alle Fragen konnten auf der Konferenz behandelt oder gar erschöpfend beantwortet werden. Jedoch gaben die durchweg interessanten Vorträge und die angeregten Diskussionen – im Plenum und in den Tagungspausen – dafür ausreichend Anregungen. Das Tagungszentrum Franz-Mehring-Platz bot für den Landesarchivtag sehr gute Bedingungen. Das gilt ebenso für die umfangreiche Archivmesse, die im Foyer um den Veranstaltungssaal stattfand.

Um die Ergebnisse der Tagung auch einer breiteren Öffentlichkeit zugänglich zu machen, werden die Vorträge in einem Schriftenband publiziert, der im Frühjahr 2025 vorliegen soll.

Yves A. Pillep, Berlin, Torsten Musial, Berlin

HEFT 2/2024 DER „BERLINER ARCHIVRUNDSCHAU“ ERSCHIENEN



Cover

Bereits Anfang November 2024 ist die aktuelle Ausgabe der „Berliner Archivrundschau“ erschienen, diesmal mit dem Hauptthema „Kolonialismus und Archive“. Vor 140 Jahren, am 15. November 1884, besprachen europäische Mächte auf einer Konferenz in der Berliner Wilhelmstraße 77 die Aufteilung eines ganzen Kontinents. Dazu und der Geschichte des Kolonialismus finden sich viele Dokumente in Berliner Archiven. Daher bot dieses Jubiläum den geeigneten Anlass, danach zu fragen, wie es um deren Zugänglichkeit bestellt ist und welche neuen Projekte es in diesem Zusammenhang gibt. Aber auch zu überlegen, ob diese Quellen, die ja die Sicht der Kolonialverwaltungen widerspiegeln und deren Sprache voll rassistischer Begrifflichkeiten ist, einer Kontextualisierung bedürfen, um die ehemals Kolonisierten nicht erneut zu stigmatisieren. Das Thema ist hochaktuell und so wollten sich ungewöhnlich viele Autorinnen und Autoren dazu äußern.

Zugleich wurden im April 2024 erste Stufen des Berliner Aufarbeitungs- sowie Erinnerungskonzeptes zu Geschichte und Folgen des Kolonialismus vorgestellt, dessen Entwicklung das Berliner Abgeordnetenhaus 2019 beschlossen hatte. Im Kern geht es um neu zu schaffende Lern- und Erinnerungsorte, um Wissen über den deutschen Kolonialismus zu vermitteln. Zu fragen ist, wie auch Archive, die derzeit von diesem Prozess ausgeschlossen scheinen, sich dabei einbringen können.

Das Thema der nächsten Ausgabe der „Berliner Archivrundschau“ im Frühjahr 2025 ist Sportarchive.

Torsten Musial, Berlin

LANDESVERBAND SACHSEN

FORTBILDUNGSVERANSTALTUNG „IM LABYRINTH DES RECHTS – ARCHIVE ZWISCHEN OPEN ACCESS UND VERSCHLUSSSACHEN“



Blick in den Vortragsraum. Foto: Lena Angelstein

Am 4. Juni 2024 lud der Landesverband Sachsen im VdA zu der von Paul Klimpel geleiteten Fortbildung „Im Labyrinth des Rechts – Archive zwischen Open Access und Verschluss-sachen“ in das Stadtarchiv Leipzig ein. Die Veranstaltung zeichnete sich durch ihren Vortragscharakter und rege Diskussionen aus.

Paul Klimpel ist u. a. als Jurist und Kulturmanager tätig. Seine Expertise umfasst insbesondere das Urheberrecht sowie die organisatorischen, technischen und rechtlichen Veränderungen in Gedächtnisorganisationen infolge der Digitalisierung.

Klimpel vermittelte im Rahmen dieser Veranstaltung auf anschauliche Weise die Entstehung des deutschen Urheberrechts im 19. Jahrhundert und schlug die Brücke zu aktuellen

juristischen Entwicklungen, unter anderem zum Zitatrecht, Ausstellungsrecht, Persönlichkeitsrecht oder Leistungsschutzrecht. Der Referent gab ferner einen Einblick in die rechtlichen Aspekte im Umgang mit Archivalien im digitalen Zeitalter. Neben Open-Access-Strategien und der Definition von verwaisten Werken standen die allgemeinen Grundlagen des Zugangs zu Informationen in Archiven, sowohl in digitaler als auch in analoger Form, im Fokus. Dies umfasste die entsprechenden Rechtsfragen bei Digitalisierungsprojekten und der Archivierung von Social-Media-Inhalten. Weitere Themen waren Datenschutz, das Kunsturhebergesetz, Archivrecht und Schutzfristen. Der Redner stützte seine Argumentation dabei auf praktische Fallbeispiele, die einen Bezug zur Arbeit in Archiven haben, und rief zu Mut und Pragmatismus auf, denn die Digitalisierung im Archiv bilde laut Klimpel Chance und Risiko zugleich.

Insgesamt nahmen über 40 Teilnehmerinnen und Teilnehmer unter anderem aus Staats-, Landes-, Kreis-, Stadt- und Universitäts-Archiven vor allem aus Sachsen, aber auch aus ganz Deutschland teil. Die Veranstaltung bot allen Interessierten eine ausgezeichnete Gelegenheit, vorhandenes Wissen zu vertiefen und aktuelle Herausforderungen zu diskutieren. Sowohl innerhalb der Themenblöcke als auch am Ende gab es ausreichend Zeit für Fragen und Anmerkungen. Es entstand ein reger und gewinnbringender Gedankenaustausch, in dem neue Lösungsansätze für rechtliche Problematiken innerhalb der Archivarbeit diskutiert wurden. Fortbildungen dieser Art erhöhen nicht nur das Bewusstsein für das Thema, sondern vermitteln Archivmitarbeiterinnen und -mitarbeitern auch praxisorientierte Maßnahmen, welche die Erhaltung, Vermittlung und den Zugang zu unserem vielfältigen Kulturgut fördern.

Lena Angelstein, Leipzig

GEDENKEN

Der VdA gedenkt an dieser Stelle seiner verstorbenen Mitglieder, deren Tod uns seit der letzten Ausgabe angezeigt wurde:



Nina Matuszewski M.A. (Köln) im Alter von 57 Jahren

Olaf Weißbach (Berlin) im Alter von 68 Jahren

PERSONALNACHRICHTEN

*Zusammengestellt vom
VdA – Verband deutscher Archivarinnen und Archivare e. V.*

STAATLICHE ARCHIVE

GEHEIMES STAATSARCHIV PREUSSISCHER KULTURBESITZ

Ernannt

Diana Kittelmann zur Archivinspektorin (1.1.2024) – **Julian Schulenburg** zum Archivinspektor (1.1.2024).

Versetzt

Wissenschaftlicher Rat **Dr. Ramon Voges** vom Deutschen Buch- und Schriftmuseum der Deutschen Nationalbibliothek Leipzig an das Geheime Staatsarchiv Preussischer Kulturbesitz (16.12.2024).

In den Ruhestand getreten

Archivoberamtsrätin **Christiane Brandt-Salloum** (30.9.2024) – Archivangestellte **Kornelia Lange** (31.12.2024).

BADEN-WÜRTTEMBERG

Eingestellt

Simon Ernst als Archivoberinspektor beim Landesarchiv Baden-Württemberg, Generallandesarchiv Karlsruhe (28.9.2024) – **Ellen Kaiser** als Archivoberinspektorin beim Landesarchiv Baden-Württemberg, Staatsarchiv Sigmaringen (28.9.2024).

Ernannt

Regierungsüberinspektorin **Stefanie Hårdter** zur Regierungsamtsfrau beim Landesarchiv Baden-Württemberg, Zentrale Dienste (1.9.2024).

BAYERN

Ernannt

Regierungsrätin **Barbara Monot** zur Regierungsberrätin bei der Generaldirektion der Staatlichen Archive Bayerns (1.6.2024) – Archivrat **Dr. Hubert Seliger** zum Archivoberrat bei der Generaldirektion der Staatlichen Archive Bayerns (1.6.2024) – Archividirektor **Dr. Julian Holzapfl** zum Leitenden Archividirektor beim Staatsarchiv München (1.7.2024) –

Archivrätin **Dr. Sarah Hadry** zur Archivoberrätin beim Bayerischen Hauptstaatsarchiv (1.10.2024) – Archivrätin **Christine Kofer** zur Archivoberrätin beim Bayerischen Hauptstaatsarchiv (1.10.2024) – Archivoberrat **Dr. Markus Schmalzl** zum Archividirektor beim Staatsarchiv München (1.11.2024).

Versetzt

Archivoberrätin **Dr. Ulrike Hofmann** vom Staatsarchiv München an das Stadtarchiv München (1.7.2024) – Archivsekretär **Lukas Ortner** vom Staatsarchiv München an das Stadtarchiv Augsburg (1.11.2024).

In den Ruhestand getreten

Archividirektor a. D. **Dr. Martin Rütth** beim Staatsarchiv Landshut (30.9.2024) – Archivamtsrätin a. D. **Dr. Monika Ruth Franz** beim Staatsarchiv Landshut (31.12.2024).

Verstorben

Archivhauptsekretär a. D. **Klaus Losert** im Alter von 86 Jahren beim Staatsarchiv Würzburg (4.8.2024).

Bayrische Archivschule

Archivreferendarin **Dr. des. Cordula Bauer** wurde als Archivrätin eingestellt, unter gleichzeitiger Zuweisung zur Generaldirektion der Staatlichen Archive Bayerns (29.11.2024) – Archivreferendar **Dr. Matthias Baumgartl** wurde als Archivrat eingestellt, unter gleichzeitiger Zuweisung zum Bayerischen Hauptstaatsarchiv (29.11.2024) – Archivreferendar **Gregor Jakob** wurde als Archivrat eingestellt, unter gleichzeitiger Zuweisung zur Generaldirektion der Staatlichen Archive Bayerns (29.11.2024) – Archivreferendar **Georg Kolb** wurde als Archivrat eingestellt, unter gleichzeitiger Zuweisung zum Bayerischen Hauptstaatsarchiv (29.11.2024).

Hochschule für den öffentlichen Dienst in Bayern, FB Archiv- und Bibliothekswesen, Fachrichtung Archiv

Mireille Aschenbrenner wurde als Archivinspektoranwärterin eingestellt (4.11.2024) – **Raphael Grünbeck** wurde als Archivinspektoranwärter eingestellt (4.11.2024) – **Marie Hempel** wurde als Archivinspektoranwärterin eingestellt (4.11.2024) – **Elena Hiemer** wurde als Archivinspektoranwärterin eingestellt (4.11.2024) – **Amelie Knittel** wurde als

Archivinspektoranwärterin eingestellt (4.11.2024) – **Felicia Knobloch** wurde als Archivinspektoranwärterin eingestellt (4.11.2024) – **Julia Landmann** wurde als Archivinspektoranwärterin eingestellt (4.11.2024) – **Susanne Schroer** wurde als Archivinspektoranwärterin eingestellt (4.11.2024).

BERLIN

Eingestellt

Aisha Reiche als Angestellte beim Politischen Archiv des Auswärtigen Amtes (1.8.2024) – **Nataliia Levchenko** als Auszubildende zur Fachangestellten für Medien- und Informationsdienste, Fachrichtung Archiv, beim Landesarchiv Berlin (1.9.2024) – **Karoline Otto** als Auszubildende zur Fachangestellten für Medien- und Informationsdienste, Fachrichtung Archiv, beim Landesarchiv Berlin (1.9.2024).

Ernannt

Stellvertretender Archivleiter **Dr. Martin Kröger** zum Archivleiter beim Politischen Archiv des Auswärtigen Amtes (1.8.2024).

HESSEN

Eingestellt

Dr. Thomas Fuchs als Archivrat auf Probe beim Hessischen Landesarchiv, Präsidialbüro, Staatsarchiv Marburg (1.10.2024) – Archivbeschäftigter **Brandon-Collin Van Camp** beim Hessischen Landesarchiv, Abteilung Staatsarchiv Marburg (1.11.2024).

Versetzt

Archivrat **David Gniffke** vom Hessischen Landesarchiv, Präsidialbüro, Staatsarchiv Darmstadt an den Landschaftsverband Westfalen-Lippe, Münster (1.9.2024).

Ausgeschieden

Dr. Nadine Hofmann beim Hessischen Landesarchiv (31.3.2024).

Archivschule Marburg

Petra Hoffmann wurde eingestellt (1.9.2024).

Der. 62. Fachhochschullehrgang wurde am 1.10.2024 mit folgenden Teilnehmerinnen und Teilnehmer eröffnet:

Frieda Alberti (Niedersächsisches Landesarchiv) – **Matthias Albig** (Landesarchiv Schleswig-Holstein) – **Jonas Beinhoff** (Geheimes Staatsarchiv Preußischer Kulturbesitz) – **Lea Bischoff** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Nathalie Durot** (Geheimes Staatsarchiv Preußischer Kulturbesitz) – **Lena Eckernkemper** (Niedersächsisches Landesarchiv) – **Friederike Ester** (Staatsarchiv Marburg, Hessen) – **Heinz Philipp Embacher** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Simon Heeke** (Staatsarchiv Hamburg) – **Helena Höss** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Cedric Howeg** (Geheimes

Staatsarchiv Preußischer Kulturbesitz) – **Sabrina Jorewitz** (Staatsarchiv Marburg (Stadtarchiv Bad Homburg), Hessen) – **Lina Langer** (Niedersächsisches Landessarchiv) – **Helen Lutz** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Leonie Mohr** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Regina Reisch** (Bundesarchiv) – **Elsa Schacke** (Staatsarchiv Marburg, Hessen) – **Robert Schittko** (Staatsarchiv Hamburg) – **Saskia Schlusche** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Saskia Schmidt** (Bundesarchiv) – **Cosmo Schmitter** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Sebastian Spitz** (Hauptstaatsarchiv Stuttgart, Baden-Württemberg) – **Simeon Türk** (Landesarchiv Schleswig-Holstein) – **Robert Wolff** (Staatsarchiv Marburg, Hessen).

Hans Grün ist in den Ruhestand getreten (31.7.2024).

NIEDERSACHSEN

Ernannt

Archivrat **Dr. Bernhard Homa** zum Archivoberrat beim Niedersächsischen Landesarchiv, Abteilung Hannover (6.12.2024).

NORDRHEIN-WESTFALEN

Eingestellt

Marcel Remmers als Tarifbeschäftigter beim Landesarchiv Nordrhein-Westfalen, Dienstort Münster-Coerde, Abteilung Westfalen (W1) (1.9.2024) – **Celina Zenker** als Tarifbeschäftigte beim Landesarchiv Nordrhein-Westfalen, Dienstort Duisburg, Abteilung Fachbereich Grundsätze (F3) (1.10.2024) – **Ann Christine Wecker** als Tarifbeschäftigte beim Landesarchiv Nordrhein-Westfalen, Dienstort Detmold, Abteilung Fachbereich Grundsätze (F3) (18.11.2024) – **Lara Tesching** als Tarifbeschäftigte beim Landesarchiv Nordrhein-Westfalen, Dienstort Detmold, Abteilung Ostwestfalen-Lippe (OWL1) (1.12.2024).

SACHSEN

Eingestellt

Dr. Sebastian Müller als Referent beim Sächsischen Staatsarchiv, Abteilung 4, Staatsarchiv Chemnitz (1.9.2024) – **Kristina Paul** als Sachbearbeiterin beim Sächsischen Staatsarchiv, Abteilung 3, Staatsarchiv Leipzig (1.9.2024).

Versetzt

Referent **Dr. Thomas Fuchs** vom Sächsischen Staatsarchiv an das Hessische Landesarchiv, Standort Marburg (1.10.2024).

Ausgeschieden

Sachbearbeiter **Bernd Scheperski** beim Sächsischen Staatsarchiv, Hauptstaatsarchiv Dresden (30.09.2024).

KOMMUNALE ARCHIVE

Aachen: Stadtarchiv Aachen

Städtische Archivamtsrätin und stellvertretende Archivleiterin **Angelika Pauels** ist in den Ruhestand getreten (30.11.2023) – **Dorothea Oelze** hat die stellvertretende Archivleitung übernommen (1.12.2023) – Städtischer Archivoberinspektor **Michael Jerusalem** wurde eingestellt (1.09.2024) – **Laurena Müller** wurde als Fachangestellte für Medien- und Informationsdienste eingestellt (1.1.2025).

Augsburg: Stadtarchiv Augsburg

Archivreferendarin **Madeleine Zier** wurde als Archivrätin eingestellt (1.12.2024).

Bremen: Staatsarchiv Bremen

Fachangestellte für Medien- und Informationsdienste **Anna Czech** wurde versetzt zum Senator für Finanzen (1.10.2024).

Dorsten: Stadtarchiv Dorsten

Melek Türksev wurde als Fachangestellte für Medien- und Informationsdienste eingestellt. (1.8.2024) – **Melina Wendland** wurde als Fachangestellte für Medien- und Informationsdienste eingestellt (9.9.2024).

Dresden: Stadtarchiv Dresden

Maren Gündel M.A. wurde als Sachbearbeiterin Vorfeld/Erschließung/Bestandsverwaltung eingestellt (1.1.2025).

Erfurt: Stadtarchiv Erfurt

Dr. Nadine Hofmann wurde als Direktorin eingestellt (1.4.2024).

Heidenheim an der Brenz: Kreisarchiv Heidenheim

Archivleiter Kreisarchivamtsrat **Markus Baudisch** ist in den Ruhestand getreten (31.12.2024).

Heilbronn: Stadtarchiv Heilbronn

Archivleiter **Prof. Dr. Christhard Schrenk** ist in den Ruhestand getreten (30.11.2024) – **Miriam Eberlein M.A.** hat die Leitung des Stadtarchivs übernommen (1.12.2024) – **Dr. Saskia David-Gaubatz** hat die stellvertretende Leitung des Stadtarchivs übernommen (1.12.2024).

Köln: Historisches Archiv der Stadt Köln mit Rheinischem Bildarchiv

Henrik Bruns wurde als Fachangestellter für Medien- und Informationsdienste, Fachrichtung Archiv, eingestellt (1.6.2024) – **Anika Engelen** wurde als Archivbeschäftigte umgesetzt (1.9.2024) – **Daniela Wagner** wurde als Wissenschaftliche Archivarin umgesetzt (1.9.2024) – **Helena Weber** wurde als Sachgebietsleiterin beim Rheinischen Bildarchiv eingestellt (1.12.2024).

Krefeld: Stadtarchiv Krefeld

Sven Theemann B.A. wurde als Sachbearbeiter eingestellt (1.11.2024).

München: Stadtarchiv München

Rita Nikolaisen wurde als Archivinspektoranwärterin eingestellt (1.10.2024).

Münster: LWL-Archivamt für Westfalen

Dr. Lutz Engelskirchen wurde als Fachangestellter für Medien- und Informationsdienste eingestellt (1.10.2023) – **Yasin Ceran** wurde als Verwaltungsmitarbeiter eingestellt (15.5.2024) – **Vincent Knoke** wurde als Auszubildender zum Fachangestellten für Medien- und Informationsdienste eingestellt (1.8.2024) – **Ronja Magdalena Moj** als Auszubildende zur Fachangestellten für Medien- und Informationsdienste eingestellt (1.8.2024) – Landesoberarchivrat **Dr. Gunnar Teske** ist in den Ruhestand getreten (31.8.2024) – **Sebastian Bojdo** wurde zum Landesarchivinspektoranwärter ernannt, in Auftragsausbildung für den Kreis Coesfeld (1.9.2024) – **David Gniffke** wurde vom Hessischen Landesarchiv als Landesarchivrat an das LWL-Archivamt für Westfalen versetzt (1.9.2024) – **Julia Hellert** wurde zur Landesarchivinspektoranwärterin ernannt (1.9.2024) – **Johanna Kießler** wurde zur Landesarchivinspektoranwärterin ernannt (1.9.2024) – Landesarchivdirektor **Dr. Alfred Bruns** ist im Alter von 89 Jahren verstorben (16.9.2024) – Landesarchivrat **Dr. Daniel Droste** wurde zum Landesoberarchivrat ernannt (1.10.2024).

Schwäbisch Hall: Kreisarchiv Schwäbisch Hall

Matthias Dörr wurde nach Beendigung der Ausbildung zum Fachangestellten für Medien- und Informationsdienste, Fachrichtung Archiv, als Sachbearbeiter eingestellt (8.7.2024) – **Hendrik Rudolf Stucke** wurde als Auszubildender zum Fachangestellten für Medien- und Informationsdienste, Fachrichtung Archiv, eingestellt (1.9.2024).

Ulm: Stadtarchiv Ulm

Mareikje Mariak wurde als Stadtarchivrätin eingestellt und hat die Leitung des Sachgebiets Reichsstädtische Überlieferung übernommen (1.5.2024).

KIRCHLICHE ARCHIVE

Rottenburg am Neckar: Diözesanarchiv Rottenburg

Janina Beseler wurde eingestellt (1.12.2024).

Stuttgart: Landeskirchliches Archiv Stuttgart

Noah-Joshua Veit ist als FSJler ausgeschieden (30.8.2024) – **Daniel Miller Martínez** wurde als FSJler eingestellt (1.09.2024) – Archivreferendar **Dr. des. Gregor Hofmann** wurde als Kirchenarchivrat mit Schwerpunkt archivarische Grundsätze eingestellt (1.12.2024).

ARCHIVE DER PARLAMENTE, POLITISCHEN PARTEIEN, STIFTUNGEN UND VERBÄNDE

Bonn: Archiv der sozialen Demokratie

Bastian Rausch wurde als Sachbearbeiter im Referat Infrastrukturen und digitale Grundsatzfragen eingestellt (1.1.2025) – **Henning Wiekhorst** wurde als Sachbearbeiter im Referat Infrastrukturen und digitale Grundsatzfragen eingestellt (1.1.2025).

ARCHIVE DER HOCHSCHULEN SOWIE WISSENSCHAFTLICHER INSTITUTIONEN

Geisenheim: Stadt- und Hochschularchiv der Hochschulstadt Geisenheim

Karsten Welcher wurde als Stadtarchivar eingestellt (1.11.2024).

Die hier veröffentlichten Personalnachrichten beruhen auf den Meldungen und Angaben der archivischen Ausbildungseinrichtungen, der Archiveinrichtungen bzw. der zuständigen Verwaltungen. Der VdA übernimmt keine Gewähr für die Richtigkeit, Vollständigkeit und Aktualität der hier veröffentlichten Personalnachrichten!

Die Meldungen sind direkt an die Geschäftsstelle des VdA – Verband deutscher Archivarinnen und Archivare e.V., Wörthstraße 3, 36037 Fulda, E-Mail: info@vda.archiv.net unter Angabe des Einsenders (Dienststelle, Archiv, Institution) und des Bearbeiters (Name, Vorname, Telefon, E-Mail) zu senden.

REDAKTIONSSCHLUSS

PERSONALNACHRICHTEN in Heft 02, 78. Jg.
(Erscheinungstermin Mai 2025):

1. März 2025

Elisabeth Brachmann-Teubner †

Geb. 14. April 1935 Tuhanzel (Sudetenland)
Gest. 9. Oktober 2024 Potsdam



*Elisabeth Brachmann-Teubner
1980er Jahre*

Elisabeth Brachmann-Teubner wurde 1935 in Tuhanzel (Sudetenland) geboren. Ihr Abitur legte sie 1954 in Köthen (Anhalt) ab. Von 1954 bis 1956 studierte sie in Halle (Saale) Geschichte, von 1956 bis 1958 setzte sie ihr Studium mit dem Schwerpunkt Archivwissenschaft an der Humboldt-Universität zu Berlin fort. Ihre Ausbildung zur Diplomarchivarin absolvierte sie von 1959 bis 1960 am Institut für Archivwissenschaft in Potsdam.

Von 1960 bis 1990 war Elisabeth Brachmann-Teubner im Deutschen Zentralarchiv bzw. Zentralen Staatsarchiv der DDR in Potsdam beschäftigt, wo sie sich in den 1960er und 1970er Jahren besonders um die sogenannte Vorfeldarbeit verdient gemacht hat. Die systematische archivische Betreuung der Einrichtungen der Regierung der DDR wurde unter ihrer Leitung als Arbeitsbereich Bewertung und Bestandsergänzung im ZStA aufgebaut. Von 1985 bis 1990 war Elisabeth Brachmann-Teubner dann Direktorin des Zentralen Staatsarchivs der DDR. Als diese erwarb sie sich in den Jahren 1989 und 1990 besondere Verdienste bei der Sicherung der Überlieferung der in Auflösung begriffenen Ministerien und Behörden der DDR. Bereits Mitte Oktober veranlasste sie konsequent und ohne Zeitverzug Maßnahmen für den Erhalt der Unterlagen, die sich in den Verwaltungsarchiven oder sogar noch in der laufenden Schriftgutverwaltung befanden. Besondere Erwähnung verdient auch ihr persönlicher Anteil an der gelungenen Zusammenführung von Bundesarchiv und Zentralem Staatsarchiv. Mit hoher Sachkenntnis, viel menschlicher Empathie und immer ansprechbar hat sie den Integrationsprozess beratend begleitet und damit sehr befördert.

Nach der Zusammenführung der Archive übernahm Elisabeth Brachmann-Teubner wissenschaftliche Aufgaben. Sie beschäftigte sich insbesondere mit der archivalischen Überlieferung zur Geschichte der Juden in Deutschland, einer Thematik, die sie bereits 1989 nach Israel und zur Gedenkstätte Yad Vashem geführt hatte. Wesentlichen Anteil hatte sie an der Vorbereitung des vollständig neu bearbeiteten Gedenkbuchs „Opfer der Verfolgung der Juden unter

der nationalsozialistischen Gewaltherrschaft in Deutschland 1933 – 1945“, das unter den Bedingungen des wiedervereinigten Deutschlands auf eine völlig neue Quellenbasis gestellt werden konnte. Auch hierbei hat sie großes Engagement und viel Einfühlungsvermögen eingebracht.

Nachdem Elisabeth Brachmann-Teubner mit Ablauf des 31. Dezember 1998 in den Ruhestand getreten ist, hat sie zunächst neue Urlaubsziele und kulturelle Bereicherungen nutzen können. In den letzten Jahren haben ihr jedoch gesundheitliche Einschränkungen und pflegende Aufgaben allerhand abverlangt. Am 9. Oktober 2024 ist sie in Potsdam verstorben.

Michael Hollmann

Ein Merkblatt mit Kriterien für die Veröffentlichung eines Nachrufes ist auf der Website der Zeitschrift unter <https://www.archive.nrw.de/landesarchiv-nrw/wir-ueber-uns/der-archivar-abrufbar>.

Der Nachruf ist direkt an die Geschäftsstelle des VdA – Verband deutscher Archivarinnen und Archivare e.V., Wörthstraße 3, 36037 Fulda, E-Mail: info@vda.archiv.net unter Angabe des Einsenders (Dienststelle/Archiv, Institution) und des Verfassers (Name, Vorname, Telefon, E-Mail) zu senden.

Die redaktionelle Verantwortung liegt beim VdA.

REDAKTIONSSCHLUSS

NACHRUFE in Heft 02, 78. Jg.
(Erscheinungstermin Mai 2025):

4. März 2025

GRÜNDUNG: ARBEITSKREIS ARCHIVE IN DER SOZIALWIRTSCHAFT

Die Vereinigung der Wirtschaftsarchivarinnen und Wirtschaftsarchivare e. V. (VdW) ist seit 1957 der Fachverband für das Archivwesen der Wirtschaft in Deutschland und des deutschsprachigen Auslands. Die große Vielfalt des Wirtschaftslebens spiegelt sich im auch Archivwesen der Wirtschaft. So haben sich im Laufe der Zeit innerhalb der VdW Arbeitskreise für den näheren Austausch gebildet. Der Arbeitskreis „Archive in der Sozialwirtschaft (AK ASW)“ setzt sich aus Archivar*innen von Archiven des Dritten bzw. Non-Profit-Sektors zusammen. Unternehmen des Dritten Sektors müssen einerseits gewinnorientiert arbeiten und ihre Tätigkeit generell über den Markt finanzieren. Andererseits werden die erzielten Gewinne für die Erbringung gemeinnütziger und sozialer Leistungen aufgewendet. Im Umfeld der bundesdeutschen sozialen Marktwirtschaft nehmen diese Unternehmen eine wachsende Stellung ein. Archive der Sozialwirtschaft sind als Unternehmensarchive nicht nur für das eigene Unternehmen relevant, sondern dokumentieren die Geschichte des Sozialen und sozialer

Bewegungen und damit den Wandel des Umgangs mit sozialen und (zivil-)gesellschaftlichen Themen. Ihre Bestände zeigen sowohl die Perspektive verschiedener eingebundener Akteure als auch die institutionellen Rahmenbedingungen und damit verbundene Handlungsspielräume. Archive aus diesem Sektor wurden 2022 in einem Themenheft der Zeitschrift „Archiv und Wirtschaft“ vorgestellt (Link: <https://www.wirtschaftsarchive.de/publikationen/archiv-und-wirtschaft/2022-heft-1/>).

Der AK ASW beabsichtigt, die oft weniger bekannten und häufig kleineren Archive der Sozialwirtschaft in der Archivar*innen-Community, der Forschung und der Öffentlichkeit bekannter und sichtbarer zu machen, zum fachlichen Austausch anzuregen und gemeinsame Projekte anzustoßen. Wir freuen uns auf eine rege Beteiligung und Interessenten. Bitte wenden Sie sich an Dr. Marion Baschin, Institut für Geschichte der Medizin, Stuttgart, marion.baschin@igm-bosch.de oder Dr. Dominik Erdmann, Fürst Donnersmarck-Stiftung, Berlin, D.Erdmann@fdst.de.

VORSCHAU

Das nächste Heft befasst sich im Schwerpunkt mit dem Thema „Zukunft der Nutzung“.

U. a. sind folgende Beiträge geplant:

- Die Digitale Transformation unterstützen: Nutzer- und Nutzungsforschung für Archive und andere Informationseinrichtungen
von Sebastian Mundt
- From access to content. Recent developments in the digitisation of archives in The Netherlands and at the Gelders Archief
von Petra Links und Hedzer Uulders
- Liest das jemand? Ein Plädoyer, Digitalisierung konsequent an den Interessen der Nutzenden auszurichten
von Peter Worm
- Prozesse, Prognosen, Perspektiven: zukünftige Benutzung von Archivgut im Bundesarchiv
von Andrea Hänger
- Mehr Kooperation gewünscht? Überlegungen zu Nutzererwartungen an moderne Archive
von Christoph Brüll
- Auf die Mischung kommt es an - wie Befragungen mit neuen User Experience Methoden kombiniert werden können
von Simone Fühles-Ubach

IMPRESSUM

Herausgeber: Landesarchiv Nordrhein-Westfalen, Schifferstr. 30, 47059 Duisburg, Postanschrift: 40208 Düsseldorf, VdA -Verband deutscher Archivarinnen und Archivare e.V., Wörthstr. 3, 36037 Fulda

Gesamtredaktion: Kathrin Pilger in Verbindung mit Ralf Jacob, Frank M. Bischoff, Torsten Musial, Bettina Joergens und Bernhard Homa

Mitarbeiterin Gesamtredaktion: Laura Bremer

Mitarbeiter VdA (Personalnachrichten und VdA-Teil): Thilo Bauer

ISSN 0003-9500 / ISSN 2199-9252 (Internet)

Kontakt: Landesarchiv Nordrhein-Westfalen, Redaktion „ARCHIV. theorie & praxis“, Schifferstr. 30, 47059 Duisburg, Postanschrift: 40208 Düsseldorf, Tel. 0203-98721-0, -119 (Kathrin Pilger), Fax 0203 /98721-111, E-Mail: archiv.t&p@lav.nrw.de

Druck und Vertrieb: Franz Schmitt, Kaiserstraße 99-101, 53721 Siegburg, Tel. 02241/62925, Fax 02241/53891, E-Mail: archivtheorieundpraxis@verlagfranzschmitt.de

Gestaltung: ENGEL UND NORDEN, Wuppertal, Mitarbeit: Ruth Michels, www.engelundnorden.de

Anzeigenverwaltung: Verlag Franz Schmitt (Preisliste 24, gültig ab 1. Januar 2022)

Zuständig für Anzeigen: Sabine Schmitt im Verlag Franz Schmitt

Die Verlagsrechte liegen beim Landesarchiv Nordrhein-Westfalen. Amtliche Bekanntmachungen, Mitteilungen und Manuskripte bitten wir, an die Redaktion zu senden, Personalnachrichten und Veranstaltungshinweise dagegen an die Geschäftsstelle des VdA. Für unverlangt eingesandte Beiträge übernehmen wir keine Haftung, unverlangt eingesandte Rezensionsexemplare werden nicht zurückgesandt. Zum Abdruck angenommene Arbeiten gehen in das unbeschränkte Verfügungsrecht des Herausgebers über. Dies schließt auch die Veröffentlichung im Internet ein. Die Beiträge geben die Meinungen ihrer Verfasser, nicht die der Redaktion wieder.

„ARCHIV. theorie & praxis“ erscheint viermal jährlich. Der Bezugspreis beträgt für das Einzelheft einschl. Porto und Versand 10,00 EUR im Inland, 10,50 EUR im Ausland, für das Jahresabonnement im Inland einschl. Porto und Versand 40,- EUR, im Ausland 42,- EUR.

Hinweise für VdA-Mitglieder: Alle Personalnachrichten, geänderte Anschriften und Bankdaten sind ausschließlich an folgende Adresse zu melden: VdA-Geschäftsstelle, Wörthstr. 3, 36037 Fulda, Tel. 0661/2910972

E-Mail: info@vda.archiv.net, Internet: www.vda.archiv.net

Bankverbindung: Konto für Mitgliedsbeiträge VdA: Sparkasse Fulda, SWIFT-BIC: HELADEF1FDS,

IBAN: DE18 5305 0180 0043 0464 47

Konto für Spenden an den VdA: Sparkasse Fulda, SWIFT-BIC: HELADEF1FDS, IBAN: DE20 5305 0180 0043 0500 00.



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung – Share Alike 4.0 International Lizenz (<https://creativecommons.org/licenses/by-sa/4.0/deed.de>)

KLUG
CONSERVATION

Werte bewahren



Kulturgut bestens geschützt
Präventive Konservierung

Mehr zu unserem Produktsortiment unter:
www.klug-conservation.de/produkte



GANZ EINFACH **ZUVERLÄSSIG.**

- Scannen von Dokumenten von **DIN A7 bis DIN A0**
- Scannen von **Mikrofiches und Mikrofilm-Jackets**
- Scannen von Büchern bis zum Format DIN A2
- Frühes Scannen Ihrer Eingangsrechnungen und -post
- Digitale Bearbeitung gescannter Dokumente
- Scannen und **Weiterverarbeitung der Daten in Kundensystemen (DMS)**
- Klassifizierung, Validierung und Datenerfassung (bspw. Extraktion von Rechnungsdaten, Auswertung von Fragebögen und Formularen)
- Erstellen von **digitalen Dokumentationen**
- Transport und Logistik Ihrer Unterlagen
- Zertifizierte Aktenvernichtung nach DIN 66399; in Schutzklasse 2; Sicherheitsstufe 5



Unsere Zertifizierung nach DIN ISO 9001:2015 garantiert eine dokumentierte Qualitätssicherung aller Prozesse

Haus Freudenberg GmbH

Anerkannte Werkstatt für behinderte Menschen gemäß §142 SGB IX

Digitale Archivierung & Dokumenten-Management

Johann-van-Aken-Ring 12 · 47551 Bedburg-Hau

Ralf Katemann · Fon 02821 7620-52

r.katemann@haus-freudenberg-gmbh.de

Michael Wenzel · Fon 02821 7620-29

m.wenzel@haus-freudenberg-gmbh.de

FAUST 10

Für ein modernes Archiv.

- Mit FAUST können Sie alles archivieren:
Akten, Bilder, Briefe, Bücher, Dateien,
Audios, Videos, Zeichnungen
- Digitale Archivierung
- Bestandsverwaltung, Tektonik
- Flexible Datenbanken
- Umfassende
Rechercheleistungen
- Findbuch, Reports
und umfangreiche
Exportmöglichkeiten
- Integriertes Lesesaalmodul
- Infos unter: www.land-software.de

Land Software-Entwicklung
Tel: 0049-(0)911-69 69 11
info@land-software.de



LAND
Software
Entwicklung

HANS  SCHRÖDER

Kulturgut sicher aufbewahren

mit unseren zertifizierten Schutzverpackungen
zur präventiven Langzeitarchivierung



Entdecken Sie unser Produktsortiment
mit über 1.000 Archivprodukten unter

www.archivbox.com

AUGIAS-IntraFind

Digitaler Lesesaal

Das Gemeinschaftsprojekt von AUGIAS-Data und der IntraFind AG kombiniert die bewährte Funktionalität von FINDBUCH.Net mit den KI-gestützten Recherche-Möglichkeiten von IntraFind.

- Derzeit über 700.000 Einzelwortlexeme (Wortkerne) im deutschen Lexikon
- Über 100 prozedurale Regeln, z.B. für die Behandlung von Straßennamen wie Berliner Straße oder Berlinerstraße
- 50.000 Fachbegriffe und Eigennamen für Deutsch, optionale Suche nach Eigennamen (Personen, Organisationen, Orte, Adressen, etc.)
- Jährliche Aktualisierung mit Wortneubildungen (Neologismen)
- Unknown Word Recognition: unbekannte Wörter (meist seltene Eigennamen) werden als solche markiert und einer speziellen „Plural-s“-Normalisierung unterzogen
- Unterstützt neue und alte deutsche Orthographie



DATENSCHÄTZE IM GRIFF

SORI

*Die Software für
Langzeitarchivierung.*



Einfache Bedienung

- Benutzerfreundliche Weboberfläche
- Vorkonfigurierte Workflows
- Sofort einsetzbar

OAIS-konforme Langzeitarchivierung

- Erweiterbarer Ingest-Workflow
- Intelligenter DIP-Creator
- Anbindbar an unterschiedliche Erfassungssoftware

Service

- Kostengünstige Lösung
- Einfach zu installieren und zu administrieren
- Sichere Datenhaltung, Hardware nach Wahl
- Begleitende Beratung und Schulung



Kennenlern-Produktdemos:

Per QR-Code-Scan oder Linkeingabe gelangen Sie zu den Terminen.

produkt demos.starttext.de

starttext GmbH
Riemenschneiderstr. 11 • 53175 Bonn
Telefon: +49 228 959 96-0
info@starttext.de • www.starttext.de

starttext

Datenschätze im Griff