

***Nachgefragt: Digitalisate von analogem Archivgut online
Professionelle Bereitstellung von Digitalisaten in Archivischen
Fachinformationssystemen (AFIS) – ein KLA-Diskussionspapier¹***

1) Einleitung

Menschen, die im Netz nach (historischen) Informationen suchen, erwarten schnelle, zielgenaue Treffer zu ihrer Recherche sowie sinnvolle Informationen und qualifizierte Kontextdaten, die sie schnell und richtig verstehen und einordnen können. Um die Rechercheergebnisse sinnvoll weiterverarbeiten zu können, müssen die bereitgestellten Daten dauerhaft referenzierbar sein. Ferner wird häufig davon ausgegangen, dass alle vorhanden Daten und Informationen, etwa von Archiven, online zur Verfügung stehen und recherchierbar sind.

Archive sind auf die Organisation und Bereitstellung² von Informationen spezialisiert. Archivar*innen sind Profis für authentische und qualitativ hochwertige, d.h. vertrauenswürdige, referenzierbare und nachvollziehbare Informationen. Archive sind ungeachtet unterschiedlicher Profile und Träger für unikales Kulturgut zuständig. Sie bilden eine wichtige Säule der Informationsinfrastruktureinrichtungen³ und müssen sich mit ihrem Angebot in diesem Umfeld behaupten und attraktiv bleiben. Die Bereitstellung von archivischen Informationen ist eine der Kernaufgaben von Archiven.

Diese Aufgabe muss vor allem angesichts veränderter technischer Möglichkeiten und sich wandelnder Erwartungen permanent weiterentwickelt und professionalisiert werden.⁴ Ausgelöst durch die Corona-Pandemie ist der Bedarf an der Online-Bereitstellung von (archivischen) Informationen nochmals erheblich gewachsen. Die Verfügbarkeit von Digitalisaten von analogem Archivgut spielt dabei eine enorme Rolle. Die Bereitstellung von Digitalisaten ist also eine dynamische Gegenwarts- und Zukunftsaufgabe von Archiven im Kontext der Digitalisierung von Arbeitsprozessen. Deshalb ist sie erneut Gegenstand eines Diskussionspapiers der KLA.

¹ Der vorliegende Beitrag wurde vom KLA-Ausschuss „Archivische Fachinformationssysteme“ (AFIS) unter Federführung von Bettina Joergens erarbeitet und von der KLA (Konferenz der Leiterinnen und Leiter der Archivverwaltungen im Bund und in den Ländern) in der 133. Sitzung am 27.09.2021 zur Veröffentlichung im „Archivar“ freigegeben. Folgenden Kolleg*innen des AFIS-Ausschusses der KLA und aus verschiedenen staatlichen Archivverwaltungen danken wir für hilfreiche und inspirierende Hinweise: Daniel Fähle, Sabine Fees, Bastian Gillner, Johannes Haslauer, Martin Hoppenheit, Nathalie Huck, Franziska Klein, Benjamin Kram, Benedikt Nientied, Natascha Noll, Nils Reichert, Peter Sandner, Stephan Switaiski, Nils Volkening, Christian Wortmann.

² Mit Bereitstellung sind in diesem Text alle Prozesse und Vorgänge gemeint, die für die Nutzung und Veröffentlichung von archivischen Informationen, Digitalisaten und genuin elektronischen Archivalien erforderlich sind, also etwa auch die Verzeichnung oder die Verknüpfung von Erschließungsdaten mit Digitalisaten.

³ Robert Kretzschmar hat als einer der ersten im Archivwesen Archive als Teil der Informationsinfrastruktur diskutiert, s. Robert Kretzschmar, Archive als digitale Informationsinfrastrukturen. Stand und Perspektiven, in: Archivar, 66 (2013), H. 2, S. 146-153, s.a.

https://www.archive.nrw.de/sites/default/files/media/files/ARCHIVAR_02-13_internet.pdf (aufgerufen am 17.02.2021).

⁴ Vgl. zum Wechselspiel zwischen (potentiellen) Nutzer*innen und Archiven in Bezug auf Nutzungsmöglichkeiten aus historischer Perspektive: Philipp Müller, Geschichte machen. Historische Forschung und die Politik der Archive, Göttingen 2019.

Im Folgenden wird die Bereitstellung von Digitalisaten im Zusammenhang der Erschließungsmetadaten, der Archivinformationssysteme, der Bereitstellung von Archivgut und archivischen Informationen insgesamt und der Dateninfrastruktur betrachtet. Dabei spielen Datenspeichersysteme und Datenverarbeitungsprozesse ebenso eine Rolle wie das Informationssicherheitsmanagement⁵ und die Trennung von interner und externer Nutzung sowie die Pflege der Daten. Deshalb wird in dieser „Bestandsaufnahme mit Ausblick“ zur Bereitstellung von Digitalisaten von analogem Archivgut in Archivinformationssystemen das Themenfeld entsprechend weit gesteckt, ohne die einzelnen Aspekte vertiefen zu können. Rechtliche Aspekte sowie Fragen der Öffentlichkeitsarbeit, der Organisation oder der Ressourcen stehen dabei nicht im Mittelpunkt, sondern werden eher cursorisch gestreift, wenngleich sie eine enorme Rolle bei der Bereitstellung und Veröffentlichung von archivischen Informationen spielen.⁶

Vorliegendes Diskussionspapier ist vor allem als Zusammenfassung der aktuellen Erkenntnisse zu einer optimalen Bereitstellung von Digitalisaten in Verbindung mit Erschließungsmetadaten im Archivwesen zu verstehen. Perspektivisch sollen Erfordernisse zur weiteren Professionalisierung formuliert und eine grobe Handlungsempfehlung gegeben werden. Auf eine „Best-Practice-Checkliste“ wird verzichtet, da das hier zu beleuchtende Aufgabenfeld einer enormen Dynamik unterliegt und der Austausch zwischen den Archiven i.S. eines weiteren Erkenntnisgewinns erst noch intensiviert werden muss.

Im Folgenden werden zunächst 2) die Ausgangssituation und die Genese des aktuellen Stands dargelegt. Im Anschluss wird der gesamte Themenkomplex in folgende Aspekte gegliedert: 3) Daten, 4) Datenorganisation, 5) Datenverarbeitungsprozesse. Abschließend 6) folgen eine Zusammenfassung und ein Ausblick.

2) Ausgangssituation und Genese

Als etwa Mitte der 1990er Jahre das Internet einem größeren Teil der Menschen in Deutschland bekannt und zugänglich wurde, waren nur wenige „Institutionen des deutschen Archivwesens im Netz vertreten“⁷. Erst allmählich drangen Fragen zur Nutzung der relativ neuen Technologie für die Aufgaben von Archiven in die Fachdiskussion ein.⁸ Im Jahr 1999

⁵ Nach BSI – Richtlinie (Grundschutz-Katalog, BSI-100-1 bis 200-4 Standard)

(https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/BSI-Standards/bsi-standards_node.html, aufgerufen am 29.07.2021).

⁶ Die Archivgesetze, aber auch das Urheberrecht und die Datenschutzgesetze regeln die Bereitstellung und Veröffentlichung von archivischen Informationen und Archivalien. Vgl. etwa das KLA-Empfehlungspapier des Ausschusses „Archive und Recht“ (Cindy Braun und Jörn Brinkhus) „Die bisherigen Reformen des Urheberrechts und die Onlinestellung digitalisierter oder digitaler Archivguts“(2019)

(<https://www.bundesarchiv.de/DE/Content/Downloads/KLA/aur-gutachten-onlinestellung-digitalisate.pdf?blob=publicationFile>, aufgerufen am 29.07.2021). Beim DFG-Rundgespräch zum Umgang mit urheberrechtlichem Material am 28.04.2021 wurden wegweisende Ansätze hinsichtlich einer digital gesteuerten Bereitstellung von in diesem Fall urheberrechtlich geschützten Materialien vorgestellt, s. z.B. den Beitrag von Simon Herrmann: „Impuls mit anschließender Diskussion: Rechtklärung in der Deutschen Nationalbibliothek: Workflows und Tools – Möglichkeiten und Perspektiven in Massengeschäftsgängen (GND, vergriffene Werke, Dokumentation)“.

⁷ Karsten Uhde, Das Internet-Archiv – Möglichkeiten und Grenzen der Arbeit im und mit dem Internet für Archive, in: Frank M. Bischoff / Wilfried Reininghaus (Hg.), Die Rolle der Archive in Online-Informationssystemen. Beiträge zum Workshop im Staatsarchiv Münster, 8.-9. Juli 1998, Münster 1999, S. 19-39, hier: 19.

⁸ Ebda.

waren „Informationen zu über 1000 Häusern [Archiven] im Internet vertreten“.⁹ Das Archivportal archive.nrw.de ging 1998 ans Netz und war seinerzeit eines der ersten und größten Portale, die archivische Informationen bündelten.¹⁰ Inzwischen wurde es mehrfach erneuert, zuletzt 2020,¹¹ und steht längst nicht mehr alleine in der archivischen Informationslandschaft. 2009 veröffentlichte die Archivreferentenkonferenz (heute KLA) „Empfehlungen zur Weiterentwicklung der Präsentationen von Erschließungsinformationen im Internet“.¹²

Im Jahr 2014 wurde das Archivportal-D online geschaltet.¹³ Mit diesem Meilenstein wurde nicht nur ein bundesweites Archivinformationssystem geschaffen, sondern auch ein Aggregator für die spartenübergreifende Deutsche Digitale Bibliothek und die internationalen Plattformen Archivportal Europa und die Europeana – mit dem Ziel, möglichst viele Informationen aus Kulturerbeinstitutionen zu bündeln und zugänglich zu machen,¹⁴ wie auch in dem KLA-Positionspapier „Archive und Online-Portale. Thesen des ARK-Ausschusses ‚Archivische Informationssysteme‘“ von 2014 dargelegt¹⁵.

Die gemeinsame Bereitstellung von digitalen Daten und Metadaten zu Archivgut verlangte eine internationale Standardisierung von Erschließungsdaten und Austauschformaten, also einen Abschied von analogen Findbüchern, Karteikarten und spezifischen nationalen und lokalen Erschließungstraditionen. Teilweise ist dieser in den 1990er Jahren in der gesamten, auch internationalen, Fachcommunity vorangetriebene Prozess immer noch nicht abgeschlossen.¹⁶ Wie noch zu zeigen sein wird, wird künftig weiter über Standardisierungen für Daten, Metadaten und Datenverarbeitungsprozesse nachgedacht werden müssen, um u.a. angesichts von stetig wachsenden Datenmengen sowie aus Effizienzgründen Verarbeitungs- und Bereitstellungsprozesse (weiter) zu automatisieren. Parallel zu diesen Entwicklungen forcierten die Archive die Digitalisierung von Archivgut. Digitalisate dienten zunächst und in erster Linie als Schutzmedium, um die bislang in den Lesesälen bereitgestellten Mikrofilme und Mikrofiches allmählich abzulösen oder zumindest zu ergänzen. Wenngleich die Digitalisierung nach wie vor auch aus der Bestandserhaltung motiviert ist, so spielen Nutzungsinteressen und Fragen der Online-Fähigkeit (hinsichtlich der Daten-/Metadatenqualität und rechtlicher Bedingungen) eine (immer) größere, wenn nicht

⁹ Frank M. Bischoff, Archivische Informationsvermittlung im Wandel: Internetverbund, Rechterservice und Datenpflege in Nordrhein-Westfalen, in: Ders. / Reininghaus 1999, S. 58-83, hier S. 58.

¹⁰ Ebd.

¹¹ Vgl. aktuell mit weiteren Referenzen: Bettina Joergens, Das neue Portal für alle Archive in NRW – der Relaunch von archive.nrw.de, in: Archivar, 73 (2020), H. 4, S. 384-387; dies., Das neue Portal für alle Archive in NRW, in: Archivpflege in Westfalen-Lippe, (2021), H. 93/94, S. 29-44.

¹² <https://www.bundesarchiv.de/DE/Content/Artikel/Ueber-uns/Partner/KLA/kla-ausschuss-afis.html> (aufgerufen am 23.08.2021).

¹³ Die Veröffentlichung fand am 24.09.2014 beim 84. Deutschen Archivtag in Magdeburg statt: <https://www.archivportal-d.de/content/archivportal-d/startschuss-fuer-das-archivportal-d-deutschlands-archive-jetzt-online-leichter-zugaenglich> (aufgerufen am 17.02.2021).

¹⁴ Vgl. Kerstin Arnold / Susanne Waidmann, Vernetzte Präsentation. Erfahrungen mit Portalen, in: Archivar, 66 (2013), H. 4, S. 431-438; s.a.

https://www.archive.nrw.de/sites/default/files/media/files/Archivar_2013_4_Internet.pdf (aufgerufen am 17.02.2021); vgl. <https://www.ica.org/en/kerstin-arnold> (aufgerufen am 29.07.2021).

¹⁵ S. https://www.bundesarchiv.de/DE/Content/Downloads/KLA/archive-und-onlineportale.pdf?__blob=publicationFile (aufgerufen am 17.02.2021).

¹⁶ S. aktuell auch zur Genese: Silke Jagodzinski, Internationale Standards für archivische Erschließung. Von ISAD(G) zu Records in Contexts, in: Archivar, 73 (2020), H. 2, S. 134-141, s.a. https://www.archive.nrw.de/sites/default/files/media/files/Archivar_2020-2_Internet.pdf (aufgerufen am 17.02.2021).

inzwischen maßgebliche Rolle bei der Auswahl der zu digitalisierenden Archivalien. Hierfür ist außerdem inzwischen zunehmend das Potential zur automatisierten Erschließung durch Handschriftenerkennung, Keyword Spotting oder vergleichbare Technologien ausschlaggebend. Viele Impulse kamen und kommen aus der Drittmittelförderung. Insbesondere die DFG beschleunigte diese Entwicklung, v.a. durch den DFG-Produktivpiloten „Digitalisierung von archivalischen Quellen“ in den Jahren 2013-2015.¹⁷ Im Zuge dessen wurden der DFG-Viewer zur Online-Anzeige von Digitalisaten entwickelt und der Metadatenstandard METS/MODS (Metadata Encoding & Transmission Standard¹⁸) als Containerformat zur Verbindung von Erschließungs- und Bilddaten (Digitalisaten) festgelegt.¹⁹

Gleichzeitig stießen, so Andrea Wettmann in der Einführung zum Workshop „Digitalisierungs- und Dateninfrastruktur in der Portallandschaft“ am 27. Januar 2021, nach dem bemerkenswerten „Digitalisierungsboom“ der 2010er nahezu alle staatlichen Archivverwaltungen von Bund und Ländern bei der Speicherung und Bereitstellung von Digitalisaten an die Grenzen ihrer Infrastruktur und Ressourcen.²⁰ Für eine nachhaltige Langzeitspeicherung fehlten anfangs die Erfahrungen, Ressourcen und Strukturen; auch bei der Bereitstellung der Digitalisate mangelte es zunächst an der notwendigen Anpassung organisatorischer Strukturen und archivischer Prozessabläufe.

Im September 2018 formulierte die KLA daher in einem Positionspapier mit dem Titel „Das digitale Gedächtnis nachhaltig aufbauen. Digitalisierung archivischer Quellenbestände, ihre Speicherung und Bereitstellung im Netz“²¹ die Anforderung, funktionierende Strukturen zu schaffen bzw. vorhandene Strukturen weiterzuentwickeln, um digitalisiertes Archivgut für eine breite Nutzung durch Wissenschaft und Gesellschaft im Internet nachhaltig anbieten zu können. Um diese Anforderung zu erfüllen, sieht das Papier das Erfordernis, „von Seiten der Archivträger nicht nur temporär Mittel für die Digitalisierung bereitzustellen, sondern dauerhaft in den Etats zu verankern“. Dies gilt in noch stärkerem Maß auch für die erforderlichen Investitionen in „den Aufbau und den nachhaltigen Betrieb der technischen Infrastruktur“, also der Digitalisierungs- und Dateninfrastruktur sowie den damit verbundenen Speicherlösungen und Datennetzen. Die KLA betont in dem 2019 veröffentlichten Positionspapier „Zum Umgang mit Digitalisaten“ außerdem die Notwendigkeit der kontextualisierten und vernetzbaren Bereitstellung von Digitalisaten auf der Basis von Standard-Metadaten.²²

¹⁷ Vgl. <https://www.lwl-archivamt.de/de/fachinformationen/digitalisierung/dfg-projekt/> (aufgerufen am 17.02.2021); vgl.

https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/digitalisierung_erschliessung/index.html (aufgerufen am 17.02.2021).

¹⁸ https://www.loc.gov/standards/mets/METSOverview.v2_de.html (aufgerufen am 17.02.2021).

¹⁹ <http://dfg-viewer.de/metadaten> (aufgerufen am 17.02.2021).

²⁰ Ergebnisbericht über den Workshop: Das digitale Gedächtnis nachhaltig aufbauen. Empfehlungen zur Weiterentwicklung der Digitalisierungs- und Dateninfrastruktur in der Portallandschaft. Bericht über die Ergebnisse des vom KLA-Ausschuss AFIS durchgeführten Workshops am 27. Januar 2021 zur Vorlage an die 132. KLA am 22./23. März 2021 in Schwerin. Vgl. in diesem Zusammenhang Empfehlungspapier des Fototechnischen Ausschusses der KLA „Wirtschaftliche Digitalisierung in Archiven“ aus dem Jahr 2015 (<https://www.bundesarchiv.de/DE/Content/Downloads/KLA/wirtschaftliche-digitalisierung.pdf?blob=publicationFile>, aufgerufen am 17.02.2021).

²¹ <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/positionspapier-digitalisierung-archivischer-quellen.pdf?blob=publicationFile> (aufgerufen am 17.02.2021).

²² <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/sachstand-umgang-mit-digitalisaten.pdf?blob=publicationFile> (aufgerufen am 17.02.2021).

An diesen Stand der Fachdiskussion und der Positionierung der KLA knüpft die vorliegende Abhandlung an und möchte einen Beitrag leisten, um den gesamten Komplex der Bereitstellung von archivischen Informationen weiter zu professionalisieren, nachhaltig zu organisieren und nutzerorientiert auszurichten – i.S. des Rechtsstaatsprinzips und der Aufgabe der Archive als Informationsinfrastruktureinrichtungen und des gesetzlichen Auftrags: Die E-Government-Gesetze (EGovG)²³ und das Online-Zugangsgesetz (OZG)²⁴ beinhalten letztlich auch an Archive und ihre Träger die Aufforderung, die Entwicklung und Bereitstellung der Archivinformationssysteme als Teil des E-Government-Auftrags zu betrachten – auch im Hinblick auf die Ressourcen. Die Präsentationen und Diskussionen beim Online-Workshop des KLA-Ausschusses AFIS „Digitalisierungs- und Dateninfrastruktur in der Portallandschaft“ am 27. Januar 2021 sind dabei wichtige Referenzen.

3) Daten und Metadaten

Daten bzw. Metadaten sind nur sinnvolle Daten, wenn sie zur Erfüllung von Aufgaben, in diesem Fall zur Speicherung und Bereitstellung durch die Archive, genutzt, weiterverarbeitet und referenziert werden können. Ansonsten hat man es nämlich mit unbrauchbaren Daten zu tun. Insbesondere für den Bereich der Forschungsdaten wurden mit den FAIR-Prinzipien Qualitätskriterien für Daten entwickelt. Den FAIR Data Principles zufolge müssen Daten und Metadaten „Findable, Accessible, Interoperable, and Re-usable“ sein. Dies wird insbesondere erreicht durch Persistent Identifiers, ausreichend aussagekräftige Metadaten (unter Einbeziehung kontrollierter Vokabulare und übergreifender Normdateien), Durchsuchbarkeit mit Indexen, standardisierte Datenbeschreibungen und interoperable Schnittstellen, weitgehende Zugänglichkeit bzw. Lizenzen zur rechtssicheren Nachnutzung, (Maschinen-)Lesbarkeit und Kontextualisierung.²⁵ Archive sollten die FAIR Data Principles als Qualitätsmerkmale für die eigenen Daten und Metadaten dort festlegen, wo nicht ohnehin schon geschehen. In Erweiterung der FAIR-Prinzipien wurden mittlerweile die CARE-Prinzipien entwickelt, die der Datenverwendung eine ethische Komponente hinzufügen.²⁶

In diesem Abschnitt ist ohne Anspruch auf Vollständigkeit zunächst zu klären, welche Daten und Metadaten in archivischen Informationssystemen bereitgehalten und bereitgestellt werden. Dabei wird aus systematischen Gründen zwischen Daten und Metadaten unterschieden. Die Unterscheidung ist oft eine Frage der Perspektive darauf, welche Daten Primärdaten sind und welche Daten diese Primärdaten beschreiben und damit Metadaten sind. Erschließungsdaten von Archivgut (z.B. in EAD (DDB)) oder beschreibende Daten für

²³ Gesetz zur Förderung der elektronischen Verwaltung (E-Governmentgesetz, EGovG), https://www.gesetze-im-internet.de/egovg/inhalts_bersicht.html, aufgerufen am 23.08.2021.

²⁴ S. z.B. auf der Seite des Bundesministeriums des Innern: <https://www.onlinezugangsgesetz.de/Webs/OZG/DE/startseite/startseite-node.html> (aufgerufen am 27.07.2021).

²⁵ S. https://www.forschungsdaten.org/index.php/FAIR_data_principles; <http://www.go-fair.org/fair-principles/> und vor allem: https://www.ruhr-uni-bochum.de/researchdata/de/fair_data.html (zuletzt aufgerufen am 09.07.2021).

²⁶ Vgl. The CARE Principles for Indigenous Data Governance (<https://www.vad-ev.de/care-principles/>; aufgerufen am 17.08.2021). Vgl. Arvid Deppe: FAIR, CARE und mehr. Prinzipien für einen verantwortungsvollen Umgang mit Forschungsdaten, in: Matthias Schulze (Hrsg.): Historisches Erbe und zeitgemäße Informationsinfrastrukturen: Bibliotheken am Anfang des 21. Jahrhunderts. Kassel 2020, S. 299-312.

Bilder / Images (z.B. METS/MODS) sind somit Metadaten, denn sie beschreiben jeweils als Primärdaten angesehene Archivalien bzw. Images/Digitalisate.²⁷

3.1 Daten – welche Daten?

Für eine gute – archivfachliche – Organisation (Speicherung, Verarbeitung und Bereitstellung sowie dauerhafte Verfügbarkeit) von Daten ist es zunächst essentiell festzulegen, welche Daten-Inhaltstypen oder Elemente i.S. des Datenmodells (z.B. EAD(DDB) oder EAD3) in der Hierarchie archivischer Erschließung nach ISAD(G)²⁸ berücksichtigt werden sollen und in welchem Format:²⁹ Handelt es sich um Basiserschließungsangaben, um zusätzliche Informationen aus Transkriptionen oder Indexierungsprojekten, um Normdaten, um Digitalisate analogen Archivguts oder um genuin elektronisches Archivgut, um die Beschreibung von Einzelblättern innerhalb einer Verzeichnungseinheit usw.? Dabei ist es wichtig, sich auf eine begrenzte Auswahl von Daten-Formaten zu einigen.

Ferner muss schon aus Datenschutzgründen definiert werden, ob es sich um Daten handelt, die ausschließlich intern genutzt (z.B. Sicherungen) oder für die Nutzung herausgegeben bzw. veröffentlicht³⁰ werden. Darüber hinaus wird insbesondere im Archivwesen unterschieden, woher Daten stammen (selbst erstellt, aus abgebenden Behörden oder von Nutzer*innen generiert) und auf welcher Ebene der Informationshierarchie sich die Daten befinden (Tektonik oder Bestand, Klassifikation oder Verzeichnungseinheit). In diesem Text stehen die Digitalisate von analogem Archivgut als Primärdaten im Zentrum.

Der Fototechnische Ausschuss der KLA hat im Jahr 2016 ein Empfehlungspapier zu den Speicherformaten für Digitalisate³¹ insbesondere mit Blick auf die Wirtschaftlichkeit und Folgekostenberechnung erstellt.³² Hierin wird zwischen dem TIFF-Format und dem JPEG-Format (auch komprimiert in PDF/A) in unterschiedlich hohen Auflösungen unterschieden. Weitergehende Erfahrungen und fachliche Diskussionen zeigen inzwischen, dass es aus wirtschaftlichen Gründen und i.S. der Nachhaltigkeit sinnvoll sein kann, die Parameter „verlustfrei“ und „verlustbehaftet“ differenzierter und mit Blick auf die Ziele von Archiven

²⁷ Vgl. Martin Warneke, Daten und Metadaten – Online-Ressourcen für die Bildwissenschaft, in: zeitenblicke, 2 (2003) Nr. 1 (<http://www.zeitenblicke.de/2003/01/warneke/warneke.pdf>); S.a.

<https://www.ionos.de/digitalguide/websites/web-entwicklung/metadaten/> (aufgerufen am 07.07.2021);

<https://de.wikipedia.org/wiki/Metadaten> (aufgerufen am 07.07.2021).

²⁸ Vgl. https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_DE.pdf; <https://landesarchiv-bw.de/de/landesarchiv/fachaufgaben---jahresberichte/erschliessung---digitalisierung/49600> (aufgerufen am 21.07.2021).

²⁹ An dieser Stelle sei Nils Meyer für hilfreiche Hinweise zur Ausgestaltung von EAD 3 und wie dieses Schema auf ISAD(G) basiert ist gedankt. Zusatzinformationen zu Verzeichnungseinheiten, etwa Indexierungsdaten, Transkriptionen, Erschließungen, können auf der Ebene der Verzeichnungseinheit (die nicht unbedingt eine einzelne Akte beschreiben, sondern ggf. auch einen ganzen Bestand) erfasst und präsentiert werden. EAD 3 berücksichtigt für die Darstellung dieser Informationen unterschiedliche Repräsentationen.

³⁰ Nutzung ist hier i.S. der Archivnutzung nach den archivrechtlichen Regelungen zu verstehen. Nutzung erfolgt also durch benannte Einzelpersonen in einzelnen Nutzungsvorgängen. Dagegen ist die Bereitstellung von digitalem oder digitalisiertem Archivgut im Internet für eine unbestimmte Öffentlichkeit eine Veröffentlichung.

³¹ Digitalisate sind Reprographien von analogem Archivgut und nicht zu verwechseln mit genuin elektronischen Unterlagen, zu denen kein analoges Pendant existiert. Zu genuin elektronischen Unterlagen sind in diesem Fall auch die in Behörden angefertigten Ersatzscans zu zählen, die dem Archiv anstatt des analogen Schriftguts angeboten werden. Denn auch sie kennen – systematisch betrachtet – kein analoges Gegenstück (mehr).

³² S. <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/wirtschaftliche-digitalisierung.pdf?blob=publicationFile> (aufgerufen am 18.02.2021).

(Langzeitspeicherung und Präsentation bei akzeptabler Performanz) zu diskutieren.³³ Auch die Frage des Primärdatenformats für Digitalisate wird sicherlich die Fachwelt noch weiter beschäftigen, soll hier aber nicht vertieft werden.

Für die Klassifikation und Organisation von Daten interessiert hier vielmehr die Differenzierung von Nutzerreprographien und Schutzdigitalisaten für den Langzeiterhalt: Im Unterschied zu den Digitalisaten, die dauerhaft erhalten werden, wählen Archive bisher in der Regel für die von Nutzer*innen individuell angeforderten Reproduktionen von Archivgut einen deutlich vereinfachten Standard hinsichtlich des Dateiformats, der Dateibenennung und des (vorübergehenden) Speicherortes. Dieser vereinfachte Standard reicht allerdings nicht mehr aus, wenn „on Demand“ erstellte Digitalisate im Anschluss dauerhaft erhalten werden sollen.

Folglich werden an „Digitalisate on Demand“ doppelte Anforderungen gestellt: Sie werden nach individuellem Bedarf erstellt und bereitgestellt sowie gleichzeitig nach den technischen Standards für dauerhaft abzuspeichernde Digitalisate hergestellt, mit Metadaten versehen und systematisch eingeordnet. Damit bilden sie eine Sonderform der systematischen Schutzdigitalisierung (neben den dauerhaft zu sichernden Schutzdigitalisaten (Master-Kopien) und den Nutzungsdigitalisaten), die nach eigenem Verfahren verarbeitet werden und das Prozessmanagement herausfordern.

Um die Abläufe einfach und effizient zu gestalten, empfiehlt es sich, die von Nutzer*innen individuell angestoßenen Digitalisierungen in den Gesamtprozess der Digitalisierung zu integrieren, wie dies im Bundesarchiv der Schweiz auf der Basis der Prozesssoftware Kitodo erfolgt.³⁴ Voraussetzung dafür ist, dass das komplette Archivale digitalisiert wird und nicht nur Einzelblätter daraus. Folglich müssen für diese Reproduktionen und deren Speicherung sowie ‚Verknüpfung‘ mit Metadaten dieselben hohen Standards gelten wie für die vom Archiv initiierten Digitalisierungen.

Voraussetzung für eine sinnvolle Nutzung, für das Verständnis von Daten und deren Auswertung sowie Weiterverarbeitung ist eine hinreichend hochwertige Datenqualität. Der Rede von „Daten als dem neuen Öl“ muss zumindest teilweise widersprochen werden: Nur qualitativ hochwertige Daten sind wertvoll, da nachnutzbar³⁵. Inzwischen wurden (schlecht genutzte) Daten schon als das neue CO² bezeichnet, da sie geeignet sind, Schaden anzurichten.³⁶ Die Qualität von Daten basiert nicht zuletzt auf der Qualität der jeweiligen

³³ Vgl. Benjamin Kram / Kristian Peters, Die Sichtbarmachung zerstörter Schrift auf frühneuzeitlichen Hadernpapieren durch die Multispektraldigitalisierung. Ergebnisse eines Pilotprojekts im Landesarchiv NRW an Beispielen der sogenannten „Kahnakten“, in: Zeitschrift für Bibliothekswesen und Bibliografie, Bd. 68 (2021), H. 4, S. 196-209, hier S. 197 (erscheint im Dezember 2021).

³⁴ Vgl. hierzu das Schweizer Bundesarchiv: Matthias Finck / Stefan Kwasnitza, Die Ablösung des Lesesaals: Hilfe von Kitodo.Production 3.0, in: b.i.t. online, 23 (2020) Nr. 2, S. 165-169, hier S. 166f. (<https://b-i-t-online.de/heft/2020-02-fachbeitrag-finck.pdf>, aufgerufen am : 29.07.2021). Außerdem bezieht sich die Ausführung auf das vom KLA-Ausschuss AFIS initiierte Fachgespräch mit dem Schweizer Bundesarchiv, Herrn Stefan Nellen, am 24.03.2021. Herr Nellen referierte zu der strategischen Neuausrichtung des Bundesarchivs und die Digitalisierung des Services. Vgl. Stefan Kwasnitza, Auf dem Weg zum komplett digitalen Archiv. Online-Zugang und Digitalisierung „on demand“ im Schweizer Bundesarchiv, in: Kultur und Politik (2020) Nr. 3 (s. <https://www.kulturrat.de/themen/erinnerungskultur/archiv/auf-dem-weg-zum-komplett-digitalen-archiv/?print=print>, aufgerufen am 29.07.2021).

³⁵ Vgl. Vortrag von Vilde Ronge, Business Needs. The Ace of Spades, Vortrag bei der RIMPA Convention 2020 (virtuell): <https://www.youtube.com/watch?v=aUfp00eWlwk> (aufgerufen am 29.07.2021).

³⁶ Vgl. <https://luminategroup.com/posts/blog/data-isnt-the-new-oil-its-the-new-co2> (aufgerufen am 17.08.2021).

Metadaten: „Metadata is at the heart of information systems intended to provide broad access to cultural heritage material which is often heterogeneous and published in diverse forms.“³⁷

3.2 Metadaten

Die digitale Bereitstellung von Archivgut und von Informationen über Archivgut zeichnet sich besonders durch eine strikte Struktur, also Ordnung der Informationen, aus. Diese Strukturinformationen sind wesentlich für das Verständnis und die Einordnung der archivischen Informationen. Hierbei müssen bereitgestellte Digitalisate des analogen Archivguts klar unterscheidbar sein von genuin elektronischen Unterlagen. Notwendig ist auch die Angabe zum Ursprung und damit möglicherweise zur Verbindlichkeit von Daten; das kann etwa davon abhängen, ob diese vom Archiv selbst erstellt wurden, in Drittmittelprojekten erzeugt wurden oder aus Crowdsourcing-Projekten stammen. Denn Erschließungsmetadaten machen (digitale) Informationen überhaupt erst effizient zugänglich und erfüllen damit eine ganz entscheidende – auch archivrechtlich normierte – archivische Aufgabe. Es kommt also auf die Qualität der Metadaten an.

Die Anwendung von standardisierten Metadaten ist essentiell für die grundsätzliche Identifikation³⁸, die systematische Speicherung, die automatisierte Bearbeitung, die Verknüpfung von Daten untereinander (z.B. mit dem METS-Container die „Verknüpfung“ von Digitalisaten mit Erschließungsdaten), die Nachnutzung bzw. Weiterleitung von Daten, etwa von einem Archivportal in ein anderes, sowie für die Nutzung von Recherchertools, KI-Anwendungen³⁹, Linked Open Data im Semantic Web oder den nachhaltigen Erhalt der Daten. Ausgehend von der Auseinandersetzung mit Digitalisaten, dem Hauptanliegen dieses Textes, interessieren hier die Daten, die Digitalisate beschreiben.

Das zumindest in Deutschland wichtigste Digitalisate beschreibende Metadatenformat METS/MODS wurde bereits mehrfach erwähnt.⁴⁰ Zur Verbindung von Erschließungsdaten

³⁷ Juliane Stiller / Péter Király, Multilinguality of Metadata. Measuring the Multilingual Degree of Europeana's Metadata, in: Maria Gäde / Violeta Trkulja / Vivien Petras (Hg.), Everything's Changes, Everything Stays the Same? Understanding Information Space. Proceedings of the 15th International Symposium of Information Science (ISI 2017) (Schriften zur Informationswissenschaft 70), Glückstadt 2017, S. 164-176, hier: S. 165 (http://isi2017.ib.hu-berlin.de/ISI_17_ONLINE_FINAL.pdf, aufgerufen am 07.07.2021).

³⁸ Worauf hier nicht eingegangen werden kann, was aber an anderer Stelle vertieft werden muss, sind technische Parameter bzgl. der Digitalisierung. Diese können entscheidend sein, wie von Nils Volkening berichtet: So gab es z.B. bei Xerox Business Scannern über acht Jahre lang einen Bug, der beim Digitalisieren ein Patternmatching betrieben hat, was dazu führen konnte, dass Zahlen oder gar ganze Textbausteine ausgetauscht worden sind, und so zu fehlerhaften Digitalisaten geführt hat. Da dieses im Nachgang nicht mehr nachgewiesen werden kann, wenn kein analoges Original mehr vorhanden ist, können solche Informationen im Nachhinein wichtige Rückschlüsse auf die Integrität und Verlässlichkeit der Quellen schließen. Wenn nun in den Metadaten die Informationen über das Gerät zur Digitalisierung und bestmöglich noch die Firmware-Version vermerkt ist, kann man so nachträglich noch Daten bewerten bzgl. verschiedener Kriterien.

³⁹ Vgl. zu KI: Künstliche Intelligenz. Wirtschaftliche Bedeutung, gesellschaftliche Herausforderungen, menschliche Herausforderungen, hg. v. Bitkom e.V. (Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.) / Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Berlin 2017 (https://www.dfki.de/fileadmin/user_upload/import/9744_171012-KI-Gipfelpapier-online.pdf); Andreas Wichert, Künstliche Intelligenz, in: Spektrum. Lexikon der Neurowissenschaft (<https://www.spektrum.de/lexikon/neurowissenschaft/kuenstliche-intelligenz/6810>); beides aufgerufen am 21.07.2021.

⁴⁰ An dieser Stelle wird lediglich auf den auf internationaler Ebene entwickelten Standard IIIF hingewiesen werden, der auf ein schlankes Datenaustauschformat setzt und zugleich über eine neue Generation von

mit Digitalisaten wird aus EAD(DDB) auf die METS/MODS-Dateien verlinkt. Basis sind also die standardisierten Erschließungsdaten nach EAD(DDB), einem Format, auf das man sich im Zuge der Konzipierung der Deutschen Digitalen Bibliothek bzw. des Archivportals D im deutschen Archivwesen geeinigt hat und das in dem internationalen XML-Austauschformat EAD (Encoded Archival Description) nach dem International Standard of Archival Description (General) (ISAD (G)) begründet ist.⁴¹ Die EAD-AG⁴² strebt nun an, die insbesondere für das Archivportal-D spezifizierte Ableitung von EAD, nämlich EAD(DDB) (aktuell 1.2), nicht weiter zu entwickeln, sondern sich dem internationalen Standard für EAD 3 anzuschließen.⁴³ Dies soll den internationalen Datenaustausch befördern, v.a. aber den nationalen Sonderweg beenden und dauerhafte Kompatibilität zum internationalen EAD-Standard herstellen. Die Hinweise auf diese beiden Standards METS/MODS und EAD (DDB) werden hier knappgehalten, weil sie sich in Deutschland weitgehend als selbstverständlich durchgesetzt haben. Allerdings ist hervorzuheben, dass immer noch nicht alle Softwareprodukte für die archivische Erschließung standardgemäße Daten liefern.⁴⁴ Obgleich eine strikte Validierung beim Upload oder Export/Import von Erschließungsdaten sinnvoll wäre, wird mit Rücksicht auf die „Datenlieferanten“ vereinzelt davon abgewichen. Dies führt dann leider zu nicht ausreichend validen Daten. In der Folge müssen nach wie vor in der Praxis teilweise händisch Anpassungen an den Daten vorgenommen werden, was effiziente und (teil-)automatisierte Verfahren verhindert.⁴⁵ Oder die Metadaten müssen aufgrund von Änderungen des Standards angepasst werden, wie dies durch Spezifikationsänderungen beim DFG-Viewer etwa im September 2020 für die METS-Dateien erforderlich war.⁴⁶ Folglich erfordert die Nutzung der Standards deren stetige Weiterentwicklung und Anpassung bei gleichzeitiger Erhaltung bzw. Herstellung von Konformität und Validität der eigenen Daten. Dafür wiederum bedarf es Fach- und Spezialwissens, personeller und finanzieller Ressourcen, organisatorisch-struktureller Anpassungen sowie institutionenübergreifender Kommunikation. Ziel muss es sein, Theorie und Praxis bei der Anwendung von Metadatenstandards i.S. der Standardanforderungen noch weiter zueinander zu bringen.

Persistente Referenzen auf archivische Informationen und Digitalisate

Ähnlich verhält es sich bei dem für die Bereitstellung von Archivinformationen essentiellen Thema „Permalinks“. Wesentlich ist – neben der Archivsignatur – die eindeutige Adressierung von Archivinformationen mittels eines permanenten Links, um Quellen

Digitalisate-Viewern (Mirador, Universal Viewer) vielseitige und innovative Nutzungsszenarien ermöglicht (<https://iiif.io/> / <https://www.digitale-sammlungen.de/de/p/20aec4c0-dc8a-4e40-84ac-49cfcac9a96c>

(aufgerufen am 25.08.2021). Diese Entwicklungen sollten im Archivwesen beobachtet und diskutiert werden.

⁴¹ S. z.B. <https://wiki.deutsche-digitale-bibliothek.de/pages/viewpage.action?pageId=19010180> (aufgerufen am 18.02.2021).

⁴² S. zur EAD AG: <https://wiki.deutsche-digitale-bibliothek.de/display/DFD/8.+EAD%28DDB%29-Arbeitsgruppe> (aufgerufen am 25.08.2021).

⁴³ Nils Meyer / Tobias Schröter-Karin, Stand und Perspektiven von EAD (DDB), in: Archivar, 74 (Mai 2021), H. 2, S. 97f.

⁴⁴ Vgl. den offenen Brief der Archivberatungsstellen der Landschaftsverbände Rheinland und Westfalen-Lippe an die Softwarehersteller: <https://archivamt.hypotheses.org/7018> (aufgerufen am 18.02.2021).

⁴⁵ S.z.B. <https://archivamt.hypotheses.org/tag/online-findbuecher> (aufgerufen am 18.02.2021).

⁴⁶ S. <https://archivamt.hypotheses.org/14566> (aufgerufen am 18.02.2021).

dauerhaft referenzieren und nachweisen zu können, insbesondere für den wissenschaftlichen Nachweis von historischen Quellen.⁴⁷

Die eindeutige Referenzierung und Identifizierung eines von einem Archiv bereitgestellten Digitalisats erfolgt standardgemäß über die METS-Datei, die einen Identifier in der Qualität einer GUID⁴⁸ enthalten muss, und auf die von der dazu gehörigen EAD-Datei referenziert wird. Der EAD (DDB)-Standard enthält auf mehreren Ebenen des jeweiligen Objekts einen eindeutigen Identifier in der Qualität einer GUID.⁴⁹ Viele, aber nicht alle Softwarehersteller für Archivinformationssysteme erfüllen diese Anforderungen mit ihren Systemen.

Der für die Referenzierung erforderliche Permalink besteht im Wesentlichen aus den in den Metadaten enthaltenen Identifikatoren (nach GUID/UUID) und der URL des Bereitstellers der Informationen. Die Herausforderungen für die Archive und ihrer Softwarehersteller liegt auch hier in der konsequenten Umsetzung der Anforderungen an Persistent Identifiers und somit eine eindeutige Referenzierung.

Daten aus „Tiefenerschließungen“

Traditionell stellen Archive Erschließungsdaten zur Beschreibung einzelner Archivalien und Bestände bereit. Inhalte zu Archivalien werden nach dem o.g. ISAD(G) nur im Titel und ggf. in einem nach bestimmten Regeln zu verfassenden Enthältvermerk beschrieben.

Insbesondere im Zuge des verstärkten Interesses an personenbezogenen Informationen, nicht zuletzt für die weit verbreitete genealogische Familienforschung, wird die Bereitstellung von recherchierbaren Inhaltsinformationen etwa in Form von Indizes, Transkriptionen oder Editionen nachgefragt.⁵⁰ Archive können diese vertiefte Erschließung i.d.R. nicht alleine leisten. Deshalb werden diese gewünschten Informationen häufig in Crowdsourcing-Projekten, mit Freiwilligen und häufig in Kooperation mit genealogischen Vereinen erarbeitet und bereitgestellt. Archivinformationssysteme stoßen hierbei jedoch an ihre Grenzen, da die Erschließungsstandards und das Datenschema EAD(DDB) nicht für diese Informationstypen vorgesehen sind. Sie sind strukturell nicht mitgedacht. Man kann aber davon ausgehen, dass bei der künftigen Entwicklung von EAD, genauer: in EAD 3, diese Bedarfe berücksichtigt werden.⁵¹

In Bezug auf die Präsentation von Digitalisaten im Zusammenhang mit Informationen aus einer wie auch immer gearteten „Tiefenerschließung“ besteht darüber hinaus die Herausforderung in der strukturellen Zuordnung dieser Erschließungsdaten zu den Bilddateien: Bislang bezieht sich ein „Container“ mit Digitalisaten einer Verzeichnungseinheit

⁴⁷ Frank M. Bischoff, Überlegungen zur Zukunftsfähigkeit archivischer Webangebote, in: Archivpflege in Westfalen-Lippe, 90 (2019), S. 36-47, hier: S. 45f.

⁴⁸ GUID steht für Globally Unique Identifier. Teilweise werden GUID und UUID (Universally Unique Identifier) synonym verwendet, s. z.B. <https://tools.ietf.org/html/rfc4122>; <http://guid.one/guid>; <https://www.webopedia.com/definitions/guid/>; und https://de.wikipedia.org/wiki/Globally_Unique_Identifier. In diesen Beiträgen werden die Anforderungen für eine GUID bzw. eine UUID dargelegt (aufgerufen am 24.02.2021).

⁴⁹ S. z.B. <https://wiki.deutsche-digitale-bibliothek.de/display/DFD/4.+EAD%28DDB%29+Tag+Library> (aufgerufen am 24.02.2021).

⁵⁰ Tiefenerschließung ist insbesondere auch im Kontext des Großprojekts „Transformation der Wiedergutmachung“ relevant. Auch hier wird überlegt, wie diese Daten inkl. der ermittelten Volltexte in die jeweilige Datenhaltung integriert und in Portale „transportiert“ werden kann. S. https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Oeffentliche_Finzen/Vermögensrecht_und_Entschädigungen/2019-11-08-Historiker-Symposium.html (aufgerufen am 25.08.2021).

⁵¹ Vgl. Meyer/Schröter-Karin 2021, S. 98 (s. Anm. 43).

auf den gesamten Datensatz dieser Verzeichnungseinheit. Stellt man künftig einen Index beispielsweise eines Personenstandsregisters zur Verfügung, sollten die Indexeinträge jeweils auf das Digitalisat der betreffenden Seite des Registers verlinkt sein. Gleiches gilt beispielsweise auch für digitalisierte Karteikarten (keine Findmittelkartei, sondern Archivalien wie z.B. Gefangenenkarteien), die jeweils keine eigene Verzeichnungseinheit bilden. Der strukturelle Bezug zwischen Digitalisat und Erschließungsdaten wird komplexer; es wird künftige Mehrfachbezüge geben, die sich auch an dem Informationsaufbau innerhalb des Archivale orientieren. Hier zeichnet sich – je nach Herangehensweise – ein Paradigmenwechsel hinsichtlich der Struktur von archivischen Informationen und damit der Anforderungen an Archivinformationssystemen ab. Jedenfalls wird dies Auswirkungen auf die Organisation von Digitalisaten und von Informationen über Digitalisate haben.

Normdaten und die Anbindung an das Semantic Web

Wenngleich das Thema „Normdaten“ im deutschen Archivwesen noch verhältnismäßig jung ist, so ist es inzwischen fast unstrittig, dass die Nutzung von Normdaten bei der Erschließung von Archivgut prinzipiell angestrebt und empfohlen wird.⁵² Normdaten verbessern den ‚Zugang‘ zu Archivgut bzw. erhöhen die Aufmerksamkeit auf Archivgut, da sie die Anschlussfähigkeit von archivischen Informationen an das Semantic Web herstellen, die Vernetzung archivischer Informationen gewährleisten und das Wissen über Archivgut verbreiten und anreichern helfen.

Auch hier wirkt die DFG mit den Praxisregeln „Digitalisierung“ beschleunigend und Normgebend. Verlangt wird, „wo immer möglich mit publizierten Normdaten zu verknüpfen. Für die Erfassung personenbezogener, biografischer und geografischer Informationen wird die Verwendung der Gemeinsamen Normdatei (GND) der Deutschen Nationalbibliothek erwartet.“⁵³ Die GND ist die bedeutendste Normdatei im deutschsprachigen Raum.

Bislang sind die meisten Archive noch nicht ausreichend auf diese Anforderung vorbereitet. Die wenigen Pilotprojekte⁵⁴ täuschen ebenfalls nicht darüber hinweg, dass die Mehrheit der Archive aus technischen oder/und organisatorischen Gründen noch weit weg von einem umfassenden und alltäglichen Einsatz von Normdaten ist: Vielfach bietet die jeweils genutzte Erschließungssoftware noch keine (ausreichende) Möglichkeit, Normdaten auch auf der Basis eines Datenabgleichs mit der GND zu verzeichnen und sie so als Vernetzungsinstrument von Daten wirksam werden zu lassen.⁵⁵ Darüber hinaus ist die Beschreibung oder Verknüpfung von Erschließungsinformationen mit Normdaten (wenn auch beschränkt auf Personen, Körperschaften und Geografika) mit einem hohen personellen Aufwand verbunden, weshalb verstärkt maschinelle Verfahren zur Normdatenanreicherung in den Blick genommen werden. Der Einsatz von Normdaten erfordert außerdem fachliche Konzepte und Regeln, etwa zu den Fragen, welche Normdaten (Personen, Geografika oder Körperschaften) auf welcher Erschließungsebene verwendet werden.

⁵² S. dazu aktuell: Krauth, Wolfgang / Sandner, Peter / Westphal, Sina, Archivische Erschließung mit Normdaten, in: Archivar, 73 (2020), H. 2, S. 142-144.

⁵³ DFG Praxisregeln „Digitalisierung“, S. 6 (https://www.dfg.de/formulare/12_151/12_151_de.pdf (16.04.2021)).

⁵⁴ Beispielsweise bietet das Archivportal-D die Möglichkeit, Erschließungsinformationen mit Normdaten nach der GND zu versehen (<https://www.archivportal-d.de/content/hilfe/#themenportal>, aufgerufen am 21.07.2021); s.a. das Portal Leo BW (<https://www.leo-bw.de/>, aufgerufen am 23.08.2021).

⁵⁵ Krauth / Sandner / Westphal 2020, 143f.

Gleichzeitig sind Archive erst verhältnismäßig spät in den Gremien der GND aktiv geworden, so dass archivische Anliegen, etwa Normdaten für Personen und Geografika mit maximal regionaler Bedeutung, bislang noch eine untergeordnete Rolle bei der Definition von Normdaten spielen. Um die Bedarfe an Normdaten für Kulturerbe-Einrichtungen, über die Bibliotheken hinaus, zu entwickeln, wurde das Projekt GND4C / GND für Kulturdaten ins Leben gerufen, und eine von der KLA initiierte „Interessengruppe Archive“ hat sich bei der GND gebildet.⁵⁶ Daneben ist die Mitarbeit bzw. die Gründung von GND-Agenturen, etwa auf regionaler Ebene, ratsam, um spezifische Bedarfe zu definieren und damit sich die beteiligten Institutionen gegenseitig bei der Umsetzung unterstützen.⁵⁷

Mit der Verwendung von Normdaten wird die Bereitstellung von Digitalisaten und archivischen Informationen qualitativ i.S. der Zugänglichkeit zu Archivgut verbessert und gleichzeitig die Fachaufgabe „Bereitstellung“ um zusätzliche, ressourcenintensive Aufgaben – auf der Ebene der fachlichen und technischen Konzeption und auf der Ebene der Umsetzung – erweitert. Auch deshalb ist zu empfehlen, dass hierfür Best Practices und Erfahrungen im Archivwesen, etwa mit Pilotprojekten, unter Berücksichtigung archivischer Standards kommuniziert, ausgetauscht und systematisiert werden. Dabei ist der fachliche Austausch mit Bibliotheken nicht zu vernachlässigen.

Archive operieren also mit einer zunehmenden, aber bisher überschaubaren Anzahl von Formaten und Standards von Daten und Metadaten. Vor allem macht ihnen eine zunehmende und potentiell überfordernde Menge von teilweise nicht standardkonform strukturierten Daten zu schaffen, die in verschiedenen Prozessen miteinander in Verbindung stehen bzw. aufeinander verweisen. Eine klare, streng standardisierte Kategorisierung der unterschiedlichen Daten und Informationsobjekte ist nicht zuletzt aus Gründen der Verständlichkeit der von Archiven bereit gestellten Informationen zentral. So ist etwa i.S. einer strukturierten und kontextorientierten Präsentation von archivischen Informationen empfehlenswert, die für die Quellenkunde und die Auswertung relevante Unterscheidung zwischen einem digitalen Objekt („born digital“) und einem digitalisierten Objekt (einer Reprographie) kenntlich zu machen. Hierbei kann die Anwendung des Repräsentationen-Modells ein probates Mittel sein.⁵⁸ Verschiedenartige Daten bzw. Informationsobjekte müssen klar unterscheidbar sein. Dies ist eine Frage der Datenpräsentation und des Designs, die wiederum auf die bereitgestellten archivischen Informationen zurückspiegelt. Bei der Präsentation von Hybridakten, die aus einem genuin elektronischen und einem analogen sowie ggf. digitalisierten Teil bestehen und unter einer Signatur bzw. in einem Datensatz einer Verzeichnungseinheit firmieren, müssen darüber hinaus die weiteren Archivalienteile und deren Repräsentationen eindeutig erkennbar und bezeichnet sein. Grundlage für eine verständliche und strukturierte Repräsentation von archivischen

⁵⁶ <https://www.dnb.de/DE/Professionell/ProjekteKooperationen/Projekte/GND4C/gnd4c.html> (aufgerufen am 21.07.2021); https://gnd.network/Webs/gnd/DE/UeberGND/Organisation/organisation_node.html (aufgerufen am 17.08.2021); <https://wiki.dnb.de/display/GNDCON/MiniCon++Archiv++GND+4+Archives> (aufgerufen am 17.08.2021).

⁵⁷

https://gnd.network/Webs/gnd/SharedDocs/Downloads/DE/materialien_Info3.pdf?blob=publicationFile&v=3 (aufgerufen am 21.07.2021).

⁵⁸ Peter Sandner, Das Repräsentationenmodell als Konzept für die archivische Arbeit. KLA verabschiedet Thesenpapier, in: Archivar, 69. Jg. (2016), H. 1, S. 18-21 (https://www.archive.nrw.de/sites/default/files/media/files/Archivar_1_2016.pdf#page=15, aufgerufen am 24.08.2021).

Informationen (inklusive Digitalisaten) ist die Einhaltung von Daten- und Metadatenstandards.

Veränderte Nutzungserwartungen und technische Möglichkeiten führen zu einer Dynamisierung der im Gesamtsystem zu organisierenden Daten und Prozesse. Wollen Archive künftig (weiterhin) qualitativ hochwertige Daten und Metadaten professionell und effizient sichern und kund*innengerecht für die Nutzung und Veröffentlichung bereitstellen, bedarf die der Organisation von Daten / Metadaten in Archiven einer erhöhten Aufmerksamkeit. Darum wird es im folgenden Abschnitt gehen.

4) Organisation: Datenstrukturen / Datenbanken / Dateninfrastruktur

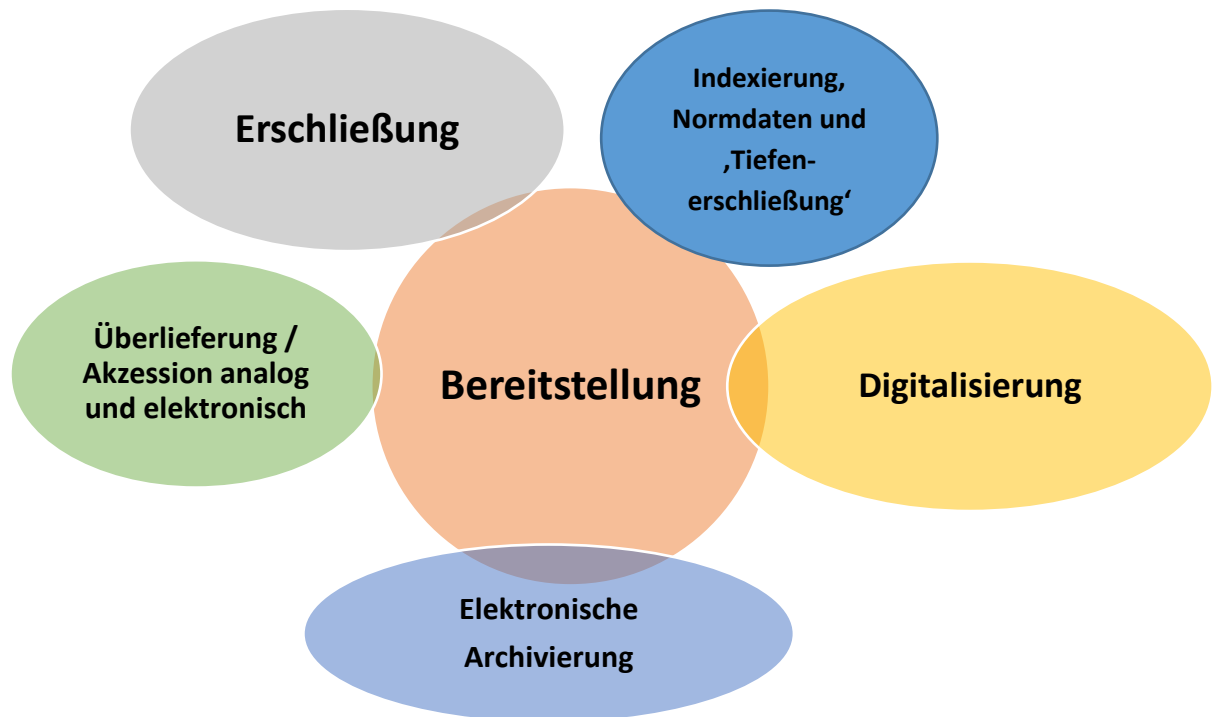
Archive sind seit jeher Orte der Informationsfülle, der Informationssicherheit sowie der Informationsbereitstellung. Inzwischen sind sie auch Institutionen des Datenmanagements, der Datensicherung bzw. des Datenschutzes und des Datenzugangs. Bisweilen werden Archive sogar als Einrichtungen betrachtet, die mit „Big Data“ operieren, wie etwa von Emily Larson: „Big Data is becoming a key part of transaction and decision making processes, and archivists are increasingly called to intervene in its management.“⁵⁹ Man kann diskutieren, ob Big Data ein zu großes Schlagwort für die Arbeit in Archiven ist. Aber unbestritten ist zum einen die zunehmende Menge an Daten, Datenbanken und Metadaten in Archiven sowie zum anderen die zunehmende Anzahl von Prozessen, die anders als zu analogen Zeiten, ineinandergreifen (müssen):

Die Arbeitsschritte der Archivierung von der Übernahme bis zur Bereitstellung von Archivalien waren vor dem Einsatz der IT zur Prozessunterstützung zwar aufeinander aufbauende, aber voneinander getrennte und nur bedingt voneinander abhängige Schritte. Wurde beispielsweise bei der Akzessionierung die Behördenbezeichnung nicht korrekt vermerkt, konnte dies bei der Verzeichnung leicht korrigiert werden. Ein Archivale kannte i.d.R. jeweils nur eine Repräsentationsform. Im „digitalen Zeitalter“ kann beispielsweise eine mangelhafte elektronische Schriftgutverwaltung zu großen Überlieferungsverlusten führen, etwa wenn keine Standardformate verwendet wurden, keine strukturierte, aktenplanbasierte Ablage erfolgte oder wenn die Schnittstellen nicht den archivischen Anforderungen entsprechen. Wurden in der Verzeichnung im elektronischen archivischen Fachinformationssystem zum Beispiel für eine skriptgesteuerte Verarbeitung der Erschließungsdaten und deren Verknüpfung mit Digitalisaten unbrauchbare Signaturen oder Titelangaben verwendet, kann sich die Planungsgrundlage für ein Digitalisierungsprojekt als Fehl kalkulation erweisen, da Management- und Organisationsdefizite durch aufwändige Nacharbeiten ausgeglichen werden müssen.

Schließlich werden die Datensätze einzelner Verzeichnungseinheiten in den AFIS inzwischen nicht mehr nur mit Digitalisaten „angereichert“ bzw. darauf via Link verwiesen, sondern auch mit Normdaten oder etwa Daten aus Tiefenerschließungs- oder Indexierungsprojekten wie oben dargestellt. Zumindest als Zwischenlösung können diese Anreicherungsdaten (und darüber hinausgehende Forschungsdaten) aber auch in einem externen Repositorium gesichert und im AFIS verlinkt werden. Diese Anreicherung ist das Ergebnis mehrerer Datengenerierungs- und -verarbeitungsprozesse, etwa der

⁵⁹ Emily Larson, Big Questions: Digital Preservation of Big Data in Government, in: The American Archivist, 83 (Spring/Summer 2020), Nr. 1, S. 5-20, hier S. 5.

Digitalisierungsprozesse, der Normdatenindizierung oder der Indexierung, um nur Stichworte zu nennen. Diese Prozesse erfolgen nicht geordnet hintereinander, sondern prozessabhängig gleichzeitig bzw. zeitlich unabhängig voneinander. An diesen Prozessen arbeiten i.d.R. unterschiedliche Organisationseinheiten und Personen. Sie müssen daher (aufwändig) gesteuert werden.



Deshalb sind Archive zunehmend auf ein professionelles (Daten-)Prozessmanagement und eine tragfähige und zumindest nach innen transparente IT-Systemarchitektur angewiesen und damit auf die Anwendung von Erkenntnissen der Informatik und Wirtschaftsinformatik. Die wesentlichen Ziele sind in den Archivgesetzen begründet: kund*innenorientierte Bereitstellung von archivischen Informationen und Archivalien sowie Sicherung der Daten und Informationen für den Langzeiterhalt, auch i.S. des Datenschutzes und der Informationssicherheit. Beide Ziele sind wesentlich für die Erfüllung der Funktionen von Archiven im demokratischen Rechtsstaat: Zugang der Bürger*innen zu Unterlagen der öffentlichen Verwaltung sowie gesicherte und nachvollziehbare Dokumentation staatlichen Handelns.

Zudem verlangen und ermöglichen die E-Government-Gesetze und das Onlinezugangsgesetz die Digitalisierung dieser Leistungen und Services sowie damit der dahinter stehenden Prozesse und Systemarchitekturen. In diesem Kapitel liegt dabei der Fokus auf der Erstellung, Verarbeitung und Bereitstellung von Digitalisaten in archivischen Fachinformationssystemen. Zu betrachten sind dabei 1) die Systemarchitektur zur Strukturierung der Daten und Module und 2) die Prozesse.

4.1 Systemarchitektur und Datenorganisation

Die Voraussetzung für effiziente und professionelle Prozesse ist eine Gesamtsystemarchitektur, die die einzelnen Komponenten, Module und Speicherorte aufzeigt und sie zueinander in Beziehung setzt sowie eine adäquate, nachhaltige und gesicherte Serverstruktur vorsieht. Dabei wird der Begriff „IT-Systemarchitekturen“ einerseits für die Struktur der Komponenten und Datenflüsse innerhalb einer Software verwendet und andererseits für die Gesamtsystematik der IT in einer Organisation, Firma oder Behörde. Folglich müssen sowohl bei der (Weiter-)Entwicklung einer Software als auch bei der Einpassung dieser Software die Systemarchitektur der einzelnen Software und die übergeordnete Architektur der Systeme innerhalb eines Archivs aufeinander abgestimmt werden.

Dabei können Verwaltungssysteme (z.B. für die Haushaltsführung, die E-Akte) von den Fachsystematiken getrennt werden. Allerdings sind – je nach Konzept – Schnittstellen z.B. zwischen Benutzerdaten und dem Rechnungswesen etwa für die Abrechnung von Gebühren oder E-Payment-Tools ebenso erforderlich wie zwischen einem AFIS, das Nutzerdaten verwaltet, und dem E-Aktensystem. Es bestehen folglich Abhängigkeiten voneinander. Außerdem ist zu entscheiden, ob Digitalisate getrennt von genuin elektronischen Unterlagen, also dem digitalen Langzeitarchiv gespeichert werden, oder gemeinsam mit diesem. In beiden Fällen können sie nach dem OAIIS-Konzept⁶⁰ verarbeitet, gesichert und ausgegeben werden. Das Schweizer Bundesarchiv sieht hierfür keine Trennung vor und sichert Digitalisate und genuin elektronische Unterlagen im „Digitalen Archiv“.⁶¹ Auch in Deutschland, etwa im DiPS-Nutzerkreis, wurde dieses Szenario bereits diskutiert. Abzuwägen sind hier einerseits u.a. Kostenfragen (z. B. Lizenzgebühren), unterschiedliche Sicherheitsstandards und andererseits der Anspruch, die Datensicherungssysteme so wenig komplex wie möglich zu halten.

Darüber hinaus sind aus Gründen der Informationssicherheit und des Datenschutzes zu sichernde Bereiche (etwa Daten zu schutzbewehrtem Archivgut) strikt von Systemmodulen zu trennen, auf die ein befugter Zugriff von außen möglich ist (etwa über Angebote im Internet). Auch innerhalb des Archivs sind Systemelemente zu definieren, die nur einem ausgewählten Kreis zugänglich sind, z.B. elektronische VS-Archive. Folglich ist im Zusammenhang mit dem Prozessmanagement auch ein solides Rechte- und Rollenkonzept zur IT-Systemarchitektur erforderlich.

Dies stellt Archive bei der Einrichtung eines „Virtuellen Lesesaals“ vor enorme Herausforderungen, da – so die Definition des „Virtuellen Lesesaals“⁶² – authentifizierte

⁶⁰ S. zum Referenzmodell Open Archival Information System z.B. die deutsche Übersetzung der nestor AG von 2013 (<https://d-nb.info/104761314X/34>, aufgerufen am 29.07.2021).

⁶¹ Finck / Kwasnitza (2020), S. 166 (Anm. 36) (<https://b-i-t-online.de/heft/2020-02-fachbeitrag-finck.pdf>, aufgerufen am 29.07.2021).

⁶² Unter „Virtuellem Lesesaal“ (VL) wird ein Online-Angebot von Archiven verstanden, das authentifizierte Nutzer*innen den selben Service bietet wie bei einem Besuch in einem Lesesaal eines Archivs, also auch die Einsicht in gesperrtes Archivgut nach einer individuellen Schutzfristenverkürzung. Nicht gemeint ist der Online-Service von Archiven für anonyme User und User mit Kennung / Registrierung. Vgl. Peter Sandner, „Virtueller Lesesaal“. Originär archivspezifische Anforderungen an einen virtuellen Nutzungsbereich, in: Neue Wege ins Archiv – Nutzer, Nutzung, Nutzen. 84. Deutscher Archivtag in Magdeburg, hg. vom VdA – Verband deutscher Archivarinnen und Archivare e.V. (Tagungsdokumentation zum Deutschen Archivtag, Bd. 19), Fulda 2016, S. 37-45; Martina Wiech, Veränderungen im Zugang zu Archivgut im Spiegel der rechtlichen Rahmenbedingungen, in: Neue Wege ins Archiv – Nutzer, Nutzung, Nutzen. 84. Deutscher Archivtag in Magdeburg, hg. vom VdA – Verband deutscher Archivarinnen und Archivare e.V. (Tagungsdokumentation zum Deutschen Archivtag, Bd. 19), Fulda 2016, S. 201-210.

Nutzer*innen online dieselben Zugangsrechte auf Archivgut erhalten können wie im analogen Lesesaal, also auch auf schutzbewehrte Archivalien nach individueller Genehmigung einer Schutzfristenverkürzung. Denn die Zugänge müssen so zielgenau, individualisiert und gesichert sein, dass kein Zugriff von Dritten auf gesperrte Informationen möglich ist. Dies erfordert ein engmaschiges Informationssicherheitsmanagement⁶³ und regelmäßige Sicherheits-Updates. Die dafür einzusetzenden Ressourcen sind basal für die Umsetzung dieses umfassenden Online-Services und die Aufrechterhaltung der Vertrauenswürdigkeit der jeweiligen Archivinstitution.

Die Vertrauenswürdigkeit eines Archivs basiert ferner auf der langfristigen Sicherung von Digitalisaten, Erschließungsdaten (Metadaten) und genuin elektronischen Archivalien, ohne auf Daten der Eigenverwaltung einzugehen. Folglich muss eine Systemarchitektur in Archiven zwischen Speicher- und Nutzdaten sowie Nutz- und Back-up-Systemen unterscheiden. Die dafür erforderliche Serverstruktur und -sicherung ist dabei elementar. Die Auslagerung der IT-Infrastruktur eines Archivs oder Teile davon an einen (insbesondere außereuropäischen) Cloud-Dienstleister ist unter Maßgaben der Archivgesetze⁶⁴ und der Datenschutzgesetze sowie dabei der EU-DSGVO (hier insbesondere Kapitel 5, Art. 44, sowie Erwägungsgrund 100) äußerst kritisch zu prüfen, auch wenn ein derartiges Outsourcing aus Kostengründen zunächst attraktiv erscheinen mag⁶⁵. Welche zuverlässigen und sicheren Cloudlösungen künftig für die Nutzung durch Archive in Frage kommen, wird zu prüfen und zu diskutieren sein.

4.2 Prozessmanagement

Untrennbar von der Systemarchitektur sind die (Arbeits-)Prozesse und Datenflüsse, die für die Bereitstellung erforderlich sind. Diese Prozesse müssen effektiv, transparent, nachvollziehbar und klar definiert sein sowie einerseits den Anforderungen der Informationssicherheit und des Datenschutzes (inkl. Archivrecht) und andererseits fachlichen Standards entsprechen. Insofern ist das IT-Prozessmanagement eine technische und fachliche sowie organisatorische Aufgabe.

Wie oben geschildert, haben sich in den Archiven die (IT-)Prozesse alleine für den Aufgabenkomplex „Bereitstellung“ u.a. aufgrund der Herstellung und Sicherung von Digitalisaten sowie deren Verknüpfung mit archivischen Informationen und Online-Stellung vervielfacht. Diese Prozesse müssen effizient gesteuert, mit anderen Prozessen harmonisiert und ständig technisch, fachlich und organisatorisch optimiert werden. Die Umsetzung von Anforderungen der E-Government-Gesetze und des Onlinezugangsgesetzes, d.h. die Digitalisierung von Dienstleistungen der Verwaltung, wie etwa die Bereitstellung von

⁶³ Vgl. den Beitrag von Nils Volkening zum Workshop des KLA-Ausschusses AFIS am 27.01.2021: „Impulsvortrag Speicherlösungen, IT-Infrastruktur und Workflows. IT-Sicherheit und Speicherbedarf (FHH)“. Volkening sieht vor allem einen Nachholbedarf in der Frage der Integrität und Vertraulichkeit im digitalen Bereich. Archive nehmen seiner Ansicht nach die Gefahren im digitalen Bereich häufig nicht ausreichend wahr, obwohl hier deutlich mehr Angriffsvektoren bestehen.

⁶⁴ S. z.B. § 7 VII Satz 5 ArchivG NRW.

⁶⁵ Vgl. das Schweizer Beispiel: Isabel Pfaff, Schweiz will Daten bei chinesischem Konzern lagern, in: Süddeutsche Zeitung am 02.07.2021 (<https://www.sueddeutsche.de/politik/schweiz-alibaba-cloud-1.5339784>; aufgerufen am 09.07.2021). Vgl. Fatemi Moghaddam, Faraz, Philipp Wieder, Süleyman Berk Çemberci, and Ramin Yahyapour: Cloud Security Distributary Set (CSDS): A Policy-Based Framework to Define Multi-Level Security Structure in Clouds, in: Firouzi, Farshad, Krishnendu Chakrabarty, Bahar Farahani, Fangming Ye, and Vasilis F. Pavlidis: Proceedings of the International Conference on Omni-Layer Intelligent Systems (COINS 2019), Crete, Greece, May 5-7, 2019, pp. 74-79, (DOI: 0.1145/3312614.3312633).

archivischen Informationen, erfolgt nicht zuletzt durch eine Geschäftsprozessoptimierung.⁶⁶ Geschäftsprozessoptimierungen (GPO) spielen eine grundlegende Rolle in der Digitalen Transformation.⁶⁷ Die Optimierung von Verwaltungs- oder – wie hier – archivischen Dienstleistungen, etwa bei der Bereitstellung von Digitalisaten, erfolgt letztlich nur durch ein professionelles Prozessmanagement⁶⁸. Ziel der Geschäftsprozessoptimierung in dem hier betrachteten Themenfeld muss die möglichst weitgehende Automatisierung von Prozessen sein, sofern dadurch die Servicequalität für die Nutzer*innen erhöht wird oder das Archiv die Aufgaben quantitativ besser bewältigen kann, wenn massenhaft und auf fachlichen Standards basierend Digitalisate mit archivischen Kontextinformationen online bereitgestellt werden.

Dabei erfolgt die GPO nicht als 1:1-Übertragung analoger Abläufe in digitale Prozesse oder automatisierte Verfahren. Archive wissen dies etwa aus der Einführung von elektronischen Fachsystemen in der Verwaltung. Gerade da in Transformationsprozessen Vieles nicht automatisierbar ist, ist ein hohes Maß an Agilität gefragt, etwa durch „agile Verfahren“ im Geschäftsprozessmanagement.⁶⁹ Derartige grundlegende Change-Prozesse sind eine Führungsaufgabe.⁷⁰ Folglich sind die Online-Bereitstellung von Digitalisaten und die dafür erforderlichen Prozesse im Vorfeld (oder im Hintergrund – je nach Perspektive) unter den Gesichtspunkten des professionellen Prozessmanagements zu betrachten. Dabei muss – so das Plädoyer – auf eine Steigerung der Effizienz abgezielt werden – nicht zuletzt wegen der begrenzten Ressourcen und des Wirtschaftlichkeitsgebots. Letztlich ist die Optimierung und zunehmenden Automatisierung der Prozesse zur Bereitstellung von Digitalisaten ein Baustein der Digitalen Transformation in Archiven.⁷¹

Vor einer Prozessanalyse und -optimierung, d.h. auch Prozessdigitalisierung, sind die strategischen und operativen Ziele bzw. Kriterien zu definieren. Konkret bezogen auf die Verarbeitung und Bereitstellung von Digitalisaten sind insbesondere folgende Zielformulierungen zu empfehlen:

⁶⁶ Vgl. z.B. Gaby Beckmann / Marc Groß, Erfolgsfaktoren für das Online-Zugangsgesetz September 2019 (<https://kommunal.de/erfolgsfaktoren-online-zugangsgesetz>; s.a. Seite des Bundesministeriums des Innern: <https://www.onlinezugangsgesetz.de/Webs/OZG/DE/startseite/startseite-node.html>; beide aufgerufen am 27.07.2021).

⁶⁷ Vgl. dazu die Beiträge bei der ICA-Konferenz „Designing the Archive“ (Oktober 2019) (<https://www.ica.org/en/4-design-and-transformation>; aufgerufen am 29.07.2021); s.a. Frank M. Bischoff / Bettina Joergens, Designing the Archive. Bericht von der Jahreskonferenz des International Council on Archives in Adelaide (Süd-Australien), in *Archivar*, 73 (Mai 2020), H, 2, S. 162-16; vgl. für Unternehmen: Thomas Allweyer, Untersuchung aktueller Ansätze des Geschäftsprozessmanagements als Enabler für die digitale Unternehmenstransformation, Kaiserslautern 2020, S. 11-25, insb. S. 25.

⁶⁸ S. zur Definition z.B. <https://www.bwl-lexikon.de/wiki/prozessmanagement/> (aufgerufen am 27.07.2021).

⁶⁹ Vgl. z.B. Thomas Allweyer, Untersuchung aktueller Ansätze des Geschäftsprozessmanagements als Enabler für die digitale Unternehmenstransformation, Kaiserslautern 2020, S. 13; vgl. Martin Bartonitz, Veronika Lévesque, Thomas Michl, Wolf Steinbrecher, Cornelia Vonhof, Ludger Wagner (Hg.), *Agile Verwaltung. Wie der Öffentliche Dienst aus der Gegenwart die Zukunft entwickeln kann*, Wien 2018, und <https://agile-verwaltung.org/> (aufgerufen am 27.07.2021); Susanne Nickel / Gunhard Keil, *So geht Agilität. Die besten agilen Methoden im Job*, Freiburg 2020. S. insb. für das Land NRW: <https://www.wirtschaft.nrw/mwidedigital2022> (aufgerufen am 27.07.2021) und im Bund

(https://www.onlinezugangsgesetz.de/Webs/OZG/DE/umsetzung/servicestandard/prinzip-8/prinzip-8-node.html;jsessionid=6DF3F3F299B649DA6D2B83A393A1A915.1_cid295, aufgerufen am 27.07.2021).

⁷⁰ S. z.B. <https://kommunal.de/onlinezugangsgesetz-kommunen-sind-verunsichert> (aufgerufen am 27.07.2021).

⁷¹ Die betriebswirtschaftliche Perspektive reicht zur Erfassung des Gesamtkomplexes nicht aus (s. z.B. <https://www.enzyklopaedie-der-wirtschaftsinformatik.de/lexikon/technologien-methoden/Informatik--Grundlagen/digitalisierung/digitale-transformation> (aufgerufen am 29.07.2021)). Vgl. Jan-Felix Schrape, *Die Digitale Transformation*, Stuttgart 2021.

- Vermeidung von Systembrüchen und hybriden Prozessen, etwa E-Mail-Benachrichtigungen jenseits des Systems: der „medienbruchfreie Informationskreislauf“;⁷²
- Vermeidung des „Handanlegens“ bei Datenverarbeitungsprozessen durch weitgehende Automatisierung der Prozesse; die Vermeidung von Ausnahme- und Sonderfällen ist dafür Voraussetzung;
- Einhaltung von Daten- und Metadatenstandards; Vermeidung von Datenmanipulationen, z.B. um Metadaten „passend zu machen“;
- Vermeidung personenabhängiger Prozesse, von Wissensmonopolen und Prozessunterbrechungen bei Abwesenheiten der Expert*innen;
- kritische Prüfung vor eigener Softwareentwicklung sowie den Folgeaufwänden dafür; möglichst Beteiligung an vorhandenen Softwareprojekten in stabilen Entwicklungsgemeinschaften⁷³.

Schließlich steht über diesen Zielen das Grundanliegen, interessierten Bürger*innen orts- und zeitunabhängig, komfortabel und umfassend Zugang zu Archivgut bzw. archivischen Informationen anzubieten. Die Bereitstellung von unterschiedlichen Informationskategorien zu einer Verzeichnungseinheit, insbesondere von Digitalisaten, Erschließungsdaten, Daten aus Indexierungsprojekten oder von Transkriptionen (via OCR-Tools) ist daher genauso in den Gesamtprozess systematisch zu integrieren wie Digitalisate-on-demand-Vorgänge – aus Gründen der Kundenorientierung, der Wirtschaftlichkeit und der fachlichen Effizienz. Die Umsetzung dieser Ziele ist nur möglich, wenn die notwendigen zusätzlichen Ressourcen an Sachmitteln und an Personal – mit entsprechenden IT-fachlichen Kenntnissen – zur Verfügung stehen.

Fazit und Ausblick

Die Bereitstellung von Digitalisaten ist sowohl aus Kund*innensicht als auch mit Blick auf die archivinternen Prozesse eines der Herzstücke archivischer Aufgaben und Herausforderungen, wenn man dabei alle dafür erforderlichen fachlichen, technischen und organisatorischen Vorgänge berücksichtigt und wenn man sie als einen Baustein in der Digitalen Transformation von Archiven, inklusive der Managementherausforderungen, begreift.⁷⁴ Dabei ist zu betonen, welch immenses Potential der Einsatz von Digitalisaten bietet – von der ortsunabhängigen Erschließung, auch mit Crowdsourcing, über die automatische Handschriftenerkennung bis hin zu Möglichkeiten der digitalen Öffentlichkeitsarbeit. Zugleich aber erfordern die empfohlenen Optimierungen enorme Anstrengungen, da technische, fachliche und organisatorische Veränderungsprozesse integriert werden müssen; sie tragen allerdings – so die These – wesentlich zur Zukunftsfähigkeit von Archiven bei.

Aus Sicht des Ausschusses der Konferenz der Leiterinnen und Leiter der Archivverwaltungen im Bund und in den Ländern für Archivische Fachinformationssysteme

⁷² Finck/Kwasnitza (2020), S. 165 f. (Anm. 36) <https://b-i-t-online.de/heft/2020-02-fachbeitrag-finck.pdf> (aufgerufen am 29.07.2021).

⁷³ Vgl. die (Weiter-)Entwicklung von Kitodo (Key to Digital Objects): <https://www.kitodo.org/> (aufgerufen am 29.07.2021).

⁷⁴ Vgl. Die Digitalstrategie des Landesarchivs Hessen, 2019 (https://landesarchiv.hessen.de/sites/landesarchiv.hessen.de/files/digitalstrategie_hla.pdf, aufgerufen am 29.07.2021).

können solche Veränderungen nur bedingt von einzelnen Einrichtungen alleine bzw. isoliert zum Erfolg gebracht werden. Es wird auch nicht mehr genügen, im fachlichen Austausch in erster Linie über Erfolge zu referieren, sondern es wird zunehmend erforderlich sein, gemeinsam „unter die Motorhaube“ oder „in die Werkstatt“ zu schauen sowie von der Offenlegung von Fehlern⁷⁵ und Problemen gemeinsam zu profitieren. Denn die Verwendung von Ressourcen für Parallelentwicklungen ist meist unökonomisch und für die fachliche Entwicklung und das Standing als wichtige Sparte innerhalb der Informationsgesellschaft ungünstig. Beispielgebend sind etwa länderübergreifende Kooperationen bei der Archivierung digitaler Archivalien. Aber auch die Notfallverbände zur gemeinsamen Notfallvorsorge bei Havarien in Archiven können hier durchaus vorbildhaft sein: Denn man schafft nicht nur gemeinsam eine umfassende Notfallausrüstung an, z.T. mit Drittmitteln, sondern bringt sich gegenseitig auf den gleichen Sachstand, entwickelt gemeinsam Standards, führt zusammen Übungen durch und packt schließlich gemeinsam bei beschädigten Archiven an. Auch bei der Weiterentwicklung von archivischen Fachinformationssystemen und damit der Infrastrukturen für die Bereitstellung von archivischen Informationen und Digitalisaten tragen heute schon übergreifende AFIS-Kooperationen zwischen Archivinstitutionen zur Zukunftsfähigkeit von Archiven bei – eine Entwicklung, die durchaus verstärkt werden könnte.

Wanted: Digital Copies of Analog Archival Material Online. Professional Provision of Digital Copies in Archival Information Systems – a Discussion Paper of the Conference of the Directors of the Federal and State Archives in Germany

This paper discusses criteria for a professionalization of provision of digital copies of analog archival material. It is not only about the online presentation, but focuses on data quality and standards, persistent identifiers, extended archival description data, normdata and linked data and – last but not least – about optimizing the process management.

⁷⁵ Vgl. als Beispiel für einen offengelegten Fehler: <https://fotoarchiv.hypotheses.org/215> (aufgerufen am 20.08.2021).