H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth (Hrsg.)

nestor Handbuch

Eine kleine Enzyklopädie der digitalen Langzeitarchivierung

Version 2.3

Kapitel 8

Digitale Erhaltungsstrategien

nestor

nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung
hg. v. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth
im Rahmen des Projektes: nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland
nestor – Network of Expertise in Long-Term Storage of Digital Resources
http://www.langzeitarchivierung.de/

Kontakt: editors@langzeitarchivierung.de
c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen,
Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter http://www.d-nb.de/ abrufbar.

Neben der Online Version 2.3 ist eine Printversion 2.0 beim Verlag Werner Hülsbusch, Boizenburg erschienen.

Die digitale Version 2.3 steht unter folgender Creative-Commons-Lizenz:
„Namensnennung-Keine kommerzielle Nutzung-Weitergabe unter gleichen Bedingungen 3.0 Deutschland“
http://creativecommons.org/licenses/by-nc-sa/3.0/de/

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen, Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und als solche den gesetzlichen Bestimmungen unterliegen.

URL für Kapitel 8 „Digitale Erhaltungsstrategien“ (Version 2.3): urn:nbn:de:0008-2010062472
http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-2010062472

Gewidmet der Erinnerung an Hans Liegmann (†), der als Mitinitiator und früherer Herausgeber des Handbuchs ganz wesentlich an dessen Entstehung beteiligt war.

8 Digitale Erhaltungsstrategien

8.1 Einführung

Stefan E. Funk

Wie lassen sich die Dinge bewahren, die uns wichtig sind, Objekte, die wir der Nachwelt am allerliebsten in genau dem Zustand, in dem sie uns vorliegen, erhalten wollen?

Handelt es sich bei diesen Objekten um Texte oder Schriften, wissen wir, dass Stein- und Tontafeln sowie Papyri bei geeigneter Behandlung mehrere tausend Jahre überdauern können. Auch bei Büchern haben wir in den letzten Jahrhunderten Kenntnisse darüber gesammelt, wie diese zu behandeln sind bzw. wie diese beschaffen sein müssen, um nicht der unfreiwilligen Zerstörung durch zum Beispiel Säurefraß oder Rost aus eisenhaltiger Tinte anheim zu fallen. Auch Mikrofilme aus Cellulose mit Silberfilm-Beschichtung sind bei richtiger Lagerung viele Jahrzehnte, vielleicht sogar Jahrhunderte, haltbar. Alle diese Medien haben den Vorteil, dass sie, wenn sie als die Objekte, die sie sind, erhalten werden können, von der Nachwelt ohne viele Hilfsmittel interpretiert

werden können. Texte können direkt von Tafeln oder aus Büchern gelesen und Mikrofilme mit Hilfe eines Vergrößerungsgerätes recht einfach lesbar gemacht werden.

Bei den digitalen Objekten gibt es zwei grundlegende Unterschiede zu den oben genannten analogen Medien: Zum einen werden die digitalen Informationen als Bits (auf Datenträgern) gespeichert. Ein Bit ist eine Informationseinheit und hat entweder den Wert „0“ oder den Wert „1“. Eine Menge dieser Nullen und Einsen wird als Bitstream bezeichnet. Die Lebensdauer der Bits auf diesen Datenträgern kennen wir entweder nur aus Laborversuchen oder wir haben noch nicht genug Erfahrungswerte für eine sichere Angabe der Lebensdauer über einen langen Zeitraum hinweg sammeln können. Schließlich existieren diese Datenträger erst seit einigen Jahren (bei DVDs) oder Jahrzehnten (bei CDs). Eine Reaktion auf die Unsicherheit über die Lebensdauer dieser Medien ist die Bitstreamerhaltung sowie die Mikroverfilmung. Zum anderen ist keines der digitalen Objekte ohne technische Hilfsmittel nutzbar. Selbst wenn wir die Nullen und Einsen ohne Hilfsmittel von den Medien lesen könnten, dann könnten wir wenig bis gar nichts mit diesen Informationen anfangen. Da diese konzeptuellen Objekte digital kodiert auf den Medien gespeichert sind, bedarf es spezieller Hilfsmittel, die diese Informationen interpretieren können. Als Hilfsmittel dieser Art ist einerseits die Hardware zu sehen, die die Daten von den Medien lesen kann (beispielsweise CD- bzw. DVD-Laufwerke) und natürlich die Computer, die diese Daten weiterverarbeiten. Andererseits wird die passende Software benötigt, die die Daten interpretiert und so die digitalen Objekte als konzeptuelle Objekte erst oder wieder nutzbar macht.

Kann der Bitstream nicht mehr interpretiert werden, weil das Wissen um eine korrekte Interpretation verloren ging, ist der Inhalt des konzeptuellen Objektes verloren, obwohl die eigentlichen Daten (der Bitstream) noch vorhanden sind. Lösungsansätze für dieses Problem sind die Migration und die Emulation. Eine weitere Idee ist es, in einem so genannten Computermuseum die originale Hard- und Software bereitzustellen und so die konzeptuellen Objekte zu erhalten.

8.2 Bitstream Preservation

Dagmar Ullrich

Grundlage aller Archivierungsaktivitäten ist der physische Erhalt der Datenobjekte, die Bitstream[1] Preservation. Es wird eine Speicherstrategie vorgeschlagen, die auf einer redundanten Datenhaltung auf mindestens zwei unterschiedlichen, marktüblichen und standardisierten Speichertechniken basiert. Die eingesetzten Speichermedien sollten regelmäßig durch aktuelle ersetzt werden, um sowohl dem physischen Verfall der Speichermedien als auch dem Veralten der eingesetzten Techniken vorzubeugen. Es werden vier Arten von Migrationsprozessen vorgestellt. Das sind: Refreshment, Replication, Repackaging und Transformation. Als Medienmigration im engeren Sinne werden nur die beiden ersten, Refreshment und Replication, angesehen. Sie bezeichnen das Auswechseln einzelner Datenträger (refreshing) oder eine Änderung eingesetzter Speicherverfahren (replication). Durch die kurzen Lebenszyklen digitaler Speichermedien erfolgt ein Erneuern der Trägermedien oft im Rahmen der Aktualisierung der eingesetzten Speichertechnik.

Physischer Erhalt der Datenobjekte

Um digitale Daten langfristig verfügbar zu halten, muss an zwei Stellen angesetzt werden. Zum einen muss der physische Erhalt des gespeicherten Datenobjekts (Bitstreams) auf einem entsprechenden Speichermedium gesichert werden. Zum anderen muss dafür Sorge getragen werden, dass dieser Bitstream auch interpretierbar bleibt, d.h. dass eine entsprechende Hard- und Software-Umgebung verfügbar ist, in der die Daten für einen menschlichen Betrachter lesbar gemacht werden können. Ohne den unbeschädigten Bitstream sind diese weiterführenden Archivierungsaktivitäten sinnlos. Der physische Erhalt der Datenobjekte wird auch als „Bitstream Preservation“ bezeichnet. Für den physischen Erhalt des Bitstreams ist eine zuverlässige Speicherstrategie erforderlich.

---

1 Der Begriff „Bitstream“ wird hier als selbsterklärend angesehen. Eine Erläuterung des Begriffs findet sich in: Rothenberg, Jeff (1999): *Ensuring the Longevity of Digital Information*. http://www.clir.org/pubs/archives/ensuring.pdf
Bei diesem Text handelt es sich um eine ausführlichere Fassung eines gleichnamigen Artikels, der 1995 in der Zeitschrift „Scientific American“, Band 272, Nummer 1, Seiten 42-47 erschienen ist.
Alle hier aufgeführten URLs wurden im Mai 2010 auf Erreichbarkeit geprüft .

Verfahrensvorschläge für eine Bitstream Preservation

Die nachstehenden vier Verfahrensvorschläge können als Grundlage für eine zuverlässige Speicherstrategie zur Sicherstellung des physischen Erhalts der archivierten Datenobjekte verwendet werden:[2]

1. *Redundante Datenhaltung*: Die Daten sollten in mehrfacher Kopie vorliegen. Zur Sicherung gegen äußere Einflüsse empfiehlt sich auch eine räumlich getrennte Aufbewahrung der unterschiedlichen Kopien.
2. *Diversität eingesetzter Speichertechnik*: Die Daten sollten auf mindestens zwei unterschiedlichen Datenträgertypen gesichert werden.
3. *Standards*: Die verwendeten Speichermedien sollten internationalen Standards entsprechen und auf dem Markt eine weite Verbreitung aufweisen.
4. *Regelmäßige Medienmigration*: Die verwendeten Speichertechniken bzw. Datenträger müssen regelmäßig durch neue ersetzt werden.

Redundanz, Speichertechniken und Standards

Eine mehrfach *redundante Datenhaltung* ist in vielen Bereichen der Datensicherung üblich. Bei wertvollen, insbesondere bei nicht reproduzierbaren Daten wird man sich nicht auf eine einzige Kopie verlassen wollen. Um das Risiko äußerer Einflüsse wie Wasser- oder Brandschäden zu verringern, bietet sich die räumlich getrennte Aufbewahrung der Kopien an. Um auch die Gefahr eines Datenverlusts durch menschliches Versagen oder Vorsatz einzuschränken, kann eine Aufbewahrung bei zwei unabhängigen organisatorischen Einheiten in das Redundanzszenario mit einbezogen werden. Zusätzliche Sicherheit lässt sich gewinnen, indem die jeweiligen Kopien *auf unterschiedlichen Speichertechniken* gehalten werden. Dies mindert das Risiko eines Datenverlusts durch Veralterung einer der eingesetzten Techniken. Sofern vorhanden, sollten Fehlererkennungs- und Korrekturmechanismen zur Sicherung der Datenintegrität eingesetzt werden. Weiter sollte die Funktionstüchtigkeit der Speichermedien und Lesegeräte anhand von Fehlerstatistiken überwacht werden. Die sachgerechte Handhabung von Datenträgern und Lesegeräten ist in jedem Fall vorauszusetzen. Alle

---

2 Die Auflistung erhebt keinen Anspruch auf Vollständigkeit. Ähnliche Aufstellungen finden sich z.B. in: Rathje, Ulf (2002): Technisches Konzept für die Datenarchivierung im Bundesarchiv. In: Der Archivar, H. 2, Jahrgang 55, S.117-120, http://www.bundesarchiv.de/imperia/md/content/abteilungen/abtb/1.pdf und: o.V. (o.J.) Digital preservation. Calimera Guidelines. S.3. http://www.calimera.org/Lists/Guidelines%20PDF/Digital_preservation.pdf

verwendeten Speichertechniken bzw. -medien sollten auf internationalen *Standards* basieren und über eine möglichst breite Nutzergruppe verfügen.

Regelmäßige Medienmigration

Als Medienmigration kann jeder Vorgang betrachtet werden, bei dem das physische Trägermedium eines Datenobjekts innerhalb eines Archivs geändert und der Vorgang mit der Absicht durchgeführt wird, das Datenobjekt zu erhalten, indem die alte Instanz durch die neue ersetzt wird. Eine entsprechende Definition von „Digital Migration" findet sich im OAIS-Referenzmodell[3]:

> *Digital Migration is defined to be the transfer of digital information, while intending to preserve it, within the OAIS. It is distinguished from transfers in general by three attributes:*
>
> *- a focus on the Preservation of the full information content*
>
> *- a perspective that the new archival implementation of the information is a replacement for the old; and*
>
> *- full control and responsibility over all aspects of the transfer resides with the OAIS.*

Im OAIS-Referenzmodell werden vier Arten der Migration genannt: Refreshment, Replication, Repackaging und Transformation.[4]

Refreshment: Als Refreshment werden Migrationsprozesse bezeichnet, bei denen einzelne Datenträger gegen neue, gleichartige Datenträger ausgetauscht werden. Die Daten auf einem Datenträger werden direkt auf einen neuen Datenträger gleichen Typs kopiert, der anschließend den Platz des alten in der Speicherinfrastrukur des Archivs einnimmt. Weder an den Daten noch an der Speicherinfrastruktur werden also Änderungen vorgenommen, es wird lediglich ein Datenträger gegen einen gleichartigen anderen ausgetauscht.

Replication: Eine Replication ist ein Migrationsprozess, bei dem ebenfalls Daten von einem Datenträger auf einen neuen kopiert werden. Bei der Replica-

---

3 Consultative Committee for Space Data Systems (CCSDS) (2002): *Reference Model for an Open Archival Information System (OAIS).* Blue Book. Washington DC. Seite 5-1. vgl. auch Kapitel 4.
http://public.ccsds.org/publications/archive/650x0b1.pdf

4 Consultative Committee for Space Data Systems (CCSDS) (2002): *Reference Model for an Open Archival Information System (OAIS).* A.a.O. Seite 5-4.

tion jedoch kann es sich bei dem neuen Datenträger auch um einen andersartigen, z.B. aktuelleren, handeln. Andersartige Datenträger erfordern eine entsprechende Anpassung der Speicherinfrastruktur. Der neue Datenträger kann in der Regel nicht unmittelbar den Platz des alten einnehmen. Der wesentliche Unterschied zum Refreshment liegt daher in den mit dem Kopierprozess einhergehenden Änderungen der verwendeten Speicherinfrastruktur.

Repackaging: Ein Repackaging ist ein Migrationsprozess, bei dem ein sogenanntes Archivpaket verändert wird. Diese Änderung betrifft nicht die eigentlichen Inhaltsdaten, sondern die Struktur des Archivpakets.

Transformation: eine Transformation ist ein Migrationsprozess, bei dem auch die Inhaltsdaten des Archivpakets verändert werden.

Refreshment und Replication können als Medienmigrationen im engeren Sinne angesehen werden. Der Umkopierprozess erfolgt in beiden Fällen mit der Absicht, das Trägermedium zu ersetzen, unabhängig davon, welche Inhalte auf ihm abgelegt sind. Die Replication wird im Folgenden im Sinne eines Technologiewechsels interpretiert.[5] Ein Refreshment beschränkt sich dagegen auf den Wechsel einzelner Datenträger innerhalb einer Speichertechnik, z.B. einer Magnetbandgeneration. Bei Repackaging und Transformation dagegen werden auch die Datenobjekte selbst umgeschrieben. Ein Beispiel für ein Repackaging ist die Änderung des Packformats von ZIP zu TAR. Eine Formatmigration, z.B. von JPG zu TIFF, ist dagegen eine Transformation, da die Inhalte des Archivpakets verändert werden. Die Unterscheidung dieser vier Arten von Migrationen erleichtert die begriffliche Abgrenzung einer Medienmigration von einer Formatmigration. Eine Formatmigration umfasst letztlich immer auch eine Medienmigration, da ein neues Datenobjekt erstellt und auf einem eigenen Trägermedium abgelegt wird. Die Formatmigration erfolgt aber mit Blick auf die künftige Interpretierbarkeit des Bitsreams, die Medienmigration im engeren Sinne hingegen dient dessen Erhalt. Für die Bitstream Preservation sind nur die beiden ersten, Refreshment und Replication, wesentlich, da die beiden anderen den Bitstream verändern. Ein Refreshment ist in der Regel weniger aufwändig als eine Replication, da nicht das Speicherverfahren, sondern nur einzelne Datenträger erneuert werden.

5 Eine Replication muss nach der zitierten Definition nicht notwendig von einem veralteten Medium auf ein aktuelleres erfolgen, sondern ggf. auch auf ein gleichartiges. In der Praxis wird das aber eher selten der Fall sein.

Refreshment und Replication

Ein Erneuern (refreshing) einzelner Datenträger kann aufgrund von Fehlerraten oder auf der Basis bestimmter Kriterien wie Zugriffshäufigkeit oder Alter erfolgen. Der Aufwand solcher Maßnahmen ist gegen die Wahrscheinlichkeit eines Datenverlusts durch einen fehlerhaften Datenträger abzuwägen. Auf der einen Seite können zusätzliche Kontrollverfahren eine sehr hohe Systemlast erzeugen, die den aktiven Zugriff auf die Daten beträchtlich einschränken kann. Zudem sind die Beurteilungskriterien wie Zugriffshäufigkeit, Alter und ggf. die tolerierbare Fehlerrate oft strittig und zum Teil nur mit teurer Spezialsoftware oder auch gar nicht feststellbar. Nicht selten können sie auch im Einzelfall durch Unterschiede in Produktionsablauf oder Handhabung zwischen Datenträgern desselben Typs stark variieren. Auf der anderen Seite wird die Haltbarkeit von Trägermedien aufgrund des raschen Technologiewandels meist gar nicht ausgereizt. Die Wahrscheinlichkeit schadhafter Datenträger durch altersbedingten Verfall ist daher eher gering. Um diesen Zusammenhang deutlich zu machen, kann die durchschnittliche Lebensdauer eines Datenträgers von seiner durchschnittlichen Verfallszeit unterschieden werden.[6]

„*Medium Expected Lifetime (MEL)*: The estimated amount of time the media will be supported and will be operational within the electronic deposit system."

"*Medium Decay Time (MDT)*: The estimated amount of time the medium should operate without substantial read and write errors."

Die Definition der durchschnittlichen Lebensdauer enthält zwei durch „und" verbundene Zeitangaben. Die eine bezieht sich auf die Dauer der Unterstützung eines Speichermediums durch den Hersteller, die andere auf die Dauer des Einsatzes eines Speichermediums im digitalen Archiv. Diese beiden Zeitspannen können durchaus differieren. Nicht selten zwingt die wegfallende Unterstützung durch den Hersteller zur Migration, auch wenn die vorhandenen Systeme voll funktionsfähig sind und noch weiter betrieben werden könnten. Für Speichertechniken, die vom Hersteller nicht mehr unterstützt werden, können Ersatzteile oder technische Betreuung nicht mehr garantiert werden. Ihr Weiterbetrieb ist daher nicht ratsam. Der Begriff der durchschnittlichen Lebensdauer wird aus diesen Gründen hier als die durchschnittlich zu erwartende Hersteller-Unterstützung interpretiert. Solange diese durchschnittliche Lebensdauer unter der durchschnittlichen Verfallszeit liegt, ist ein Ausfall einzelner

6 Van Diessen, Raymond J. und van Rijnsoever, Ben J. (2002): *Managing Media Migration in a Deposit System. IBM/KB Long-Term Preservation Study Report Series Nr. 5.* Amsterdam: IBM Niederlande. S.4.
http://www-5.ibm.com/nl/dias/resource/migration.pdf

Datenträgern selten zu erwarten. Statt aufwändiger Kontrollen der Datenträger kann es in diesem Fall einfacher sein, auf eine redundante Datenhaltung zu vertrauen, im konkreten Fehlerfall einzelne Datenträger oder Laufwerke zu ersetzen und den gesamten Bestand im Rahmen eines Technologiewechsels (Replication) komplett auszutauschen.

Eine Replication im Sinne eines Technologiewechsels umfasst Änderungen in der bestehenden Speicherinfrastrukur. Erforderliche Technologiewechsel können sehr unterschiedlich ausfallen. Sie können von einer Magnetbandgeneration zur nächsten reichen oder einen vollständigen Wechsel z.B. von Magnetbändern zu optischen Medien bedeuten. Im ersten Schritt muss die neue Speichertechnik in die bestehende Infrastruktur integriert werden. Anschließend müssen die betroffenen Datenbestände von der alten Technik auf die neue umkopiert werden. Bei großen Datenmengen mit ggf. hohen Sicherheits- oder Verfügbarkeitsansprüchen können diese Umkopierprozesse aufwändig und langwierig sein. Die Lesegeschwindigkeit der älteren Speichermedien wird in der Regel langsamer sein als die Schreibgeschwindigkeit der neuen. Beide müssen für einen Kopierprozess koordiniert werden, ggf. über Zwischenspeicher. Der Übertragungsvorgang muss abgeschlossen sein, bevor die alte Speichertechnik unbrauchbar wird. An diesem Punkt sei auf die oben ausgeführte Interpretation von „Medium Expected Lifetime“ hingewiesen. Dass der Migrationsprozess abgeschlossen sein muss, bevor eine Speichertechnik nicht mehr auf dem Markt ist, wäre ein sehr hoher Anspruch, da viele Speichermedien nur drei bis fünf Jahre lang angeboten werden. Unter Umständen kann ein solcher Anspruch je nach Wert der betroffenen Daten gerechtfertigt sein. Häufig bieten Hersteller die Unterstützung von Speichermedien einige Jahre länger an, als diese Technik aktiv vertrieben wird. Dies verlängert die zuverlässige Einsatzdauer von Speichertechniken. Eine zusätzliche Sicherheit kann in diesem Kontext auch der Verfahrensvorschlag unterschiedliche Speichertechniken einzusetzen bieten.

Zusammenfassung

Ein Langzeitarchiv muss über zuverlässige Speicherstrategien verfügen, die nicht nur ein „Refreshment“ eingesetzter Datenträger innerhalb einer Speichertechnik ermöglichen, sondern darüber hinaus auch die Erneuerung ganzer Speichertechniken. Solche Strategien müssen sicherstellen, dass zu keinem Zeitpunkt Datenbestände unzugänglich werden, weil ihre Trägermedien nicht mehr lesbar sind.

Literatur

Rothenberg, Jeff (1999), *Ensuring the Longevity of Digital Information*. http://www.clir.org/pubs/archives/ensuring.pdf
Bei diesem Text handelt es sich um eine ausführlichere Fassung eines gleichnamigen Artikels, der 1995 in der Zeitschrift „Scientific American“, Band 272, Nummer 1, Seiten 42-47 erschienen ist.

Rathje, Ulf (2002): *Technisches Konzept für die Datenarchivierung im Bundesarchiv*. In: Der Archivar, H. 2, Jahrgang 55, S.117-120. http://www.archive.nrw.de/archivar/hefte/2002/Archivar_2002-2.pdf

o.V. (o.J.) *Digital preservation*. Calimera Guidelines. http://www.calimera.org/Lists/Guidelines%20PDF/Digital_preservation.pdf

Consultative Committee for Space Data Systems (CCSDS) (2002): *Reference Model for an Open Archival Information System (OAIS)*. Blue Book. Washington DC. Seite 5-1. http://public.ccsds.org/publications/archive/650x0b1.pdf

Van Diessen, Raymond J. und van Rijnsoever, Ben J. (2002): *Managing Media Migration in a Deposit System*. IBM/KB Long-Term Preservation Study Report Series Nr. 5. Amsterdam: IBM Niederlande. http://www-5.ibm.com/nl/dias/resource/migration.pdf

8.3 Migration

Stefan E. Funk

Migration und Emulation

Wenn die Archivierung des Bitstreams sichergestellt ist (siehe Bitstreamerhaltung), kann man beginnen, sich über die Archivierung und vor allem über die Nutzung von digitalen Objekten Gedanken zu machen. Bei nicht digitalen Medien wie Büchern und Mikrofilmen hat man in den letzten Jahrzehnten und Jahrhunderten sehr viel Erfahrung mit deren Erhaltung gesammelt, das heißt, auf physikalischer Ebene konnten und können diese Medien sehr lange verfügbar gehalten werden. Ein Buch braucht als zu erhaltendes Objekt auch nur auf der physischen Ebene betrachtet zu werden, denn zum Benutzen eines Buches reicht es aus, das Buch selbst zu erhalten und so die Lesbarkeit zu gewährleisten.

Zwei Strategien, welche die Lesbarkeit der archivierten digitalen Dokumente über lange Zeit (Long Term) garantieren sollen, sind zum einen die Migration und zum anderen die Emulation. „Long term“ wird vom Consultative Committee for Space Data Systems (CCSDS) definiert als:

> *„Long Term is long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely.”*

Die Migration passt die digitalen Objekte selbst einem neuen Umfeld an, die Dokumente werden zum Beispiel von einem veralteten Dateiformat in ein aktuelles konvertiert. Mit der Emulation wird das originäre Umfeld der digitalen Objekte simuliert, das neue Umfeld also an die digitalen Objekte angepasst. Diese Strategien können alternativ genutzt werden; sie sind unabhängig voneinander.

Um ein digitales Dokument archivieren und später wieder darauf zugreifen zu können, sind möglichst umfassende Metadaten nötig, also Daten, die das digitale Objekt möglichst genau beschreiben. Dazu gehören in erster Linie die technischen Metadaten. Für die Migration sind weiterhin die Provenance Metadaten wichtig, die wie erläutert die Herkunft des Objekts beschreiben. Deskriptive Metadaten sind aus technischer Sicht nicht so interessant. Sie werden benötigt, um später einen schnellen und komfortablen Zugriff auf die Objekte zu ermöglichen. Rechtliche Metadaten können schließlich genutzt werden, um Einschränkungen für die Migration, die Emulation und den Zugriff auf die digitalen Objekte festzulegen.

Migration

Mit dem Stichwort Migration werden innerhalb der Langzeitarchivierungs-Community unterschiedliche Prozesse bezeichnet. Dies sind sowohl die Datenträgermigration als auch die Daten- oder Formatmigration.

Bei der Datenträgermigration werden Daten von einem Träger auf einen anderen kopiert, z.B. von Festplatte auf CD, von DVD auf Band etc. Diese Art der Migration ist die Grundlage der physischen Erhaltung der Daten, der Bitstream Preservation.

Bei einer Datenmigration (auch Formatmigration genannt) werden Daten von einem Datenformat in ein aktuelleres, möglichst standardisiertes und offen gelegtes Format überführt. Dies sollte geschehen, wenn die Gefahr besteht, dass archivierte Objekte aufgrund ihres Formates nicht mehr benutzt werden können. Das Objekt selbst wird so verändert, dass seine Inhalte und Konzepte erhalten bleiben, es jedoch auf aktuellen Rechnern angezeigt und benutzt werden kann. Problematisch ist bei einer Datenmigration der möglicherweise damit einhergehende Verlust an Informationen. So ist es zum Beispiel möglich, dass sich das äußere Erscheinungsbild der Daten ändert oder - noch gravierender - Teile der Daten verloren gehen.

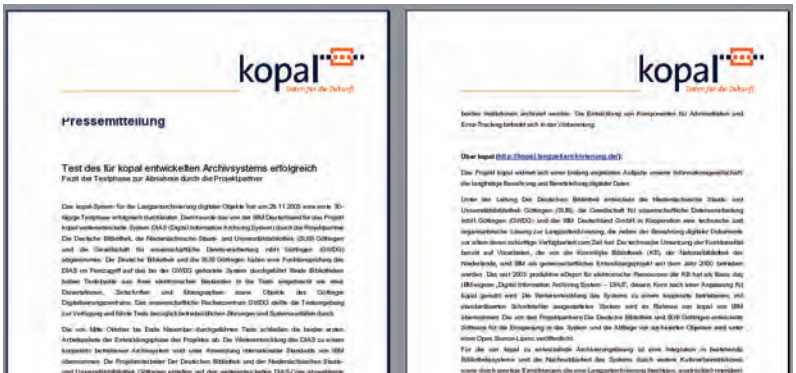
Eine verlustfreie Migration ist dann möglich, wenn sowohl das Original-Format wie auch das Ziel-Format eindeutig spezifiziert sind, diese Spezifikationen bekannt sind UND eine Übersetzung von dem einen in das andere Format ohne Probleme möglich ist. Hier gilt: Je einfacher und übersichtlicher die Formate, desto größer ist die Wahrscheinlichkeit einer verlustfreien Migration. Bei der Migration komplexer Datei-Formate ist ein Verlust an Informationen wahrscheinlicher, da der Umfang einer komplexen Migration nicht unbedingt absehbar ist. Eine Migration eines Commodore-64 Computerspiels in ein heute spielbares Format für einen PC ist sicherlich möglich, jedoch ist es (a) sehr aufwändig, (b) schlecht bzw. gar nicht automatisierbar und (c) das Ergebnis (sehr wahrscheinlich) weit vom Original entfernt.

Beispiel: Alte und neue PCs

- Sie haben einen recht alten PC, auf dem Sie seit langem Ihre Texte schreiben, zum Beispiel mit einer älteren Version von Word 95 (Betriebssystem: Windows 95). Sie speichern Ihre Daten auf Diskette.
- Ihr neuer Rechner, den Sie sich angeschafft haben, läuft unter Windows XP mit Word 2003 und hat kein Diskettenlaufwerk mehr.
- Nun stehen Sie zunächst vor dem Problem, wie Sie Ihre Daten auf den neuen Rechner übertragen. Wenn Sie Glück haben, hat Ihr alter Rechner

schon USB, sodass Sie Ihre Daten mit einem USB-Stick übertragen können. Vielleicht haben Sie auch noch ein Diskettenlaufwerk, auf das Sie zurückgreifen können. Oder aber Ihr alter Rechner kann sich ins Internet einwählen und Ihre Daten können von dort mit dem neuen Rechner heruntergeladen werden. Hier ist unter Umständen ein wenig zu tun. Es gibt jedoch noch genügend Möglichkeiten, Ihre Daten zu übertragen.

- Nehmen wir an, Ihre Daten sind sicher und korrekt übertragen worden. Wenn Sie Glück haben, meldet sich Word 2003 und sagt, Ihre Dateien seien in einem alten .doc-Format gespeichert und müssen in das aktuelle Format konvertiert werden. Diese Konvertierung ist dann eine Migration in ein neues, aktuelleres .doc-Format. Wenn die Migration erfolgreich abläuft, sieht Ihr Dokument aus wie auf dem alten Rechner unter Word 95. Es besteht jedoch die Möglichkeit, dass Ihr Dokument sich verändert hat (Formatierung, Schriftart, Schriftgröße etc.).
- Sollten Sie Pech haben, erkennt Word das alte Format nicht und eine Migration ist nicht automatisch möglich. Dann bleibt noch die Möglichkeit, die alten Dateien mit einem Zwischenschritt über ein anderes Textformat, das beide Textprogramme beherrschen, zu konvertieren. Sicherlich können beide Programme einfache Textdateien verarbeiten (.txt), vielleicht auch Dateien im Rich-Text-Format (.rtf). Sie müssen nun Ihre Dokumente mit dem alten Word alle als Text- oder RTF-Datei neu speichern, diese erneut (wie oben beschrieben) auf den neuen Rechner übertragen und dann mit dem neuen Word (als Text- oder RTF-Datei) wieder öffnen. Sehr wahrscheinlich sind dann sehr viele Formatierungen (Inhaltsverzeichnisse, Überschriften, Schriftdicken, Schriftarten, etc.) verloren gegangen, da eine .txt-Datei keinerlei solcher Dinge speichern kann. Nur der Text entspricht dem originalen Dokument. Mit einer RTF-Datei haben Sie sicherlich weniger Informationsverlust. Sie führen also praktisch zwei Migrationen durch: .doc (Word 95) – .txt (bzw. .rtf) – .doc (Word 2003), siehe hierzu die Abbildungen 1 und 2.



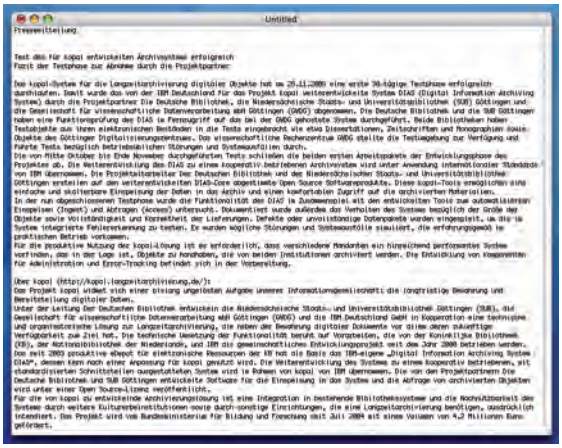
kopal

Pressemitteilung

Test des für kopal entwickelten Archivsystems erfolgreich
Fazit der Testphase zur Abnahme durch die Projektpartner

kopal

Abbildung 1: Ein Word-Dokument mit Grafiken, Formatierungen, Link, etc.



Untitled

Pressemitteilung

Test des für kopal entwickelten Archivsystems erfolgreich
Fazit der Testphase zur Abnahme durch die Projektpartner

Das kopal-System für die Langzeitarchivierung digitaler Objekte hat am 20.11.2003 eine weitere 36-tägige Testphase erfolgreich durchlaufen. Damit wurde das von der IBM Deutschland für das Projekt kopal weiterentwickelte System DIAS (Digital Information Archiving System) durch die Projektpartner Die Deutsche Bibliothek, die Niedersächsische Staats- und Universitätsbibliothek (SUB) Göttingen und die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) abgenommen. Die Deutsche Bibliothek und die SUB Göttingen haben eine Funktionsprüfung des DIAS im Fernzugriff auf das bei der GWDG gehostete System durchgeführt. Beide Bibliotheken haben Testobjekte aus ihren elektronischen Beständen in die Tests eingebracht. Die etwa Dissertationen, Zeitschriften und Monographien sowie Objekte des Göttinger Digitalisierungszentrums. Das wissenschaftliche Rechenzentrum GWDG stellte die Testumgebung zur Verfügung und führte Tests bezüglich betriebsbedingter Störungen und Systemausfällen durch.
Die von Mitte Oktober bis Ende November durchgeführten Tests schlossen die beiden ersten Arbeitspakete der Entwicklungsphase des Projektes ab. Die Weiterentwicklung des DIAS zu einem kooperativ betriebenen Archivsystem wird unter Anwendung internationaler Standards von IBM übernommen. Die Projektpartner der Deutschen Bibliothek und der Niedersächsischen Staats- und Universitätsbibliothek Göttingen erstellen auf den weiterentwickelten DIAS-Core angepasste Open Source Softwareprodukte. Diese kopal-Tools ermöglichen eine einfache und skalierbare Einspeisung der Daten in das Archiv und einen komfortablen Zugriff auf die archivierten Materialien.
In der nun abgeschlossenen Testphase wurde die Funktionalität des DIAS im Zusammenspiel mit den entwickelten Tools zum automatisierten Einspeisen (Ingest) und Abfragen (Access) untersucht. Dokumentiert wurde außerdem das Verhalten des Systems bezüglich der Größe der Objekte sowie Vollständigkeit und Korrektheit der Lieferungen. Defekte oder unvollständige Datenpakete wurden eingespielt, um die im System integrierte Fehlererkennung zu testen. Es wurden mögliche Störungen und Systemausfälle simuliert, die erfahrungsgemäß im praktischen Betrieb vorkommen.
Für die produktive Nutzung der kopal-Lösung ist es erforderlich, dass verschiedene Mandanten ein hinreichend performantes System vorfinden, das in der Lage ist, Objekte zu handhaben, die von beiden Institutionen archiviert werden. Die Entwicklung von Komponenten für Administration und Error-Tracking befindet sich in der Vorbereitung.

Über kopal (http://kopal.langzeitarchivierung.de/):
Das Projekt kopal widmet sich einer einzig angelegten Aufgabe unserer Informationsgesellschaft: die langfristige Bewahrung und Bereitstellung digitaler Daten.
Unter der Leitung Der Deutschen Bibliothek entwickeln die Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB), die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) und der IBM Deutschland GmbH in Kooperation eine technische und organisatorische Lösung zur Langzeitarchivierung, die neben der Bewahrung digitaler Dokumente vor allem deren zukünftige Verfügbarkeit zum Ziel hat. Die technische Umsetzung der Funktionalität bereits auf Vorarbeiten, die von der Koninklijke Bibliotheek (KB), der Nationalbibliothek der Niederlande, und IBM als gemeinschaftliches Entwicklungsprojekt seit dem Jahr 2000 betrieben wurden. Das seit 2003 produktive eDepot für elektronische Ressourcen der KB hat als Basis das IBM-eigene „Digital Information Archiving System (DIAS)", dessen Kern nach einer Anpassung für kopal genutzt wird. Die Weiterentwicklung des Systems zu einem kooperativ betriebenen, mit standardisierten Schnittstellen ausgestatteten System wird im Rahmen von kopal von IBM übernommen. Die von den Projektpartnern Die Deutsche Bibliothek und SUB Göttingen entwickelte Software für die Einspeisung in das System und die Abfrage von archivierten Objekten wird unter einer Open Source-Lizenz veröffentlicht.
Für die von kopal zu entwickelnde Archivierungslösung ist eine Integration in bestehende Bibliothekssysteme und die Nachnutzbarkeit des Systems durch weitere Kulturarbeitsinstitutionen sowie durch sonstige Einrichtungen, die eine Langzeitarchivierung benötigen, ausdrücklich intendiert. Das Projekt wird vom Bundesministerium für Bildung und Forschung seit Juli 2004 mit einem Volumen von 4,2 Millionen Euro gefördert.

Abbildung 2: Das selbe Dokument im .txt-Format ohne Formatierungen

Beispiel: Zeichenkodierungen

- Eine Organisation, die in den 80er Jahren ihre Daten mit IBM Mainframes bearbeitet hat, möchte diese Daten auch auf späteren Systemen nutzen können. Die IBM Mainframes nutzten einen Zeichenstandard namens EBCDIC.[7]
- In den 90er Jahren installierte Rechner nutzten den ASCII Zeichencode (American National Standard Code for Information Interchange), welcher nicht alle Zeichen des EBCDIC abdeckte. Die Organisation musste sich nun entscheiden, ob sie alle Dokumente nach ASCII konvertierten (und einen permanenten Verlust von Daten hinnahmen) oder sie nur bei Bedarf in ASCII umwandelten und die Originaldaten in EBCDIC beließen. So hatte man den gleichen Verlust beim Umwandeln, jedoch für spätere Zeit die Originaldaten erhalten.
- Zum Jahrtausendwechsel begann UNICODE[8] die Welt zu erobern und tatsächlich enthält UNICODE alle Zeichen des EBCDIC, sodass nun alle Dokumente 1:1 von EBCDIC in UNICODE konvertiert werden konnten (sofern die Originaldateien noch existierten!). Bei einer sofortigen Konvertierung in ASCII wären tatsächlich Daten verloren gegangen.

Zusammenfassung: Vor- und Nachteile von Migration

Vorteile von Migration

- Migration ist technisch (verglichen mit Emulation) gut zu realisieren.
- Migration kann in vielen Fällen automatisiert werden.
- Die migrierten Dokumente sind unabhängig von weiteren Komponenten (abgesehen von der aktuellen Darstellungssoftware).
- Die originalen Objekte können aufbewahrt werden, um evtl. später darauf zurückgreifen zu können.

Nachteile von Migration

- Jedes Objekt muss einzeln migriert werden.
- Die Wahrscheinlichkeit von Datenverlust bzw. Datenveränderung ist (besonders über mehrere Migrationsschritte) sehr hoch.
- Jede Version (Migration) eines Objekts inklusive des Original-Dokuments sollte gespeichert werden. Damit ist unter Umständen ein hoher Speicherplatzbedarf verbunden.

---

7 Extended Binary Coded Decimal Interchange Code: http://www.natural-innovations.com/computing/asciiebcdic.html

8 http://www.unicode.org

- Für jedes Format und für jeden Migrations-Schritt muss es ein Migrations-Werkzeug geben.
- Migration ist nicht für alle Formate realisierbar.

Literatur

Consultative Committee for Space Data Systems (2001): *Reference Model for an Open Archi-val Information System (OAIS),* CCSDS 650.0-B-1, BLUE BOOK, http://public.ccsds.org/publications/archive/650x0b1.pdf

Jenkins, Clare (2002): *Cedars Guide to: Digital Preservation Strategies*, http://www.webarchive.org.uk/wayback/archive/20050409230000/http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html

8.4 Emulation

Stefan E. Funk

Mit Emulation (Nachbildung, Nachahmung, von lat. aemulator = Nacheiferer) versucht man die auftretenden Verluste einer Datenformatmigration zu umgehen, indem man die originale Umgebung der archivierten digitalen Objekte nachbildet. Emulation kann auf verschiedenen Ebenen stattfinden:

- zum einen auf der Ebene von Anwendungssoftware,
- zum anderen auf der Ebene von Betriebssystemen und zu guter Letzt
- auf der Ebene von Hardware-Plattformen.

So kann zum Beispiel die originale Hardware des digitalen Objekts als Software mit einem Programm nachgebildet werden, welches das archivierte Betriebssystem und die darauf aufbauenden Softwarekomponenten laden kann (Emulation von Hardware-Plattformen). Ein Beispiel für die Emulation von Betriebssystemen wäre ein MS-DOS-Emulator[9], der die Programme für dieses schon etwas ältere Betriebssystem auf aktuellen Rechnern ausführen kann. Ein Beispiel für den ersten Fall wäre etwa ein Programm zum Anzeigen und Bearbeiten von sehr alten Microsoft Word-Dateien (.doc), die das aktuelle Word nicht mehr lesen kann. Auf diese Weise wird die Funktionalität dieser alten und nicht mehr verfügbaren Soft- oder Hardware emuliert und die Inhalte bzw. die Funktionalität der damit erstellten Dokumente erhalten. Im Gegensatz zur Migration, bei der jeweils eine neue und aktuellere Version des digitalen Objektes selbst erzeugt wird, werden die originalen Objekte bei der Emulation nicht verändert. Stattdessen muss man für jede neue Hardwarearchitektur die Emulationssoftware anpassen, im schlechtesten Fall muss diese jedes Mal neu entwickelt werden. Wenn das aus irgendeinem Grund nicht geschieht, ist der komplette Datenbestand der betroffenen Objekte unter Umständen nicht mehr nutzbar und damit für die Nachwelt verloren.

Emulation von Anwendungssoftware

Da es um die Darstellung der digitalen Dokumente geht, die wir vorhin beschrieben haben, ist die Emulation der Software, die mit diesen Dokumenten arbeitet, eine erste Möglichkeit. So kann auf einem aktuellen System ein Pro-

9 DOS – Disc Operating System, näheres unter: http://www.operating-system.org/betriebssystem/_german/bs-msdos.htm

gramm entwickelt werden, das archivierte digitale Objekte in einem bestimmten Format öffnen, anzeigen oder bearbeiten kann, auf die mit aktueller Software auf diesem System nicht mehr zugegriffen werden kann, weil vielleicht die Originalsoftware nicht mehr existiert oder auf aktuellen Systemen nicht mehr lauffähig ist.

Wenn wir zum Beispiel eine PDF-Datei aus dem Jahr 1998, Version 1.2, darstellen möchten, und der aktuelle Acrobat Reader 7.0 stellt das Dokument nicht mehr richtig dar, müssen wir einen PDF-Reader für diese PDF-Version auf einem aktuellen Betriebssystem programmieren, sprich: einen alten PDF-Reader emulieren. Dieser sollte dann alle PDF-Dateien der Version 1.2 darstellen können. Für jeden Generationswechsel von Hardware oder Betriebssystem würde so ein PDF-Reader benötigt, um den Zugriff auf die PDF-Dokumente in Version 1.2 auch in Zukunft zu gewährleisten. Die genaue Kenntnis des PDF-Formats ist hierzu zwingend erforderlich.

Emulation von Betriebssystemen und Hardware-Plattformen

Bei einigen Anwendungen kann es sinnvoll sein, eine komplette Hardware-Plattform zu emulieren, zum Beispiel wenn es kein einheitliches Format für bestimmte Anwendungen gibt. Hier ist der Commodore-64 ein gutes Beispiel. Die Spiele für den C-64 waren eigenständige Programme, die direkt auf dem Rechner liefen, soll heissen, es wird direkt die Hardware inklusive des Betriebssystems[10] benötigt und nicht ein Programm, das diese Spiele ausführt (wie ein PDF-Viewer).

Es muss also ein Commodore-64 in Software implementiert werden, der sich genau so verhält wie die Hardware und das Betriebssystem des originalen Commodore-64 und auf einem aktuellen Computersystem lauffähig ist. Diese C-64-Emulatoren gibt es für nahezu alle aktuellen Computersysteme und auch weitere Emulatoren für andere ältere Systeme sind erhältlich.[11]

Die Emulation eines Betriebssystems oder einer Hardware-Plattform ist eine sehr komplexe Sache, die schon für einen C-64-Emulator sehr viel Arbeit bedeutet. Man kann jedoch auch die Hardware eines PC in Software nachbilden, um dann auf einem solchen virtuellen PC beliebige Betriebssysteme und die auf ihnen laufenden Anwendungsprogramme oder auch Spiele zu starten (die Be-

10 Eine Trennung von Hardware und Betriebssystem ist beim Commodore-64 nicht nötig, da diese beiden Komponenten sehr eng zusammenhängen. Auch andere „Betriebssysteme" des C-64, wie zum Beispiel GEOS, setzen direkt auf das Betriebssystem des C-64 auf.

11 Hier einige Adressen im Internet zum Thema Emulatoren: http://www.aep-emu.de/, http://www.homecomputermuseum.de/

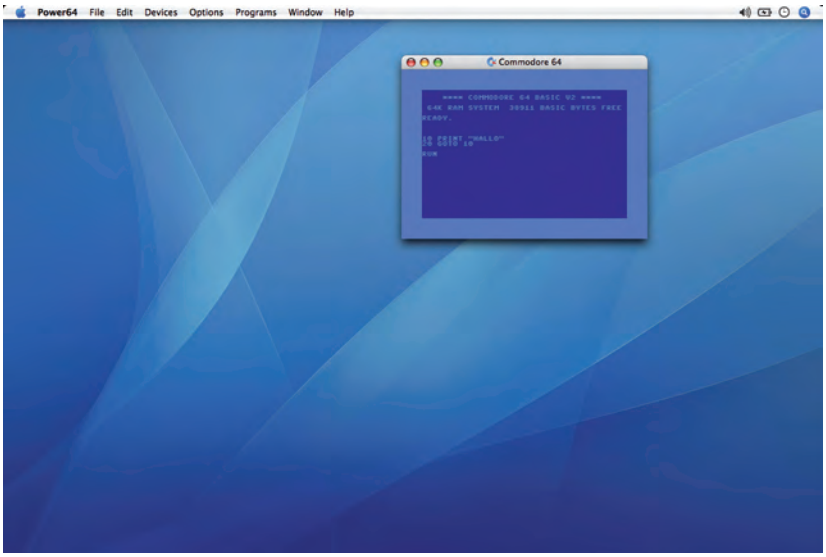


Abbildung 1: Power 64, ein Commodore-64 Emulator für Mac OS X

triebssysteme wie auch die Programme bleiben dann im Originalzustand). Dies bedeutet im Allgemeinen, dass eine gute Performanz auf der aktuellen Hardware vorhanden sein muss. Eine Emulation eines Commodore-64 auf einem aktuellen PC ist jedoch keine performanzkritische Anwendung. Für zukünftige Computersysteme, die unsere heutigen emulieren sollen, wird im Allgemeinen davon ausgegangen, dass deren Performanz weitaus höher ist als heute, sodass auch hier die Performanz für eine erfolgreiche Emulation ausreichen dürfte.

Beispiel: Migration und Emulation alter C-64 Programme

- Da der Commodore 64 ein sehr beliebter und weit verbreiteter Homecomputer war, gibt es sehr viele Emulatoren für nahezu alle aktuellen Computersysteme. Viele Videospiele, die es für den C-64 gab, sind im Internet als C-64 Disk-Image zu finden. Die darin enthaltenen Programme können dann mit den Emulatoren geladen und genutzt werden. Als alter C-64 Nutzer stand ich also nicht vor dem Problem, meine Spiele von alten 5,25-Zoll Disketten auf neuere Datenträger migrieren zu müssen. Ein Emulator für den Apple unter Mac OS X ist Power64[12], siehe Abbildung 1.
- Anders sah es hingegen für die Programme aus, die ich vor mehr als 20 Jahren auf dem C-64 selbst programmiert habe. Es handelt sich hier

12 http://www.infinite-loop.at/Power64/index.html

um viele Programme in Commodore-64 BASIC. Die Frage, die sich mir stellte, war nun die, ob und wie ich diese Daten von meinen alten (auf dem Original C-64 noch laufenden) 5,25 Zoll-Disketten von 1982 bis 1987 auf die Festplatte meines PC kopieren und ich diese Daten auch für den C-64-Emulator nutzen kann.

- Der erste Versuch, einfach ein vor einigen Jahren noch gebräuchliches 5,25 Zoll-Laufwerk[13] an den PC anzuschließen und die C-64 Daten am PC auszulesen, schlug zunächst einmal fehl. Grund hierfür waren die unterschiedlichen Dichten und die unterschiedlichen Dateisysteme der 5,25 Zoll-Disketten. Auf eine Diskette des C-64 war Platz für 170 KB, damals einfache Dichte (single density). Die Disketten für den PC hatten jedoch doppelte Dichte (double density) oder gar hohe Dichte (high density), sodass das mit zur Verfügung stehende Diskettenlaufwerk die C-64 Disketten nicht lesen konnte.
- Nach kurzer Recherche entdeckte ich eine Seite im Internet (die Community für den C-64 ist immer noch enorm groß), die Schaltpläne für einige Kabel abbildete, mit denen man seinen PC mit den Diskettenlaufwerken seines C-64 verbinden konnte. Mit Hilfe des Programmes Star Commander[14], das unter DOS läuft, kann man damit seine Daten von C-64 Disketten auf seinen PC kopieren und auch gleich Disk-Images erstellen. Inzwischen kann man solche Kabel auch bestellen und muss nicht selbst zum Lötkolben greifen (Für die Nutzung dieses Programms muss natürlich eine lauffähige DOS-Version zur Verfügung stehen, ist keine verfügbar, kann evtl. eine emuliert werden :-)
- Nach diesen Aktionen kann ich nun meine alten selbst erstellten Programme auf vielen C-64 Emulatoren wieder nutzen, weiterentwickeln und spielen, wie in Abbildung 2 und 3 zu sehen ist (und das sogar auf mehreren virtuellen Commodore-64 gleichzeitig).

13 Den ersten Versuch unternahm ich vor etwa vier Jahren, 5,25-Zoll-Diskettenlaufwerke waren nicht mehr wirklich gebräuchlich, aber noch erhältlich. Heute werden selbst die 3,5-Zoll-Laufwerke schon nicht mehr mit einem neuen Rechner verkauft. Neue Medien zum Datenaustausch und zur Speicherung sind heute USB-Stick, DVD, CD-ROM und Festplatte.

14 http://sta.c64.org/sc.html

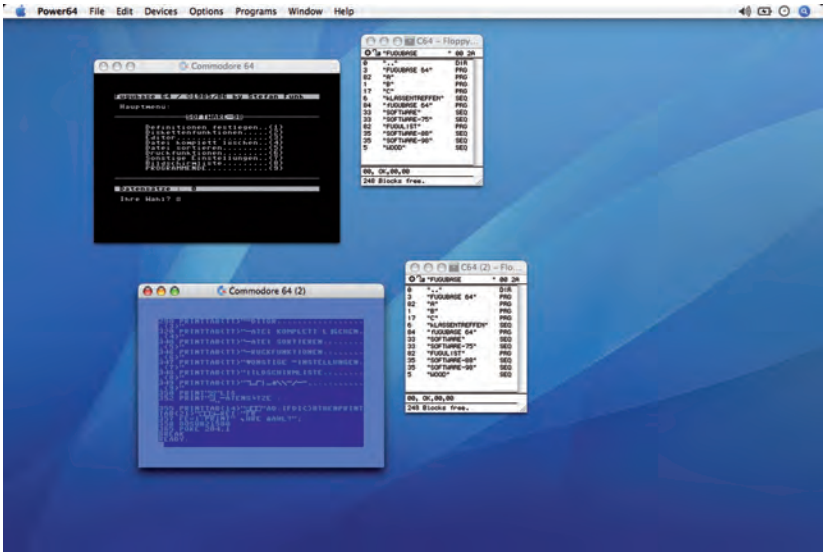


Abbildung 2: Fugubase 64, ein Datenverwaltungs-Programm in Basic für den C-64, emuliert unter Mac OS X (S. E. Funk, 1985/86)

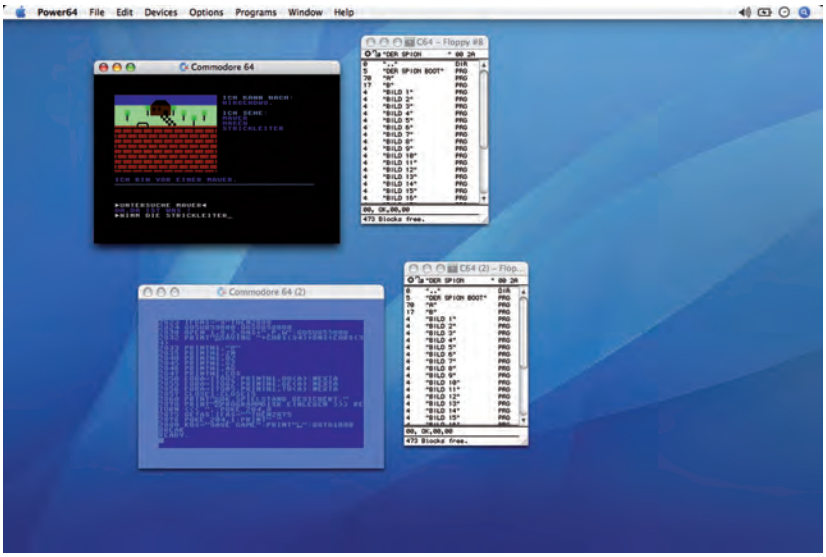


Abbildung 3: Der Spion, ein Adventure in Basic für den C-64, emuliert unter Max OS X (S. E. Funk, 1987)

Beispiel: Eine Emulation in der Emulation

- Es ist nun auch möglich, einen Emulator wiederum zu emulieren, wenn ein weiterer Generationswechsel einer Hardwareplattform ansteht. Ein praktisches Beispiel ist ein Apple Notebook, das unter Mac OS X, einem Unix-basierten Betriebssystem, arbeitet. Auf diesem werden zwei Emulatoren und ein weiteres originales Betriebssystem gestartet.
- Auf diesem Rechner wird das Programm Q gestartet[15], das eine Hardware-Plattform emuliert (einen Pentium x86 mit diversen Grafik-, Sound- und weiteren Hardwarekomponenten). Es basiert auf dem CPU-Emulator QEMU.[16]
- Auf dieser virtuellen Hardwareplattform kann nun ein originales Windows 98 installiert werden, so dass man ein reguläres, altbekanntes Windows 98 auf diesem nicht-Windows-Rechner nutzen kann. Das installierte Windows 98 kann selbstverständlich alle Programme für Windows 98 ausführen, da es sich tatsächlich um ein originales Windows 98 handelt. Sogar ein Windows-Update über das Internet ist möglich.
- Jetzt kann natürlich auch ein C-64 Emulator für Windows, hier der VICE[17], gestartet werden. Darauf laufen nun alle altbekannten und beliebten Commodore-64 Programme.
- Probleme kann es bei dieser Art von Emulation zum Beispiel bei der Performanz geben und je nach Qualität der Emulatoren auch mit hardwarespezifischen Dingen wie Grafik, Sound und angeschlossener Peripherie (Mäuse, Joysticks, etc.). Der C-64 Emulator muss schließlich durch Windows über die virtuelle Hardware (Emulation QEMU) auf die reale Hardware des Notebooks zugreifen. Bei steigender Komplexität solcher Emulationsszenarien wird die Anzahl der möglichen Fehler stark ansteigen. Als Beispiel siehe Abbildung 4.

Der Universal Virtual Computer (UVC)

Mittlerweile gibt es einen elaborierteren Ansatz der Emulation, den Universal Virtual Computer (UVC) von IBM. Der UVC ist ein wohldokumentierter virtueller Computer**,** der auf unterschiedlichen (auch zukünftigen) Architekturen nachgebildet werden kann. Aufgebaut ist er ähnlich wie ein heute existierender Computer, der beispielsweise Speicherzugriff ermöglicht. Mit Hilfe dieser Dokumentation ist es einem Programmierer auch auf zukünftigen Systemen

15 http://www.kju-app.org/
16 http://www.nongnu.org/qemu/
17 http://www.viceteam.org/

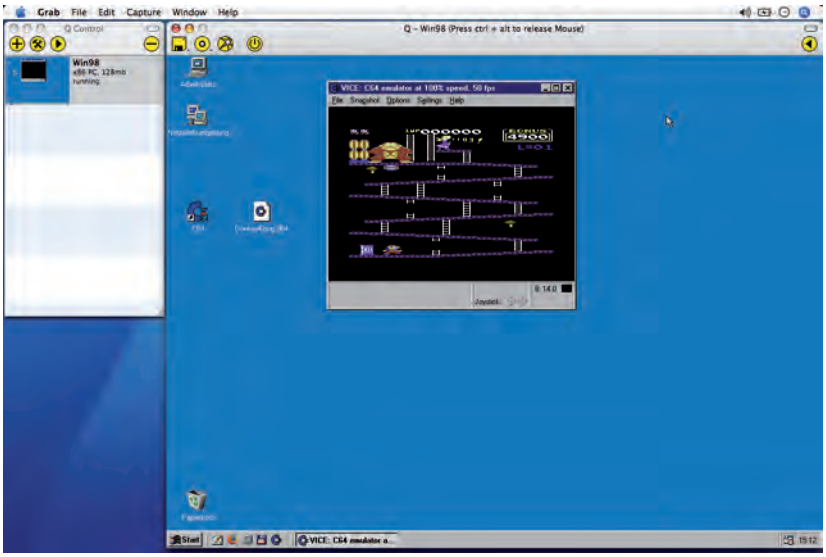


Abbildung 4: Das Videospiel Donkey Kong auf einem C-64 Emulator auf einem Windows 98 auf einem virtuellen Pentium auf einem Apple PowerBook unter Mac OS X

möglich, diesen virtuellen Computer zu implementieren. Auf diesem virtuellen Computer aufbauend können nun Programme geschrieben werden, die zum Beispiel eine PDF-Datei lesen oder Grafiken darstellen können.

Archiviert wird jetzt der PDF-Reader (der Bildbetrachter), der für den UVC programmiert wurde, sowie das originale PDF-Dokument (oder die originale Grafik) selbst. Ein zukünftiger Nutzer kann dann auf einer zukünftigen und wahrscheinlich hoch entwickelten Hardware auch in ferner Zukunft noch mit Hilfe der Dokumentation des UVC einen solchen implementieren und mit Hilfe dieses virtuellen Computers den PDF-Reader starten, mit dem das archivierte PDF-Dokument dargestellt wird. Die Dokumentation muss selbstverständlich erhalten bleiben und lesbar sein.

Ein Problem dieser Idee ist sicherlich, dass bei zunehmendem Anspruch an die Emulation, die auf dem UVC laufen soll, eine Programmierung derselben immer schwieriger wird. Es wird sehr kompliziert, wenn für den UVC ein Betriebssystem wie Linux oder Windows programmiert werden soll, mit dessen Hilfe dann die Applikationen von Linux oder Windows genutzt werden können. Schon eine nachprogrammierte Version eines Textverarbeitungsprogrammes wie zum Beispiel Word, mit dem später alte Word-Dokumente (.doc) auf dem UVC gelesen und bearbeitet werden können, ist ein höchst umfangreiches Un-

ternehmen. Zumal hier nicht nur die Formatbeschreibung, sondern auch alle Programmfunktionen bekannt sein müssen.

Zusammenfassung: Vor- und Nachteile von Emulation

Vorteile von Emulation

- Bei der Emulation bleiben die Originalobjekte unverändert.
- Eine Konvertierung der Objekte ist nicht nötig.
- Für die Emulation wird weniger Speicherplatz benötigt, da keine Migrationen gespeichert werden müssen.

Nachteile von Emulation

- Für komplizierte Objekte/Systeme (wie Betriebssysteme oder Anwendungsprogramme) sind Emulatoren technisch schwer zu implementieren.
- Es entsteht ein hoher Aufwand pro Hardware-Generationswechsel. Es müssen für jede Plattform neue Emulatoren entwickelt werden.
- Die Spezifikationen für die zu emulierenden Objekte/Systeme sind nicht immer hinreichend bekannt.

Literatur

Lorie, Raymond (2002): *the UVC: a method for preserving digital documents – proof of con-cept*, http://www-5.ibm.com/nl/dias/resource/uvc.pdf

Nationaal Archief (2005): *Technical Description of the Universal Virtual Computer (UVC) - Data preservation process for spreadsheets,* http://www.digitaleduurzaamheid.nl/bibliotheek/docs/TDUVCv1.pdf

Oltmans, Eric; Nanda Kol (2005): *A Comparison Between Migration and Emulation in Terms of Costs*, http://www.worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file959.html#article0

8.5 Computermuseum

Karsten Huth

Die Erhaltung von Hardware ist auf lange Sicht keine vielversprechende Strategie zur Bewahrung von digitalen Objekten. Solange aber keine technischen Möglichkeiten zum Transfer von Daten aus den alten Systemen auf aktuelle Plattformen sowie zur Migration oder Emulation zur Verfügung stehen, ist der Erhalt des originalen Systems ein notwendiger erster Schritt zur Erhaltung eines digitalen Bestandes. Zudem gibt es auch Museen, die ihren Besuchern den Eindruck der historischen Software mitsamt der historischen Hardware vermitteln wollen. Dieser Artikel gibt eine kurze Einführung in die Probleme, die sich vor allem durch die Auflösungserscheinungen der Werkstoffe ergeben.

Definition

Auch wenn man die Strategie der „Hardware Preservation", also der Erhaltung von Hardware, als Methode zur Langzeitarchivierung auf keinen Fall empfehlen sollte, so ist es leider alltägliche Praxis, dass digitale Langzeitarchive auch obsolete Hardware vorhalten müssen, zumindest bis sie in der Lage sind, besser geeignete Strategien durchzuführen. Aber gerade in den Anfängen eines digitalen Archivs, wenn es noch über keinen geregelten Workflow verfügt, werden digitale Objekte oft auf ihren originalen Datenträgern oder mitsamt ihrer originalen Hardware/Software-Umgebung abgeliefert. Dies betrifft vor allem digitale Objekte, die technologisch obsolet geworden sind. Deshalb sind in der Praxis, wenn auch ungewollt, Computermuseen eher die Regel als eine Ausnahme.

Leider hat sich der Begriff „Computermuseum" im deutschen Sprachraum verfestigt. Passender wäre der Begriff „Hardware-/Software-Konservierung", denn die konservierten Computer müssen nicht unbedingt nur im Rahmen eines Museums erhalten werden. Man muss vielmehr differenzieren zwischen:

1. Hardware Preservation als Strategie zur Archivierung von digitalen Objekten:
   Eigentliches Ziel ist die Erhaltung der digitalen Objekte. Zu diesem Zweck versucht man die ursprüngliche Hardware/Software-Plattform so lange wie möglich am Laufen zu halten.

2. Hardware Preservation im Rahmen eines Technikmuseums:
   Wird im ersten Fall die Hardware/Software-Plattform nur erhalten, um den Zugriff auf die digitalen Objekte zu ermöglichen, so ist hier die

ursprüngliche Hardware/Software Plattform das zentrale Objekt der konservatorischen Bemühungen. Während im ersten Fall Reparaturen an der Hardware einzig der Lauffähigkeit der Rechner dienen, so fallen im Rahmen eines Technikmuseums auch ethische Gesichtspunkte bei der Restaurierung ins Gewicht. Die Erhaltung der Funktion ist bei einer Reparatur nicht mehr das einzige Kriterium, es sollten auch möglichst die historisch adäquaten Bauteile verwendet werden. Diese Auflage erschwert die beinahe unmögliche Aufgabe der Hardware-Konservierung noch zusätzlich.

Bei einem technischen Museum liegt die Motivation zur Konservierung von Hardware auf der Hand. Die historische Hardware zusammen mit der originalen Software sind die Sammelobjekte und Exponate des Museums. Deswegen müssen sie solange wie möglich in einem präsentablen Zustand erhalten werden. Daneben gibt es aber auch noch weitere Gründe, die für die Hardware Preservation als Archivierungsstrategie sprechen.

Gründe zur Aufrechterhaltung eines Computermuseums

- Keine andere Strategie erhält soviel vom intrinsischen Wert der digitalen Objekte (Look and Feel). An Authentizität ist dieser Ansatz nicht zu übertreffen.[18]
- Bei komplexen digitalen Objekten, für die Migration nicht in Frage kommt, und eine Emulation der Hardware/Software Umgebung noch nicht möglich ist, ist die Hardware Preservation die einzige Möglichkeit, um das Objekt zumindest für einen Übergangszeitraum zu erhalten.[19]
- Zur Unterstützung von anderen Archivierungsstrategien kann die zeitweise Erhaltung der originalen Plattformen notwendig sein. Man kann z.B. nur durch einen Vergleich mit der ursprünglichen Hardware/Software-Plattform überprüfen, ob ein Emulatorprogramm korrekt arbeitet oder nicht.[20]

---

18 Borghoff, Uwe M. et al. (2003): *Methoden zur Erhaltung digitaler Dokumente*. 1. Aufl. Heidelberg: dpunkt-Verl., 2003: S. 16-18

19 Jones, Maggie/ Beagrie, Neil (o.J.): *Preservation Management of Digital Materials: A Handbook*. Digital Preservation Coalition.

20 Rothenberg, Jeff (1998): *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation: A Report to the Council on Library and Information Resources*. Washington D.C.: Council on Library and Information Resources: S. 12-13 http://www.clir.org/pubs/reports/rothenberg/inadequacy.html

Probleme der Hardware Preservation

Ob man ein Hardware-Museum aus dem ersten oder dem zweiten Grund führt, in beiden Fällen hat man mit den gleichen Problemen zu kämpfen. Zum einen ergeben sich auf lange Sicht gesehen große organisatorische und zum anderen rein technische Probleme der Konservierung von Hardware und Datenträgern.

1. Organisatorische Probleme:

- Die Menge an zu lagerndem und zu verwaltendem Material wird stetig wachsen. Da nicht nur die Rechner, sondern auch Peripheriegeräte und Datenträger gelagert werden müssen, steigen Platzbedarf und Lagerungsaufwand enorm an. „Selbst heute schon erscheint es unrealistisch, sämtliche bisher entwickelten Computertypen in einem Museum zu versammeln, geschweige denn dies für die Zukunft sicher zu stellen.“[21]
- Techniker und Experten, die historische Computer bedienen und gegebenenfalls reparieren können, werden über kurz oder lang nicht mehr zur Verfügung stehen. Mit wachsendem Bestand müssten die Mitarbeiter des Museums ihr Fachwissen ständig erweitern, oder der Bedarf an Technikexperten und neuen Mitarbeitern würde ständig wachsen.[22]
- Die Nutzung der digitalen Objekte ist nur sehr eingeschränkt möglich. Da die obsoleten Computersysteme von der aktuellen Technologie abgeschnitten sind, könnte der Nutzer nur im Computermuseum auf die Objekte zugreifen.[23]

2. Technische Probleme:

- Die technischen Geräte und Bausteine haben nur eine begrenzte Lebenserwartung. Da für obsolete Systeme keine Ersatzteile mehr produziert werden, ist die Restaurierung eines Systems irgendwann nicht mehr möglich.[24]
- Neben der Hardware muss auch die originale Softwareumgebung erhalten und archiviert werden. Diese muss natürlich auf den entsprechenden Datenträgern vorgehalten werden. Da Datenträger ebenso wie die Hardware nur eine begrenzte Lebensdauer haben, müssen die Software und die Daten von Zeit zu Zeit auf neue, frischere Datenträger des gleichen

21 s. Borghoff (2003); a.a.O.

22 Dooijes, Edo Hans (200): *Old computers, now and in the future*. Department of Computerscience/University of Amsterdam. http://www.science.uva.nl/museum/pdfs/oldcomputers_dec2000.pdf

23 s. Rothenberg (1998), a.a.O.

24 s. Borghoff (2003), a.a.O.

Typs, oder zumindest auf passende Datenträger des gleichen Computersystems umkopiert werden. Da jedoch Datenträger eines obsoleten Systems nicht mehr hergestellt werden, stößt diese Praxis zwangsläufig an ihre Grenze und Software und Daten gehen verloren.[25]

Auftretende Schäden bei der Lagerung

Es gibt wenig Literatur über die tatsächlich in der Praxis auftretenden Schäden. Der folgende Abschnitt bezieht sich auf eine Umfrage in Computermuseen. Diese Umfrage war Teil einer Abschlussarbeit an der San Francisco State University im Fach Museum Studies. Die folgende Aufzählung ist eine vorläufige Rangliste der auftretenden Probleme.[26]

- Zerfall von Gummiteilen: Gummi wird für viele Bauteile der Hardware verwendet. Riemen in Motoren, Rollen in Magnetbänderlaufwerken, Lochkartenleser und Drucker, um nur einige Beispiele zu nennen. Gummi ist anfällig für Oxidation. Harte Oberflächen werden durch Oxidation weich und klebrig. Mit fortschreitendem Zerfall kann der Gummi wieder verhärten und dabei brüchig werden.
- Zerfall von Schaumstoffisolierungen: Schaumstoff wird hauptsächlich zur Lärmisolierung und Luftfilterung in Computern verwendet. Vor allem Schaumstoff aus Polyurethan ist sehr anfällig für eine ungewollte Oxidation. Das Material verfärbt sich zunächst und zerfällt dann in einzelne Krümel.
- Verfärbung von Plastikteilen: UV-Licht verändert die chemische Zusammensetzung der Plastikgehäuse. Die Funktion des Geräts wird dadurch zwar nicht beeinträchtigt, aber die Farbe des Gehäuses verändert sich merklich ins Gelb-bräunliche.
- Schäden durch Staub: Staub greift sowohl das Äußere der Hardware als auch ihr Innenleben an. Staub ist nur eine grobe Umschreibung für eine Vielzahl an Schadstoffen, wie z.B. Ruß, Ammoniumnitrat, Ammoniumsulfat und Schwefelsäure. Mit dem Staub lagert sich Salz und Feuchtigkeit an den Bauteilen ab. Dadurch wird die Anfälligkeit für Rost oder Schimmel erhöht. Lüfter mit Ventilatoren zur Kühlung von Prozessoren ziehen den Staub in das Gehäuse des Rechners.
- Zerfall der Batterien: Leckende Batterien können das Innenleben eines Rechners zerstören. Batterien sind Behälter bestehend aus Metall und

25 s. Rothenberg (1998), a.a.O.

26 Gibson, Mark A. (2006): *The conservation of computers and other high-tech artifacts . Unique problemes and long-term solutions*: Thesis M.A. San Francisco : San Francisco State University

Metalloxid, eingetaucht in eine Flüssigkeit oder ein Gel aus Elektrolyten. Batterien sind sehr anfällig für Rost. Bei extrem unsachgemäßer Behandlung können sie sogar explodieren. Austretende Elektrolyte können Schaltkreise zersetzen.

- Korrosion: Metall ist ein häufiger Werkstoff in elektronischen Geräten. Metall wird vor allem für das Gehäuse sowie für Klammern, Schrauben und Federn verwendet.
- Beschädigte Kondensatoren: Ähnlich wie bei einer Batterie ist ein Elektrolyt wesentlicher Bestandteil eines Kondensators. Der Elektrolyt kann eine Flüssigkeit, eine Paste oder ein Gel sein. Problematisch wird es, wenn der Elektrolyt austrocknet, da dann der Kondensator nicht mehr arbeitet. Trocknet der Elektrolyt nicht aus, kann der Kondensator lecken, so dass der Elektrolyt austritt und ähnlichen Schaden anrichtet, wie eine kaputte Batterie. Kondensatoren, die lange ungenutzt bleiben, können explodieren.
- Zerfall des Plastiks: Plastik löst sich über einen längeren Zeitraum hinweg auf. Der sogenannte Weichmacher, ein chemischer Stoff, der bei der Produktion beigemengt wird, tritt in milchartigen Tropfen aus dem Material aus. Bei bestimmten Plastiksorten riecht die austretende Feuchtigkeit nach Essig. Der Prozess beeinträchtigt auch die Haltbarkeit von anderen Materialien, die mit dem zerfallenden Plastik verbunden sind.
- Schimmel: Bei einigen Monitoren aus den siebziger und achtziger Jahren kann Schimmel an der Innenseite der Mattscheibe auftreten.

Stark gefährdete Geräte und Bauteile

Von den oben genannten möglichen Schäden sind die folgenden Bauteile am häufigsten betroffen:

- Schaltkreise, die auf Dauer ausfallen.
- Kondensatoren, die ausfallen oder explodieren.
- Ausfall von batteriebetriebenen Speicherkarten und EPROMs, sowie damit einhergehender Datenverlust.
- Durch kaputte Gummirollen zerstörte Kartenleser und Magnetbandlaufwerke.
- Verstaubte und verschmutzte Kontakte.
- Gebrochene oder verloren gegangene Kabel.[27]

27 s. Dooijes (2000), a.a.O.

Gesundheitsschädliche Stoffe und Risiken

Zu beachten ist, dass Restauratoren mit gesundheitsgefährdenden Stoffen am Arbeitsplatz in Kontakt kommen können. Welche Stoffe in Frage kommen, hängt vom Alter und der Bauart der Hardware ab. Dokumentiert ist das Auftreten von:

- Quecksilber
- Blei (auch bleihaltige Farbe)
- Polychloriertem Biphenyl (PCB)
- Thorium und anderen radioaktiven Substanzen
- Asbest
- Cadmium

Besondere Vorsicht ist beim Umgang mit Batterien (vor allem defekten, leckenden Batterien) und Kondensatoren geboten. Abgesehen davon, dass Kondensatoren oft gesundheitsgefährdende Stoffe enthalten, können sie auch in stillgelegtem Zustand über Jahre hin eine hohe elektrische Spannung aufrechterhalten. Wenn Kondensatoren nach längerer Zeit wieder unter Strom gesetzt werden, können sie explodieren.[28]

Empfehlung zur Lagerung und Restaurierung:

Die Hardware sollte bei der Lagerung möglichst vor Licht geschützt werden. Ideal ist ein Helligkeitswert um 50 Lux. Fensterscheiben sollten die ultraviolette Strahlung herausfiltern. Dadurch wird der Zerfall von Plastik und Gummi verlangsamt. Ebenso ist eine möglichst niedrige Raumtemperatur, unter 20°C, sowie eine relative Luftfeuchtigkeit von unter 50% ratsam. Beides verlangsamt den Zerfall von Gummi und Plastik. Die niedrige Luftfeuchtigkeit verringert die Wahrscheinlichkeit von Korrosion. Vor der Inbetriebnahme eines Rechners sollte abgelagerter Staub durch vorsichtiges Absaugen entfernt werden. Dabei ist erhöhte Sorgfalt geboten, damit keine elektrostatische Energie die Schaltkreise beschädigt und keine wichtigen Teile mit eingesaugt werden. Mit einer zuvor geerdeten Pinzette können gröbere Staubknäuel beseitigt werden. Batterien sollten während der Lagerung möglichst aus der Hardware entfernt werden. Weitverbreitete Batterietypen sollten nicht gelagert werden. Wenn die Hardware in Betrieb genommen wird, werden frische Batterien des betreffenden Typs eingesetzt. Seltene, obsolete Batterietypen sollten separat gelagert werden. Alle genannten Maßnahmen können den Zerfall der Hardware jedoch

28 s. Gibson (2006), a.a.O.

nur verlangsamen. Aufzuhalten ist er nicht. Defekte Bauteile werden oft durch das Ausschlachten von Hardware gleicher Bauart ersetzt. Dabei werden alle intakten Teile zu einer funktionierenden Hardwareeinheit zusammengefügt. Natürlich stößt dieses Verfahren irgendwann an seine Grenzen.

Bereits eingetretene Schäden sollten durch Restaurierungsarbeiten abgemildert werden. Auslaufende Flüssigkeiten aus Kondensatoren oder Batterien sollte man umgehend mit Isopropanol-Lösung entfernen.

Dokumentation

Ein Computermuseum kommt natürlich um die korrekte Verzeichnung seiner Artefakte (Hardware und Software) nicht herum. Zusätzlich werden Informationen über den Betrieb, die Bedienung und die verwendete Technik der Hardware und Software benötigt. Des Weiteren sollten Informationen über den Erhaltungszustand und potentiell anfällige Bauteile der Hardware erhoben und gesammelt werden. Wie bei anderen Erhaltungsstrategien fallen auch hier Metadaten an, die gespeichert und erschlossen werden wollen. Schon bei der Aufnahme eines obsoleten Systems in das Archiv sollte darauf geachtet werden, dass die notwendigen Zusatzinformationen verfügbar sind (z.B. Betriebshandbücher über die Hardware/Software, technische Beschreibungen und Zeichnungen usw.). Da diese Informationen bei älteren Systemen meistens nur in gedruckter Form vorliegen, sollte auch hier Raum für die Lagerung mit einkalkuliert oder eine Digitalisierung der Informationen erwogen werden.[29]

Beispieldaten des Computerspiele Museums Berlin

Die Softwaresammlung umfasst zurzeit 12.000 Titel über eine Zeitspanne von 1972 bis heute. Die Software wird getrennt von der Hardware in normalen Büroräumen gelagert und hat einen Platzbedarf von ca. 70 qm.

In der Hardwaresammlung des Computerspiele Museums befinden sich augenblicklich 2180 Sammlungsstücke. Sie sind in einer Datenbank inklusive Foto erfasst und inventarisiert. Die Sammlung besteht aus Videospielautomaten, Videospiele Konsolen, Heimcomputern, Handhelds, technischen Zusatzteilen (Laufwerke, Controller, Monitore etc.). Des Weiteren besitzt das Museum eine umfangreiche Sammlung gedruckter Informationen wie Computerspiele, Magazine und Handbücher. Diese sind in einer gesonderten Datenbank erfasst. Die Hardwaresammlung ist auf ca. 200 qm an der Peripherie Berlins untergebracht. Der Hauptgrund dafür ist die günstigere

29 s. Dooijes (2000), a.a.O.

Miete für die Räume, als das in zentralerer Lage möglich wäre. Die Räume sind beheizbar und entsprechen größtenteils ebenfalls Bürostandard.[30]

30 Daten stammen von Herrn Andreas Lange, Kurator des Computerspielemuseums Berlin (2006)

8.6 Mikroverfilmung

Christian Keitel

Die Eignung des Mikrofilms als analoger oder digitaler Datenträger für digital vorliegende Bildinformation wird diskutiert, notwendige Rahmenbedingungen werden benannt.

Ein ungelöstes Problem bei der langfristigen Archivierung digitaler Informationen ist die begrenzte Haltbarkeit digitaler Datenträger. Künstliche Alterungstests sagen CDs, DVDs und Magnetbändern nur eine wenige Jahre währende Haltbarkeit voraus, während herkömmliche Trägermedien wie z.B. Pergament oder Papier mehrere Jahrhunderte als Datenspeicher dienen können. Hervorragende Ergebnisse erzielt bei diesen Tests insbesondere der Mikrofilm. Bei geeigneter (kühler) Lagerung wird ihm eine Haltbarkeit von über 500 Jahren vorausgesagt. Verschiedene Projekte versuchen daher, diese Eigenschaften auch für die Archivierung genuin digitaler Objekte einzusetzen. Neben der Haltbarkeit des Datenträgers sind dabei auch Aspekte wie Formate, Metadaten und Kosten zu bedenken.

In Anlehnung an die Sicherungs- und Ersatzverfilmung herkömmlicher Archivalien wurden zunächst digitale Informationen auf Mikrofilm als Bild ausbelichtet und eine spätere Benutzung in einem geeigneten Lesegerät (Mikrofilmreader) geplant. Das menschliche Auge benötigt zur Ansicht dieser Bilder nur eine Lupe als optische Vergrößerungshilfe. Erinnert sei in diesem Zusammenhang an das in den Anfängen des EDV-Einsatzes in Bibliotheken übliche COM-Verfahren (Computer Output on Microfilm/-fiche) zur Produktion von Katalog-Kopien. In letzter Zeit wird zunehmend von einer Benutzung im Computer gesprochen, was eine vorangehende Redigitalisierung voraussetzt. Dieses Szenario entwickelt die herkömmliche Verwendung des Mikrofilms weiter, sie mündet in einer gegenseitigen Verschränkung digitaler und analoger Techniken. Genuin digitale Daten werden dabei ebenso wie digitalisierte Daten von ursprünglich analogen Objekten/Archivalien auf Mikrofilm ausbelichtet und bei Bedarf zu einem späteren Zeitpunkt über einen Scanner redigitalisiert, um dann erneut digital im Computer benutzt zu werden. Eine derartige Konversionsstrategie erfordert im Vergleich mit der Verwendung des Mikrofilms als Benutzungsmedium einen wesentlich höheren Technikeinsatz.

Neben der Haltbarkeit des Datenträgers liegt ein zweiter Vorteil darin, dass die auf dem Mikrofilm als Bilder abgelegten Informationen nicht regelmäßig wie bei der Migrationsstrategie in neue Formate überführt werden müssen. Völ-

lig unabhängig von Formaterwägungen ist der Mikrofilm jedoch nicht, da er über die Ablagestruktur von Primär- und v.a. Metadaten gewisse Ansprüche an das Zielformat bei der Redigitalisierung stellt, z.B. die bei den Metadaten angewandte Form der Strukturierung. Die Vorteile im Bereich der Formate verlieren sich, wenn der Mikrofilm als digitales Speichermedium begriffen wird, auf dem die Informationen nicht mehr als Bild, sondern als eine endlose Abfolge von Nullen und Einsen binär, d.h. als *Bitstream*, abgelegt werden. Es bleibt dann allein die Haltbarkeit des Datenträgers bestehen, die in den meisten Fällen die Zeit, in der das verwendete Dateiformat noch von künftigen Computern verstanden wird, um ein Vielfaches übersteigen dürfte. Auf der anderen Seite entstehen für diese Zeit verglichen mit anderen digitalen Speichermedien nur sehr geringe Erhaltungskosten.

Bei der Ausbelichtung der digitalen Objekte ist darauf zu achten, dass neben den Primärdaten auch die zugehörigen Metadaten auf dem Film abgelegt werden. Verglichen mit rein digitalen Erhaltungsstrategien kann dabei zum einen die für ein Verständnis unabdingbare Einheit von Meta- und Primärdaten leichter bewahrt werden. Das von OAIS definierte archivische Informationspaket (AIP) wird hier physisch konstruiert. Zum anderen verspricht die Ablage auf Mikrofilm auch Vorteile beim Nachweis von Authentizität und Integrität, da die Daten selbst nur schwer manipuliert werden können (die Möglichkeit ergibt sich nur durch die erneute Herstellung eines Films).

Vor einer Abwägung der unterschiedlichen Erhaltungsstrategien sollten sowohl die Benutzungsbedingungen als auch die Kosten beachtet werden, die bei der Ausbelichtung, Lagerung und erneuten Redigitalisierung entstehen. Schließlich ist zu überlegen, in welcher Form die Informationen künftig verwendet werden sollen. Während der Einsatz des Mikrofilms bei Rasterbildern (nichtkodierten Informationen) naheliegt, müssen kodierte Informationen nach erfolgter Redigitalisierung erneut in Zeichen umgewandelt werden. Die Fehlerhäufigkeit der eingesetzten Software muss dabei gegen die zu erwartenden Vorteile aufgewogen werden.

Literatur

Projekt ARCHE, s. http://www.landesarchiv-bw.de >>> Aktuelles >>> Projekte

Eine Bibliographie findet sich beim Forum Bestandserhaltung unter http://www.uni-muenster.de/Forum-Bestandserhaltung/konversion/digi.html