



Digitale Archive

Warnke, Martin

Published in:
Archivprozesse

Publication date:
2002

Document Version
Verlags-PDF (auch: Version of Record)

[Link to publication](#)

Citation for published version (APA):
Warnke, M. (2002). Digitale Archive. In H. Pompe, & L. Scholz (Hrsg.), *Archivprozesse: die Kommunikation der Aufbewahrung* (S. 269-281). (Mediologie; Band 5). DuMont Literatur und Kunst Verlag. <http://nbn-resolving.de/urn:nbn:de:hbz:38-23722>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**ARCHIVPROZESSE:
DIE KOMMUNIKATION DER AUFBEWAHRUNG**

Mediologie

Band 5

Eine Schriftenreihe des Kulturwissenschaftlichen Forschungskollegs

»Medien und kulturelle Kommunikation«

Herausgegeben von Ludwig Jäger

ARCHIVPROZESSE: DIE KOMMUNIKATION DER AUFBEWAHRUNG

**Herausgegeben von
Hedwig Pompe und
Leander Scholz**

DuMont

Diese Publikation ist im Sonderforschungsbereich/Kulturwissenschaftlichen
Forschungskolleg 427 »Medien und kulturelle Kommunikation«, Köln, entstanden
und wurde auf seine Veranlassung unter Verwendung der ihm von der Deutschen
Forschungsgemeinschaft zur Verfügung gestellten Mittel gedruckt.

Erste Auflage 2002

© 2002 DuMont Literatur und Kunst Verlag, Köln

Alle Rechte vorbehalten

Ausstattung und Umschlag: Groothuis, Lohfert, Consorten

Gesetzt aus der DTL Documenta und der DIN Mittelschrift

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Satz: Greiner & Reichel, Köln

Druck und Verarbeitung: B.o.s.s Druck und Medien GmbH, Kleve

Printed in Germany

ISBN 3-8321-6005-1

INHALTSVERZEICHNIS

Hedwig Pompe/Leander Scholz
Vorbemerkung der Herausgeber 9

Jürgen Fohrmann
»Archivprozesse« oder Über den Umgang mit der Erforschung
von ›Archiv‹. Einleitung 19

I. PRODUKTIVITÄT (UND MACHT)

Klaus Militzer
Entstehung und Bildung von Archiven in Köln
während des Mittelalters 27

Helmut Zedelmaier
Buch, Exzerpt, Zettelschrank, Zettelkasten 38

Wolfgang Ernst
ROM 1881: Die Medialität des Vatikanischen Geheimarchivs
als Gesetz der Historie 54

Urs Stäheli
Die Wiederholbarkeit des Populären:
Archivierung und das Populäre 73

Heike Behrend
Flugapparate aus dem Jenseits.
Geisterarchive im interkulturellen Vergleich 84

Michael Thompson
Benji the Binman and his Anti-archive 100

II. UNIVERSALITÄT (UND REST)

Dagmar Börner-Klein

Assoziation mit System: Der Talmud,
die »andere« Enzyklopädie

113

Hedwig Pompe

Botenstoffe – Zeitung, Archiv, Umlauf

121

Friedrich Balke

Die Enzyklopädie als Archiv des Wissens

155

Markus Krajewski

Restlosigkeit: Wilhelm Ostwalds Welt-Bildungen

173

Dörte Wittig

Ein Medium für das Wissens- und Content-Management
von akademischen Gemeinschaften

186

III. TRANSFORMATION (UND VERLUST)

Michael Fehr

Zur Bedeutung des Ankaufs der Briefe aus dem Nachlass
Adalbert Colsmann 209

Eckhard Siepmann

Der performative turn erreicht das Museum 226

Joachim W. Schmidt/Hans-Werner Sehring/Martin Warnke

Der Bildindex zur Politischen Ikonographie in der Warburg
Electronic Library. Einsichten eines interdisziplinären Projektes 238

Martin Warnke

Digitale Archive 269

Erika Linz

»The warehouse theory of memory is wrong« –
Zur Performativität semantischer Wissensstrukturen 282

Hartmut Winkler

Das Modell. Diskurse, Aufschreibesysteme, Technik, Monumente –
Entwurf für eine Theorie kultureller Kontinuierung 297

Autorenverzeichnis

316

Bildnachweise

320

Martin Warnke
DIGITALE ARCHIVE

Archive haben Konjunktur. Dass dies so ist, liegt sicher auch daran, dass mit Hilfe digitaler Speicher- und Distributionstechniken – CD-ROM, DVD, das Internet, speziell das World Wide Web – umfangreiche Datensammlungen auch in die Reichweite von Privatleuten oder kleineren Institutionen gelangt sind. Denn einen PC, einen Scanner, eine große Festplatte und einen CD-Brenner kann sich heutzutage so ziemlich jede und jeder leisten. Warum also nicht endlich einmal alles aufheben, was wertvoll zu sein scheint?

Ob dabei Computertechnik tatsächlich die Lösung des archivarischen Problems, ob sie vielleicht sogar alte Probleme in neuem Ausmaß und darüber hinaus ganz neue erzeugt und welches das Verhältnis von digitaler Speicherung zu Gedächtnis und Erinnerung ist, soll im Folgenden zur Sprache kommen.

**1. DIGITALE RESURREKTION ODER BEERDIGUNG ERSTER KLASSE:
 DIGITALE ARCHIVE UNTERGEGANGENER KULTUREN**

Vordringlichste und vornehmste Aufgabe aller Archivierung ist es, so die landläufige Auffassung, das zu Archivierende vor dem Verfall und dem Vergessen, aber auch vor unautorisiertem Zugriff zu schützen. Es handelt sich bei den Archivalien also um Güter, um Bestandteile der Kultur, die nicht mehr lebendig genug sind – hier durchaus buchstäblich *und* figurativ gemeint –, um selbst für ihren Erhalt oder ihren Schutz sorgen zu können.

Archive bewahren Totes, haben demnach viel mit Mausoleen gemein. Denn wie auch Mausoleen bewahren sie das Gedenken und die Erinnerung an Totes am Leben, sie sind Mahnmale für die Macht, die die Toten noch immer besitzen.

Berühmte Beispiele kennen wir alle. Da wäre etwa das Taj Mahal, das monumental die Macht der Liebe über den Tod hinaus symbolisiert, oder der ausgestopfte Lenin am Roten Platz in Moskau, bei dem Konservierungsspezialisten immer wieder Hand anlegen müssen, um die Folgen des Verfalls auszubessern.

Hier soll es im Folgenden um Dokument-Archive gehen, die allerdings ihre sepulkrale Natur ebenfalls nicht verleugnen können.

Beispiele sind etwa: Die DDR im WWW¹, eine der ersten Fundstellen, wenn man in der Suchmaschine Google den Begriff »Archiv« eingibt.

Ob es sich dabei tatsächlich um ein Archiv handeln kann, ob vielleicht eher mit Mitteln der Informationstechnologie ein WWW-Mausoleum implementiert werden soll, scheint mir noch nicht ausgemacht.

Zweites Beispiel: Mit dem unverkennbaren Gestus einer schuldbewussten Annäherung an begangene Verbrechen liegen die zerstörten jüdischen Sakralbauten in Frankfurt am Main auf dem Web.² Sie sind die Stars eines Artikels in der Süddeutschen Zeitung vom 5.9.2000. Er trägt den Untertitel: »Das Internet entwickelt sich zur Ausstellungsstätte für Kostbarkeiten, die längst zerstört oder Experten vorbehalten sind«. Es scheint hierfür keinen großen Unterschied zu machen, ob die Kostbarkeiten zerstört oder in Expertenhand sind, sie sind in jedem Falle entzogen. Dem Internet wird nun zugetraut, Aufschluss-Stätte für solcherart Kostbarkeiten zu sein.

In der Süddeutschen heißt es weiter: »Nach den Überlieferungstechniken Sprache und Schrift steht in der Form der digitalen Information ein Medium zur Verfügung, das den Zugang zu und die Präsentation von Kulturgütern – dem Gedächtnis der Menschheit – revolutionieren [wird].«

Ob das – aus technischer Sicht – vielleicht zu viel der Ehre ist, soll sich im Folgenden erweisen. Denn es ließe sich dann, nach Enttäuschung technischer Heilserwartungen, auch mutmaßen, bei der Idee des Zugänglich-Machens von Kulturgütern mit digitalen Medien sei vielleicht ein ganz anderer Wunsch Vater des Gedankens gewesen: Möglicherweise haben digitale Archive ja sogar die Funktion, den Deckel über dem Material endlich zu schließen, für eine letzte Ruhestätte zu sorgen, an der auch die Erinnerung letztlich an ihr Ende kommen kann. Derrida hat das das Archivübel genannt: *Mal d'Archive*.³

Was haben wir, könnte die Frage sein, wenn der virtuelle Deckel des Archivs über dem Bewahrten zuklappt: digitales ewiges Leben oder letzte Daten-Ruhestätte?

Dieser Frage werden wir jetzt nachgehen.

2. EIN DIGITALES ARCHIV VON ÜBERHAUPT ALLEM

Kommen wir zunächst zu Aspekten der reinen Quantität, zu technischen Randbedingungen digitaler Speicherung als Teil einer Archivierung: Was ist davon zu halten, einmal grundsätzlich überhaupt alles archivieren zu wollen?

Fangen wir klein an, mit dem Gedächtnis eines Menschen etwa. Ein gewisser kognitiver Psychologe mit dem Namen Landauer aus Colorado hat geschätzt, dass der Informationsgehalt eines typischen menschlichen Langzeitgedächtnis-

ses zwischen 150 und 225 MB umfasst.⁴ Das ist immerhin der Text von etwa ein paar hundert Paperbacks. Ein gängiger PC müsste nur etwa ein Zehntel seines Speichers opfern, um das Gedächtnis einer typischen Kleinfamilie abzuspeichern. Sofern dieses Bild überhaupt stimmt, wohlgemerkt, sofern also Erinnerung und Speicher vergleichbar wären. Rechnen wir diesen Wert auf die gesamte Menschheit hoch, so ergeben sich 1.350 PetaByte, also kein Problem, wie wir gleich sehen werden. Auch was ein PetaByte ist, wird gleich geklärt.

Nehmen wir uns als Nächstes vor, was, anders als ein menschliches Gedächtnis, ohnehin schon als Datenmasse externalisiert ist: die Texte, Bilder und Töne, die die Menschheit über technische Medien ständig absondert.

Sehen wir einmal nach.

Die Library of Congress enthält⁵

20 Millionen Bücher, macht 20 TeraByte (Tera ist 10^{12} , eine Million Millionen).

13 Millionen Fotos, macht 13 TB.

4 Millionen Karten und Pläne, macht 200 TB,

500.000 Filme zu etwa 500 TB,

3,5 Millionen Klangdokumente: 2000 TB.

Zusammen ergibt das etwa 3000 TB oder 3 PetaByte, Peta heißt 10^{15} , eine Milliarde Millionen.

Nehmen wir noch mehr hinzu: alles Schriftliche, alle Fotos, alle neu ausgestrahlten Fernsehsendungen, der Hörfunk, alle publizierte Musik und vor allem: alle geführten Telefonate, die einige spezielle staatliche Stellen sowieso gern aufzeichnen würden, machen auf der ganzen Erde pro Jahr ca. 4.600 PB oder 4,6 ExaByte, Exa: 10^{18} , eine Milliarde Milliarden. Eine ganze Menge. Aber: Das ist weniger als der jährlich von der Industrie gefertigte und verkaufte elektronische Speicher.

Mithin: wenn wir wollten, könnten wir alles speichern. Wenngleich kritische Stimmen angesichts der enormen menschlichen Mitteilbarkeit, vor allem am Telefon, fragen: »So who wins the war here – a handful of cybrarian archivists, or the entire chattering human race?«⁶

3. DAS INTERNET ALS ARCHIV

Doch man muss gar nicht zu den Sternen, also zum gesamten Weltgeist, greifen, um auf erhebliche Probleme zu stoßen, nehmen wir uns vor, was ohnehin digital da und weltweit verfügbar ist: das Internet, speziell das World Wide Web.

Ist es selbst vielleicht schon ein digitales Archiv? Es liegt ja bitweise zugreifbar vor.

Technisch umfasst das Internet die Gesamtheit aller Datenleitungen, aller Computer, die Relais-Funktionen übernehmen, aller Datenpakete, die transportiert werden, und aller Server und Clients, die Informationen anbieten und abfordern.

Wären Internet-Pakete solche wie bei der Paketpost, so würden sie, zusammengekommen, den Bestand des Internet ausmachen. Doch: Internet-Pakete sind vergänglich. Damit alle die Daten, die zu übertragen sind und die gelegentlich verschiedene Wege ausprobieren müssen, um dann schließlich hoffentlich beim Empfänger anzugelangen, damit diese Daten nicht sinnlos herumliegen und alles verstopfen, hat man ihnen einen Selbstzerstörungs-Mechanismus eingebaut, das TTL-Feld, das heißt »Time To Live«. Normalerweise trägt diese digitale Lebenserwartung anfangs den Wert 255, und bei jeder Passage über eine Computer-Relais-Station, bei jedem *hop*, wird dieser Wert um eins dekrementiert. Ist er null, wird das Paket gelöscht.⁷

Das Internet ist offenbar eine Mischung aus reinem Transport und temporärer Speicherung, die bei den Paketen nie länger als ein paar Handvoll Sekunden dauert. Es kommt für die Frage nach der Lebensdauer von Internet-Dokumenten also auf die Endgeräte an, denn: die Übertragungspakete verschwinden von selbst. Was nicht mehr auf den Servern liegt, kann nicht mehr erreicht werden, es tritt der berühmte »Error 404, document not found« auf.

Das Netz selbst ist also offenbar als Archiv *untauglich*.

Lassen Sie uns nachsehen, wie es um die Daten auf den Servern selbst bestellt ist.

Ein Blick auf die Statistik unseres WWW-Servers an der Universität Lüneburg besagt, dass rund 2,5 % aller Anforderungen an Web-Seiten nicht bedient werden konnten, weil es sie nicht gab, aber wahrscheinlich schon einmal gegeben hat. Sie sind also seit der letzten Bereinigung der Verweise auf sie verschwunden. Gelöscht. Nicht mehr für wichtig oder erhaltenswert befunden, vielleicht durch Neueres ersetzt.

Das wirft die Frage auf, wie hoch die Lebenserwartung einer WWW-Seite ist, bis sie von den Servern verschwunden ist. Tage, Monate, Jahre?

Brewster Kahle weiß, wie viel Websites stehen, und er weiß auch, wie lange sie am Himmelszelt des Internet glänzen: noch Ende 1996 sollten es 400.000 Websites mit insgesamt 1,5 TeraByte Umfang sein, im März 2000 waren es 13,8 TB⁸, also wie vorhersehbar etwa 10-mal so viele. Und wie lange bleibt so ein Dokument typischerweise auf seinem Server erreichbar? Auch hier weiß Brewster

Kahle Antwort. Seine Verfahren registrieren natürlich, wenn sich etwas auf einem Server verändert, wenn Daten verschwinden. Die Veränderungsrate der Dateien 1996 von 600 GB im Monat lässt den Schluss⁹ zu, dass die mittlere Lebenserwartung eines Dokuments im WWW ganze 75 Tage beträgt. Nach 75 Tagen sind die meisten Seiten also nicht mehr dort, wo sie einmal vorzufinden waren. Referenzen, Zitate, Bezüge auf sie liefern dann den »Error 404, document not found«. Zweieinhalb Monate sind wahrhaftig nicht viel, vergleicht man das mit der Langlebigkeit analoger Medien, etwa den fünftausend Jahre alten frühesten Texten der Menschheit, den Proto-Keilschrift-Tafeln aus Mesopotamien.¹⁰ Selbst Bücher auf säurehaltigem Papier kommen gut dagegen weg.

Was heißt das für die Inkunablen des Internet, die »Wiegendrucke«, Erstlinge, Erstauflagen? Diese Frühzeit des WWW ist verloren. Es gibt die Dokumente nicht mehr, weil niemand sie bewahrt oder archiviert hat. Ständig verschwinden Dokumente im digitalen Nirwana, und niemand scheint das aufhalten zu können.

Fast niemand. Brewster Kahle versucht es. Seine Organisation heißt »The Internet Archive«, und sie macht Schnappschüsse des gesamten WWW, weshalb er auch die Kenndaten des WWW besitzt. In Form einer Skulptur ist ein solcher seiner Schnappschüsse der Library of Congress als Geschenk übergeben worden. Diese Skulptur ist eine digitale Momentaufnahme des Web von 1997, etwa 2 Tera-Byte Daten. Man kann den Internet-Archiv-Service auch als normaler User in Anspruch nehmen: unter dem Namen »Alexa« gibt es einen Browser-Zusatz, der dann, wenn der Fehler 404 zuschlägt, stattdessen die Kopie aus seinem Archiv anzeigt. Besonders ergiebig ist er übrigens nicht, oft gibt es dennoch Seiten mit dem Fehler 404, die auch Alexa schon vergessen hat.

Hinzu kommt: Vieles im Web entsteht erst bei der Abfrage, die Bahn- oder die Telefonauskunft oder aktuelle Preis- und Produktlisten, weil die Seiten dynamisch zur Laufzeit erst aus Datenbanken erzeugt werden. Diese sind den Netz-Kriechtieren, den Suchrobotern von Brewster Kahle sowieso unzugänglich und können von ihm nicht archiviert werden.

Fazit: Das Internet taugt bei einer Dokument-Lebensdauer von zweieinhalb Monaten nicht zum Archiv. Es zu archivieren ist ein heroischer Akt, der die Frühzeit des Web ohnehin nicht mehr retten kann. Und auch die meisten der zeitgenössischen Websites lassen sich so nicht dem Vergessen entreißen, sie gehen ständig unwiederbringlich verloren.

4. LEBENSDAUER DIGITALER SPEICHERMEDIEN

Nehmen wir einmal an, alles, was archiviert werden sollte, wäre tatsächlich schon von irgendeinem *cybrarian* digital gespeichert. Sind die Kisten voller Bänder, Platten und CDs dann ein sanftes Ruhekissen für den guten Menschen, dem der Erhalt der Kulturgüter so am Herzen liegt?

Lassen wir Augenzeugen berichten:

»In Taiwan habe ich Disketten gesehen, die voller Pilze und Schimmel waren (grün und haarig). In Missouri habe ich Disketten-Hüllen gesehen, die von der Hitze im Inneren eines Autos völlig verzogen waren. In Utah habe ich Disketten voller Flugsand gesehen. Der Besitzer sagte mir, es habe ein seltsames Kratzgeräusch gegeben, als sie seinem Laufwerk den Lesekopf zu Schrott geschliffen haben.«¹¹

Aber selbst vorsichtigeren Zeitgenossen, ja selbst den sagenhaften Raketenwissenschaftlern bei der NASA, mit Technologie doch auf Du und Du, ist schon Schlimmes widerfahren:

»Der Inhalt von 1,2 Millionen Magnetbändern, die drei Jahrzehnte amerikanische Raumfahrt dokumentieren, ist hinüber, so schreibt Dr. Michael Friedewald vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung in Karlsruhe.« Die Aufbewahrung in Lagerhäusern hat den wertvollen Tapes nicht gut getan. Die Trägerfolie löst sich auf, die Bänder zersetzen sich.¹² Bei Audio-Bändern kennt man das als Quietschen beim Abspielen, als ein Zeichen dafür, dass bald alles vorbei sein wird.

Wie lange Medien halten, hängt davon ab, wie sorgsam sie aufbewahrt werden. Aber auch bei größter Vor- und Umsicht ist ihnen nur eine gewisse Spanne beschieden. Etwa so:

Bänder halten zwischen 2 und 30 Jahren.

CD-ROM (die silbrigen, industriell gefertigten) 5 bis 100 Jahre, ebenso wie magneto-optische Platten. Ähnliches kann von anderen Medien gesagt werden, die nur ein Mal beschrieben werden (WORM).

Kurz und knapp: »Computerbänder, Videobänder und Tonbänder halten ungefähr so lange wie ein Chevy oder ein Pudel.«¹³ Dasselbe kann von den digitalen Medien behauptet werden. Dann schlägt *data rot* zu, die kalte Datenrotte.

Doch das ist noch nicht einmal das Schlimmste, denn nicht nur die Träger altern, auch die Formate kommen in die Jahre und sterben einfach aus. Haben Sie noch ein 5¼-Zoll-Laufwerk? Wo kann ich meine Lochkarten von vor zwanzig Jahren einlesen lassen? Kennt jemand noch das Format eines *Schneider*-Schreibcomputers?

Schätzt man somit die Lebensdauer von Datenformaten, so landet man bei noch sehr viel niedrigeren Werten, die bei fünf bis zehn Jahren liegen. Danach hilft nur noch ein Computermuseum mit geschickten Technikerinnen oder Technikern, die ohne Ersatzteile, die die Industrie natürlich nicht mehr liefern kann, durch Basteln die alten Gerätschaften am Laufen halten.

Glücklicherweise ist *retro computing* zum Sport einiger Unverzagter geworden, die Spaß daran haben, mit Computern zu spielen, in die man noch hineinblicken kann.¹⁴

Eine Interessengruppe namens »Dead Media« bringt ihre »Dead Media Working Notes« heraus,¹⁵ in denen natürlich nur von toten Medien die Rede ist, etwa der pneumatischen Post in Paris, Hummels Telediagrammen, den »Peek-a-Boo-Index-Cards«, aber auch den obsoleten Computern, die nach Gordon Moores Gesetz alle eineinhalb Jahre durch ihre schnelleren Nachfolger ersetzt werden. Dort, wo tatsächlich alte Formate und alte Hardware gebraucht werden, hilft nur noch eines: Man muss die alten Programme auf neuem Gerät simulieren, oder, im Informatik-Jargon: emulieren.

Dead-Media-Aktivist Bruce Sterling merkt an: »Diese Entwicklung ist für Dead Media Studies interessant, weil die rasche Folge, durch die elektronische Komponenten obsolet werden, immer ein Kainsmal der elektronischen Medien war. Simulation und Emulation toter Hardware wird weiter an Bedeutung zunehmen, solange der Friedhof toter Multimedien nur so wimmelt von Opfern des Mooreschen Gesetzes.«¹⁶

So sieht sich die amerikanische Air Force auch gezwungen, spezielle Vorsorge zu treffen, um beim Generationswechsel elektronischer Schaltungen, mit denen die modernen Kampffjets bestückt sind, nicht auch gleich neue Flugzeuge bauen zu müssen. Denn die Hardware und die darauf implementierte Software muss einwandfrei laufen, damit der Vogel am Himmel bleibt, auch wenn die Chips schon längst nicht mehr hergestellt werden. Also gibt es Ersatzteil-Probleme.

»Die Teileknappheit rührt größtenteils von der kurzen kommerziellen Lebensspanne digitaler elektronischer Komponenten, verglichen mit dem langen Wartungsleben von Waffensystemen. Eine digitale Komponente z. B. mag eine Lebenszeit von 18 Monaten haben, während ein Waffensystem, das diese Komponente verwendet, oft Jahrzehnte im Einsatz ist.«¹⁷ Das wird sehr teuer, und es führt uns wieder zurück zu unserem eigentlichen Thema, den digitalen Archiven, die ohne größte Anstrengungen auf dem Feld einer aufwendigen Daten-Archäologie sehr schnell digitalem Vergessen anheim fallen.

5. DATEN-ARCHÄOLOGIE

Jeff Rothenburg ist durch ein Diktum bekannt geworden, das da lautet: »Digital documents last forever – or five years, whichever comes first.«¹⁸

Und er weiß, wovon er spricht, denn sein in die Ewigkeit, das heißt in die nächsten fünf Jahre, greifendes Urteil ist Frucht einer ausführlichen und sehr überzeugenden Studie zum Thema digitaler Dokumentarchivierung. Das Resultat lautet: »[...] there is – at present, no way to guarantee the preservation of digital information.«

Wenngleich Garantien nicht abzugeben sind, so gibt es doch eine Strategie, die, wenn verfolgt, Hilfe verspricht: »The best way to satisfy the criteria for a solution is to run the original software under emulation on future computers.«

Wer hätte das gedacht? Die Originalsoftware unter einer Emulation hoffnungslos veralteter Betriebssysteme längst verrotteter Hardware muss immer wieder zum Laufen gebracht werden.

Zunächst ist also der Bit-Strom der digitalen Daten zu erhalten, umzukopieren auf je neue Speichermedien, als Maßnahme gegen den *data rot*, zu ergreifen etwa alle ein bis zwei Jahre. Anschließend hat man dafür zu sorgen, dass die Daten auch korrekt interpretiert werden. Wenn man nicht weiß, wie der Inhalt eines Mediums zu interpretieren ist, ist man noch nicht viel weiter. Neben dem *data rot* war auch fehlende Beschriftung eine der Ursachen für die massiven Datenverluste der NASA.

Metadaten sind anzubringen. Sie beschreiben, was wie zu interpretieren ist. Die digitalen Schubladen sind zu beschriften, digital, versteht sich. Die Metadaten müssen auch beschreiben, unter welchem Betriebssystem, auf welcher Hardware die Software läuft, die die Daten interpretiert. Rothenburg schreibt dazu: »This point cannot be overstated: in a very real sense, digital documents exist only by virtue of software that understands how to access and display them; they come into existence only by virtue of running this software.«

Und damit man sich nicht auf Computer-Oldtimer verlassen muss, die irgendwann trotz liebevollster Pflege auseinander fallen müssen, muss das Ganze auf jeweils neuen Rechnern funktionieren, die wegen ihres eigenen, neuen Betriebssystems dann also das alte emulieren müssen, unter dem die originale Software abläuft, die man ebenfalls beilegen muss. Anzulegen ist also auch ein Archiv von Betriebssystem-Emulationen in allen relevanten Versionen und eine Sammlung von Software, die die Dokumente interpretieren kann.

Ständig frisch umkopierte Dokumente könnten so auf neuesten Computern unter Betriebssystem-Emulationen von Originalsoftware angezeigt, mithin ar-

chiviert, werden. Nur dann, wenn wir diesen Aufwand treiben, werden digitale Dokumente archiv-fähig. Man sieht ein, das ist nichts mehr für Privatleute, hier sind staatliche Institutionen gefragt, die eine solche außerordentlich aufwendige Arbeit kontinuierlich leisten. Der Archiv-Begriff, der ja auf das Amtshaus des Archonten zurückgeht, der die Macht über die Regierungsdokumente ausübt, zeigt seine ursprüngliche Bedeutung.

Derrida schreibt in *Mal d'Archive*:

»... archive, sein einziger Sinn, vom griechischen *archeion*: zuerst ein Haus, ein Wohnsitz, eine Adresse, die Wohnung der höheren Magistratsangehörigen, die *archontes*, diejenigen, die geboten. Jenen Bürgern, die auf diese Weise politische Macht innehatten und bedeuteten, erkannte man das Recht zu, das Gesetz geltend zu machen oder darzustellen. Ihrer so öffentlich anerkannten Autorität wegen deponierte man zu jener Zeit bei ihnen zu Hause, an eben jenem Ort, der ihr Haus ist (ein privates Haus, Haus der Familie oder Diensthaus), die offiziellen Dokumente. Die Archonten sind zunächst Bewahrer. Sie stellen nicht nur die physische Sicherheit des Depots und des Trägers sicher. Man erkennt ihnen auch das Recht und die Kompetenz der Auslegung zu. Sie haben die Macht, die Archive zu interpretieren.«¹⁹

Und weiter:

»... die technische Struktur des *archivierenden* Archivs bestimmt auch die Struktur des *archivierbaren* Inhalts schon in seiner Entstehung und in seiner Beziehung zur Zukunft. Die Archivierung bringt das Ereignis im gleichen Maße hervor, wie sie es aufzeichnet. Das ist auch unsere politische Erfahrung mit den sogenannten Informationsmedien.«²⁰

Insbesondere wird das die Erfahrung mit digitalen Archiven sein. Nur Macht und Geld können sie vor dem Verfall retten, die so viel anfälliger sind als ihre analogen Vorläufer. Und wer die überkommenen Dokumente so unter seiner Ägide hat, kann sie nach Belieben einsetzen, interpretieren, vorenthalten.

Sehr real vernichtet der technische Fortschritt, der unabdingbar ist, um immer mehr Dokumente in digitale Archive einstellen zu können, das Archiv selbst: Digitale Archive sind Schauplätze eines *Mal d'Archive*, eines digitalen Archiv-Übels.

Ein wenig kann man das technische Problem entschärfen, indem man möglichst digitale Dokumente in solchen Formaten abspeichert, dass wenigstens der Inhalt noch lange richtig interpretierbar sein wird, dass also schon der Erhalt des Bit-Stroms der Daten Wesentliches rettet. Kandidaten für solche Formate sind etwa »html«, das Format des jetzigen World Wide Web, das allein schon deshalb nicht so einfach von der Bildfläche verschwinden wird, weil es so viele Doku-

mente gibt (13,8 TeraByte im März 2000), die mit dieser Konvention kodiert wurden.

Bei den Metadaten, die Identifizierung und Suche erleichtern, gibt es ebenfalls gute Vorschläge: *Dublin Core* und die *Open Archive Initiative*.²¹ Diese Konventionen könnten, wenn sie weite Verbreitung fänden, die Suche und das Auffinden digitaler Dokumente erheblich vereinfachen. Die Verwendung solcher Standards macht Dokumentarchivierung damit noch nicht zu einem Kinderspiel, aber doch wenigstens zu einem, an dem auch akademische Institutionen teilnehmen können.

6. DAS GENOM ALS ARCHIV

Gestatten Sie mir einen kurzen Exkurs in die Humangenetik, denn wenn es um große digitale Dokumente geht, um Archive des Lebens gar, muss einem auch das menschliche Genom in den Sinn kommen.

Das Genom ist zweifellos der Träger digital kodierter Information. Das Alphabet des Kodes besteht aus diskreten Zeichen, die in den Basen Adenin, Cytosin, Guanin und Thymin ihren materiellen Träger gefunden haben. Die Länge der Zeichenkette entspricht etwa 5000 Buch-Bänden, einer recht stattlichen Privatbibliothek, die wir in jeder unserer Zellen mit uns herumtragen.²²

Doch natürlich stimmt die Text- und damit auch die Buch-Metapher nicht.

Denn das meiste, was in der Zeichenkette steht, ist Schrott, Datenmüll. Nur ca. 3 % aller Zeichen kodieren die Gene, unsere Erbanlagen, auf die es eigentlich ankommt. Beim Umkopieren der Daten auf jeweils frische Datenträger, unsere Kinder, – einem Vorgang, der, wie wir wissen, mindestens zum Erhalt einer digitalen Informationssammlung erforderlich ist –, bei diesem Umkopieren verändert sich das Genom, und es schleichen sich Fehler ein, die evolutionären Fortschritt, aber auch Krankheit bedeuten können. Das Umkopieren geschieht auch bei jeder Zell-Neubildung in einem Organismus, wobei es einem dem *data rot* entsprechenden Prozess gibt: Mit zunehmendem Alter und unter ungünstigen äußeren Bedingungen erhöht sich die Zahl der Kopierfehler, und was dabei entstehen kann, heißt bei Mensch und Tier: *Krebs*, eine Krankheit zum Tode. Will man den Vergleich mit dem Internet wagen, dann wären das Zellen mit einer TTL, einer TimeToLive, von unendlich: eben eine bösartige Wucherung.

Ein weiterer Grund, warum die Schrift-Metapher in die Irre führt, liegt darin, dass die Zeichen und ihre Anordnung zwar durchaus für den Menschen lesbar sind, nämlich für die Ribosomen, die Erzeuger der Proteine, für die das Genom

die Bauanleitung ist, aber durchaus *nicht* für das menschliche Bewusstsein, das doch ansonsten für Textinterpretation zuständig ist. Die vier Basen und ihre Kombinationen bilden keine Symbolschrift, wie wir sie aus Texten gewohnt sind. Sie wirken nur durch das Leben und Sterben selbst, nicht über Symbol-Interpretation. Denn, was vor allem dazu fehlt, das sind die *Metadaten*: nirgends steht geschrieben, welche Bedeutung einzelne Abschnitte der DNA haben und wie sie zu interpretieren sind. Wir stehen im Moment, der ja als der gefeiert wird, zu dem wir 99 % der DNA nachbuchstabieren können, vor der Situation, dass wir einen Dokumenttext ohne Dokumentation haben, der unter unbekannter Software auf einer nur schlecht bekannten Hardware läuft, um einmal eine andere, sicher auch sehr schlechte Metapher zu wählen.

Jedenfalls sind wir weit davon entfernt, den Bit-Strom der Daten etwa *verstehen* zu können, und bei der Größe und Komplexität dieses Problems sollte man eher zurückhaltend sein, dieses jemals erwarten zu können. Man kann sich erinnert fühlen an den Daten-GAU bei der NASA: ein Haufen von Daten, das meiste davon Abfall, alles unbeschriftet, entzifferbar nur *per trial and error*.

Die DNA ist ein digitales Archiv des Lebens, eines, das in je neuen Versionen die Evolution dokumentiert, das aber nicht von Menschen lesbar zu sein scheint, ein Geheimarchiv größter Bedeutung, aber ohne Zutritt für uns Sterbliche, was dafür aber seine Integrität noch ein kleines Weilchen sichern kann.

7. GEDÄCHTNIS VERSUS SPEICHER

Die Geschichte der Computertechnik ist die Geschichte von schrägen Metaphern und Anthropomorphismen: Charles Babbage nannte Bestandteile seiner gebauten und geplanten Maschinen in Anlehnung an die Landwirtschaft noch *mill* und *store*, aber aus dem an ein Getreidesilo erinnernden *store* wurde im öffentlichen Sprachgebrauch, wie wir wissen, *memory*, Gedächtnis eines künstlichen Gehirns.

Dabei ist der Neurophysiologie noch gar nicht klar, wie das Gedächtnis eines Lebewesens mit Zentralnervensystem funktioniert. Es hat etwas mit Hirnmaterie zu tun, wie die Folgen von Gehirnverletzungen zeigen, aber niemand kann unter dem Mikroskop irgendwelche Speicherplätze zeigen, an denen Gedächtnisinhalte zu lokalisieren wären. Gerhard Roth schreibt: »Das Gedächtnis ist [...] unser wichtigstes ›Sinnesorgan‹. Es ist zugleich aber [...] nur ein Glied im Kreisprozeß von Wahrnehmung, Gedächtnis, Aufmerksamkeit, Erkennen, Handeln und Bewerten.«²³ Hier ist er einig mit dem operationalen Konstruktivismus eines Heinz von Foerster, aber auch mit dem späten Wittgenstein, der polemisiert:

»Ein Ereignis läßt eine Spur im Gedächtnis, das denkt man sich manchmal. [...] Der Organismus mit einer Diktaphonrolle verglichen; der Eindruck, die Spur, ist die Veränderung, die die Stimme auf der Rolle zurückläßt. Kann man sagen, das Diktaphon (oder die Rolle) erinnere sich wieder des Gesprochenen, wenn es das Aufgenommene wiedergibt?«²⁴

Gedächtnis ist mithin nicht zu isolieren, schon gar nicht technisch-konstruktiv. Gedächtnis ist weniger ein Ding als vielmehr ein Prozess, der sich als lebendiger vollzieht und sich dabei möglicherweise auf irgendwelche systemischen Zustandswechsel stützt, die im Prozess des Erinnerns eine kodifizierende Rolle spielen könnten.

Alan Turing hat ja seine Maschine als abstraktes Modell von Computern auch mit Zuständen und Symbolen bestückt, um die Arbeit eines rechnenden Menschen zu simulieren. Dabei spielen die Symbole noch die Rolle von »Gedächtnisstütze[n]«²⁵. Die eigentliche Anthropomorphisierung erfolgte erst später in der Künstliche-Intelligenz-Forschung, deren Credo die *symbol systems hypothesis* ist: Alles, was die Welt ausmacht, sei durch Symbole kodifizierbar, die nach festen Regeln manipulierbar seien, wodurch das menschliche Denken inklusive Gedächtnis nachzubilden wäre. Hier gibt es nun keinen Unterschied mehr zwischen Speicher und Gedächtnis.

Mit der tatsächlichen Situation, in der auch und gerade digital kodierte Daten dem Verfall anheim gegeben sind, kann diese KI-Metapher nicht umgehen. Weltwissen und Fakten, ihre Pendants zu Erfahrung und Gedächtnis, können zwar eventuell durch spätere Ableitungs-Prozesse obsolet und widerlegt werden (*non monotonical reasoning*), aber ein einfaches Verschwinden aufgrund von *data rot* ist für dieses Wissen und solche Fakten ebenso wenig vorgesehen wie die Einbettung des Gewussten in Handlung, Wahrnehmung und Bewertung.

Die Einsicht in die Zeitlichkeit digitaler Daten, die Notwendigkeit, digitale Archive mit hohem Aufwand über Jahrzehnte hinweg zu retten, schlägt in dieselbe Kerbe wie die von Gerhard Roth, den Konstruktivisten oder dem späten Wittgenstein bearbeitete: ein Speicher mag über lange Zeit hinweg intakt bleiben, ohne zu vergehen; Gedächtnis und Erinnerung bedürfen aber genau so wie interpretierbare Daten andauernder tätiger Erneuerung.

Archive, digitale zumal, überdauern nur, wenn sie ständig benutzt werden, wenn eine erhaltende Instanz sie stets neu kodifiziert, interpretiert und bewertet, sich ihre Dokumente handelnd aneignet, sie herausgibt oder verheimlicht, damit Wissen ermöglicht und strukturiert, Handlungen provoziert oder zu unterdrücken trachtet. Nur so überstehen digitale Archive die Jahrzehnte.

Sie leben so lange, wie eine Macht sie trägt und ihren informationellen Stoffwechsel aufrechterhält.

Danach werden sie bestenfalls Mausoleen, in deren Innerem man nichts Brauchbares mehr finden wird, deren Tore besser verschlossen bleiben, weil ihr Inhalt ohnehin Moores Gesetz oder dem zweiten Hauptsatz der Thermodynamik zum Opfer gefallen ist: Entropie, negative Information, Datenstaub, Vergessen.

- 1 Unter: <http://ddr-im-www.de>.
- 2 Unter: http://www.cad.architektur.tu-darmstadt.de/architectura_virtualis/Juedische_Sakralbauten/start.html.
- 3 S. u.
- 4 Nach Hartmut Krech: Der Weltgeist: 1350 Petabyte, in: Die Zeit, Nr. 46, 5.11.1998.
- 5 Ebd.
- 6 Dead Media Working Note 26.6, unter: <http://www.deadmedia.org>.
- 7 Andrew S. Tanenbaum: Computernetzwerke, München 2000, S. 438.
- 8 Unter: <http://www.archive.org/>.
- 9 1,5 TeraByte geteilt durch 600 GigaByte pro Monat ergibt 2,5 Monate oder 75 Tage, bis alle Dokumente ihren Platz verlassen hätten.
- 10 Hans J. Nissen/Peter Damerow/Robert K. Englund: Frühe Schrift und Techniken der Wirtschaftsverwaltung im alten Vorderen Orient. Informationsspeicherung und -verarbeitung vor 5000 Jahren, Bad Salzdetfurth 1990.
- 11 Unter: <http://www.phlab.missouri.edu/~ccgreg/tapes.html>, Übersetzung MW.
- 12 Digital-Alzheimer, in: macmagazin 10 (2000), S. 134.
- 13 Unter: <http://www.phlab.missouri.edu/~ccgreg/tapes.html>.
- 14 Detlef Borchers: Der Glanz von Gestern, in: Süddeutsche Zeitung, Nr. 233, 10.10.2000, S. V2/15.
- 15 Unter: <http://www.deadmedia.org>.
- 16 Ebd.
- 17 Ebd.
- 18 Jeff Rothenberg: Avoiding Technological Quicksand, 1998, unter: <http://www.clir.org/pubs/reports/rothenberg/>.
- 19 Jacques Derrida: Dem Archiv verschrieben. Eine Freudsche Impression, Berlin 1997, S. 11 (Originaltitel: Mal d'Archive).
- 20 Ebd., S. 35.
- 21 Unter: <http://www.openarchives.org/>.
- 22 Matt Ridley: Alphabet des Lebens, München 2000.
- 23 Gerhard Roth: Das Gehirn und seine Wirklichkeit, Frankfurt/M. 1996, S. 241.
- 24 Ludwig Wittgenstein: Bemerkungen über die Philosophie der Psychologie, in: ders.: Werkausgabe, Frankfurt/M. 1984, Bd. 7, I, S. 220.
- 25 Alan Mathison Turing: On Computable Numbers, with an Application to the Entscheidungsproblem, in: Proc. of the London Math. Society, 42/2 (1937). Deutsch: Alan Mathison Turing: Über berechenbare Zahlen mit einer Anwendung auf das Entscheidungsproblem, in: Bernhard Dotzler/Friedrich Kittler (Hg.): Intelligence Service, Berlin 1987, S. 21.