

Emulation – das Archivierungskonzept der Zukunft?

FRANK M. BISCHOFF

In jüngster Zeit ist in Archivreisen ein Konzept zur Archivierung von digitalen Unterlagen rezipiert und diskutiert worden,¹ das von seinem geistigen Urheber, Jeff Rothenberg, als „Emulationslösung“ bezeichnet wird. Rothenberg hatte bereits in einem 1995 erschienenen Artikel angesichts der bestehenden Probleme geschlußfolgert, daß eine konsequente Lösung der Archivierungsfrage nur in der unveränderten Abspeicherung von Daten und Programmen bestehen könne.² Zur Realisierung dieses Ansatzes wollte er sich der Emulation bedienen. Offen blieb, wie die konkrete Umsetzung dieses Ansatzes erfolgen könnte. In einem *Report* für den *Council on Library & Information Resources* vom Januar diesen Jahres hat er sein theoretisches Konzept näher spezifiziert.³

Das Konzept wirft eine Reihe von Fragen auf und hätte – den Erfolg vorausgesetzt – unter Umständen erhebliche Konsequenzen für die bisherigen Ansätze zur Archivierung digitaler Unterlagen. Nach Auffassung Rothenbergs sind die bisherigen Ansätze kurzsichtig, arbeitsintensiv und letztlich nicht imstande, digitale Unterlagen in ihrer ursprünglichen Form zu archivieren. Mit der Realisierung der Emulationslösung wären diese unzulänglichen Ansätze obsolet.⁴

Es wird in dem vorliegenden Beitrag deshalb darum gehen, den derzeitigen Standort der Diskussion um die Archivierung digitaler Unterlagen zu bestimmen, die Probleme der angestrebten Lösungsansätze zu beleuchten, um dann das Emulationskonzept und die damit einhergehenden Konsequenzen darzustellen. Die Publikationen, auf die hier Bezug genommen wird, sind zum Teil nur im World Wide Web zugänglich, was ein Licht auf die Neuartigkeit dieses Ansatzes und anderweitiger Bemühungen zum Thema wirft.

1 Standortbestimmung

Im vorigen Jahr wurden zwei Studien fertiggestellt, die sich mit der Archivierung digitaler Unterlagen befassen. Im Auftrag des *Arts and Humanities Data Service*, eines Datenarchivs für die Wiederverwertung wissenschaftlicher Daten in London, befaßten sich Neil Beagrie und Daniel Greenstein in ihrem Bericht über ein *Strategic Policy Framework for Creating and Preserving Digital Collections*⁵ auch mit öffentlichen Archiven in Großbritannien. Sie kamen zu dem Ergebnis, daß die Archivierung digitaler Unterlagen im wesentlichen auf einer Nutzung von Standardformaten und einer Migrationsstrategie aufbaut.

Zu einer weniger optimistischen Einschätzung gelangten Margaret Hedstrom und Sheon Montgomery.⁶ Ihr im Auftrag der *Research Libraries Group* im vergangenen Dezember veröffentlichter Bericht fußt auf einer Befragung von 54 Mitgliedsinstitutionen, überwiegend Universitätsbibliotheken aus den Vereinigten Staaten. Lediglich drei öffentliche Archive haben an der Befragung teilgenommen, die *Massachusetts Archives*, die *New York State Archives* und das *Public Record Office* in London. In den

¹ Vgl. etwa Alf Erlandsson: *Electronic Records Management: A Literature Review* (International Council on Archives, Studies, 10), Paris 1997, bes. S. 82 f.; Hartmut Weber: *Archiv-Server / Server-Archive – Wie sehen die Kulturspeicher der Zukunft aus?* In: Roland Kamzelak (Hrsg.): *Computergestützte Text-Editoren* (Beihefte zu editio, 12), Tübingen 1999, S. 135-141, bes. S. 139.

² Jeff Rothenberg: *Die Konservierung digitaler Dokumente*. In: *Spektrum der Wissenschaften*, Sept. 1995, S. 66-71.

³ Ders.: *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. A Report to the Council on Library and Information Resources, Januar 1999, <http://www.clir.org/pubs/reports/rothenberg/pub77.pdf>. Vgl. außerdem ders.: *Metadata to Support Data Quality and Longevity*, 3. Juni 1996, http://www.computer.org/conferen/meta96/rothenberg_pa.../ieec.data-quality.htm.

⁴ Ders., *Avoiding Technological Quicksand*, S. 1 (Introduction).

⁵ Neil Beagrie, Daniel Greenstein: *A Strategic Policy Framework for Creating and Preserving Digitale Collections*, Version 4.0, 14. Juli 1998, <http://ahds.ac.uk/manage/framework.htm>.

⁶ Margaret Hedstrom, Sheon Montgomery: *Digital Preservation Needs and Requirements in RLG Member Institutions*. A study commissioned by the Research Libraries Group, Dezember 1998, <http://www.rlg.org/preserv/digpres.html>.

befragten Institutionen werden digitale Unterlagen in insgesamt 24 verschiedenen Datenformaten aufbewahrt, darunter nach TIFF bereits an zweiter Stelle der Häufigkeitsskala proprietäre Textformate, gefolgt von JPEG, PDF, GIF und ASCII. Befragte, die über schriftliche Richtlinien zur Aufbewahrung und Pflege digitaler Unterlagen verfügen – es handelt sich gerade um 18 Institutionen –, gehen im allgemeinen vorsichtiger mit ihrem Material um, sind häufiger an Standardformaten orientiert und verfügen oft über Migrationsstrategien, schriftlich niedergelegte aber lediglich in sieben Fällen. Aufgrund dieses alarmierenden Ergebnisses schlagen Hedstrom und Montgomery der Research Libraries Group vor, Richtlinien, Standards und *best practices* für die Aufbewahrung und Konservierung von digitalen Unterlagen zu entwickeln und an die Mitglieder zu verteilen. In ihren Empfehlungen für Servicestellen und -unternehmen weisen sie auf den dringenden Bedarf an vertrauenswürdigen Dienstleistungen und Unterstützung bei der Aufbewahrung und Pflege digitaler Unterlagen hin. Insbesondere seien Werkzeuge vonnöten, um Unterlagen aus proprietären Textformaten zu konvertieren und verschiedene TIFF-Versionen zu handhaben.

Vergleicht man diese Ergebnisse mit dem *Report of the Task Force on Archiving of Digital Information*, den John Garrett und Donald Waters 1996 ebenfalls für die Research Libraries Group erstellt haben,⁷ so erschreckt, daß sich an den Kernproblemen wenig geändert hat. Ob die Archive in ihrer Summe besser dastehen als die Bibliotheken, muß an dieser Stelle offen bleiben.

Wenn also in dieser Standortbestimmung davon ausgegangen wird, daß die Orientierung an Standardformaten und die Einbeziehung von Migrationsstrategien in Archiven derzeit die tragenden Konzepte für die Archivierung von Unterlagen aus digitalen Systemen darstellen, dann besteht eine gewisse Wahrscheinlichkeit, daß es sich dabei vor allem um eine theoretische Einsicht handelt, die in der Praxis erst noch realisiert werden muß. Ohne Risiko wird man die Behauptung aufstellen dürfen, daß nur wenige Archive über eigene schriftliche Richtlinien zum Umgang mit digitalen Unterlagen verfügen.

Um Mißverständnisse bei der zum Teil uneinheitlich verwendeten Terminologie zu vermeiden, sollen die beiden Basiskonzepte kurz erläutert werden. Eine Orientierung an Standards bedeutet, daß zur Archivierung elektronischer Unterlagen nur bestimmte Datenformate akzeptiert werden, nämlich solche, die einen anerkannten und relativ stabilen Standard darstellen, wie zur Zeit z. B. ASCII oder TIFF. Proprietäre Formate werden vermieden. In der Einsicht, daß digitale Unterlagen nur mit Programmen lesbar gemacht werden können, versucht man damit zu gewährleisten, daß ein Wechsel der Software- oder Systemgeneration nicht automatisch eine Migration in andere Formate nach sich ziehen muß, da auch die neuen Programme ältere Standardformate noch lesen können. Das Leitprinzip besteht also darin, Software nicht zu archivieren. Nur am Rande sei hier angemerkt, daß mit der Archivierung von digitalen Unterlagen in Standardformaten auch eine Archivierung von Metadaten verbunden ist.

Die Orientierung an Standardformaten allein wäre aber eine unzureichende Archivierungsstrategie. Denn auch Standards entwickeln sich weiter oder können irgendwann in eine Sackgasse führen und müssen aufgegeben werden. Für diesen Fall ist eine Konvertierung oder Migration in ein neues Standarddatenformat unerlässlich. Unter Migration wird hier folglich nicht das *Refreshing* verstanden, also der Vorgang des Umkopierens von einem Speichermedium auf ein anderes, weil die physische Haltbarkeit bzw. die Datensicherheit eines Speichermediums begrenzt ist. Solche Konservierungsmaßnahmen werden bei allen mit der Materie befaßten Autoren als selbstverständlich vorausgesetzt und im folgenden nicht weiter thematisiert. Migration ist die Übersetzung eines Formats in ein anderes, so daß sich der Bitstream der Ursprungsdatei von dem Bitstream der resultierenden Datei unterscheidet. Im angestrebten Idealfall sollte aber die Darstellung der beiden Bitstreams, d. h. ihre Interpretation

⁷ John Garret, Donald Waters: Preserving Digital Information. Report of the Task Force on Archiving of Digital Information, commissioned by the Commission on Preservation and Access and the Reserach Libraries Group, 1. Mai 1996, <http://www.rlg.org/pub/archtf/final-report.pdf>.

durch das jeweilige Programm, zu einem identischen Ergebnis führen. Von der Orientierung an Standardformaten erhofft man sich im Hinblick auf eine Migration den weiteren Vorteil, daß die Konvertierung aus einem Standard in einen anderen kostengünstiger zu realisieren sein wird als die Migration von einem proprietären Format in ein anderes.

2 Probleme von Standard und Migration

Mit dem Hinweis auf den „Idealfall“ ist bereits ein Problem dieses Konzepts berührt. Die Konvertierung in ein anderes Format kann nämlich durchaus von Verlusten begleitet sein. Die Verluste können „lediglich“ funktionaler Art sein, etwa indem man Text- oder Datenbankdateien in ein Bildformat überführt und damit alle Funktionen verliert, die mit dem Textformat ursprünglich verbunden waren (beispielsweise Recherchefunktionen). Sie können aber auch inhaltlicher Art sein, etwa wenn Informationen über die äußere Gestaltung bei der Überführung von einem Textformat (z. B. WORD) in ein anderes (z. B. eine *Markup Language*) verloren gehen oder Farbpaletten bei der Übertragung von einem Bildformat in ein anderes nicht vollständig oder korrekt abgebildet werden (im einfachsten Fall etwa bei einer Übertragung in eine Graustufendarstellung). Die im Zuge der Konvertierung auftretenden Veränderungen sind im Einzelnen nicht immer vorhersehbar. Aber selbst wenn die Informationsverluste kalkulierbar und akzeptabel erscheinen, kann man zumindest davon ausgehen, daß die Migration erhebliche Kosten verursacht und in einer Kostenspirale mündet. Denn die Migrationsstrategie setzt unabhängig von einer Benutzung der Unterlagen voraus, daß kein grundlegender Innovationszyklus ausgelassen wird, so daß die Kosten für kontinuierlich wachsende Datenbestände periodisch anfallen.⁸

Es gibt weitere Probleme, die der Migrationsstrategie anhaften. Nimmt man eine Konvertierung in ein Standardformat vor, so ist dieser Schritt in der Regel nicht mehr umkehrbar. Die bei einer Konvertierung von Texten in eine Markup Language oder in ein Bildformat auftretenden Verluste lassen sich bei einer Retrokonversion in ein Textformat – sofern dies überhaupt mit Wirtschaftlichkeitserwägungen vereinbar ist – nur mit hohem Aufwand oder gar nicht mehr rückgängig machen.

Darüber hinaus ist es unvorhersehbar, ob zukünftige Standards noch so gestaltet sein werden, daß eine Konvertierung der älteren Standards möglich ist. Man muß dabei keine Zukunftsszenarien entwerfen und fragen, wie ein Bild-Text-Dokument, das womöglich noch in ein Bitmap konvertiert wurde, in einer Umgebung dargestellt werden soll, die in erster Linie auf Sprachverarbeitung ausgerichtet ist. Bereits die in den letzten zwanzig Jahren erfolgten Paradigmenwechsel bei Datenbankformaten werfen diese Problematik auf. Die Beziehungen zwischen Datensätzen und Datenfeldern, die in hierarchischen, relationalen und objektorientierten Datenbanken dargestellt werden, lassen sich nicht ohne weiteres und verlustfrei in den einen oder anderen Datenbanktyp übertragen. Im günstigsten Falle kann die Kluft zwischen den verschiedenen Datenbanktypen durch eine umfangreiche Dokumentation der Ursprungsumgebung und des ursprünglichen Datenmodells überbrückt werden.

Gravierender ist der Verlust von Funktionen und Funktionalitäten. Was in jedem Tabellenkalkulationsprogramm zur Standardausstattung gehört, wirft für die Archivierung unüberwindbare Probleme auf. Zur Zeit existiert keine Strategie, mit der man hinterlegte Funktionen dauerhaft sichern könnte. Allenfalls lassen sich die daraus resultierenden Ergebnisse zu einem bestimmten Zeitpunkt abbilden, und mit entsprechendem Aufwand die Funktionen dokumentieren in der Hoffnung, daß sie in neuen Umgebungen simuliert werden können. Ein Teil dieser Probleme wurde 1998 auf der Ludwigsburger Tagung des Arbeitskreises *Archivierung von Unterlagen aus digitalen Systemen* im Zusammenhang mit den geografischen Informationssystemen angesprochen.⁹ Will man mehr als Momentaufnahmen

⁸ Vgl. dazu auch Weber, *Archiv-Server / Server-Archive*, S. 138.

⁹ Vgl. die einschlägigen Beiträge in dem Band Udo Schäfer, Nicole Bickhoff (Hrsg.): *Archivierung elektronischer Unterlagen* (Werkhefte

aus einem bestimmten Ausschnitt des Daten- und Funktionspools geografischer Informationssysteme sichern, gelangt man mit den bisherigen Archivierungskonzepten schnell an die Grenzen des Machbaren.

Das verdeutlichen in besonderem Maße die neuen Dokumenttypen, nämlich Hypertexte, oder allgemeiner gesprochen Hypermedien. Deren Fähigkeit, verschiedenste Formate miteinander zu verknüpfen und Informationsvernetzungen zu realisieren, stößt an die Grenzen der bisherigen Archivierungsstrategien. Bereits die verlustfreie Sicherung einer einfachen E-Mail mit den dazugehörigen Attachments ist keine triviale Aufgabe und würde einen ganzen Kranz von Metadaten erfordern, die die Beziehungen der verschiedenen Dokumente zueinander erläutern.

Die angesprochenen Probleme hängen nicht zuletzt mit einer mangelnden Standardisierung zusammen, die mit der raschen Entwicklung auf dem Markt der digitalen Medien nicht Schritt halten kann. Heutige Programme und Funktionen zeichnen sich dadurch aus, daß sie verschiedenste Formate, Text, Bild und Ton, miteinander verknüpfen und darstellen können. Ein vergleichsweise einfaches Beispiel bildet das in Nordrhein-Westfalen demnächst einzuführende System für das elektronische Grundbuch, in dem Pixelgrafiken mit Datenbankeinträgen und Vektorgrafiken kombiniert abgebildet werden sollen.¹⁰ Für eine dauerhafte Archivierung dieser Unterlagen scheint bei derzeitiger Technik eine digitale Hardcopy, also das Herunterbrechen aller Information auf ein Bild-Format, die einzige Lösung darzustellen – mit allem Funktionsverlust, der damit einhergeht.

Was in geschlossenen Umgebungen mit immerhin limitierten Funktionen noch machbar ist, versagt in offenen Netzen. Eine Archivierung von vernetzten Hyperdokumenten wird zwangsläufig an irgendeiner Stelle den Verlust eines Links zur Folge haben. Es handelt sich dabei aber nicht allein um technisch-organisatorische Probleme für die Archivierung. In dem Maße, in dem sich Hyperdokumente durchsetzen, wird die Betrachtungsweise von Akten als physisch mehr oder weniger geschlossene Einheiten, die voneinander abgrenzbare Schriftstücke umfassen, revidiert werden müssen.¹¹ Aber damit sind bereits Probleme angesprochen, die über das hier gestellte Thema hinausweisen und für die sich zur Zeit grundsätzlich keine Lösung abzuzeichnen scheint, da Archive kein digitales Universum aufbewahren, sondern lediglich die im Bewertungsprozeß als archivwürdig befundenen Unterlagen.

3 Leistungen der Emulation

Vor dem Hintergrund der aufgezeigten Probleme vergleicht Rothenberg Standardisierung und Migration mit einer Übersetzung alter Texte in eine Volkssprache und anschließender Vernichtung des Ursprungstextes. Seine Absicht ist es, eine wirkliche Langzeitlösung für die Aufbewahrung digitaler Unterlagen zu finden, die keiner permanenten Anstrengungen und wiederholten Interventionen im Zusammenhang mit geänderten Standards bedarf und sich absolut neutral gegenüber Form und Inhalt des digitalen Materials verhält. Zugleich muß die angestrebte Lösung unabhängig von zukünftigen technischen Entwicklungen tragfähig sein. Die einzige Voraussetzung ist, daß zukünftige Computer jede Art von rechnerischen Funktionen durchführen können, und daß sie schneller und/oder billiger sind als heutige Rechner.

Mit der Emulationslösung verbindet Rothenberg das Ziel, die ursprüngliche Form und Funktion eines

der staatlichen Archivverwaltung Baden-Württemberg, A 13), Stuttgart 1999.

¹⁰ Vgl. Frank M. Bischoff: Elektronisches Grundbuch in Nordrhein-Westfalen. Möglichkeiten der Überlieferungssicherung aus archivischer Perspektive. In: Schäfer, Bickhoff (Hrsg.): Archivierung elektronischer Unterlagen.

¹¹ Die auftretenden Probleme werden in dem Tagungsband der *Second Stockholm Conference on Archival Science and the Concept of Record* diskutiert. Vgl. vor allem Torbjörn Hörnfeldt: The Concept of Record – On Being Digital. In: The Concept of Record. Report from the Second Stockholm Conference on Archival Science and the Concept of Record, 30-31 May 1996 (Skrifter utgivna av Riksarkivet, 4), Stockholm 1998, S. 67-74, und Carol Couture: Is the Concept of a Record still relevant in the Information Age? In: ebd., S. 77-99.

Dokuments wiederbeleben zu können, d. h. Authentizität, Validität und Evidenzwert feststellen zu können und zu verstehen, wie der Urheber oder Bearbeiter des Dokuments es sah, was er daraus herleiten konnte, welche Einsichten es ihm erlaubt hat und welchen ästhetischen Wert es für ihn gehabt haben mag.¹² Insbesondere will er die *Kernattribute von digitalen Unterlagen* – ein zentraler Begriff in seinen Überlegungen – erhalten wissen. Darunter fallen die Fähigkeiten

- perfekt kopiert werden zu können,
- ohne räumliche Beschränkung zugänglich zu sein,
- praktisch ohne steigende Kosten verteilt werden zu können,
- und maschinenlesbar zu sein, so daß die Unterlagen zugänglich, recherchierbar und mit Programmen bearbeitbar sind.

Außerdem müssen neue, „digital geborene“ Formen von Dokumenten, wie dynamische, verteilte oder interaktive Hypertexte oder Hypermedien, ihre ursprünglichen Funktionalitäten bewahren.

Den einzig verlässlichen Weg, uneingeschränkten Zugang zur Bedeutung und zum Inhalt digitaler Dokumente sicherzustellen, sieht Rothenberg darin, diese Dokumente mit der Ursprungssoftware oder einer sehr nahe verwandten Software zu lesen und zu bearbeiten, da eine verlässliche Interpretation des Bitstreams nur dadurch zu gewährleisten ist. Im Grunde verbirgt sich hinter dieser Auffassung Rothenbergs nicht nur ein ausgeprägtes Mißtrauen gegenüber den internationalen Standardisierungsbemühungen im Bereich der Datentechnik, sondern geradezu die Verneinung, daß solche Bestrebungen erfolgreich und für eine Archivierung digitaler Unterlagen nutzbar sein könnten, wemgleich er zugesteht, daß die Strategie der Standardisierung und Migration von allen schlechten Alternativen noch die beste darstellt.

Emulatoren sind Programme, die Programme von anderen Systemen in den Maschinencode des emulierenden Systems umsetzen. Es handelt sich also um eine Art von Mimikry des emulierenden Systems. Emuliert werden kann beispielsweise ein Betriebssystem, so daß Applikationen, die dieses Betriebssystem erfordern, in der Emulationsumgebung lauffähig sind. Rothenberg setzt allerdings tiefer an. Er möchte die Hardwareumgebung emulieren, vor allem deshalb, weil sich deren Verhalten präziser beschreiben läßt als das Verhalten eines Betriebssystems. Neben der eigentlichen Applikation muß daher auch das Betriebssystem aus der Ursprungsumgebung in die emulierende Umgebung portiert werden. Die Realisierung dieses Ansatzes setzt voraus,

1. daß generalisierende Techniken zur Spezifikation von Emulatoren für Hardwareplattformen entwickelt werden, damit aufgrund dieser Spezifikationen auf unbekanntem, zukünftigen Systemen Emulatoren entwickelt werden können, die lauffähig sind und die ursprüngliche Hardwareumgebung genau imitieren;
2. daß Techniken entwickelt werden, Metadaten in lesbarer Form sichern, und zwar Metadaten, die Auskunft darüber geben, wie die Emulationsspezifikation, die Applikationen und die Daten lesbar gemacht werden können;
3. daß Techniken entwickelt werden, um Unterlagen, Metadaten, Software und Emulatorspezifikationen in einer Weise einzukapseln, daß deren Kohäsion gewährleistet wird und sie vor Korruption geschützt sind.

Abgespeichert bzw. eingekapselt werden also

- das Betriebssystem,
- die Originalsoftware,
- die Daten und Dokumente,
- die Emulatorspezifikation,
- darüber hinaus noch eventuell benötigte Erläuterungen, Metadaten und Dokumentationen zu den Unterlagen, ihrer Provenienz, dem historischen und administrativen Kontext etc.

¹² Rothenberg, *Avoiding technological Quicksand*, S. 3 (Kap. 3: *Preservation in the Digital Age*).

Das Ganze wird eingeklammert von Metadaten, die Auskunft darüber geben, welche Informationen vorhanden sind und wie die Begleitdokumentation und die Emulatorspezifikation lesbar gemacht und genutzt werden können. Es handelt sich hier um einen einfachen, fortlaufenden Text, ähnlich den *Readme*-Dateien, die man meist auf den Installationsdisketten oder –CDs von Softwareprodukten findet. An dieser Stelle sieht Rothenberg am ehesten den Bedarf nach einer Standardisierung. Denn diese Metadaten müssen dauerhaft lesbar sein, da andernfalls ein Zugang zu den benötigten Parametern für die Emulation nicht zu gewährleisten ist. Er kalkuliert deshalb ein, daß diese Metadaten im Laufe der Zeit in andere Standards konvertiert werden müssen.¹³

Emulatorspezifikationen und Emulationen müssen getestet werden, um ihre Lauffähigkeit unter Beweis zu stellen. Dazu will sich Rothenberg verschiedener, derzeit auf dem Markt befindlicher Hardwaresysteme bedienen. In diesem Zusammenhang will er aber auch auf ältere Systeme zurückgreifen und diese auf heutigen Systemen emulieren. Darin besteht für ihn übrigens der einzige praktische Nutzen von Computer Museen. Laufen die Emulationen auf den emulierenden Systemen, so sieht er darin genügend Sicherheit, daß Emulationsspezifikation und Emulation fehlerfrei sind und auch auf zukünftigen Systemen korrekt eingesetzt werden können.

Vorausgesetzt, daß Rothenbergs Konzept realisierbar ist, wird damit eine erhebliche Ökonomisierung der Archivierung elektronischer Unterlagen erreicht. Rothenberg geht davon aus, daß für jeden Typus von Hardwareumgebung die Emulationsspezifikation nur einmal erstellt werden muß. Für jeden neuen Typus von Hardwareumgebung muß auch die darauf aufbauende Emulation nur einmal erstellt werden. Darüber hinaus müssen lediglich die periodischen Sicherungskopien erstellt werden, d. h. daß der Bitstream unverändert von einem Datenträger auf einen neuen kopiert wird. Allein die Metadaten, die den Zugang erläutern, müssen gegebenenfalls konvertiert werden.

4 Offene Fragen

Rothenbergs Konzept wirft eine Reihe von Fragen auf, die von zentraler Bedeutung für die Nutzung der Emulationslösung im Rahmen einer Archivierung von digitalen Unterlagen sind.

Rechtliche Aspekte

In der Regel liegen die Rechte der Computerprogramme, die in Gerichten, Behörden und Verwaltungen eingesetzt werden, beim Hersteller. Archive dürfen diese Programme daher nicht einfach übernehmen und einsetzen. Um diesem Dilemma zu entgehen, verweist Rothenberg darauf, daß es für die meisten Textprogramme Viewer gibt, die als Freeware oder Shareware kostenlos benutzt werden können. Anstelle der originalen Editoren sollten daher solche Viewer archiviert werden. Damit verletzt er aber sein eigenes Prinzip, Unterlagen nur mit den Ursprungsprogrammen zu betrachten, weil nur diese eine korrekte Darstellung des Bitstreams gewährleisten können. In der Tat können bei der Verwendung von Viewern, die nicht genau auf die Ursprungssoftware abgestimmt sind, Modifikationen auftreten. Darüber hinaus dürfte es sich als unpraktikabel erweisen, daß Archive bei der Übernahme digitaler Unterlagen erst noch einen passenden, kostenfreien Viewer suchen und testen müssen. Bei bestimmten Anwendungen, wie etwa Datenbanken, würde das wohl auch unmöglich sein. Das Problem der Rechte stellt sich aber, wie Rothenberg selbst eingesteht, ohnehin bei den Betriebssystemen. Daraus folgert, daß Behörden in Zukunft Software nur unter der Bedingung kaufen dürfen, daß eine kostenlose Nachnutzung durch die Archive gestattet ist. Andernfalls müßten die Archive die Nutzungsli-

¹³ Die Lesbarmachung der *Instruktionen*, die für Rothenberg ein durchaus zentrales Anliegen in seinen Überlegungen darstellt, würde wohl für die archivistische Praxis die geringsten Probleme bereiten. Archive wären durchaus in der Lage, diese Informationen in gedruckter Fassung zu archivieren und über Verweise mit den entsprechenden Datenträgern zu verknüpfen. Eine derart pragmatische Hybridlösung entspricht aber nicht den Intentionen Rothenbergs.

zenzen der zu archivierenden Software – und strenggenommen muß man davon ausgehen, daß jede einmal eingesetzte Version der Software hierzu gehört – erstehen, und zwar in einer Situation, in der ihnen praktisch jeder beliebige Preis diktiert werden kann.

Technische Aspekte

Während Rothenberg auf der einen Seite allen Standardisierungsbemühungen mißtraut, muß er zur Realisierung seiner Emulationslösung letztlich doch wieder Standards aufbauen. Es handelt sich weniger um das Standardtextformat, um die Metadaten lesbar zu machen. Er ist vielmehr gezwungen, die Spezifikationen für eine Emulation zu standardisieren. Dabei räumt er ein, daß dieser Standard offen sein muß für einen weiteren Ausbau, da nicht abzusehen ist, ob zukünftige Hardwareumgebungen mit den für heutige Hardware notwendigen Spezifikationen ausreichend beschrieben werden können. Im Datenbank-Bereich hat Rothenberg auf die Inkompatibilität der Datenstrukturen von hierarchischen, relationalen und objektorientierten Datenmodellen hingewiesen. In Anlehnung daran muß man fragen, ob auch im Bereich der Hardware ein Paradigmenwechsel derart denkbar ist, daß eine Emulation älterer Systeme und der korrekte Ablauf von Programmen behindert werden könnte.

In den meisten, insbesondere für eine Archivierung relevanten Fällen handelt es sich heute um komplexe Systeme, die in Netzwerken laufen. Die Installation und Pflege dieser Systeme setzt mithin erhebliches Know-how voraus, das in zwanzig oder dreißig Jahren nicht mehr vorhanden sein wird. Insofern stellt sich die Frage, ob eine Emulation und eine ‘Wiederbelebung’ der alten Programme in der Praxis mit der Leichtigkeit durchführbar ist, die Rothenberg suggeriert. Jeder Nutzer ist in einem Netzwerk-basierten System mit bestimmten, mehr oder weniger limitierten Rechten ausgestattet. Um Zugang zu einem solchen System zu erlangen, benötigt man eine Benutzerkennung und das dazugehörige Paßwort. Vielleicht gehen die Anforderungen aber auch weiter, und das System verlangt zur Identifikation des Benutzers dessen Chipkarte zu lesen, wie beispielsweise das elektronische Grundbuch. Wie gewährleistet man den Zugang in solchen Fällen, wenn man Chipkarte und Chipkartenleser nicht archivieren will?

Es handelt sich hier um die grundsätzlichere Frage, welche Veränderungen der Peripherie welche Auswirkungen haben. Touch-Screens und Spracherkennung entwickeln sich mit einer derartigen Geschwindigkeit, daß niemand prognostizieren kann, welche Bedienungselemente Computer in 30 Jahren aufweisen werden. Kann die Emulationslösung also gewährleisten, daß eine Applikation, die mit Maus und Tastatur bedient wird, in Zukunft auch ohne diese Peripherie benutzt werden kann? Kann weiterhin sichergestellt werden, daß Dateien aus veralteten Systemen, mit denen ebenfalls nur veraltete Druckertreiber gesichert sind, auf zukünftigen Printern ausgedruckt werden können? Nicht zuletzt im Bereich der Druckersteuerung haben sich in den letzten zwölf Jahren erhebliche Veränderungen ergeben. Die äußere Gestalt eines Dokuments war in den 80er Jahren noch weitgehend von den Möglichkeiten abhängig, die der jeweilige Drucker bot. Heute kann zwar immer noch die Mindestgröße der Blattränder in Abhängigkeit vom verwendeten Drucker differieren, doch werden Fonts statt dessen in der Software bzw. mit dem Betriebssystem verwaltet. Damit stellt sich die Frage, bis zu welchem Grad die Peripherie eines Rechners zur Realisierung einer Emulationslösung beschrieben werden muß, und welchen Aufwand das in der Praxis erfordert.

Es ist allgemein bekannt, daß die fehlerfreie Software eine Wunschvorstellung ist. Im günstigsten Fall werden Fehler nicht offenbar, weil eine bestimmte Kombination von Befehlen oder Umgebungsparametern nicht eingesetzt wird. „Never touch a running system“ ist eine Devise, die im Rahmen der Archivierung digitaler Unterlagen zwangsläufig nicht gelten kann. Es stellt sich aber die Frage, wie Applikationen nach ihrer Portierung in eine neue Umgebung reagieren und ob diese Reaktionen wirklich mit naturwissenschaftlicher Exaktheit vorausbestimmt werden können. Was passiert, wenn man im Jahre 2050 ein System starten will, das mit dem Jahr 2000-Problem behaftet ist?

Archivfachliche und archivpraktische Aspekte

Rothenberg hat zu Recht ausgeführt, daß die derzeitigen Strategien bei der Archivierung von Hyperdokumenten scheitern müssen. Die Frage ist allerdings, ob die Emulationslösung letztlich nicht dieselben Probleme mit sich bringen würde. Folgt man Rothenbergs Ausführungen, so gewinnt man den Eindruck, es handele sich um die vollständige und unversehrte Archivierung von geschlossenen Systemen. Tatsächlich werden Systeme immer offener, so daß etwa Links auch nach außerhalb gesetzt werden können. Außerdem findet eine vollständige Archivierung ohnehin nur in Ausnahmefällen statt. Archive bewerten und kassieren Unterlagen und werden dies auch im digitalen Zeitalter weiterhin tun. Mit dieser Anforderung befaßt sich Rothenberg aber gar nicht. Daraus ergibt sich nämlich die Konsequenz, daß möglicherweise Dokumente vernichtet werden, auf die von anderen Dokumenten aus mit einem Link verwiesen wurde. Es handelt sich hier um ein unausweichliches Dilemma.

Daneben stellen sich noch Fragen der Benutzung. Der Einsatz der Emulationslösung wird dazu führen, daß eine prinzipiell unbegrenzte Zahl von Applikationen archiviert wird. Man darf getrost unterstellen, daß es sich dabei nur zum geringeren Teil um selbsterklärende Anwendungen handelt. Die Konsequenzen werden vielleicht anhand eines Beispiels aus der Praxis deutlich: Die Administratordokumentation der in den nordrhein-westfälischen Verwaltungs- und Finanzgerichten eingesetzten *VG/FG-Verfahrenslösung* umfaßt gut einen Regalmeter. Alle Mitarbeiter, die mit dem System arbeiten, werden zuvor mehrere Tage lang geschult. Sind diese Applikationen in die Archive gelangt, wird das Archivpersonal nicht nur die Systeme in einer Emulationsumgebung installieren, sondern auch den Benutzern Zugang gewähren und erste Hilfestellungen in der Bedienung geben müssen. Das wird bei der Summe der zu archivierenden Systeme erhebliche Personalressourcen kosten.

Ein Vorteil der Emulationslösung besteht darin, daß sie einem späteren Benutzer vermittelt, welchen Blick ein früherer Sachbearbeiter auf die Daten hatte. Das, was in der Archivwissenschaft in den letzten Jahren häufig mit dem Begriff 'Evidenz' umschrieben wird, bleibt bei der Emulationslösung folglich gewahrt, allerdings mit der Einschränkung, daß dies nur für einen bestimmten Zeitpunkt zutrifft, nämlich denjenigen des letzten Arbeitsschrittes vor der Übernahme der Daten und Systeme ins Archiv. Es ist aber eine Illusion, daß der Historiker dabei stehen bleiben möchte. Er wird in der Regel einen uneingeschränkten Blick auf das Datenmaterial haben, die Historie der Bearbeitung nachvollziehen und gegebenenfalls auch statistische Untersuchungen durchführen wollen. Mit der Emulationslösung läßt sich das nicht realisieren, da der Benutzer zwangsläufig auf die Funktionen beschränkt bleibt, die ihm die Applikation und das Betriebssystem bereitstellen. Unter Umständen wird es keine Möglichkeit geben, die Daten aus einer Anwendung zu extrahieren, um sie in einer anderen Systemumgebung unter anderen Gesichtspunkten auszuwerten. Insofern wirft David Bearman zu Recht die Frage auf, ob die Emulationslösung überhaupt die Informationen sichert, die Archive dauerhaft aufbewahren wollen.¹⁴

5 Resümee

Rothenberg hat seinem Optimismus Ausdruck verliehen, daß sein Ansatz zur Archivierung digitaler Unterlagen bislang der einzige sei, der eine wirkliche Lösung des Problems biete.¹⁵ Selbst wenn man ihm so weit nicht folgen möchte, muß man doch zugestehen, daß die Emulationslösung eine wertvolle Bereicherung für die Archivierung digitaler Unterlagen darstellen kann. Insofern ist es nur zu begrüßen, daß Rothenbergs theoretisches Konzept in Zusammenarbeit mit dem niederländischen Reichsarchivdienst auf seine Praxistauglichkeit geprüft wird.¹⁶ Sollten die Tests erfolgreich ausfallen,

¹⁴ David Bearman: Reality and Chimeras in the Preservation of Electronic Records: In: D-Lib Magazine 5 (1999), Nr. 4, April 1999 (<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/april99/bearman/04bearman.html>).

¹⁵ Rothenberg, Avoiding technological Quicksand, S. V (*Executive Summary*).

¹⁶ Vgl. den Beitrag von Hans Hofmann in diesem Band.

ist es vorstellbar, daß die Emulationslösung für bestimmte Applikationen zur Archivierung digitaler Unterlagen eingesetzt wird, in jedem Fall aber für solche Unterlagen, für die keine anderweitigen validen Konzepte existieren.

Man muß sich nur darüber im Klaren sein, daß Rothenbergs Konzept in der bisher veröffentlichten Form eine ‘Alles-oder-Nichts-Lösung’ darstellt. Hat man den Weg der Emulation einmal beschritten, so wird es ab einem bestimmten Zeitpunkt kein Zurück mehr geben, weil das Know-how und die Werkzeuge zur Konvertierung der Daten aus einer alten Applikation in einer alten Systemumgebung nicht mehr vorliegen. Im Gegensatz zur Migrationsstrategie hat man bei der Emulation, sofern sie funktioniert, keine Informationsverluste. Funktioniert sie nicht, hat man nicht nur Farbpaletten, Layoutmerkmale oder Funktionen, sondern alles verloren. Vor diesem Hintergrund würde man sich einen zusätzlichen Sicherungsmechanismus wünschen, nämlich eine Migrationsstrategie, die einerseits einer technisch bedingten Katastrophe vorbeugen könnte und andererseits der Forschung ermöglichen würde, die archivierten Daten und Dokumente unter anderen, von dem Blickwinkel des Sachbearbeiters abweichenden Gesichtspunkten zu bearbeiten.¹⁷ Das scheint aber mit Rothenbergs Konzept nicht kompatibel zu sein.¹⁸

Kritisch bleibt eigentlich nur anzumerken, daß Rothenberg sein Emulationskonzept auf den ersten Blick so überzeugend darstellt, daß der Leser über mögliche praktische Probleme und Konsequenzen bei seiner Umsetzung hinweggetäuscht wird. Es wäre fatal, wenn der Eindruck entstünde, mit der Entwicklung der Emulationslösung seien die Archive hinsichtlich der Aufbewahrung digitaler Unterlagen salviert. Wer angesichts des bisher nur theoretisch entwickelten Emulationskonzepts die Schlußfolgerung zieht, daß weitere Bemühungen um die Archivierung digitaler Unterlagen überflüssig oder auf einen späteren Zeitpunkt verschiebbar und eine Orientierung an Standards und Migrationsstrategien damit obsolet seien, handelt fahrlässig.¹⁹

¹⁷ Die Anforderungen der historischen Forschung, die nicht allein daran interessiert ist, den potentiellen Wissensstand von Sachbearbeitern in vergangenen Verwaltungen zu erlangen, werden von Rothenberg grundsätzlich nicht berücksichtigt. Bei allem Respekt vor der berechtigten Forderung nach einer authentischen Überlieferungsbildung, darf der Evidenzgedanke nicht dazu führen, die Sicht auf die Quellen einzuzengen. Wie in einer Verwaltung mit elektronischen Unterlagen umgegangen wurde, welche Recherche- und Informationsmöglichkeiten die – in einem Netzwerk doch wohl mit unterschiedlichen Rechten ausgestatteten – Mitarbeiter dieser Verwaltung hatten und wie die Arbeitsabläufe organisiert waren, wird man dokumentieren, aber nicht notwendigerweise nachbilden müssen. Man sollte sich deshalb davor hüten, mit Archivierungsmodellen der zukünftigen Forschung das Spektrum möglicher Fragestellungen vorzugeben.

¹⁸ An dieser Stelle muß aber einschränkend angemerkt werden, daß sich das Urteil allein auf die bisherigen Veröffentlichungen von Rothenberg bezieht. Während der Diskussion über seinen Vortrag zur Emulationslösung an der Archivschule in Marburg am 6. April 1999 hat Rothenberg ausgeführt, daß er auch eine Migrationsstrategie einplane, die eine spätere Konvertierung erlauben würde. – In Ergänzung zu den Arbeiten von Rothenberg sei hier der Hinweis auf einen gerade erst in Großbritannien veröffentlichten Entwurf zur Frage der Langzeitsicherung digitaler Unterlagen nachgetragen, der auf einer im Auftrag des Public Record Office in London erstellten Studie von Kenneth Tombs fußt: BRITISH STANDARDS INSTITUTION: Bundles for the Perpetual Preservation of Electronic Documents and Associated Objects, Public Draft for Comment – IDT/1/4: 99/621800DC. Tombs’ Ansatz ist dem von Rothenberg sehr ähnlich. Auch er will Unterlagen und Daten, Betriebssystem, Applikationssoftware, Viewer und Metadaten einkapseln – Tombs spricht in diesem Zusammenhang von ‘Bündeln’, die die Größe der derzeit im Rahmen einer Langzeitaufbewahrung nutzbaren Speichermedien nicht überschreiten sollten (500 MB) –, so daß sie via Emulation wieder in ihrem Ursprungszustand und mit ihren Ursprungsfunktionalitäten nutzbar sind. Allerdings scheint das – im Hinblick auf die ‘Wiederbelebung’ von Daten und Programmen nicht sehr detailliert ausgeführte – Konzept parallel zur Emulation auch eine Migrationslösung vorzusehen. Sollte die Emulationslösung tatsächlich die Möglichkeit einer auch in zukünftigen Systemumgebungen weitgehend automatisch und sicher durchführbaren Massenmigration von Unterlagen beinhalten, so würde das ihre Eignung für die Sicherung von Archivalien erheblich verbessern. Damit wäre nämlich sowohl dem Sicherheitsbedürfnis von Archiven als auch einer flexiblen Benutzung digitaler Unterlagen gedient. Insoweit darf man mit Spannung auf die Ergebnisse der praktischen Tests dieser Verfahren warten.

¹⁹ Dem kritischen Urteil von Bearman, Reality and Chimeras in the Preservation of Electronic Records, ist hier zuzustimmen.