

ARCHIV, SUPERARCHIV, METAARCHIV

In Theorie und Praxis hat das Archiv im digitalen Zeitalter Konjunktur. Dabei verschwimmen die Grenzen von Archiven ebenso wie die des Archivbegriffs zunehmend. Neben jene traditionellen, mit einem gesetzlichen Sammlungsauftrag versehenen historischen Archive, die ihre Bestände zunehmend auch online zugänglich machen, treten in der Welt digital-vernetzter Medien eine Vielzahl neuer Archive im engen und weiteren Sinn. Die historisch gewachsenen und institutionalisierten Praktiken der Versammlung von Dokumenten, ihrer Bewahrung oder Zerstörung (Kassation) sowie ihrer Erschließung und Verfügbarmachung, welche den Archivbegriff der Archivkunde informieren, werden von einem eher unspezifischen Archivverständnis überlagert, in dem die Momente der Artikulation, Distribution, Sammlung, Bewahrung und Zugänglichmachung von Informationen lose und variabel aneinander gekoppelt sind.¹ Dabei ist das Archiv, wie Knut Ebeling und Stephan Günzel konstatiert haben, »zur Metapher [...] für jede Art der Speicherung«² geworden. Vor dem Hintergrund dieses vergleichsweise weiten und unspezifischen Archivverständnisses vermag auch das World Wide Web selbst als Archiv zu erscheinen, das als ›Archiv aller Archive‹ die Imagination eines totalen und unbegrenzten Informations- und Wissensspeichers einzulösen scheint.³ Dass das Gedächtnis des WWW entgegen derartiger Attribuierungen bestenfalls löchrig ist, wurde früh erkannt.⁴ In Reaktion hierauf entstand 1996 das Internet Archive, welches es sich seither zur Aufgabe macht,

1 Zum archivkundlichen Archivbegriff und dessen historischer Genese vgl. exemplarisch JOHANNES PAPRITZ: *Archivwissenschaft: Einführung, Grundbegriffe, Terminologie, Teil 1, Bd. 1*, Marburg 1976, S. 41 ff.

2 KNUT EBELING/STEPHAN GÜNZEL: »Einleitung«, in: dies. (Hg.): *Archivologie: Theorien des Archivs in Philosophie, Medien und Künsten*, Berlin 2009, S. 7–26 (hier: S. 7).

3 Die Imagination universeller Wissens- und Informationsspeicher ist nicht spezifisch für digital-vernetzte Medien. Entlang zentraler Wegmarken zeichnet ALEX WRIGHT die Geschichte dieser Vision nach, vgl. ders.: *Glut: Mastering Information Through the Ages*, Washington D. C. 2007.

4 Siehe exemplarisch PETER LYMAN: »Archiving the World Wide Web«, in: Council on Library and Information Resources/Library of Congress (Hg.): *Building a National Strategy for Digital Preservation: Issues in Digital Media Archiving. Commissioned for and sponsored by the National Digital Information Infrastructure and Preservation Program 2002*, S. 38–51 (hier: S. 38 f.), <http://www.clir.org/pubs/reports/pub106/pub106.pdf> [eingesehen am 18.03.2010]; STEWART BRAND: »Escaping the Digital Dark Age«, in: *Library Journal*, 124/2 (1999), S. 46–48.

das *Megaarchiv* WWW zu archivieren. Trotz oder gerade wegen des Informationsüberflusses, der für die zeitgenössische Medienwelt charakteristisch ist, bedarf die Bewahrung und dauerhafte Zugänglichkeit des Digitalen einer besonderen Sorge und Sorgfalt, wie Yuk Hui in seinem *Archivist Manifesto* herausgestellt hat.⁵ Begleitet wird die Expansion des Archivischen von einer Konjunktur der Archivfrage in medien- und kulturwissenschaftlichen Debatten, welche spätestens »seit dem französischen Revolutionär Foucault in aller Munde«⁶ ist. Definiert als das »Gesetz dessen, was gesagt werden kann«⁷, trägt das Archiv für Foucault nicht in erster Linie die Signatur des Historischen oder Vergangenen. Der Archivbegriff Foucaults fokussiert vielmehr die jeweilige Gegenwart der Produktion, Distribution und Bewahrung von Medienobjekten, denn das Archiv bedingt die Artikulation, Zirkulation und Sammlung von Aussagen – seien sie nun sprachlicher oder bildlicher Natur.

Zwischen den Extrempolen des archivkundlichen Archivbegriffs einerseits und Foucaults Konzeptualisierung des Archivs als »allgemeine[s] System der Formation und der Transformation der Aussagen«⁸ situieren sich in den digitalen Medien vielfältige heterogene Praktiken, Akteure, Institutionen und Verfahren der Sammlung und Bereitstellung von digitalen Inhalten. Hier finden Datenbanken oder Archive von Hobbyist_innen, Amateuren_innen und Privatsammler_innen ebenso Platz wie jene Plattformen, auf denen Nutzer_innen Bilder, Fotos, Videos etc. teilen und auf denen sie Gefundenes sammeln und kuratieren, wie z. B. Facebook, YouTube, Instagram, Tumblr, Ffffound usw. Vergessen werden dürfen auch nicht die großen Bild- und Textarchive, die Museen und Bibliotheken online stellen, um ihre Bestände der Öffentlichkeit zu präsentieren. Vieles davon wird heute umgangssprachlich Archiv genannt oder häufig als ein solches beschrieben.

Der Begriff des *Archivs* ist im digitalen Zeitalter ambig. Daher gilt es anzuerkennen, dass es *das* digitale Archiv nicht gibt. Es gibt vielmehr ein breites Spektrum heterogener Praktiken, Akteure und Verfahren der Sammlung, Distribution und Verfügbarmachung von Medienobjekten *in*, *mit* und *durch* digitale Medien. Diese Archive im engen, weiten und weitesten Sinn verbindet eine Reihe von Familienähnlichkeiten, die ihre Funktionen, Prozesse und Praktiken nur lose miteinander verknüpfen, ohne ein festes Zentrum zu haben. Es sind Orte der Versammlung – oder, wie Derrida gesagt hat, der *Konsignation* –, d. h. es sind Orte der Bewahrung, Identifikation, Ordnung,

5 YUK HUI: »Archivist Manifesto«, in: *Mute*, 2013, <http://www.metamute.org/editorial/lab/archivist-manifesto> [eingesehen am 12.11.2015].

6 EBELING/GÜNZEL: »Einleitung« (Anm. 2), S. 9.

7 MICHEL FOUCAULT: *Archäologie des Wissens*, Frankfurt a. M. 1981, S. 187.

8 Ebd., S. 188.

Erschließung, Interpretation und Verfügbarmachung, es sind aber auch Orte der Zerstörung und des Vergessens.⁹

Wenn der folgende Beitrag die mediale Logik digitaler Archive in den Blick nimmt, dann gilt das besondere Augenmerk verschiedenen Eskalationsstufen und Lösungsansätzen für das Zugriffsproblem auf digital gespeicherte Informationen. Vor dem Hintergrund der Geschichte digitaler Datenbanken als Technologien der Sammlung, Speicherung und späteren Abfrage von großen Informationsmengen *in* und *mit* Computern wird in einem ersten Schritt nach den spezifischen Herausforderungen des computer-gestützten Suchens und Findens von Informationen im Allgemeinen und Bildern im Besonderen gefragt.¹⁰ Die Frage nach digitalen Bildsammlungen drängt sich dabei nicht zuletzt angesichts der »Image Explosion« auf, welche das Marktforschungsunternehmen IDC 2007 diagnostiziert hat.¹¹ All jene digitalen und zunehmend smarten Gadgets, die wir täglich mit uns tragen, mit denen wir hantieren und uns vernetzen, haben fraglos zu einem Boom des Bildhaften im Allgemeinen sowie des Fotografischen und Videografischen im Besonderen geführt: Wir machen Schnappschüsse oder Videos von Orten, die wir besuchen, von dem Essen, das uns im Restaurant serviert wird oder von Ereignissen, denen wir – zufällig oder nicht – beiwohnen und teilen diese Bilder und Videos massenhaft online auf den Profildaten sozialer Netzwerke, auf YouTube und auf Fotoportalen wie Flickr oder Instagram. Im zweiten Schritt wird angesichts der Entwicklung von Archiven hin zu Superarchiven gefragt, wie den beständig wachsenden digitalen (Archiv-)Beständen im Rahmen neuer Archivpraktiken begegnet werden kann. Das Ziel dieser Überlegungen ist es, eine – zumindest schlaglichtartige – Perspektive auf die Vergangenheit, Gegenwart und Zukunft digitaler (Bild-)Archive zu entwickeln und in diesem Zuge einige Chancen, aber auch Herausforderungen freizulegen, welche die Sammlung und Verfügbarmachung von digitalen Medienobjekten darstellt.

Data Banking

Insbesondere im Aspekt der Sammlung und damit einhergehend der Ordnung und Bereitstellung von Medienobjekten weisen traditionelle Archive Parallelen zu jenen medientechnischen Konfigurationen auf, die seit den 1960er Jahren *Datenbanken* ge-

9 Vgl. JACQUES DERRIDA: *Dem Archiv verschrieben: Eine Freudsche Impression*, Berlin 1997, S. 13.

10 Zur Geschichte von Bildarchiven im Allgemeinen siehe das dritte Kapitel in: BENJAMIN DRECHSEL: *Politik im Bild: Wie politische Bilder entstehen und wie digitale Bildarchive arbeiten*, Frankfurt a. M. 2005, S. 91 ff.

11 IDC: *Expanding the Digital Universe: A Forecast of Worldwide Information Growth Through 2010*, 2007, <https://web.archive.org/web/20100413111742/http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf> [vom Internet Archive archivierte Version eingesehen am 10.12.2010].

nannt werden. In der frühesten Quelle, die sich – noch in einem traditionellen Papierarchiv – finden ließ, heißt es: »Data Banking is the process of communicating between many conceivers to many receptors through a store. By providing mechanisms for decision-making the receptor selects resting information«. ¹² Das *Data Banking* wird hierbei als ein mehrstufiger Kommunikationsprozess verstanden, in dem Medienobjekte gesammelt, verzeichnet, indexiert und in einem Speicher abgelegt werden, um zu einem späteren Zeitpunkt anhand spezifischer Suchkriterien wiedergefunden werden zu können. ¹³ Diese dem Abschlussbericht eines vom Naval Ordnance Laboratory geförderten Forschungsprojekts entnommene abstrakte Funktionsbeschreibung von Technologien zur Speicherung und Abfrage von Informationen erwähnt den Computer als spezifische *Information Retrieval*-Technologie noch nicht. Die Datenbank wird vielmehr als ein Raum voller Mikrofilschränke entworfen; sie hat die Kontur einer Bibliothek bzw. eines Archivs, in dem Texte, Bilder etc. für den späteren Abruf gelagert werden. ¹⁴

Auch wenn die Genese digitaler Datenbanktechnologien aus dem Geist des Managements nicht vergessen werden darf, so findet sich hier eine Entwicklungslinie, welche die computergestützte Sammlung, Speicherung und Abfrage von Informationen in die Tradition bibliothekarischer und archivarischer Praktiken rückt. ¹⁵ Diese Entwicklungslinie setzt sich in den Debatten über *Information Storage and Retrieval* ebenso fort wie in der Entwicklung von spezifischen bibliografischen Datenbankangeboten in den 1960er und 1970er Jahren, wie z. B. *MEDLARS*, *ERIC*, *DIALOG* und *ORBIT*. ¹⁶ Derartige bibliografische Online-Datenbanken, in denen man seit den 1970er Jahren gegen eine Gebühr auch via Internet recherchieren konnte, entstanden nicht zuletzt in Reak-

12 PETER K. WORSLEY et. al.: *A Study of the Fundamentals of Information Storage and Retrieval* [Final Report], in: United States Government Computing Collection, circa 1945–1990 (CBI 63), Box 13, Charles Babbage Institute, University of Minnesota, Minneapolis 1959, S. 5 f. Die Spezifik von Modellen der Datenbankkommunikation und deren Unterschied zum nachrichtentechnischen Kommunikationsmodell Claude Shannons diskutiere ich ausführlich in MARCUS BURKHARDT: *Digitale Datenbanken: Eine Medientheorie im Zeitalter von Big Data*, Bielefeld 2015, S. 167 ff.

13 WORSLEY et. al.: *A Study of the Fundamentals of Information Storage and Retrieval* [Final Report], (Anm. 12), S. 30 f. Beschrieben wird hier das Problem des *Information Retrieval*, welches Calvin N. Mooers 1950 auf den Begriff gebracht hat; vgl. ders.: »Information Retrieval Viewed as Temporal Signalling«, in: *Proceedings of the International Congress of Mathematicians*, Bd. 1, 1950, S. 572–573 (hier: S. 572).

14 Vgl. WORSLEY et al.: *A Study of the Fundamentals of Information Storage and Retrieval* (Anm. 12), S. 91.

15 Zur Geschichte digitaler Datenbanken als Managementtechnologien siehe exemplarisch THOMAS HAIGH: »A Veritable Bucket of Facts«: Ursprünge des Datenbankmanagementsystems«, in: DAVID GUGERLI et al. (Hg.): *Nach Feierabend: Daten*, Züricher Jahrbuch für Wissensgeschichte 3, Berlin 2007, S. 57–98.

16 Siehe hierzu: CHARLES P. BOURNE/TRUDI BELLARDO HAHN: *A History of Online Information Services, 1963–1976*, Cambridge 2003; CHARLES P. BOURNE: »40 Years of Database Distribution and Use: An Overview and Observation.« 1999 Miles Conrad Memorial Lecture, 23. Februar 1999, in: *Information Services & Use*, 24/1 (2004), S. 49–57; M. LYNNE NEUFELD/MARTHA CORNOG: »Database History: From Dinosaurs to Compact Discs«, in: *Journal of the American Society for Information Science*, 37/4 (1986), S. 183–190.

tion auf den Sputnik-Schock der späten 1950er Jahre, der auch ein Problem der bibliografischen Dokumentation – heute würde man eher Informationswissenschaft sagen – darstellte. So berichtet Wolfgang Stock in seinem Einführungsbuch in das *Information Retrieval* die Anekdote, dass die US-Amerikaner_innen die Signale eines Sputnik-Satelliten aufwändig entschlüsselt haben, nur um im Nachhinein festzustellen, dass die Lösung ihres Problems schon längst auf Russisch publiziert war.¹⁷ Man hätte den Code nicht knacken, sondern schlicht lesen müssen, um verstehen zu lernen, was der Satellit zur Erde funkt. Eine ähnliche Erfahrung nahm Roger Summit zum Anlass, um Ende der 1960er Jahre die bibliografische Datenbank *DIALOG* zu entwickeln: »A common statement around Lockheed at the time was that it is usually easier, cheaper, and faster to redo scientific research than to determine whether it has been done previously«.¹⁸ Für Summit war infolgedessen die Frage des computertechnischen Findens oder genauer Wiederfindens von bereits verfügbaren, aber unsichtbar in den Datenbeständen ruhenden Informationen leitend.

Zugriffsprobleme

Mit einer ähnlichen Herausforderung sahen sich auch die Entwickler digitaler Datenbanktechnologien konfrontiert, wie an der denkbar kurzen und somit zugleich relativ unspezifischen Definition digitaler Datenbanken deutlich wird, die 1963 im Rahmen der Konferenz *Development and Management of a Computer-Centered Data Base* vorgeschlagen wurde. Im Bericht zur Tagung heißt es: »A data base consists of: 1. Data, 2. Means of access to the data«.¹⁹ Abgesehen von der anfänglichen Knappheit an Speicherkapazität erwies sich dabei vor allem die Frage des Zugriffs auf gespeicherte Informationen als medientechnologische Herausforderung, welcher auch im Kontext digitaler Archive eine zentrale Bedeutung zukommt.²⁰ Gerade weil die hierfür gefundenen

17 Vgl. WOLFGANG G. STOCK: *Information Retrieval: Informationen suchen und finden*, München 2007, S. 38 f.

18 ROGER SUMMIT: »Reflections on the Beginnings of Dialog: The Birth of Online Information Access«, in: *Chronolog*, 2002, <http://support.dialog.com/publications/chronolog/200206/1020628.shtml> [eingesehen am 11.07.2009].

19 E. W. FRANKS: »Session B: Criteria Influencing Data Base Organization or Design«, in: A. WALKER (Hg.): *Proceedings of the Symposium on Development and Management of a Computer-Centered Data Base (June 10–11, 1963)*, Santa Monica 1964 (United States National Bureau of Standards Collection of Computer Literature (CBI 32), Box 231, Folder 14, Charles Babbage Institute, University of Minnesota, Minneapolis), S. 119–121 (hier: S. 120).

20 Die erste seit 1956 von IBM verkaufte Festplatte verfügte gerade einmal über 5 Megabyte Speicherkapazität. 1981 kam die erste Festplatte mit 1 Gigabyte Speicher auf den Markt. Die flächendeckende Verbreitung von Gigabyte-Speichern dauerte jedoch nochmals 10 bis 15 Jahre. Heute leben wir – zumindest im privaten Bereich – im Zeitalter von Terabyte-Speichern. Eine Rekonstruktion der Geschichte der Festplatte als digitaler Speichertechnologie findet sich in: MATTHEW G. KIRSCHENBAUM: *Mechanisms: New Media and the Forensic Imagination*, Cambridge 2008, S. 76 ff.

computertechnischen Lösungsansätze analog zu den historischen Vorläufern archivarischer und bibliothekarischer Ordnungsverfahren und Findmittel der propositionalen Logik der Sprache folgen, stellt die Bereitstellung von Suchtechnologien für Bilder eine zusätzliche Herausforderung dar. Um diese genauer in den Blick nehmen zu können, sind zunächst allgemein die Verfahren der sprachbasierten Zugänglichmachung von Informationen mit Computern zu betrachten.

Charles Bachman, ein Pionier digitaler Datenbanktechnologien, hat die allgemeine Herausforderung, unsichtbar in Digital Speichern ruhende Informationen wiederfinden zu können, 1962 in folgendem Versprechen auf den Punkt gebracht: »[T]he mass memory's ability to retrieve any specified record is translated into the ability to retrieve exactly the information needed to solve a problem.«²¹ Demzufolge geht es also nicht darum, irgendeine Information abrufen zu können, sondern in einem Bestand diejenigen Informationen ausfindig zu machen, die ein bestimmtes Informationsbedürfnis befriedigen. Computer – so schlau sie uns in manchen ihrer Funktionen auch erscheinen mögen – tun sich damit aber gerade sehr schwer. Sie sind nicht in der Lage, Informationen zu verstehen, sondern können diese nur regelgeleitet prozessieren. Kurzum: Ihnen entzieht sich der Sinn von Nachrichten. Sie operieren nur auf der Ebene von Syntax, von Strukturen, von Mustern, die in Daten vorliegen. Hinsichtlich der Frage des Zugriffs auf digitale Informationssammlungen folgt hieraus nun keineswegs ein Verzicht *auf* oder Abschied *von* der Bedeutung von Nachrichten. Im Gegenteil, das Problem der Bedeutung steht gerade im Zentrum von Technologien des *Information Retrieval*, wobei die Vermittlung oder Übersetzung von menschlichen Interpretations- und Verstehensprozessen mit der computertechnischen Logik der Verarbeitung von Information die zu lösende Aufgabe darstellt. Eben hierauf hat der Informationswissenschaftler Robert Fairthorne in den 1960er Jahren hingewiesen, als er fragte: »[H]ow far can we go in documentation, as in computing, by using ritual in place of understanding?«²² Für Fairthorne geht es somit nicht um die Nachbildung des menschlichen Verständnisvermögens im Computer, sondern um die Verarbeitung von Semantik und Bedeutung in nicht-semantischen Routinen und somit um etwas, das in Anlehnung an Konrad Becker und Felix Stalder als Herstellung »computer-lesbare[r] Signifikanz«²³ bezeichnet werden kann.

21 CHARLES W. BACHMAN: »Integrated Data Store – The Information Processing Machine That We Need!«, Minneapolis 1962 (Charles W. Bachman Papers (CB1 125), Box 1, Folder 11, Charles Babbage Institute, University of Minnesota, Minneapolis), S. 3.

22 ROBERT ARTHUR FAIRTHORNE: »Preface«, in: ders.: *Towards Information Retrieval*, London 1961, S. IX–XI (hier: S. X).

23 KONRAD BECKER/FELIX STALDER: »Einleitung«, in: dies. (Hg.): *Deep Search: Politik des Suchens jenseits von Google*, Innsbruck 2009, S. 7–12 (hier: S. 8).

Neben Verfahren der algorithmischen Analyse und Zuschreibung von Bedeutung wird computerlesbare Signifikanz vor allem durch die Auszeichnung von digitalen Medienobjekten durch erläuternde Metainformationen erzeugt. Eine (nicht *die*) Bedeutung medialer Konstellationen wird hierbei formal expliziert. Im undifferenzierten Strom digitaler Daten werden so syntaktische Unterschiede eingeführt, die einen semantischen Unterschied machen. Daten sind für Computer zunächst nur Aneinanderreihungen von Datenpunkten. In Kombination mit *computerlesbaren* Metainformationen werden diese für Computer als eine bestimmte Information interpretierbar. Computer, die Information ausschließlich auf dem Niveau von Information *als* Realität prozessieren, können digitale Daten durch solche Metainformationen als Information über Realität verarbeiten und automatisch in Datenbanken verwalten und Nutzer_innen somit den suchenden Zugriff auf digitale Informationsbestände ermöglichen.²⁴

Die Strategie, Medienobjekte durch die Hinzufügung *von* und Ordnung *anhand* erklärender Metainformationen such- und findbar zu machen, ist medienhistorisch keineswegs neu oder spezifisch für den Kontext digitaler Medien. Papierkataloge in Bibliotheken, Museen oder Archiven beruhen in ähnlicher Weise auf der Beschreibung und Ordnung von Dokumenten beispielsweise durch Schlagworte etc. Diese Gemeinsamkeit darf jedoch nicht darüber hinwegtäuschen, dass es entscheidende Unterschiede zwischen diesen Suchtechnologien gibt, die nicht zuletzt auch die Archivpraxis irritieren. Während traditionelle Bibliotheken, Archive etc. die navigierende Suche, d. h. das Stöbern in Regalreihen sowie in Zettelkästen als Modus der Suche privilegieren, erfolgt die Suche in digitalen Datenbeständen zumeist durch die Formulierung deklarativer Suchanfragen. Weder der eigentliche Bestand noch dessen papierne Doublette wird durchlaufen oder durchblättert, vielmehr werden Suchbedingungen als Query formuliert, die automatisch mit dem Gesamtdatenbestand abgeglichen und in eine Ergebnisliste überführt werden. Die Auswahl der Ergebnisse folgt dabei entweder der semantischen Ordnung des Datenbanksystems oder einer algorithmisch berechneten Relevanzordnung, wie dies beispielsweise bei Websuchmaschinen der Fall ist. Derartige Sach- und Relevanzordnungen stehen in einem gewissen Spannungsverhältnis zum archivarischen Provenienzprinzip, d. h. der Ordnung von Archivalien nach deren Entstehungs- und Herkunftszusammenhang. Diese Ordnungsform von Archivgut, so wird gemeinhin argumentiert, ist keineswegs Selbstzweck, sondern bedeutungstragend und interpretationsleitend.²⁵ Kein Dokument steht im Archiv für sich allein. Jedes Dokument ist vielmehr in einen bedeutungstragenden und bedeutungsgebenden Kontext

²⁴ Die Unterscheidung zwischen Information *als* und Information *über* Realität wurde von Luciano Floridi vorgeschlagen, um unterschiedliche Informationsbegriffe voneinander zu unterscheiden; siehe hierzu: LUCIANO FLORIDI: *The Philosophy of Information*, Oxford 2011, S. 30 f.

²⁵ JOHANNES PAPRITZ: *Archivwissenschaft: Archivische Ordnungslehre, Teil 1, Bd. 3*, Marburg 1976, S. 8 ff.

verwoben, der wiederum für die Arbeit mit Archivbeständen von großer Wichtigkeit ist. Die Suchlogik der Query unterläuft diese kontextuelle Ordnung und ersetzt sie durch die semantische Ordnung von Datenmodellen, Taxonomien, Klassifikationen, Ontologien etc. Das digitale Pendant zur Erkundung des Archivs ist nicht die Query, sondern das navigierende Herumstöbern beim Browsen. Gerade diese Zugriffsform tritt immer mehr in den Hintergrund und es bedürfte geeigneter Interfaces, welche es Nutzer_innen erlauben, Archivalien auch bei der Suche mittels Queries im Kontext zu betrachten.

Fokussiert man nicht nur allgemein digitale Archive, sondern digitale *Bildarchive*, kommt eine weitere Herausforderung hinzu. Das Bildhafte entzieht sich nämlich tendenziell genau jener textuellen Beschreibung und propositionalen Logik, welche das Regime der Metadatierung charakterisiert. Natürlich können Bilder sprachlich beschrieben werden, doch kann man sich des Eindrucks kaum erwehren, dass Sprache die Bedeutung von Bildern nur schwer einzufangen vermag. Eben dies kommt in jenem Sprichwort zum Ausdruck, dass zwar nicht aus dem alten China, sondern aus der Feder des gewieften Werbetexters Frederick Barnard stammt: »Ein Bild sagt mehr als tausend Worte.«²⁶ In Anlehnung an den Kunsthistoriker und Bildtheoretiker Ernst H. Gombrich gilt es dieser mittlerweile zum Gemeinplatz verkommenen Aussage jedoch hinzuzufügen, dass ein Bild zwar mehr als tausend Worte, dafür aber keinen einzigen Satz sagt. Denn wie Gombrich herausgestellt hat, kann »ein Bild niemals eine Aussage [...] sein.«²⁷

Das bedeutet keineswegs, dass sich über Bilder nichts sagen ließe. Im Gegenteil: Über Bilder kann viel, wenn nicht sogar zu viel gesagt werden. Ihnen ist ein Überschuss potenzieller Bedeutungen immanent, der durch die Zuweisung weniger Schlagworte nicht eingeholt werden kann. Dann aber, wenn mehr als 1000 Schlagworte mit jedem Bild verbunden würden, ließen sich diese nur noch schwerlich finden. Gerade deshalb hat Wolfgang Ernst bereits vor 15 Jahren dafür plädiert, Verfahren zu entwickeln, mithilfe derer Bilder mit bildhaften Mitteln gesucht werden können.

Jenseits der Verschlagwortung operiert heute die Suche nach Bildern mittels Bildern, die Google seinen Nutzer_innen beispielsweise seit 2011 ermöglicht.²⁸ Derartige bildbasierte Such- und Zugriffsmöglichkeiten auf digitale Bildsammlungen sind bislang jedoch keineswegs perfekt. Dies führt der Medienkünstler Sebastian Schmiege eben-

26 Vgl. WILLIAM SAFIRE: »On Language: Worth a Thousand Words«, in: *New York Times Magazine*, 07.04.1996, <http://www.nytimes.com/1996/04/07/magazine/on-language-worth-a-thousand-words.html> [eingesehen am 19.07.2016].

27 ERNST H. GOMBRICH: *Kunst und Illusion: Zur Psychologie der bildlichen Darstellung*, Berlin 2002, S. 59.

28 Vgl. WOLFGANG ERNST: »Jenseits der Verschlagwortung? Plädoyer für ein nicht-textbasiertes Bildgedächtnis«, 2000, <http://www.westfaelischer-kunstverein.de/archiv/2000/ausstellungen/nicolai/index.html> [eingesehen am 10.05.2007]; WOLFGANG ERNST/STEFAN HEIDENREICH/UTE HOLL (Hg.): *Suchbilder: Visuelle Kultur zwischen Algorithmen und Archiven*, Berlin 2003.

falls seit 2011 in *Search by Image, Recursively* vor Augen. Schmiege erstellt im Rahmen dieses Projekts Videos, welche die Ergebnisse einer rekursiven Bildsuche mit Googles bildbasierter Suche zeigen. Am Anfang steht dabei beispielsweise eine transparente PNG-Bilddatei, die er als Startbedingung an Googles bildbasierte Suche übergeben hat und deren Ergebnis zum Ausgangspunkt einer neuen Suche wurde, deren Ergebnis wiederum als Suchbedingung an Googles bildbasierte Suche übergeben wurde usw. Der vierminütige Film, der mit einem transparenten Bild begann, endet bei dem abstrakten Diagramm einer Straßenkarte.²⁹ Schmiegs wunderbare Filmmontagen führen uns nicht nur die aktuelle Imperfektion der bildbasierten Suche vor Augen, sie zeigen vielmehr auch, dass eine solche Suchtechnologie durchaus überraschende Ergebnisse zu Tage fördern kann. In ihrer künstlerischen Form haben die Experimente mit Googles bildbasierter Suche möglicherweise nur einen eingeschränkten Erkenntniswert. Jedoch ist Ernsts Plädoyer aufgreifend noch immer zu überlegen, ob es nicht gänzlich neue Formen der »Lektüre« des Archivs geben kann, welche auf Grundlage von Digitaltechnologien andere, durchaus neue Wissensformen ermöglichen, wie dies beispielsweise von der Software Studies Initiative rund um Lev Manovich erprobt wird.³⁰

Superarchiv und Metaarchiv

Die Frage nach neuen Zugriffs- und Analyseformen drängt sich nicht nur vor dem Hintergrund der *data deluge* auf, mit der sich (forschende) Nutzer_innen im Kontext jener neuen Archive im weiten, unterterminologischen Sinn konfrontiert sehen, wie z. B. Flickr, YouTube, Instagram, Twitter, Facebook etc. Durch die intensiven Digitalisierungsbemühungen historischer Archive (aber auch von Bibliotheken und Museen) sowie den wachsenden Zufluss von *born digital content* in diese Archive, gewinnt die Frage, wie Nutzer_innen mit den potenziell verfügbaren Daten, Informationen und Dokumenten umgehen und das für sie Relevante finden können, auch in diesem Bereich zunehmend an Dringlichkeit. Zumindest teilweise in Reaktion hierauf wurden in der jüngsten Vergangenheit übergreifende Infrastrukturen und Plattformen imaginiert, konzeptualisiert und realisiert, welche man als Superarchive bezeichnen kann. Gemeint sind hiermit Angebote wie das Prometheus Bildarchiv (www.prometheus-bild-

29 Auf Sebastian Schmiegs Webseite finden sich neben dem beschriebenen Film noch weitere, die beispielsweise mit einem häufig auf Webseiten eingesetzten Spacer-Bild, einem Portrait von Schmiege oder einem Bild der Erde beginnen; <http://sebastianschmiege.com/searchbyimage/> [eingesehen am 05.12.2016].

30 Vgl. exemplarisch LEV MANOVICH: »Auf den Spuren der globalen digitalen Kulturen: Kulturanalytik für Anfänger«, in: KONRAD BECKER / FELIX STALDER (Hg.): *Deep Search: Politik des Suchens jenseits von Google*, Innsbruck et al. 2009, S. 221–236; So YAMAOKA et al.: »Cultural Analytics in Large-Scale Visualization Environments«, in: *Computer*, 44/12 (2011), S. 39–48. Siehe zudem auch die Webseite der Initiative (<http://lab.softwarestudies.com/>) sowie die Webseite zu Manovichs aktuellem Projekt *Selfiecity* (<http://selfiecity.net/>).

archiv.de), Europeana (www.europeana.eu) oder auch die Digital Public Library of America (dp.la).³¹ Digitalisate bzw. digitale Dokumente auf der Webseite eines Archivs, einer Universität, einer Bibliothek, eines Museums etc. einfach online zu stellen, reicht schon lang nicht mehr aus, denn wie sollen Nutzer_innen wissen, dass sie auf den Webseiten einzelner Institutionen potenziell fündig werden können. Da deren historische Bestände in der Informationsflut des WWW und somit auf den Ergebnisseiten von Websuchmaschinen unterzugehen drohen und die Inhalte von solchen und anderen Datenbanken aus technischen Gründen meist überhaupt nicht in den Index der Suchmaschinen gelangen, wurden und werden zentrale Plattformen und Dienste entwickelt, die es Nutzer_innen nicht nur erlauben, einzelne Sammlungen zu durchsuchen, sondern durch die Formulierung einer Suchanfrage in einer Vielzahl von verteilten Sammlungsbeständen potenziell Interessantes finden zu können. Als Superarchiv versprechen und erlauben Europeana, Prometheus und dergleichen den Zugriff auf eine unüberschaubare Zahl von Ressourcen. Ganz im Geist von Paul Otlets und Henri La Fontaines Mundaneum sowie Seite an Seite mit Google oder etwa dem Internet Archive bestärken solche Superarchive die Utopie, den Gesamtwissensbestand der Menschheit digital zugänglich zu machen.³²

Obwohl durch die Implementierung und Institutionalisierung zentralisierter Superarchive das Suchen in dezentralen Archiv- und Bibliotheksbeständen in vielerlei Hinsicht erleichtert wird, gestaltet sich das Finden in einer anderen Hinsicht zugleich schwieriger. Das Glück über die potenzielle Verfügbarkeit der Bestände von tausenden Erinnerungsinstitutionen weicht dabei der Frustration, geeignete Quellen finden zu können. Dies hat zumindest Melissa Terras zum Ausdruck gebracht, als sie schrieb:

Finding decent images that are interesting [...] is a time consuming, utterly frustrating task. I speak from a few months of chuck-my-computer-across-the-room frustration in trying to navigate (mostly unsuccessfully) what the cultural heritage sector has spent millions of pounds putting online.³³

31 Genannt werden könnten aber auch die Deutsche Digitale Bibliothek (www.deutsche-digitale-bibliothek.de), das Archivportal-D (www.archivportal-d.de) oder das Archivportal Europa (www.archives-portaleurope.net).

32 Im Anfang des 20. Jahrhunderts begründeten Mundaneum sollte eine universelle Bibliografie erstellt und das gesamte Wissen der Welt versammelt und katalogisiert werden. Es kann somit als papierner Vorläufer von Google ebenso wie von digitalen Superarchiven betrachtet werden; vgl. LENA CHRISTOLOVA: »Das Mundaneum oder das papierne Internet von Paul Otlet und Henri La Fontaine«, in: STEFAN BÖHME/ROLF F. NOHR/SERJOSCHA WIEMER (Hg.): *Sortieren, Sammeln, Suchen, Spielen: Die Datenbank als mediale Praxis*, Münster 2012, S. 31–54.

33 MELISSA TERRAS: »Reuse of Digitised Content (1): So you want to reuse digital heritage content in a creative context? Good luck with that«, in: *Melissa Terras' Blog*, 2014, <http://melissaterras.blogspot.de/2014/10/reuse-of-digitised-content-1-so-you.html> [eingesehen am 20.03.2015].

Der von Superarchiven wie Europeana geführte Kampf gegen das Verschwinden und Vergessen historischer Dokumente resultiert in einer anderen Form des Vergessens: dem Nicht-Wiederfinden im Datenbestand. Ohne effektive Zugriffsmöglichkeiten fallen die Ressourcen digitaler Archive und Superarchive somit ebenso dem Vergessen anheim. Dies hat Justin Hoffman bereits 1997 in seinem Beitrag für den Katalog der Ausstellung *Deep Storage: Arsenale der Erinnerung* auf den Punkt gebracht, indem er auf die paradoxe Verschränkung von Erinnern und Vergessen im Kontext akkumulativer Sammlungspraktiken hingewiesen hat: »Das Sammeln und das Aufbewahren entwickeln sich scheinbar paradox zu einem Vernichtungsakt.«³⁴

Dem Problem respektive der Gefahr des Vergessens in der Datenbank muss sich die künftige Entwicklung digitaler Archive und Archivpraktiken stellen und zugleich Strategien dafür entwickeln, wie die scheinbare Universalität und Vollständigkeit digitaler Superarchive nicht nur theoretisch dekonstruiert werden kann, sondern wie ihre Leerstellen, blinden Flecken und Lücken auch praktisch aufgezeigt werden können, um die Räume und Lebenswelten jenseits des Digitalen nicht aus dem Blick zu verlieren. Immer größere digitale Sammlungen und Archive anzulegen, ist unter den gegebenen digitaltechnischen Möglichkeiten relativ einfach. Als Herausforderung erweist sich vielmehr das Selektieren und Formieren kleiner Sammlungen aus jenen Superarchiven. Eine metaarchivarische Praxis, die sich diesem Problem widmet, würde es sich nicht zur Aufgabe machen, immer weitere und neue Quellen digital verfügbar zu machen, sondern aus dem Vorhandenen etwas Neues zu schaffen. Die hieraus resultierenden Metaarchive sind in erster Linie nicht additiv, sondern selektiv; sie treffen eine Auswahl aus dem Vorhandenen, ziehen Quellen zusammen, verdichten und verknüpfen sie und erzeugen auf diese Weise Kontexte, vor deren Hintergrund sich mögliche Bedeutungen, Lesarten und Interpretationen des Archivs entfalten. Diese selektive, kuratorische Archivpraxis bei der Archive aus Archiven erzeugt werden, liegt in gewisser Weise quer zur archivarischen Archivpraxis und konturiert sich eher als eine forschende Praxis, die keineswegs als Alternative, sondern als Ergänzung zu den Leistungen digitaler Archive und Superarchive fungieren könnte. Dass kulturwissenschaftliche Forschung neben der Produktion von Texten auch in der Sammlung, Ordnung und Verfügbarmachung von Medienobjekten eine genuine und wichtige Aufgabe erkennen kann, hat Moritz Baßler im Rahmen seiner an den *New Historicism* anschließenden Kulturpoetik dargelegt.³⁵

34 JUSTIN HOFFMANN: »Die Angst vor dem Verschwinden«, in: *Deep Storage: Arsenale der Erinnerung; Sammeln, Speichern, Archivieren in der Kunst*, Ausst.-Kat. hg. v. INGRID SCHAFFNER et al., München 1997, S. 269–272 (hier: S. 272).

35 Vgl. MORITZ BAßLER: *Die kulturpoetische Funktion und das Archiv: Eine literaturwissenschaftliche Text-Kontext-Theorie*, Tübingen 2005.

Beispiele für die Herausbildung solcher metaarchivarischer Praxen finden sich im Kontext der maßgeblich am Center for History and New Media entwickelten Software Omeka (omeka.org). Dabei handelt es sich um eine niederschwellige Archivsoftware, die zum Experimentieren mit alternativen Archivpraxen einlädt, wie z. B. den Aufbau kollaborativer Archive, die Präsentation von Archivbeständen in Online-Ausstellungen oder aber auch metaarchivarische Praxen. Ähnlich einfach zu installieren und zu bedienen wie die beliebte Blogging-Software Wordpress ermöglicht es Omeka auch Amateur-Archivar_innen relativ leicht Sammlungen anzulegen, die akkumulierten Medienobjekte mit Metadaten auszuzeichnen und aus den Sammlungen virtuelle Ausstellungen zu erzeugen. Mittels einer standardisierten Schnittstelle – der sogenannten OAI-PMH-Schnittstelle³⁶ – können weiterhin Datenbestände aus anderen Archiven ausgelesen (*geharvested*) und automatisch in das System integriert werden. Somit besteht die Möglichkeit, Metaarchive zu bilden, neue Metadaten zu erzeugen und Wissen zu produzieren, welches in einem weiteren Schritt selbst wieder in die ursprünglichen (Super-)Archive zurückfließen und deren Teil werden könnte.

Die metaarchivarische Geste, welche in der Bildung von Archiven aus Archiven zum Vorschein kommt, kann sich im Zeitalter digital vernetzter Medien noch auf eine andere Weise artikulieren. Wie eingangs dargestellt, sind im digitalen Zeitalter neben die historischen Archive eine Reihe neuer Archive getreten, in denen sich, wenn nicht mehr, so doch Anderes finden lässt. Flickr, YouTube, Twitter, Instagram, Facebook etc. sind scheinbar unerschöpfliche Quellen für die Erforschung von Kulturen und Gesellschaften.³⁷ Obwohl diese Plattformen riesige Datensilos sind, die als Archiv zu bezeichnen in mancherlei Hinsicht nahe liegt, widerstrebt ihre auf Popularität, Aktualität und Ephemeralität angelegte Funktions- und Ordnungslogik vielen Prinzipien konventioneller historischer Archive. Durch Techniken des *Scraping* ist es möglich, Inhalte von diesen Plattformen zu selektieren, zu deplatzierten und zu verdichten, so dass auch auf diese Weise Metaarchive des *Megaarchivs* Internet entstehen können.³⁸ Um diese zu erstellen und auszuwerten, bedarf es langfristig sicherlich nicht nur technischer Lösungen, sondern, wie dargelegt wurde, vor allem neuer Fragen, Ansätze und Methoden, die sich die Potenziale der Digitaltechnologien zu Nutze machen.

36 Das *Protocol for Metadata Harvesting* der *Open Archives Initiative* (OAI-PMH) stellt im Bereich digitaler Archive und Bibliotheken den *de facto* Standard für die Distribution maschinenlesbarer Meta- und Primärdaten dar, die auf diese Weise aggregiert werden können; vgl. <https://www.openarchives.org/pmh/> [eingesehen am 20.10.2015].

37 Vgl. MANOVICH: »Auf den Spuren der globalen digitalen Kulturen« (Anm. 30).

38 Zur Praxis des Sammelns als einer deplatzierenden Verdichtung von Dingen vgl. MANFRED SOMMER: *Sammeln: ein philosophischer Versuch*, Frankfurt a. M. 2002.

DAS FOTOGRAFISCHE DISPOSITIV

BAND 2

(POST)FOTOGRAFISCHES ARCHIVIEREN. WANDEL MACHT GESCHICHTE

Victoria von Flemming

Daniel Berndt

Yvonne Bialek

(Hg.)

Marcelina Kwiatkowski

(Mitarbeit)

JONAS VERLAG