

WERKHEFTE DES LANDESARCHIVS  
BADEN-WÜRTTEMBERG

Herausgegeben  
vom Landesarchiv Baden-Württemberg

Heft 31

2024

Jan Thorbecke Verlag

Smart und intelligent –  
Digitale Unterstützung für die Arbeit  
im Archiv

Vorträge des 82. Südwestdeutschen Archivtags  
am 22. und 23. Juni 2023

Herausgegeben von Daniel Fähle und Peter Müller

2024

Jan Thorbecke Verlag



Gedruckt auf alterungsbeständigem, säurefreiem Papier.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Alle Rechte vorbehalten

© 2024 Landesarchiv Baden-Württemberg

Kommissionsverlag: Jan Thorbecke Verlag, Verlagsgruppe Patmos in der Schwabenverlag AG, Ostfildern  
[www.thorbecke.de](http://www.thorbecke.de)

Lektorat: Lydia Christine Michel, Landesarchiv Baden-Württemberg

Umschlaglayout: Bureau Johannes Erler, Hamburg

Satz und Druck: Offizin Scheufele Druck & Medien GmbH & Co. KG, Stuttgart

Hergestellt in Deutschland

ISBN 978-3-7995-2081-2

# Inhalt

|  |    |
|--|----|
| <i>Gerald Maier</i>  |    |
| Vorwort .....  | 7  |
| <i>Daniel Fable</i>   |    |
| Einfuhrung .....  | 9  |
| <i>Matthias Razum</i>  |    |
| Chancen und Herausforderungen des digitalen Wandels fur Archive .....   | 13 |
| <i>Dorothee Huff und Regina Keyler</i>   |    |
| Das Projekt OCR-BW: Automatische Texterkennung auch fur Archive .....   | 25 |
| <i>Andreas Neuburger</i>   |    |
| Online fur alle! Ein einfaches Erschlieungssystem fur ressourcenschwache kleine Archive                         | 35 |
| <i>Lambert Kansy</i>   |    |
| Innovativ und offen. Der Digitale Lesesaal der Staatsarchive Basel-Stadt und St.Gallen ..                          | 45 |
| <i>Pascal Notz</i>   |    |
| <i>Timms for Culture</i> – Aufbau eines Streaming-Services fur Kultureinrichtungen<br>in Baden-Wurttemberg ..... | 63 |
| <i>Florian Spiess</i>  |    |
| Multimedia-Sammlungen Durchsuchen und Erkunden mit maschinellem Lernen .....                                       | 69 |
| <i>Tobias Hodel</i>  |    |
| Large Language Models, oder weshalb wir kunstliche Intelligenz im Archiv finden sollten                           | 77 |
| Autorinnen und Autoren .....   | 85 |



# Vorwort

Nachdem sich der 81. Südwestdeutsche Archivtag in Reutlingen im Jahr 2022 mit der Frage beschäftigt hatte, wo genau Archive sich im Spannungsfeld zwischen digitaler und analoger Welt in Zukunft positionieren sollten, ging es beim diesjährigen Archivtag in Balingen prioritär um die digitale Welt – und konkret um die digitale Unterstützung der archivischen Arbeit. Dabei wurde nicht nur über den Einsatz neuer Technologien wie künstlicher Intelligenz oder Machine Learning, sondern auch allgemein über Chancen und Herausforderungen des digitalen Wandels für die Arbeit der Archive und für ihre Kunden gesprochen. Dazu gehören der Ausbau neuer Kooperationsformen für die gemeinsame Entwicklung technischer Lösungen und Dienstleistungen ebenso wie die Auswirkungen der Digitalisierung in der Verwaltung, in der Wissenschaft und im Kulturbetrieb auf die Archivarbeit. Im Bereich der Verwaltung denke man nur an die Einführung der elektronischen Akte mit all ihren Konsequenzen für die Archive sowie den Ausbau von digitalen Service-Dienstleistungen wie den virtuellen digitalen Lesesaal im Archiv. Im Bereich der Wissenschaft geht es um nicht weniger als um Anschluss und Einbindung der Archive in die Nationale Forschungsdateninfrastruktur (NFDI). Im Kulturbereich müssen wir uns den Herausforderungen der digitalen Präsentation von Kulturgut in unterschiedlichen Kontexten wie in der interdisziplinären Deutschen Digitalen Bibliothek (DDB) stellen.

Dem 82. Südwestdeutschen Archivtag, am 22. und 23. Juni 2023 in Balingen, ist es in hervorragender Weise gelungen, mit seinem Thema „Smart und intelligent – Digitale Unterstützung für die Arbeit im Archiv“ und den einzelnen Beiträgen des Tagungsprogramms, diese Herausforderungen aufzugreifen. Obwohl der Archivtag an sich eine Präsenzveranstaltung ist, war auch dieses Jahr wieder eine virtuelle Teilnahme möglich. So konnten auch Kolleginnen und Kollegen, die aus zeitlichen oder räumlichen Gründen nicht nach Balingen kommen konnten, der Veranstaltung im interaktiven Livestream folgen.

Da auch der vorliegende Tagungsband sofort wieder open-access online verfügbar ist, können die Beiträge zeitnah für die weitere Beschäftigung mit dem Thema genutzt werden.

Allen am Tagungsprogramm Mitwirkenden danke ich für ihre Beiträge in Wort und Schrift ganz herzlich. Ebenso danke ich allen, die sich an der Vorbereitung und Durchführung der Tagung beteiligt haben. Der Dank gilt dabei zunächst der gastgebenden Stadt Balingen und besonders Frau Dr. Yvonne Arras, der Leiterin des Stadtarchivs, die gemeinsam mit ihren Mitarbeiterinnen und Mitarbeitern die Organisation vor Ort übernommen hat. Ebenso danke ich dem Landkreis Zollernalbkreis, dessen Kreisarchivar Herr Dr. Uwe Folwarczny sich engagiert am Programm des Archivtags beteiligt hat. Mein Dank gilt auch den Unternehmen, die die Archivmesse in Balingen ermöglicht und den Südwestdeutschen Archivtag finanziell unterstützt haben.

Der abschließende Dank gilt meinen Kollegen im Landesarchiv Baden-Württemberg. Ich danke Herrn Dr. Peter Müller, dem Leiter des Staatsarchivs Ludwigsburg, und Herrn Daniel Fährle, Leiter des Referats Informationstechnologie und Digitale Dienste, für die Übernahme der

Gesamtverantwortung und Moderation der Tagung. Daniel Fähle hat als Tagungspräsident auch das Fachprogramm zusammengestellt. Peter Müller hat seit fünfzehn Jahren ehrenamtlich die Funktion des Geschäftsführenden Präsidenten begleitet. Auf dem diesjährigen Archivtag hat er mit Dr. Marco Birn, dem Leiter des Kreisarchivs Reutlingen, nach langer Suche, dankenswerter Weise, einen Nachfolger gefunden. Ich möchte daher im Namen aller Teilnehmerinnen und Teilnehmer Peter Müller Dank aussprechen, für das was er in den letzten fünfzehn Jahren für den Südwestdeutschen Archivtag geleistet hat.

Nachdem er 2008 zum Geschäftsführenden Präsidenten gewählt worden war, organisierte Peter Müller seither fünfzehn Archivtage im ganzen Land. Einer davon fand auch im Nachbarbundesland Rheinland-Pfalz statt und der Archivtag 2021 war wegen der Corona-Krise rein virtuell. Während seiner Zeit erfolgte eine Verdoppelung der Teilnehmerzahlen von unter 100 Personen bis zu teilweise mehr als 200 Teilnehmerinnen und Teilnehmern. Er sorgte zudem für eine kontinuierliche Dokumentation der Tagungen in der Publikationsreihe des Landesarchivs. Seit 2015 erfolgt die Dokumentation zusätzlich auch durch Videomitschnitte der Vorträge auf YouTube. Seit 2021 besteht die Möglichkeit zur Online-Teilnahme mit Livestream. Ihm gelang immer wieder eine behutsame Erweiterung des Programms durch neue Formate wie Workshops, Begleitveranstaltungen oder die Newcomertreffen. Hervorzuheben ist auch sein Engagement bei der regelmäßigen Kontaktpflege zum Schweizerischen Archivverband (VSA) und zu den Kolleginnen und Kollegen in den benachbarten Bundesländern Bayern und Rheinland-Pfalz.

Peter Müller und das Landesarchiv Baden-Württemberg werden dem Südwestdeutschen Archivtag auch weiterhin verbunden bleiben. So wird das Landesarchiv die Tagungen auch zukünftig in Form einer inhaltlichen Beteiligung am Tagungsprogramm unterstützen. Zudem wird es weiterhin gerne die Vorträge in seiner Publikationsreihe und auf seinem YouTube-Kanal veröffentlichen.

Für das Lektorat und die Drucklegung des diesjährigen Tagungsbandes haben Daniel Fähle, Peter Müller und Lydia Christine Michel die Verantwortung übernommen – auch dafür vielen Dank.

Stuttgart, im August 2023

Prof. Dr. Gerald Maier  
Präsident des Landesarchivs Baden-Württemberg

# Einführung

Von DANIEL FÄHLE

Der Prozess der digitalen Transformation ist in vielen Archiven in vollem Gange – naturgemäß mit unterschiedlichen Geschwindigkeiten, die den unterschiedlichen Ressourcen und Leistungsfähigkeiten folgen. Fortschritte beim digitalen Wandel sind aber auch abhängig von Offenheit und Aufgeschlossenheit oder dem, was man heute *Mindset* nennt – und zwar der jeweiligen Archiveinrichtung als Ganzes oder auch des betreffenden Archivpersonals. Entgegen einer gesamtgesellschaftlich in weiten Teilen eher von Technologieskepsis geprägten Diskussion, formulierte der Titel des Archivtags mit den Attributen „smart und intelligent“ eine deutlich positive Konnotation der technologischen Möglichkeiten, die nicht zuletzt auch aus den neuen und innovativen KI-gestützten Techniken resultieren. Im Mittelpunkt der Fachbeiträge standen daher auch die Potenziale und Möglichkeiten, die digitale Werkzeuge und neue Technologien für die Archivarbeit mit sich bringen.

Natürlich wurde dazu auch der *intelligente* Textgenerator chatGPT befragt. Die KI antwortete so:

*Digitale Unterstützung bietet im Archivbereich zahlreiche Chancen und Herausforderungen. Einerseits können digitale Tools und Systeme die Arbeit im Archiv erleichtern und beschleunigen, indem sie beispielsweise die Suche nach bestimmten Dokumenten oder Informationen automatisieren. Andererseits erfordert die Einführung und Nutzung solcher Systeme eine gewisse Einarbeitungszeit und Schulung der Mitarbeiterinnen und Mitarbeiter. Zudem müssen Datenschutz- und Sicherheitsaspekte berücksichtigt werden, um die Vertraulichkeit und Integrität der archivierten Daten zu gewährleisten. Insgesamt bietet die digitale Unterstützung im Archivbereich jedoch viele Vorteile und kann dazu beitragen, die Arbeit effizienter und effektiver zu gestalten.*

Das kann bereits als recht qualifizierte Antwort betrachtet werden. Es bliebe hinzuzufügen, dass wir angesichts der Mengen an Daten, sei es z. B. aus der Digitalisierung analogen Archivguts oder des Umfangs und der Varianz originär digitaler Unterlagen gar nicht umhinkommen werden, Automatisierung und digitale Unterstützung zur massenhaften Organisation, Bearbeitung und Aufbereitung einzusetzen, um angesichts von *Big Data* nicht gar das Un-Wort von der Alternativlosigkeit zu verwenden. Erforderlich wird dies zumal, um den mit den technologischen Möglichkeiten wachsenden Ansprüchen unserer Kunden, Behörden und Archivnutzenden zu begegnen.

Bei der Auswahl an digitalen Assistenten und Services, die in den Fachbeiträgen aufgegriffen wurden, standen praxisorientierte und bereits oder absehbar verfügbare Produkte, Softwares oder Ansätze im Vordergrund. Einige Angebote davon wenden sich dezidiert auch an kleine Archive und Einrichtungen, indem dem Umstand unzureichender IT-Infrastruktur oder dem Fehlen von Knowhow oder Personal Rechnung getragen wurde.



Das Fachprogramm startete mit dem einführenden Keynote-Vortrag von Matthias Razum (FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur). Mit einem Blick von außen, zugleich aber mit inzwischen langjähriger Erfahrung aus der engen Kooperation mit Archiven, gab er Impulse zu zentralen Handlungsfeldern und konkreten Möglichkeiten, wie Archivarinnen und Archivare den digitalen Wandel aktiv gestalten können. Er forderte die Archive auf, ihre Bestände über ihre bisherigen Publika hinaus sicht- und auffindbar zu machen, sich zu vernetzen und damit auch einen Beitrag für die sich gerade etablierenden Methoden und Werkzeuge der digitalen historischen Forschung zu leisten.

Regina Keyler (Universitätsarchiv Tübingen) und Dorothee Huff (Universitätsbibliothek Tübingen) griffen mit ihrer Vorstellung der Ergebnisse des Projekts OCR-BW zu Texterkennungsverfahren eines der bedeutsamen Handlungsfelder für die Archive auf. Dieses kann angesichts des erheblichen Anteils von textlichen Quellen an der archivalischen Überlieferung ohne Übertreibung als regelrechte Schlüsseltechnologie gelten. Beantwortet wurde die Frage, ob und wie sich der Einsatz der entsprechenden Werkzeuge insbesondere zur Handschriftenerkennung (HTR) ganz praktisch und mit überschaubarem Aufwand für Archive umsetzen lässt.

Gegenstand des Beitrags von Andreas Neuburger (Landesarchiv Baden-Württemberg) war das DFG-finanzierte Vorhaben *EEZU*. Adressaten dieses Projekts bzw. des daraus resultierenden Angebots, eines *einfachen Erschließungs- und Zugriffssystems*, seien kleinste und kleine Archive, denen häufig die notwendige IT-Infrastruktur für die digitale Erfassung und Verwaltung sowie Onlinebereitstellung ihrer Bestände fehle. Diese könnten mithilfe von *EEZU* in die Lage versetzt werden, ohne größere technische Hürden am digitalen Wandel zu partizipieren. *EEZU* wird als Open-Source-Software zur Verfügung stehen und von FIZ Karlsruhe zusätzlich als gehosteter, kostenpflichtiger Dienst angeboten.

Das Handlungsfeld Zugänglichmachung und Nutzung bildete den Hintergrund für den Vortrag von Lambert Kansy (Staatsarchiv Basel-Stadt), der die Entwicklung des Digitalen Lesesaals in den Staatsarchiven Basel-Stadt und St. Gallen vorstellte. Er skizzierte die wesentlichen Etappen und Herausforderungen bei dessen Konzeption und Realisierung, die dem Anspruch folgten, den Zugang zu Archivgut innovativ und offen zu gestalten. Der von ihm präsentierte Entwicklungsstand und der Ausblick auf die weitere Entwicklung vermittelten eindrücklich den vollzogenen Perspektivwechsel weg von der Anbieterseite hin zu einer dezidiert nutzerorientierten Gestaltung.

Im Beitrag von Pascal Notz (Zentrum für Datenverarbeitung der Universität Tübingen, ZDV) ging es – nun bezogen auf audiovisuelle Dokumente – ebenfalls um die Frage, wie ein besserer Zugang zu diesen Quellen ermöglicht werden kann. Mit *timms* – dem am ZDV Tübingen im Aufbau befindlichen Streamingservice für Kultureinrichtungen wurde im Rahmen eines gemeinsamen Projekts mit dem Landesarchiv Baden-Württemberg eine Infrastruktur geschaffen, um digitale Ton- und Filmdokumente fachgerecht zu sichern und vor allem ein attraktives Streamingangebot bereitzustellen. Die inhaltliche Basis des Angebots bildeten in der ersten Ausbaustufe die Daten des AV-Archivs des Landesarchivs sowie der digitalisierte Teil der Landesfilmsammlung bzw. des Haus des Dokumentarfilms.

Um bildliche Quellen ging es auch im Beitrag von Florian Spiess (Universität Basel), der aktuelle und zukünftige Methoden der Multimedia-Forschung und deren Anwendung auf Archivdatenbestände an konkreten Beispielen vorstellte. Mittels dieser Verfahren lassen sich Objekte und Informationen in historischen Bildern und in Filmen identifizieren oder bildbeschreibende Zusammenfassungen in Textform generieren. Durch diese neuen Methoden unter Einsatz maschinellen Lernens können riesige Multimedia-Sammlungen nicht nur automatisch analysiert, sondern auch durch Suche und Exploration zugänglich gemacht werden.

In dem abschließenden Beitrag erläuterte Tobias Hodel (Universität Bern) die technischen Grundlagen von Large Language Models bzw. Textgeneratoren wie ChatGPT und beleuchtete die möglichen Chancen und Risiken für deren Einsatz in Archiven. In der Abwägung des Für und Wider dieser häufig auch als *Gamechanger* auf dem Feld der KI betrachteten Technologie verwies er einerseits auf die bestehenden Defizite und daraus folgend den Bedarf einer kritischen Auseinandersetzung. Andererseits betonte er die Potenziale bei der Verarbeitung und Auswertung großer Datenmengen und zugleich die Notwendigkeit, durch aktive Anwendung und Bereitstellung von offenen Datensätzen die Systeme an die Bedürfnisse von Erinnerungsinstitutionen anzupassen und gleichzeitig die Algorithmen und ihre Grundlagen besser zu verstehen.

Fazit: Wenn nicht bereits im Vorhinein bekannt, so wurde in den Fachbeiträgen des Archivtags deutlich, dass bereits ein ganzer Strauß an digitalen Werkzeugen und Methoden – wenngleich mit durchaus noch unterschiedlichen Reifegraden – vorhanden ist. Wichtig ist, die richtige Auswahl für den jeweiligen Bedarf und Einsatzzweck zu treffen, sich also einen Überblick über geeignete Werkzeuge und Methoden zu verschaffen. Deutlich wurde auch, dass bei aller angestrebten Niederschwelligkeit der digitalen Assistenten und Angebote freilich auch eine hinreichende Digitalkompetenz auf der Anbieter- wie Nutzerseite bei deren Verwendung unerlässlich ist. Und sicher wird beim Einsatz digitaler Hilfsmittel jeweils auch immer über Hindernisse und Risiken zu reden sein, was jedoch nicht den Blick auf den Nutzen und die Potentiale verstellen sollte.