

# Langzeiterhaltung digitaler Daten in Museen

## Tipps zur dauerhaften Bewahrung digitaler Daten

2

### Konzepte digitaler Langzeiterhaltung

Die Bewahrung von Objekten und ihrer Dokumentation sind ein zentraler Teil der Museumsarbeit. Beides, Objekte und Dokumentation, werden zunehmend digital. Das Bewahren dieser Daten ist für die Museumsarbeit notwendig. Es gilt sowohl sie zu erhalten, als auch ihre Interpretierbarkeit langfristig zu sichern (z. B. aus einer Reihe von Einsen und Nullen ein Foto entstehen zu lassen).

Ein Wasserschaden auf einer Karteikarte muss diese nicht vollständig unlesbar machen. Verlust einzelner Bits im Datenstrom durch Entmagnetisierung, mechanische Defekte oder Zersetzung des Trägermaterials kann jedoch den Zugriff auf die Daten vollständig unmöglich machen. Deshalb kommt Maßnahmen zur Kontrolle der Vollständigkeit, Lesbarkeit und Interpretierbarkeit der Daten große Bedeutung zu. Gleiches gilt für die sachgerechte Lagerung der Speichermedien (→ Blatt 18).

Es gab und gibt eine große Anzahl unterschiedlicher Speichertechnologien (Lochkarten, Magnetbänder, ZIP- und DAT-Kassetten, Disketten (8", 5 1/4", 3 1/2" und 3"), Festplatten, CD, DVD, USB-Speichersticks, Blu-ray disc, etc.). Oft veralten Speichertechnologien schneller als die Speichermedien. Selbst gut erhaltene Medien sind nutzlos, wenn die entsprechenden Lesegeräte nicht mehr vorhanden oder nicht mehr benutzbar sind.

In gleicher Weise sind Dateiformate durch die technische Entwicklung gefährdet. Stehen keine geeigneten Programme zur Nutzung einer Datei zur Verfügung, so sind die in ihr enthaltenen Informationen zum Teil oder ganz verloren. Beispielsweise können Texte, gespeichert in alten MS Word-Dateien, in aktuellen Versionen des Programms nicht ohne Verluste (z. B. der Formatierung) oder gar nicht mehr geöffnet werden (→ Blatt 7).

Folgende Faktoren machen aktives Handeln nötig:

- die begrenzte Haltbarkeit der Speichermedien und ihre Fragilität,
- der schnelle Wechsel der verfügbaren Speichertechnologie (Lese- und Schreibgeräte),
- das schnelle Veralten der Programme, die notwendig sind, um aus den binären Daten das ursprünglich Gemeinte zu rekonstruieren.

#### KONZEPTE

Zum jetzigen Zeitpunkt ist weder Soft- noch Hardware vorhanden, die zuverlässig über einen längeren Zeitraum nutzbar ist. Die Archivierung digitaler Objekte benötigt daher konzeptionelle Ansätze, die die oben genannten Faktoren berücksichtigen.

Die wichtigsten Konzepte zur Langzeiterhaltung digitaler Daten werden im Folgenden kurz vorgestellt:

#### Migration

Unter Migration versteht man die Übertragung von digitalen Objekten in andere Dateiformate sowie die Übertragung auf neue Speichermedien oder in neue technische Umgebungen. Unter dem sogenannten "Refreshing" wird die periodisch vorzunehmende Übertragung auf ein neues Speichermedium verstanden (z. B. von einer CD auf eine neue CD).

Die Übertragung von Daten in neue Dateiformate wird dann vollzogen, wenn das bislang verwendete Dateiformat zu veralten droht oder ein ungewöhnliches, im Archivierungsprozess nicht vorgesehenes Format vorliegt. Bei der verantwortungsvollen Migration sind vor allem zwei Aspekte zu berücksichtigen:

1. Es können inhaltliche Verluste auftreten, wenn der Aufbau des alten (originalen) und des neuen Dateiformats unterschiedlich ist. Dies ist z. B. bei der Konvertierung eines Word-Dokumentes in ein PDF-Dokument der Fall. Vorhandene, programmeigene Optionen des alten Dateiformates können nicht übernommen werden und gehen verloren. Vor allem bei audiovisuellen Objekten und Applikationen kann sich das Objekt nach mehreren Migrationsvorgängen durch Informationsveränderungen deutlich vom Ursprungsobjekt unterscheiden. Um die Gefährdungen gering zu halten, ist die Migration immer unter der Prämisse vorzunehmen, dass der ursprüngliche Inhalt so wenig wie möglich verändert wird.
2. Bei der Migration wird aus dem Referenzobjekt ein neues Objekt erzeugt, das sich von diesem z. B. durch das Herstellungsdatum unterscheidet

#### Bitstream Preservation

Die Erhaltung des originalen Datenstroms (der Rohdaten) und somit der ursprünglichen Abfolge der Einsen und Nullen ist im Rahmen der Migration aus Sicherheitsaspekten wünschenswert. Bitstream Preservation ist somit als Bestandteil des Migrations-Konzeptes für die langfristige Erhaltung sinnvoll, weil durch sie der Rückgriff auf die originalen digitalen Informationen möglich bleibt.

## Emulation

Das Emulationskonzept (lat. aemulare = nachahmen) basiert auf der technischen Möglichkeit, auf einem Computersystem ein anderes System nachzuahmen.

Mit diesem Verfahren kann sich

z. B. ein Computer mit dem Betriebssystem MS Windows XP so verhalten, als würde auf ihm das 1993 entstandene MS DOS 6.0 installiert sein. Damit wird es möglich, Programme, die nur auf älteren Betriebssystemen lauffähig sind, auch auf neuen Systemen zu nutzen.

Mit diesem Konzept wird die Verfügbarkeit vor allem von komplexen Applikationen (z. B. Datenbankanwendungen oder interaktiven Anwendungen) gewährleistet. Die Übertragung (Migration) von digitalen Objekten in andere Dateiformate und die damit einhergehende Modifikation ist in diesem Fall nicht notwendig. Emulationen werden bereits erfolgreich für den Erhalt von Unternehmenssoftware, Computerspielen oder digitaler Medienkunst eingesetzt.

Allerdings werden diese Emulatoren z. Z. noch für jedes Computersystem separat erstellt. Erfolgt eine Aktualisierung des Systems, ist in der Regel der auf ihm eingesetzte Emulator nicht mehr lauffähig und die Erstellung eines neuen Emulators notwendig. Deshalb benötigt auch die Emulation periodisch vorzunehmende Anpassungen, um die Verfügbarkeit der digitalen Daten dauerhaft zu gewährleisten.

## Hardwaremuseum

Unter "Hardwaremuseum" wird die Konservierung von Computersystemen und die Sicherung ihrer Funktionsfähigkeit verstanden. Hardware ist aber auf Grund der physikalischen Alterungserscheinungen bei den einzelnen Komponenten ebenfalls in ihrer Haltbarkeit begrenzt. Die Lebensdauer sowie die Produktionszeiträume der Bauteile (Prozessor, Hauptplatine, Grafikkarte etc.) sind in der Regel kurz. Zu einem späteren Zeitpunkt Ersatzteile zu finden oder diese nachzubauen, ist problematisch. Auf Grund nicht immer offengelegter Dokumentationen ist ein Nachbau nicht in jedem Fall möglich. Daher ist diese Sicherung der Funktionsfähigkeit nur mit großem Aufwand dauerhaft sicherzustellen.

## Analoge Sicherung

Für einige digitale Objekte ist die Möglichkeit der Sicherung in analoger Form gegeben. So können Texte auf einem Drucker "dokumentenecht" ausgegeben und archiviert werden. Eine weitere

Alternative besteht in der Sicherung von digitalen Texten oder Bildern auf (Farb-)Mikrofilmen, die eine Haltbarkeitsdauer von mehreren hundert Jahren besitzen.

Die analoge Sicherung hat den Vorteil, dass zur Nutzung der Objekte kein Computer mit entsprechender Software benötigt wird. Die Vorteile digital erhaltener Objekte, wie beispielsweise der schnelle Zugriff und die digitale Weiterverarbeitung, entfallen jedoch mit dieser Form der Erhaltung. Zudem gibt es eine Reihe von digitalen Objekten, wie Datenbanken oder interaktive Programme, die sich nicht analog sichern lassen.

## Verteiltes Speichern

Durch die breite Verfügbarkeit von leistungsfähigen Netzwerken, unter anderem auch des Internets, wurden Konzepte möglich, die die Speicherung von digitalen Informationen auf verteilten Systemen

vorsehen. Der Vorteil dieser Konzepte ist die Möglichkeit, Daten an unterschiedlichen Standorten mehrfach zu sichern.

Dazu werden entsprechende Programme bereitgestellt, die die Verwaltung der Daten im System übernehmen. Allerdings ist hiermit noch nicht die Interpretierbarkeit der digitalen Informationen langfristig gesichert. Das bekannteste Projekt ist Lockss (Lot of copies keep stuff save) der Stanford University Library.

## Fazit

Die vorgestellten Konzepte unterscheiden sich nicht allein hinsichtlich ihres technischen und finanziellen Aufwands. So kann für komplexe Applikationen zur Zeit nur die Emulation Verwendung

finden, während textbasierte Informationen mit allen Methoden bewahrt werden können. Die Entscheidung für ein Konzept zur Langzeiterhaltung digitaler Daten ist deshalb abhängig von der Art der digitalen Daten, den Anforderungen an Authentizität und Verfügbarkeit sowie den personellen, technischen und finanziellen Ressourcen der jeweiligen Institution.

Es existieren eine Reihe von technischen Konzepten zur Langzeiterhaltung digitaler Daten. Einige ermöglichen die Erhaltung und Nutzung der originalen Dateien, andere die Erhaltung der darin enthaltenen Informationen.

Die Wahl des geeigneten Konzeptes ist abhängig von den Anforderungen an die Daten und den personellen, technischen und finanziellen Ressourcen der Institution.



[http://www.kb.nl/hrd/dd/dd\\_projecten/Emulation\\_Test\\_Results\\_Document\\_KB\\_NA\\_2006.pdf](http://www.kb.nl/hrd/dd/dd_projecten/Emulation_Test_Results_Document_KB_NA_2006.pdf)

Ergebnisse von Testemulationen an der Koninklijke Bibliotheek der Niederlande.

[http://www.digitaleduurzaamheid.nl/bibliotheek/docs/White\\_paper\\_emulation\\_UK.pdf](http://www.digitaleduurzaamheid.nl/bibliotheek/docs/White_paper_emulation_UK.pdf)

Studie der Koninklijke Bibliotheek der Niederlande zur Emulation im Rahmen der Langzeitarchivierung.

<http://www.lockss.org/lockss/Home>

Homepage des LOCKSS-Projektes, eines Systems zum verteilten Speichern.

<http://tud.at/uni/diplomarbeit/diplomarbeit.pdf>

Diplomarbeit zum Thema „Informationsverlust durch Digitalisierung“.

<http://www.dpconline.org/docs/reports/uknamindthegap.pdf>

Studie der Digital Preservation Coalition (DPC) über die Notwendigkeit der Erhaltung des digitalen Kulturerbes in Großbritannien.

[http://www.planets-project.eu/docs/reports/Planets\\_PA5-D1-TestingToolsForTechnicalEnvironments-Final\\_v2\\_public.pdf](http://www.planets-project.eu/docs/reports/Planets_PA5-D1-TestingToolsForTechnicalEnvironments-Final_v2_public.pdf)

„Framework and Workflow for testing tools for different technical environments“. Bericht des Planets-Projektes über Erfahrungen mit Emulationsstrategien.

LINKS

Stand: Juni 2009