Georg Vogeler

Das Digitale Archiv

Der Computer als Mediator, Leser und Begriffsbildner

Das Digitale Archiv gibt es nicht. Dieser Satz benutzt natürlich eine Unschärfe im Bedeutungshorizont des Wortes „geben“: Es kann eine physische Existenz ebenso wie eine gedachte meinen. Bevor ich im folgenden Essay den „Schauplatz Digitales Archiv“ verhandle, möchte ich einen Versuch unternehmen, das Digitale Archiv zu definieren. Dabei möchte ich mit der gedachten Bedeutung des Einleitungssatzes beginnen, also mit dem Diskurs über den Begriff „Digitales Archiv“. Die Archivfachleute verweigern nämlich dem Lemma „Digitales Archiv“, in terminologische Verzeichnisse aufgenommen zu werden: Weder die von 2012 bis 2015 von den ReferendarInnen der deutschen Archivschule in Marburg erstellte archivwissenschaftliche Terminologie (Archivschule Marburg 2012–2015) noch das 2005 von Richard Pearce-Moses erstellte Glossar der Society of American Archivists (Pearce-Moses 2005) oder die vom „Interpares Trust“ erstellte und 2012 vom International Council of Archives übernommene mehrsprachige Terminologie (International Council on Archives 2012) haben einen Eintrag „digital archive“ bzw. „digitales Archiv“. Im erst jüngst erschienenen *Handbuch Archiv* (Lepper und Raulf 2016) widmet Heinz Werner Kramski digitalen Dokumenten einen eigenen Abschnitt, in dem auch von digitalen Archiven die Rede ist. Aber auch er sieht keinen Anlass, das „Digitale Archiv“ als festen Terminus einzuführen (vgl. Kramski 2016, 178–197). Wir können uns also nicht auf die Expertise der Archivfachleute verlassen, wenn wir nach einer Definition des „Digitalen Archivs“ suchen.

Die Praxis hat es da einfacher: Wenn Kramski über die Archivierung digitaler Dokumente schreibt, dann ist ihr Aufbewahrungsort selbstverständlich ein digitales Archiv (vgl. Kramski 2016, 191). Das Digitale Archiv ist damit also die Einrichtung zur Bewahrung von digitalen Dokumenten („Digitale Langzeitarchivierung“). Das Digitale Archiv kann mit den vielen Bedeutungen des Wortes „Archiv“ aber auch interpretiert werden als die digitale Repräsentation der Inhalte einer Institution zur Bewahrung von Dokumenten von historisch relevanter Tätigkeit, z. B. von Institutionen, oder schließlich als eine online verfügbare Ressource zur dokumentarischen Information über ein historisch relevantes Thema. Für die digitale Repräsentation eines Archivs könnte man also von einem „digitalisierten Archiv“ sprechen. Die digitale Ressource dokumentarischer Informationen würde auch als „Digitales Dossier“ bezeichnet werden können. Man sieht, dass in allen drei Definitionsvorschlägen die kulturwissenschaftliche Deutung von „Archiv“

als kulturelle Praxis der Bewahrung bzw. als Oberbegriff für Wissensspeicher durchscheint (vgl. Stingelin 2016, 21–27). Man könnte sich deshalb vielleicht auf folgende Definition zu einigen versuchen: Ein Digitales Archiv ist die Praxis, eine Auswahl von Dokumenten zu späterem Gebrauch mit Hilfe von digitalen Technologien langfristig zu speichern. Diese bewusst allgemein gehaltene Definition steht im Hintergrund der folgenden Überlegungen über die Auswirkungen dieser Praxis auf die Wahrnehmung von Archiv und Archivgut im späteren Gebrauch. Die digitalen Technologien stehen dabei im Zentrum der Aufmerksamkeit, denn es geht darum, wie diese Technologien die Wahrnehmung der Dokumente verändern.

Der einleitende Satz dieses Beitrags lässt neben seiner ideellen Interpretation aber auch eine physische zu. Hier ist es weniger leicht, die „Nicht-Existenz“ des Digitalen Archivs zu verdeutlichen. Natürlich braucht auch ein digitales Archiv eine physische Form: Magnetbänder, Festplatten, optische Datenträger, mit Strom versorgte Computer etc. Diese Form weicht deutlich von der eines „klassischen“, „analogen“, „nicht-digitalen“, „Papier“-Archivs ab, denn Menschen können die gespeicherte Information nicht unmittelbar von ihrem Träger ablesen, worauf Harry Pross mit seiner Unterscheidung zwischen sekundären und tertiären Medien hingewiesen hat (vgl. Pross 1970, 129). Das Papierarchiv kann durch eine wohlwollende Vernachlässigung der originalen Informationsträger erhalten werden, und die späteren BenutzerInnen können unmittelbar mit den originalen Informationsträgern interagieren. Im Digitalen Archiv sind die Informationsträger dagegen selbst nicht mehr „original“ in dem Sinne, dass physisches Objekt und Informationsträger zusammenfallen. Digitale Forensiker wie Matthew Kirschenbaum weisen darauf hin, dass auch die physische Grundlage digitaler Medien Einfluss auf die Information nimmt. Gleichzeitig weist Kirschenbaum auch darauf hin, dass archivierte Objekte auch formal identisch sein können, wenn sie sich bei der Übertragung nicht verändern (vgl. Kirschenbaum 2012, 133). Die Erfahrung im Umgang mit digitalen Dokumenten ist nämlich, dass sich diese Information als eine Menge von binären Zuständen darstellt, die auf beliebigen Datenträgern gespeichert werden kann. Die Infrastruktur des Papierarchivs beschränkte sich also darauf, das „Wohlwollen“ in der Vernachlässigung zu beweisen: günstige klimatische Bedingungen, Verpackungen der Archivalien, die dem Beschreibstoff nicht schaden, Vorkehrungen, um Elementarereignisse wie Feuer- oder Wasserschaden weniger wahrscheinlich zu machen, Zugriffskontrollen zur Verringerung der Gefahr von Diebstahl. In einem digitalen Archiv erweitert sich der Aufgabenbereich der Infrastruktur: Zum Erhalt der Information ist ein regelmäßiger Test notwendig, ob der Kopiervorgang noch identische Dokumente erzeugen kann. Das Digitale Archiv besteht damit nicht mehr nur aus physischen Datenträgern, sondern aus Datenstrukturen, d. h. einer Anordnung der als Zeichen interpretierbaren Bits. Es entstehen damit logische Objekte (vgl. Kirsch-

enbaum 2012, 3), die von den BenutzerInnen als „Dokumente“ wahrgenommen werden.

1 Mediator

Während Wolfgang Ernst auf die Ähnlichkeit von digitalen Medien und Archiven hinweist, da digitale Objekte kontinuierlich gespeichert werden (vgl. Ernst 2012, 451), betont die Praxis digitaler Archive eher die Unterscheidung zwischen alltäglicher Speicherung, die vielleicht in der traditionellen Archivsprache als „Registratur“ bezeichnet werden könnte, und der auf lange Zeit angelegten Archivierung. Für diese „Langzeitarchivierung“ – die Anführungszeichen stehen als Hinweis auf den aus archivischer Sicht enthaltenen Pleonasmus – ist theoretisch am besten vom Open Archival Information Systems Referenzmodell (OAIS) beschrieben, das sich inzwischen als Standard der digitalen Archivierung etabliert hat – und auch ein vom ISO empfohlener Standard ist (ISO 14721:2012).¹ Das Modell beschreibt die logischen Komponenten, die in einem digitalen Archivsystem notwendig sind, um dauerhafte Aufbewahrung und Zugänglichkeit sicherzustellen. Im Kern steht das Archivpaket, das neben den eigentlichen Daten auch Informationen zur Bearbeitungsgeschichte im Archiv und Angaben enthält, die Integrität und Authentizität gewährleisten können. Um die Kontrolle des Archivpakets sicherzustellen, ist es konzeptionell von den zu archivierenden Daten ebenso klar getrennt wie von der Form, in der die BenutzerInnen es zur Verfügung gestellt bekommen. Die „Submission“ in das Archiv ist ein in sich geschlossener Vorgang, der sicherstellt, dass das aufgenommene Datenpaket mit den notwendigen Informationen versehen ist, die daraus ein „Archival Information Package“ machen. Im „Ingest“-Vorgang werden die archivisch relevanten Informationen angelegt und das archivierte Paket für die langfristige Aufbewahrung und Verfügbarkeit aufbereitet. Aus dem „Submission Information Package“ wird ein „Archival Information Package“. Das Archivpaket kann dann auf der „anderen“ Seite über definierte Zugriffsweisen abgerufen werden. Aus dem „Archival Information Package“ wird ein „Dissemination Information Package“. Das Archivpaket ist also weder identisch mit dem Datenpaket, das vor der Archivierung erzeugt und bearbeitet wurde, noch mit dem Datenpaket, das die Benutzerin/der Benutzer bei sich auf den Rechner kopiert erhält. In der Benutzungspraxis von digitalen Archiven fällt das nicht auf, denn das Benutzerinterface realisiert die definierten Zugriffsweisen. Benutzungseinschränkungen, notwendige Formattransformationen, Versionen

¹ Zu diesem Modell existiert eine deutschsprachige Einführung (Brübach 2012).

oder die Auswahlmöglichkeiten für Teile des archivierten Datenpakets sind Teil dieser Benutzungsoberfläche.

Eine Konsequenz der maschinellen Vermittlung zwischen Daten und Benutzern ist, dass die Organisation der Daten ihre Benutzbarkeit festlegt. Man muss also die Datenmodelle verstehen, um den Daten Sinn zu geben. So unterscheidet das OAIS auch zwischen den reinen Bitströmen und der Repräsentationsinformation, die im Archivinformationspaket zusammengefasst sind. Erst mit der Information, ob man eine PDF/A-Datei oder ein JPG-Bild im Archiv abgelegt hat, wird der Datenstrom sinnvoll benutzbar. Das Archivobjekt ist also kein Gegenstand oder eine bestimmte physische Konfiguration auf einem Datenträger, sondern eine abstrakte Datenstruktur, deren Logik in einem menschenlesbaren Modell beschrieben ist.

Als Beispiel für eine praktische Implementation kann das „Geisteswissenschaftliche Asset Management System“ (GAMS) dienen, das an der geisteswissenschaftlichen Fakultät der Universität Graz für die Realisierung von digitalen Forschungsressourcen verwendet wird (Steiner und Stigler 2014–2017; Stigler und Steiner 2018). Es ist ein System, in dem das Digitale Archiv eine zentrale Komponente ist. Das System ist um eine Software gruppiert, die als Speichersystem an vielen Universitäten in Gebrauch ist: das FEDORA Commons Repositorium. GAMS hat z. B. mit dem Cirilo-Client verschiedene Workflows etabliert, die von GeisteswissenschaftlerInnen erzeugte Forschungsdaten in unterschiedlichen Formaten in Archivobjekte umwandeln. Diese Workflows münden in sogenannte „Content Models“ für die Archivobjekte. Das Archiv geht z. B. davon aus, dass sich Textdaten gut als XML-Dokumente nach den Richtlinien der Text Encoding Initiative (TEI) archivieren lassen, die inhaltliche Strukturen des Textes ebenso abbilden kann wie die Beziehung zwischen digitalem Volltext und Repräsentation des Textes in digitalen Bildern. Eine MS-Word-Datei wird deshalb in dieses Format umgewandelt. Was in der Word-Datei in Layoutelementen kodiert ist, ist in der archivierten XML-Datei in eine formale Datenstruktur überführt, die nach den Regeln der TEI bestimmte Textphänomene beschreibt und damit abstrahiert. Eine Fußnote mit einer alternativen Lesart kann so z. B. zu einem Code werden, der beide Lesarten gleichwertig macht: `<app><rdg>Lesart</rdg><rdg>Lesart</rdg></app>`. Das Archiv kennt darüber hinaus Datenmodelle für die Repräsentation von Museumsobjekten, deren Metadaten nach dem für Museen entwickelten Datenaustauschstandard LIDO (= Lightweight Information Describing Objects; <http://network.icom.museum/cidoc/working-groups/lido>) erfasst werden (vgl. Coburn et al. 2010). Reine Datenobjekte werden nach dem W3C-Standard des Semantic Web, im Resource Description Framework RDF, gespeichert. Beim Ingest müssen also die Ausgangsdaten – Word-Dokumente, Excel-Tabellen, Bilder in Ordnern im Dateisystem etc. – in die Archivformate umgewandelt werden.

Diese Archivformate wurden gewählt, weil die Forschung über digitale Archivierung ihnen die größten Chancen zuschreiben, dass zukünftige Computersysteme sie ohne unnötigen Informationsverlust wieder einlesen können.

Zu den Content Models der Archivobjekte beschreibt GAMS nun Zugriffswege, auf denen die Archivobjekte den BenutzerInnen angezeigt werden. Diese können von Projekt zu Projekt unterschiedlich sein. Das Cantus-Netzwerk-Projekt erzeugt z. B. aus mehreren XML/TEI-Dateien einzelner Transkriptionen eine gemeinsame Ansicht (vgl. Praßl 2015–2019; s. Abb. 1). Natürlich ist auch die einfache Anzeige des archivierten Datenstroms in seinem Ursprungsformat möglich. Die konkrete Anzeige ist also nur je ein Angebot, wie die Daten den BenutzerInnen präsentiert werden können, nicht notwendige Voraussetzung, die Daten zu nutzen. Die existierenden Zugriffswege sind als „Dissemination Packages“ auch im Archivsystem dokumentiert.

Abb. 1: Screenshot aus der *Liber ordinarius Pataviensis*.

Eine Archivale ist also nicht mehr das Original, das in einer ihm wohlgesonnenen Umgebung darauf wartet, wieder gelesen zu werden, sondern eine Datenstruktur, die für die Archivierung nicht nur ausgewählt, sondern auch angemessen vorbereitet wurde. Sie kann dabei so verändert werden, dass sie gut archivierbar ist. Zusätzlich gibt es Angaben zu den Benutzungsweisen des digitalen Archivals. Als Datenstruktur ist das digitale Archivale reicher an Informationen über sich selbst als eine mechanische Kopie eines Papierarchivals, denn viele Eigenschaften sind explizit gemacht, so dass sie und ihnen inhärente Funktionalitäten von Maschinen verarbeitet werden können. Der Computer ist also nicht nur das Instrument, das zwischen Bitstrom und Mensch vermittelt, er vermittelt auch zwischen einem „Original“-Dokument aus der Praxis vor der Archivierung und der Benut-

zerin/dem Benutzer des Archivals, indem er vom Ausgangsdokument eine eigenständige Archivdatenstruktur erzeugt, die über ihr eingeschriebene Methoden benutzt wird.

Die Geschichtswissenschaft hat aufbauend auf den Forschungsergebnissen von Hayden White darauf hingewiesen,² dass die wissenschaftliche Erforschung von Vergangenheit, also auch die Nutzung von Archivmaterial, mit der Konstruktion von Narrationen vonstattengeht (White 1973; White 1987). Damit stellt sich die Frage, wie diese Narrative konfiguriert sind, wenn die digitalen Archivalien sich den BenutzerInnen nicht mehr als „unberührtes“ Relikt der Vergangenheit präsentieren, sondern als abstrahierter Informationsspeicher. Es sind also nicht mehr nur Überlieferungszufall und -chance (vgl. Esch 1985) oder die archivische Selektion,³ die dazu führen, dass ein Informationsträger als Quelle wissenschaftlicher Forschung dienen kann, sondern es sind Algorithmen und formale Datenmodelle, mit denen aus einem elektronischen Dokument im praktischen Gebrauch ein Archivobjekt wird und mit denen die ForscherInnen auf diese Archivobjekte zugreifen.

2 Leser

Das GAMS definiert unterschiedliche Zugriffswege zu den digitalen Archivalien. Wie aber schon immer in der Archivierung können die möglichen Zugriffswege nicht vorhergesagt werden. Ein gutes Beispiel für die Änderungen des Zugriffs auf die digital abstrahierten Archivalien ist die automatische Handschriftenentzifferung. Sie hat in den letzten Jahren nennenswerte Fortschritte gemacht. Dazu haben insbesondere die Archive beigetragen, in denen offensichtlich ist, dass eine zunehmende Bilddigitalisierung die Benutzung nur bedingt erleichtert, denn ein digitales Faksimile reproduziert nur den Informationsgehalt des visuellen Eindrucks. Die ForscherInnen müssen die Bilder erst in „Text“ oder gar „Information“ umwandeln; sie müssen sie lesen.

Das hat unter anderem dazu geführt, dass unter ArchivarInnen lange das Postulat galt, dass kein Archivale digitalisiert werden sollte, das nicht nach archivischen Standards ausreichend erschlossen ist, damit das Digitalisat ebenfalls auffindbar wird. ArchivarInnen wissen jedoch um die Schwächen ihrer Erschlie-

² Grundlegende theoretische Überlegungen zum historischen Erzählen auch von Jörn Rüsen (1985).

³ Zur archivischen Bewertungslehre vgl. Kretzschmar (1999), Schellenberg (1990) sowie Bischoff und Kretzschmar (2005).

ßung: Einen Archivbestand gegen die von den ArchivarInnen angelegten Erschließungsstrukturen zu benutzen, ist nur unter erheblichem Zeitaufwand möglich. Große serielle Bestände sind dafür besonders ungünstig: Eine summarische Beschreibung „Rechnungen 1489–1804, 400 Bde.“ ist für eine inhaltliche und formale Kategorisierung ausreichend: Der Aufbau jedes Bandes ist den übrigen so ähnlich, dass eine Einleitung zum Findbuch die zu erwartenden Inhalte gut beschreiben kann. Ungünstig wird es, wenn es um individuelle Informationen geht – und in seriellen Quellen sind es oft diese Erwähnungen einer Person, eines Ereignisses, die für die Forschung von Interesse sind. Ein digitales Faksimile ändert damit nichts an der grundsätzlichen Benutzungsweise „Lesen“.

Angesichts des Umfangs der Akten in den Archiven erscheint es jedoch sinnvoll, die BenutzerInnen dabei zu unterstützen. Das französische HIMANIS-Projekt (<https://himanis.org>) hat nun an der Serie der königlichen Kanzleiregister von 1302 bis 1483 vorgeführt, wie eine solche Unterstützung durch den Computer aussehen könnte: Die Pixelstrukturen der digitalen Faksimiles der Bilder können mit Hilfe von „Computer Vision“ als Repräsentanten von Text verstanden werden. Statistische Methoden auf Basis von bereits existierenden Transkriptionen – im Fall von HIMANIS sind das die Editionen von Guérin und Celier (1881–1958) – erlauben es dann, algorithmisch Vorschläge zu erzeugen, wie das Textbild in Einzelbuchstaben übertragen aussehen könnte (vgl. Bluche et al. 2017).

Das in unserem Kontext Interessante an der Methode ist nun nicht so sehr die Möglichkeit, den Bestand nach Stichwörtern überhaupt durchsuchen zu können, sondern wie sich dieses Suchen von jenem interessierter und/oder forschender BenutzerInnen unterscheidet. Für das Verständnis des digitalen Archivs führt dabei nicht so sehr die Geschwindigkeit weiter, mit der die über 400 Bände „gelesen“ werden können, sondern wie das „Lesen“ selbst aussieht: Die Maschine „liest“ nämlich in einem Raum von numerisch bestimmten Möglichkeiten. Im Modell der angewendeten Handwritten Text Recognition (HTR) gibt es keine eindeutige Entscheidung, welches Wort an welcher Stelle steht, sondern nur Wahrscheinlichkeitswerte für die Deutung. Diese beruhen auf Training mit bekanntem Material und integrieren so auch Wissen über die in den Texten verwendete Sprache. Das Ergebnis ist damit einem menschlichen Lesevorgang einerseits sehr ähnlich – andererseits auch vollkommen unähnlich, denn statt Sinnbildung durch Verstehen wird ein Wahrscheinlichkeitsgrad für jede Lesung angegeben. Das hat zur Folge, dass mit der Veränderung der „vertrauenswürdigen“ Wahrscheinlichkeitsstufe das Stichwort „Bawarie“ nicht nur einmal, sondern zweimal gefunden wird, sich der Text also für die BenutzerInnen verändert. Das Digitale Archiv ist damit – auf die Spitze getrieben – eine prozentuale Wahrscheinlichkeit von vergangener Existenz. Der Computer ist als Leser digitalisierter Archivalien ein mathematischer.

Diese Eigenschaft wird auch erkennbar, wenn man den Zugriff auf die Archivalien in den Erschließungsinstrumenten beschränkt, also nur Metadaten betrachtet. Das gilt schon mit den geläufigen Angaben einer Zusammenfassung mit „3 Treffer“. Das gilt aber noch viel mehr, wenn man sich ansieht, wie Archiverschließungsinformationen visualisiert werden können. Mitchell Whitelaw hat hierfür mit seinen Arbeiten zum Australischen Nationalarchiv ein besonders augenfälliges Beispiel geschaffen (vgl. Abb. 2).

Abb. 2: Mitchell Whitelaw, Screenshot aus The Visible Archive, 2009.

Die Grafik symbolisiert jeden Archivbestand als Quadrat. Der Umfang der Archivbestände bestimmt die Größe der Quadrate, dabei stellt der äußere Rand den Umfang in Regalmetern und das innere Quadrat die Anzahl von archivalischen Einheiten. Die Quadrate sind in chronologischer Reihenfolge angeordnet. Die Farbe der Quadrate repräsentiert die Identifikationsnummer der ca. 9.000 Bestandsbildner, d. h. niedrige Nummern sind mit roten Quadraten dargestellt, während für die Bestände von Bestandsbildnern mit höheren Nummern blaue oder violette Farbe verwendet wird. Orange, Gelb und Grün liegen dazwischen (vgl. Whitelaw 2009). Auch diese Darstellung beruht auf numerischen und mathematischen Eigenschaften der Archivbestände: Umfang, Datierung, Num-

merisierung der Bestandsbildner. Führen diese numerischen Eigenschaften zu numerischen Narrativen?

3 Begriffsbildner

Der Computer führt eine weitere Perspektive auf die in den Archivalien überlieferten Informationen ein. Wenn ArchivarInnen die ihnen anvertrauten Bestände zunehmend mit elektronischen Hilfsmitteln beschreiben, dann wirken die digitalen Systeme, mit denen sie arbeiten, auch auf die Beschreibungen zurück. Für die Erschließung hat z. B. der internationale Archivverband ISAD am Ende des letzten Jahrtausends Kategorien entwickelt: ISAD(G) (International Council on Archives 2000). Zu den Kernkategorien gehören die Signatur, ein Titel, die Provenienzstelle, die Laufzeit, der Umfang und die Verzeichnungsstufe. Mit letzterem ist eine Angabe gemeint, welche Art von Aggregation der Schriftstücke beschrieben ist: ganze Registraturen, freie Sammlungen, sachliche Aufteilungen, Aktenfaszikel/Dossiers oder Einzelstücke. Diese Formen der Aggregation werden als Hierarchie verstanden. Die digitale Umsetzung von ISAD(G), der von der Society of American Archivists gepflegte Standard „Encoded Archival Description“ (EAD; vgl. Technical Subcommittee for Encoded Archival Standards of the Society of American Archivists 2018) nimmt darauf Rücksicht und erwartet, dass jede Beschreibungseinheit ein Attribut besitzt, das die Verzeichnungsstufe (@LEVEL) benennt. Der Vorstellung, dass Archivalien hierarchisch zu organisieren sind, kommt das XML-Datenformat entgegen, in dem die EAD-Daten gespeichert und ausgetauscht werden: XML ist eine formale Sprache, in der die Beschreibungselemente strikt ineinander verschachtelt sein müssen (vgl. Bray et al. 2008). Die digitale Repräsentation von Archivgut in EAD legt also eine Vorstellung fest, in der jedes Aktenstück genau einem Aktenheft, jedes Aktenheft genau einer Sachgruppe in genau einer Behördenregistratur zugeordnet werden kann. Im Verständnis vieler ArchivarInnen und in den meisten online verfügbaren Archivinformationssystemen ist diese Hierarchie ein dominierendes Mittel zur Organisation der Archivbeschreibungen. In der archivfachlichen Diskussion wird diese Form der Wissensorganisation als Mittel verstanden, den zum Verständnis der archivischen Einheiten notwendigen Kontext herzustellen (vgl. Riley und Shepherd 2009).⁴ Das ist auch schlüssig, denn natürlich sind Einzelstücke als Teil des Vorgangs, zu dem sie vom Sachbearbeiter oder der Registratur zugeord-

⁴ Menne-Haritz (2016, 210) geht noch selbstverständlich von einer hierarchisch organisierten Erschließung aus.

net wurden, besser verständlich. Aktentitel wie „Generalia“ sind äußerst informationsarm, wenn sie nicht als Teil einer Behördenregistratur erkennbar sind, aus der sich der Aufgabenbereich der Behörde und damit die möglichen „generellen“ Themen ableiten lassen.

Diese Annahme einer strikten Hierarchie als bestes archivisches Ordnungsprinzip ist schon länger theoretisch angezweifelt worden, insbesondere wenn die Instabilität von Behördenstrukturen dazu führt, dass Akten ihren Kontext verändern (z. B. Yeo 2014). Einen technischen Niederschlag hat diese Diskussion erst vor kurzem gefunden. In den letzten zehn Jahren wird XML als Basistechnologie der Modellierung von Metadaten zunehmend vom Resource Description Framework (RDF) verdrängt. RDF ist, wie XML, ein vom W3C vorgeschlagenes Format zum Datenaustausch über das Internet. Es ist der Grundbaustein des sogenannten „Semantic Web“, in dem Informationen über Objekte als Verweise auf andere Informationen ausgedrückt werden, also ein Netz an Informationen entsteht (vgl. Schreiber und Raimond 2003–2014). Diese Entwicklung haben ArchivarInnen aufgegriffen und einen neuen Vorschlag für die digitale Formalisierung von Archivbeschreibungen gemacht: „Records in Context“ (RiC) (vgl. International Council on Archives, Expert Group on Archival Description 2016). Die Bezeichnung des Datenmodells ist insofern besonders aussagekräftig, als es eine gänzlich neue Vorstellung von „Kontext“ erzeugt: Während mit EAD Kontext entsteht, in dem eine archivische Einheit durch die ihr benachbarten und ihr übergeordneten Verzeichnungseinheiten besser verständlich wird, wird in RiC eine archivische Einheit („record“) nicht mehr über „Nachbarschaft“ und „Überordnung“, sondern durch Verweise auf Informationsknoten beschrieben, die wiederum mit anderen „records“ und Informationsknoten in Verbindung stehen können: Die Akteure („agents“), ihre Aufgaben („functions“) und Zuständigkeiten („mandate“) bestimmen die Dokumente ebenso wie betroffene Gegenstände, Ereignisse, abstrakte Begriffe und Orte. So kann die Autorisierung einer Handlung in einem anderen Archivbestand verwahrt werden als die Dokumente der Handlung selbst. Personen können Funktionen in einer Organisation innehaben, die eine von den Personen unabhängige Registratur führt, eine Registratur, die aber natürlich auch die Handlungen dieser Person dokumentiert. In RiC ist also der Begriff „Kontext“ nicht eine Hierarchie von Aggregationsformen, sondern das Netz an Akteuren, Handlungen und Themen, die in den Archivalien dokumentiert sind.

Die digitale Form der Archivbeschreibungen unterstützt das eine oder andere Verständnis von „Kontext“. Die auf diesen Beschreibungsstandards aufbauenden Archivinformationssysteme konfigurieren das aus den Archiven extrahierbare Wissen unterschiedlich. Der Computer scheint die Begriffe vorzuformen, mit denen die digitalen Archivalien verstanden werden.

Aus der Perspektive des interpretativen Narrativs, das eine zentrale Operation historischer Interpretation der Dokumente ist, stellt sich also abschließend die Frage, ob das Digitale Archiv als Vermittler zwischen Entstehungszusammenhang und forschender Benutzung nicht die Anschauung vom Archivgut substantiell verändert. Paul Ricœur hatte den HistorikerInnen noch einen spontanen „Realismus“ unterstellen können, die in ihrem Selektions-, Lese- und Schreibvorgang gewissermaßen die Entstehung der Akten nachvollzogen (Ricœur 1985, 183–184). Das Digitale Archiv hebt diese Analogie nun entweder auf, denn die Vorgänge, mit denen das Digitale Archivalie als Transformation in eine abstrakte Datenstruktur entsteht und durch formalisierte Algorithmen und numerische Operationen den ForscherInnen bereitgestellt wird, entsprechen nicht den Handlungsmustern der ForscherInnen. Oder die Arbeit mit digitalen Archivalien lässt sich mit formalen Algorithmen, numerischen Operationen und Begriffssystemen erzählen, die denen der Digitalen Archive vergleichbar sind – einem digitalen Schauplatz Archiv.

Literaturverzeichnis

- Archivschule Marburg. *Terminologie der Archivwissenschaft*. 2012–2015. <http://www.archivschule.de/uploads/Forschung/ArchivwissenschaftlicheTerminologie/Terminologie.html> (24.10.2018).
- Bischoff, Frank M., und Robert Kretschmar (Hg.). *Neue Perspektiven archivischer Bewertung. Beiträge zu einem Workshop an der Archivschule Marburg, 15. November 2004*. Marburg: Archivschule Marburg, 2005.
- Bluche, Théodore et al. „Preparatory KWS Experiments for Large-Scale Indexing of a Vast Medieval Manuscript Collection in the HIMANIS Project“. *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 1 (2017) 311–316. DOI: 10.1109/ICDAR.2017.59.
- Bray, Tim, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler und François Yergeau (Hg.). *Extensible Markup Language (XML) 1.0 (Fifth Edition), W3C Recommendation, 26 November 2008*. World Wide Web Consortium, 2008. <https://www.w3.org/TR/2008/REC-xml-20081126/> (23.10.2018).
- Brübach, Nils. „OAIS. Das ‚Open Archival Information System‘. Ein Referenzmodell zur Organisation und Abwicklung der Archivierung digitaler Unterlagen“. *6. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ am 5.–6. März 2002*. Dresden. 2002. https://www.staatsarchiv.sg.ch/home/auds/06/_jcr_content/Par/downloadlist_0/DownloadListPar/download_9.ocFile/Text%20Bruebach.pdf (23.10.2018).
- Coburn, Erin, Richard Light, Gordon McKenna, Regine Stein und Axel Vitzthum. *LIDO – Lightweight Information Describing Objects Version 1.0*. Hg. International Council of Museums/Conseil International des Musees. November 2010. <http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf> (23.10.2018).

- Ernst, Wolfgang. *Digital Memory and the Archives*. Hg. Jussi Parikka. Minneapolis, London: University of Minnesota Press, 2012.
- Esch, Arnold. „Überlieferungschance und Überlieferungszufall als methodisches Problem des Historikers“. *Historische Zeitschrift* 240.3 (1985): 529–570.
- Guérin, Paul, und Leonce Celier (Hg.). *Recueil des documents concernant le Poitou contenus dans les registres de la chancellerie de France*. 14 Bde. Poitiers: Société des archives historiques du Poitou, 1881–1958.
- International Council on Archives (Hg.). *ISAD(G): General International Standard Archival Description*. Second Ed. Adopted by the Committee on Descriptive Standards of the International Council on Archives, Stockholm, Sweden, 19–22 September 1999. Ottawa 2000. [https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD\(G\)_Second-edition_EN.pdf](https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD(G)_Second-edition_EN.pdf) (23.10.2018).
- International Council on Archives. Multilingual Archival Terminology. 2012. <http://www.ciscra.org/mat/> (24.10.2018).
- International Council on Archives, Expert Group on Archival Description (Hg.). *Records in Context. A Conceptual Model for Archival Description, Consultation Draft v0.1, September 2016*. International Council on Archives, 2016. URL: <https://www.ica.org/sites/default/files/RIC-CM-0.1.pdf> (23.10.2018).
- Kirschenbaum, Matthew G. *Mechanisms. New Media and the Forensic Imagination*. Cambridge/ Mass., London: MIT Press 2012.
- Kramski, Heinz Werner. „Digitale Dokumente im Archiv“. *Handbuch Archiv. Geschichte, Aufgaben, Perspektiven*. Hg. Marcel Lepper und Ulrich Raulff. Stuttgart: Metzler, 2016. 178–197.
- Kretzschmar, Robert. „Die ‚neue archivistische Bewertungsdiskussion‘ und ihre Fußnoten. Zur Standortbestimmung einer fast zehnjährigen Kontroverse“. *Archivalische Zeitschrift* 82.1 (1999): 7–40.
- Lepper, Marcel, und Ulrich Raulff (Hg.). *Handbuch Archiv. Geschichte, Aufgaben, Perspektiven*, Stuttgart: Metzler, 2016.
- Menne-Haritz, Angelika. „Erschließung“. *Handbuch Archiv. Geschichte, Aufgaben, Perspektiven*. Hg. Marcel Lepper und Ulrich Raulff. Stuttgart: Metzler, 2016. 207–217.
- Pearce-Moses, Richard. *A Glossary of Archival and Records Terminology*. Chicago: The Society of American Archivists, 2005. <http://files.archivists.org/pubs/free/SAA-Glossary-2005.pdf> (24.10.2018).
- Praßl, Franz Karl (Hg.). *Cantus Network. Libri ordinarii of the Salzburg metropolitan province*. Graz: Zentrum für Informationsmodellierung, 2015–2019. <http://gams.uni-graz.at/cantus> (23.10.2018).
- Pross, Harry. *Publizistik: Thesen zu einem Grundcolloquium*. Neuwied, Berlin: Luchterhand, 1970.
- Ricœur, Paul. *Temps et récit*. Bd. 3: *Le temps raconté*. Paris: Éditions du Seuil, 1985.
- Riley, Jenn, und Kelcy Shepherd. „A Brave New World: Archivists and Shareable Descriptive Metadata“. *American Archivist* 72.1 (2009): 91–112.
- Rüsen, Jörn. „Die vier Typen des historischen Erzählens“. *Formen der Geschichtsschreibung*. Hg. Reinhart Koselleck, Heinrich Lutz und Jörn Rüsen. München: dtv, 1982. 514–605.
- Schellenberg, Theodore R. *Die Bewertung modernen Verwaltungsschriftguts*. Hg. u. übers. Angelika Menne-Haritz. Marburg: Archivschule Marburg, 1990.
- Schreiber, Guus, und Yves Raimond (Hg.). *RDF 1.1-Primer, W3C Working Group Note 23 June 2014*. 2003–2014. <https://www.w3.org/TR/rdf11-primer> (23.10.2018).

- Steiner, Elisabeth, und Johannes Stigler. *GAMS and Cirilo Client. Policies, documentation and tutorial*. Graz 2014–2017. <http://gams.uni-graz.at/docs> (23.10.2018).
- Stigler, Johannes, und Elisabeth Steiner. „GAMS – an Infrastructure for the Long-Term Preservation and Publication of Research Data from the Humanities“. *Mitteilungen der VÖB* 71.1 (2018): 207–216.
- Stingelin, Martin. „Archivmetapher“. *Handbuch Archiv. Geschichte, Aufgaben, Perspektiven*. Hg. Marcel Lepper und Ulrich Raulff. Stuttgart: Metzler, 2016. 21–27.
- Technical Subcommittee for Encoded Archival Standards of the Society of American Archivists (Hg.). *Encoded Archival Description. Tag Library Version EAD3, 1.1.0, Ed. 2018*. Chicago: Society of American Archivists, 2018. https://www.loc.gov/ead/EAD3taglib/tl_ead3.pdf (23.10.2018).
- White, Hayden. *Metahistory. The Historical Imagination in Nineteenth-Century Europe*. Baltimore: Johns Hopkins University Press, 1973.
- White, Hayden. „Das Problem der Erzählung in der modernen Geschichtstheorie“. *Theorie der modernen Geschichtsschreibung*. Hg. Pietro Rossi. Frankfurt a. M.: Suhrkamp, 1987. 57–106.
- Whitelaw, Mitchell. *The Visible Archive*. 2009. <http://mtchl.net/folio/the-visible-archive/> (23.10.2018).
- Yeo, Geoffrey. „Contexts, Original Orders, and Item-Level Orientation. Responding Creatively to Users’ Needs and Technological Change“. *Journal of Archival Organization* 12.3-4 (2014): 170–185.