Andreas Weisser

Digitale Speicher

Langzeitarchivierung audiovisueller Inhalte – ein Überblick

Analoge Datenträger haben in den letzten Jahren bei der Speicherung audiovisueller Inhalte merklich an Bedeutung verloren. Doch auf die Frage, welche digitalen Alternativen für eine langfristige Speicherung geeignet sind, gibt es viele Antworten. Zur Verfügung stehen optische Datenträger, Festplatten und Datenbänder sowie die sogenannte „Cloud". Sie werden in dem Beitrag knapp vorgestellt. Doch so unterschiedlich diese Speichermedien sind – allen gemeinsam ist, daß sie eine konstante Pflege und Überprüfung der Inhalte benötigen, um sicherzustellen, daß die Daten lesbar sind – und bleiben. Eine einzelne Kopie reicht für den Langzeiterhalt nicht aus.

Digital Storage: Long-Term Archiving of Audio-Visual Content – An Overview

In recent years, analogue storage media have significantly lost importance for the storage of audio-visual content. However, there are many answers to the question of which digital alternatives are suitable for long-term storage. The options include optical disks, hard disks, magnetic tapes, and the so-called 'cloud'. These are described in brief in this article. Despite the differences between these storage media, common to all is the fact that they require constant care and inspection of the contents in order to ensure that the data can be read—and will remain readable. A single copy is insufficient for long-term preservation.

Spätestens seit dem langsamen Verschwinden von Film- und Magnetbandherstellern und dem Überlebenskampf traditionsreicher Unternehmen wie Kodak, zeichnet sich das Ende der Ära analoger Datenträger ab [1]. Gleichzeitig wächst die Datenmenge digital entstandener „Objekte" exponentiell. Derzeit befinden sich geschätzt 2.500.000 Petabyte [2] Daten auf unterschiedlichsten Datenträgern und müssen bewahrt werden [3]. Nicht nur im fotografischen Bereich sondern gerade im Audio- und Videobereich wie auch beim Kinofilm hat sich eine parallele Welt aus neuen Speichermedien, Formaten und Codecs entwickelt. Durch die schrittweise Digitalisierung ganzer Arbeits- und Lebensbereiche wächst auch die zu speichernde Datenmenge zudem immer schneller [4]. Und damit die Frage, wie und auf welchem Medium eine sichere und kostengünstige digitale Langzeitarchivierung (dLZA) sinnvoll und möglich ist. Neben der wichtigen Fragestellung, in welchem Format und/oder Codec ein digitales Objekt aufbewahrt werden soll, um es langfristig zu sichern, wird vor allem die Frage nach dem „Wie" immer drängender.

Die Anfälligkeit moderner Medien gegenüber mechanischen oder strukturellen Fehlern und Schäden sowie Gefahren, die durch Alterung und Obsoleszenz von Soft- und Hardware entstehen, sind weitere Gesichtspunkte, die bei einer sicheren Speicherung digitaler Daten berücksichtigt werden müssen [5]. Zudem gilt es, die wachsende Datenmenge nicht nur zu verwalten, sondern auch den Zugriff und die Nutzbarkeit zu vereinfachen und zu ermöglichen. Gerade für Archive spielt dieser Punkt eine immer wichtigere Rolle: Die Auffindbarkeit im Internet entscheidet häufig über den Grad der Wahrnehmung und Bedeutung der verwahrten Bestände. Die Präsenz und Langzeitarchivierung müssen jedoch unter dem Gesichtspunkt knapper Budgets möglichst kostengünstig und personalschonend bewerkstelligt werden – ohne dabei wichtige Sicherheitsaspekte zu vernachlässigen. Wobei Sicherheit hierbei nicht nur den Schutz vor unbefugtem Kopieren der Daten meint, sondern auch und in zunehmendem Maße den Schutz vor Datenkorruption. Denn der „bit rot" [6], der schleichende Verlust der Lesbarkeit, ist eines der gefährlichsten und am wenigsten vorhersehbaren Risiken der digitalen Langzeitarchivierung. Methoden zur Qualitätskontrolle und Validierung sollten daher mitbedacht werden [7].

Gleichzeitig wird die Generierung, Pflege und Erhebung von Metadaten bei digitalen Objekten zu einer Kernaufgabe bei der digitalen Langzeitarchivierung. Ohne technische, strukturelle und beschreibende Metadaten wird es in der Zukunft schwierig, ein File langfristig zu identifizieren, zu lokalisieren sowie es angemessen zu verarbeiten und korrekt darzustellen [8]. Die Kernpunkte eines idealen digitalen Langzeitspeichers beinhalten daher folgende Kriterien:

- redundante und räumlich getrennte Speicherung der Daten;
- hohe Verfügbarkeit für Zugriff;
- langfristige Sicherheit der gespeicherten Inhalte, z.B. gegen externe Manipulation und Datenkorruption;
- Sicherheit gegen fremden Zugriff;
- eigenständiges Medium, keine unbekannte „Daten-Nachbarschaft";
- langfristige Überprüfbarkeit der Integrität, Authentizität;
- regelmäßige Überprüfung von Speichermedium und File auf Konsistenz und Funktionsfähigkeit;
- Vertrauenswürdigkeit.

Übersicht Speichermedien

Zur Speicherung audiovisueller Inhalte stehen im Wesentlichen zwei Mediengattungen zur Verfügung: optische sowie magnetische Datenträger. Zu den optischen Datenträgern gehören CDs, DVDs und Blu-ray Discs. Die magnetischen Speichermedien gliedern sich in Festplatten und Datenbänder. Für die Speicherung großer Datenmengen scheiden Flash Speicher (SSD = Solid State Drive) derzeit aufgrund ihrer deutlich höheren Kosten (noch) aus und sollen deshalb hier nicht näher betrachtet werden [9].

Optische Datenträger: CD, DVD, Blue-ray

Aufgrund ihrer relativ einfachen Handhabung und weiten Verbreitung finden sich auch heute noch optische Datenträger im aktiven Gebrauch bei der Archivierung. Sie können meist unaufwendig

und schnell erstellt werden und haben sich als vielseitiges Speichermedium für Daten und audiovisuelle Inhalte etabliert.

Was im privaten Bereich vielleicht noch sinnhaft erscheint, wird im institutionellen Bereich jedoch schnell unbrauchbar. Zu geringe Speicherkapazitäten gepaart mit langsamen Zugriffszeiten und der hohen Anfälligkeit gegenüber Alterungsreaktionen [10] und chemischer Instabilität machen optische Datenträger unattraktiv für die Speicherung großer Datenmengen. Dies gilt ausnahmslos für CDs, DVDs und Blu-ray Discs. Letztere sind insbesondere aufgrund ihrer hohen Datendichte kritisch zu sehen [11].

Selbst die kürzlich vorgestellten „Archival Grade"-DVDs sind bei genauer Betrachtung keine sichere Lösung. Denn es ist heute nicht seriös vorherzusagen, ob in 20 bis 30 Jahren überhaupt noch funktionierende Abspielgeräte für optische Datenträger vorhanden sein werden. Die leidvollen Erfahrungen mit alternden LaserDisc-Playern sollten eine Mahnung sein. Doch selbst wenn funktionsfähige Geräte noch vorhanden sein sollten, wäre die Kompatibilität der Anschlüsse zum Computer damit noch längst nicht garantiert. Denn wer kann heute sagen, ob es in wenigen Jahrzehnten überhaupt noch „klassische" Computer mit USB-Anschlüssen geben wird. Als letzte Hürde bliebe dann die Software: Sollten die Daten in 20–30 Jahren tatsächlich noch vorhanden sein, müßten sie auch mit einer zugehörigen Software les- und darstellbar sein.

Denn nur dann, wenn die Software auch tatsächlich die codierten Nullen und Einsen in ein Foto oder ein Video umrechnen kann, sind die Daten auch wirklich verwertbar. Jeder, der einmal versucht hat, eine 5 1/4-Zoll-Diskette aus den 1980er Jahren auszulesen, weiß, wie schwierig es sein kann, tatsächlich an die gespeicherten Informationen zu gelangen.

Insgesamt überwiegen bei optischen Datenträgern die Risiken, deshalb scheiden sie für die Speicherung großer Datenmengen aus.

Festplatten

Als Zielträger für die File-basierte Speicherung großer Datenmengen haben sich Festplatten (HDD) sowie Datenbänder (früher auch als „Streamer" bezeichnet) etabliert. Beide Systeme sind in Wirtschaft und Industrie Standardanwendungen zur Speicherung großer Datenmengen geworden. Hinsichtlich Sicherheit, Implementierung, Pflege und der damit verbundenen Kosten bestehen jedoch teilweise deutliche Unterschiede. Datenbänder werden meist in sogenannten robotergestützen Tape-Librarys betrieben, Festplatten hingegen können einzeln oder im Verbund Anwendung finden.

> *Externe Festplatten*

Externe Festplatten sind vor allem bei Privatanwendern beliebte Datenträger zur Speicherung überschaubarer Datenmengen. Sie sind handlich, meist kostengünstig und verfügen – im Vergleich zu optischen Datenträgern – über eine deutlich höhere Speicherkapazität. Zudem lassen sie sich per Plug-and-Play einfach in ein bestehendes System einbinden und ebenso leicht wieder separieren. Externe Festplatten können dauerhaft angeschlossen sein, und so bestehende Datenspeicher erweitern oder aber als Back-up-Speicher dienen. Geht die zu speichernde Datenmenge jedoch über die Kapazität einer externen Festplatte hinaus, kann die Verwaltung der Daten schnell unübersichtlich werden. Doch trotzdem finden sich in einzelnen Archiven immer wieder nummerierte Festplatten, die mit „Inhaltsangaben" auf einem Regalbrett gelagert werden. Abgesehen von der mangelnden Praktikabilität beim Zugriff auf die Daten besteht bei dieser Speicherform keinerlei Sicherheit gegenüber Datenkorruption oder Verlust durch schleichende Defekte.

Dies sind jedoch nicht die einzigen Risiken. Denn auch Festplatten sind nicht immun gegenüber Alterungserscheinungen und mechanischem Verschleiß. Insbesondere die Arme der Schreib- und Leseköpfe sowie die Lager können bei der Alterung für Probleme sorgen. Zudem können durch die technische Weiterentwicklung auch bei Festplatten Software- oder Schnittstellen-Inkompatibilitäten dazu führen, daß die Kommunikation zwischen Rechner und Festplatte problematisch wird.

Ein weiterer Aspekt wird häufig unterschätzt: Gerade in kostengünstigen externen Festplatten werden oftmals qualitativ minderwertige Bauteile verbaut. Diese no-name-Produkte bergen gewisse Risiken, denn von außen ist schwer feststellbar, wie intensiv die Qualität der gefertigten Platten überprüft wurde – und wie langlebig die verbauten elektronischen Komponenten, die eine Festplatte steuern, tatsächlich sind.

Aus diesen Gründen sind einzelne externe Festplatten für die langfristige Speicherung größerer Datenmengen zu unsicher.

> *Festplattenverbünde*

Um diesen Problemen zu begegnen, werden zur Speicherung größerer Datenmengen mehrere Festplatten in einem gespiegelten Verbund betrieben. Dabei werden die Daten – je nach Konfiguration des Systems – auf mehreren Festplatten redundant gespeichert. Ein solches Redundant Array of Independent Disks (RAID) kann in verschiedenen Sicherheits- und Performance-Levels betrieben werden und bietet eine relativ große Sicherheit gegen Datenverlust durch Hardware-Fehler.

Der größte Vorteil besteht jedoch darin, daß es – bei professionellen Systemen – genügend Kontrollmechanismen gibt, um Zustand und Fehlerraten der verbauten Platten zu überwachen. Somit besteht eine deutlich größere Sicherheit, mechanische Schäden oder Schreib-Lesefehler früher zu detektieren. Denn diese sind es, die über die Zuverlässigkeit einer Speicherlösung letztendlich entscheiden: Ist das gewählte Medium nicht zuverlässig, ist es als Lösung zur Langzeitarchivierung ungeeignet.

Nachteilig ist jedoch der relativ große Aufwand, der betrieben werden muß, um ein komplexes System dauerhaft sicher zu betreiben. Permanent laufende Server erfordern konstante Pflege, da es immer wieder zum Ausfall einzelner Platten kommen kann. Während Festplattenhersteller von einer Fehlerhäufigkeit von durchschnittlich 2 % ausgehen, sprechen zwei Untersuchungen von Google [12] und dem Computer Science Department der Carnegie Mellon University [13] von deutlich höheren Fehlerraten. Zudem wurde beobachtet, daß die Ausfallraten ab einer 2- bis 5jährigen Nutzungsdauer stark ansteigen.

Neben dem permanenten Monitoring, ob einzelne Festplatten im Verbund auf Störung gehen und getauscht werden müssen, muß bei solchen Systemen überwacht werden, ob die Daten noch korrekt sind. Dies geschieht z.B. über Checksummenvergleiche.

Ein weiterer Faktor sind die Energiekosten. Diese müssen bei der Speicherung größerer Datenmengen mit beachtet werden. Denn oft macht es bei der Langzeitarchivierung keinen Sinn, Daten, die nur selten nachgefragt werden, permanent in Bewegung zu halten. Und nichts anderes bedeutet die Datenspeicherung auf Festplatten: Diese müssen sich permanent drehen und „on" sein, benötigen deshalb permanent Energie.

Als gespiegelte Variante im Verbund betrieben, sind Server oder Network Attached Storage (NAS) Systeme eine Option für die langfristige Datenspeicherung. Allerdings benötigen sie konstante Pflege und Betreuung. Spätestens nach fünf Jahren sollten die Festplatten getauscht werden.

Datenbänder

Die Alternative zu Festplatten sind Systeme, die auf Bandtechnologien basieren. Datenbänder sind robuste und stabile, im großen Maßstab kostengünstig zu betreibende Speichermedien für große Datenmengen. Ein Datenband wird in mehreren Spuren linear magnetisch beschrieben und ausgelesen. Lineare Band-Aufzeichnungssysteme sind weniger anfällig gegen Stoßbelastungen und besser gegen mechanische Beschädigungen geschützt als Festplatten. Nachteile ergeben sich durch längere Zugriffszeiten, die je nach gewähltem System und der angestrebten Nutzung jedoch nicht ins Gewicht fallen. Einzelne neuere Systeme sind jedoch schon so performant, daß sie sogar Streaming-Anwendungen zulassen.

Datenbänder waren bis vor kurzer Zeit reine IT-Speicherlösungen, die mittels spezieller Archivierungssoftware beschrieben werden mußten. Eine Anbindung an die bestehende Infrastruktur war daher nicht immer einfach zu bewerkstelligen. Dies lag unter anderem daran, daß auf Datenbändern keine Dateisysteme wie z.B. auf Personal Computern eingerichtet werden konnten. Mittlerweile ist es möglich, Verzeichnisstrukturen anzulegen [14]. Datenbänder können nun wie externe Festplatten oder Laufwerke eingebunden und beschrieben werden, was die Nutzbarkeit stark vereinfacht.

Marktführer bei Datenbändern ist das Format LTO. Linear Tape Open wurde von den Firmen Hewlett Packard (HP), International Business Machines (IBM) und Seagate entwickelt. Das Konsortium, dem auch weitere Firmen beigetreten sind, entwickelte das Bandformat als offenen Standard. Dies gewährleistet einen breiten Support und die Kompatibilität von Bändern und Laufwerken innerhalb der beteiligten Firmen. Für den Verbraucher ergeben sich zudem ökonomische Vorteile durch den Wettbewerb der beteiligten Unternehmen.

LTO funktioniert mit einem Spulenkörper und einem linearen Aufzeichnungsverfahren. Dies bedeutet, daß das Band nur auf einer Spule aufgewickelt ist und im Laufwerk automatisch eingefädelt wird.

Es existieren mehrere Entwicklungsstufen (LTO-1 bis LTO-6), die kontinuierlich erweitert werden. Für die jeweiligen Weiterentwicklungen garantiert das LTO-Konsortium eine Rückwärtskompatibilität beim Lesen von zwei Generationen (Schreiben: eine Generation [15]). Momentan bietet die Generation LTO-6 im Native-Modus (= unkomprimiert) eine Speicherkapazität von 1,2 Terabyte pro Band [16].

Die Bandhersteller gehen von einer Haltbarkeit der Bänder von ca. 30 Jahren aus [17]. Aber auch hier scheint eine kritische Haltung sinnvoll, denn die Haltbarkeit der Daten ist damit nicht gleichzusetzen. Ohne funktionierende Laufwerke, in denen die Bänder gelesen werden können, ist ein historisches Datenband ebenfalls obsolet. Für Laufwerke kann – je nach Wartung und Nutzungsintensität – eine Lebensdauer von maximal 5-7 Jahren angenommen werden [18].

Wenn Datenbänder in einer gespiegelten, robotergestützten Tape Library betrieben werden, stellen sie eine kostengünstige, sichere und effektive Option für die langfristige Speicherung von Daten dar. Allerdings sollte im Abstand von zwei Generationen eine Migration der Daten auf die aktuelle Generation stattfinden.

Cloud-basierte Lösungen

Für viele Institutionen und Archive stellt sich ab einer gewissen Datenmenge die Frage, ob der erhebliche Aufwand zur Betreuung und Pflege von Infrastruktur und Datenspeicher intern geleistet werden kann – und soll. Neben den geschilderten Herausforderungen und Risiken stehen dem Aufbau einer eigenen Speicherlösung auch erhebliche Kosten entgegen.

Doch selbst wenn diese Punkte gelöst werden können, ist die räumlich redundante Speicherung der Daten nicht immer einfach zu lösen. Zwar kann die interne Redundanz eines Servers oder einer Tape Library als niederschwellige Sicherheit gegen Datenverlust fungieren. Bei größeren Havarien jedoch ist ohne eine räumliche Redundanz der Daten keine Sicherheit gewährleistet.

Aus diesen Gründen entscheiden sich Archive und Sammlungen häufig für die Ausgliederung der Speicherung an einen externen Dienstleister. Obwohl es sich hierbei um klassische IT-Rechenzentren mit dem Fokus auf Speicherdienstleistungen handelt, wurde diese Dienstleistung unter dem Begriff „Cloud" bekannt und vermarktet [19]. Die scheinbar ephemere Datenwolke existiert jedoch ausschließlich in den Hochglanzbroschüren der Betreiber. Tatsächlich handelt es sich meist um Rechenzentren, in denen sowohl Festplatten als auch Datenbänder eingesetzt werden. Je nach Marktausrichtung, Sicherheitsaspekt und Performance, schwankt der prozentuale Anteil der beiden Speichermedien.

Im Normalfall erfährt ein Kunde nur wenig über Standort, Sicherungsmedium und Infrastruktur des Cloud-Dienstleisters. Das optische Erscheinungsbild des User-Interfaces bestimmt oftmals die Wahrnehmung.

Bei der Betrachtung von Cloud-Modellen zur Speicherung größerer Datenmengen sollte nicht allein der Preis den Ausschlag bei der Auswahl eines Dienstleisters geben. Datensicherheit, deren Verfügbarkeit, gewählte Speichertechnologie, Handhabung sowie der Standort des Unternehmens sind wichtige Punkte, die bei einer Entscheidung beachtet werden sollten. Professionelle Business-Lösungen sind hier eindeutig den einfacheren Lösungen für den Endkundenmarkt vorzuziehen.

Viele Dienstleister haben nicht nur ihren Firmensitz im Ausland, sondern in der Regel auch die dazu gehörenden Rechenzentren, in denen die Daten gelagert werden. Sie unterliegen somit im Falle von Komplikationen ausländischer Rechtsprechung, was zumindest finanziell unkalkulierbare Folgen haben kann. Durch „Prism" [20], den gezielten Lauschangriff auf Datencenter sowie den Email-

und Telefonverkehr, stehen amerikanische Anbieter derzeit zudem unter besonders kritischer Beobachtung.

Gleichzeitig kann nie sicher gesagt werden, wo und in welcher „Nachbarschaft" sich die eigenen Daten tatsächlich befinden. Denn die Daten werden nach Nutzungswahrscheinlichkeit und systeminternen Algorithmen optimiert im Speichersystem abgelegt. Eine logische Zuordnung zusammengehörender Files (oder deren Bestandteile) erfolgt allein über die Verknüpfung einer Datenbank. Konkret kann das bedeuten, daß die eigenen Daten auf Laufwerken gemeinsam mit Daten aus Versicherungen, Industrie oder aber sonstigen Kunden gespeichert werden. Im Jahr 2011 wurde ein Fall in den USA publik [21], bei dem das FBI aus Ermittlungsgründen den Server eines Speicheranbieters beschlagnahmte, weil sie illegales Datenmaterial einer Hackergruppe dort vermutete. Durch die Stilllegung waren auch alle anderen Nutzer ohne jegliche Vorwarnung von ihren gespeicherten Daten abgeschnitten.

In solchen Fällen kann die virtuelle „Nachbarschaft" bei der Datenlagerung zu ungeahnten Komplikationen führen – und zum Risiko werden. Auch die Möglichkeit einer Insolvenz oder Pleite eines Speicheranbieters sollte berücksichtigt werden. Denn was geschieht in einem solchen Fall mit den Daten, die sich im Ausland auf den Laufwerken des Dienstleisters befinden? Werden sie der Konkursmasse zugeschlagen oder können sie problemlos aus dem Gesamtspeicher herausgelöst werden? Und generell ist von Interesse, wie die Besitzverhältnisse eines Speicheranbieters aussehen. Gehört die Hardware einer Leasing-Gesellschaft, oder befindet sie sich tatsächlich im Besitz des Anbieters?

Es gibt nur wenige Dienstleister, die solchen Risiken durch gezielte Maßnahmen und Vorkehrungen entgegentreten. Mittels maßgeschneiderter Lösungen können somit auch Anforderungen von audiovisuellen Archiven erfüllt werden. Doch es sind nicht die großen Anbieter, die – zumeist mit internationalem Background – die Szenerie beherrschen. Lösungen im deutschsprachigen Raum sind bisher rar gesät, aber vorhanden.

Neben den bereits genannten Fragestellungen sollten Performance, Stabilität und Skalierbarkeit der Speicherlösungen beachtet werden. Während die Performance bei der Langzeitarchivierung oftmals eine geringere Rolle spielt, sollten bei Sicherheit, Stabilität und Erreichbarkeit („Uptime") keinerlei Kompromisse eingegangen werden.

Prinzipiell eignen sich Cloud-basierte Lösungen zur Speicherung großer Datenmengen sehr gut, da die technische Infrastruktur in einem großen Maßstab betrieben und gewartet wird. Allerdings ist die Auswahl eines geeigneten Anbieters nicht einfach und sollte wohl überlegt sein.

Fazit

Vermutlich mehr als zuvor werden sich Sammlungen und Archive mit der Fragilität von digitalen Daten und Langzeitspeichern beschäftigen müssen. Die Wahl der Speicherlösung und Dokumentation der Metadaten in Verbindung mit automatisierten Checks der Datenintegrität sind die entscheidenden Punkte für den Langzeiterhalt der Daten. Optische Datenträger – gleich welcher Ausprägung – sind hierfür eher ungeeignet. Skalierbare, redundante Speichersysteme bieten die größten Sicherheitsreserven und sind besser an den jeweiligen Bedarf adaptierbar. Insgesamt scheint sich zu bestätigen, daß die Speicherung von Daten auf Festplatten nicht nur mit höheren Kosten und Energieaufwand [22], sondern auch mit größeren Risiken verbunden ist als auf Datenbändern, da bei ersteren eine schleichende Datenkorruption häufiger auftritt [23]. Für Archive, Sammlungen und Museen bieten sich deshalb Datenbänder in automatisierten Bandspeichersystemen als kosteneffektive und verläßliche Speicherlösung an. Nachteilig ist bei allen Systemen, daß eine „Speichern und Liegenlassen"-Strategie nicht (mehr) funktioniert: Aktuelle Speichersysteme benötigen permanente Betreuung, Sicherung, Pflege und Überwachung. So wird es denn vermutlich bald zu einem Paradigmenwechsel kommen: Nicht mehr die absolute „Haltbarkeit" in Jahren ist das ausschlaggebende Argument für die Wahl des Trägermediums, sondern seine temporäre Zuverlässigkeit, Nutzbarkeit und Kosteneffizienz für eine begrenzte Zeitspanne. Denn keines der aktuell erhältlichen Speichersysteme ist derzeit für eine 30- oder mehrjährige Verwendung konzipiert. Und selbst wenn die Medien diese Zeitspanne überdauern würden, wären die Inhalte mit großer Sicherheit wegen fehlender Anschlüsse oder Software verloren.

Deshalb dürfen nicht nur die eigentlichen Speichermedien betrachtet werden. Auch die Abspielgeräte und Software sowie die Stabilität von Files und Formaten gehören ins Zentrum der Betrachtung. Hier hilft nur Obsoleszenz-Monitoring und eine regelmäßige Migration der Daten auf neue Träger [24]. Beim absehbaren Aussterben von Fileformaten hingegen müssen veraltete Bestände in neue Formate transkodiert werden.

Diese Punkte alle permanent zu überwachen und adäquat darauf zu reagieren, wird für viele Sammlungen, Archive und Museen personell und finanziell eine große Herausforderung werden. Deshalb werden professionelle Speicheranbieter zukünftig eine größere Rolle beim Langzeiterhalt audiovisueller Daten spielen (müssen). Wahrscheinlich können nur so die vielfältigen Aufgaben bei der Pflege der Daten und der wachsenden Datenmenge bewältigt werden. Es wird deshalb notwendig sein, Regularien zu entwickeln, die es ermöglichen, auch mit externen Anbietern langfristig vertrauensvoll zusammenzuarbeiten. Insbesondere dann, wenn klar ist, daß regelmäßig neue Daten hinzukommen und der Speicherbedarf kontinuierlich wächst.

Eines ist jedoch sicher: Der alte Archiv-Leitsatz „Eine Kopie ist keine Kopie" ist im digitalen Zeitalter noch wichtiger geworden.

Anmerkungen

[1] Im IT-Bereich betrifft dies die Diskette sowie optische Datenträger wie CD und DVD. Im audiovisuellen Bereich Foto- und Kinefilm, Ton- und Videobänder.

[2] 1 Petabyte (PB) entspricht 1.000 Terabyte (TB).

[3] Owens Trevor: Blog-Post „The Signal" 01.08.2012 <http://blogs.loc.gov/digitalpreservation/2012/08/prognosticating-digital-preservation-infrastructure-finals-results-from-the-ndsa-storage-survey> (abgerufen am 22.07.2013).

[4] So werden beispielsweise pro Minute ca. 100 Stunden Material auf YouTube hochgeladen. Quelle: <https://www.youtube.com/yt/press/de/statistics.html> (abgerufen am 22.07.2013).

[5] Oermann, Andrea; Jäschke, Gerald; und Dittmann, Jana: Vertrau-

enswürdige und abgesicherte Langzeitarchivierung multimedialer Inhalte. In: nestor-materialien, Vol. 14 (2009). Frankfurt am Main, S. 3.
[6] Unter „bit rot" versteht man die Veränderung einzelner Bits ohne äußere Einwirkung.
[7] Addis, Matthew: Threats to data integrity from use of large-scale management environments. Hilversum 2010, S. 4f.
[8] Oermann et al., wie Anm. 5, S. 25.
[9] Allerdings ist zu beobachten, daß hybride Systeme, also großvolumige Festplatten kombiniert mit SSD-Speichern, insbesondere bei Servern an Beliebtheit gewinnen. Analyse-Software lagert häufig nachgefragte Daten auf den SSDs, so daß diese sehr viel schneller bereitgestellt und gelesen werden können. Daten, die nur begrenzt nachgefragt werden, verbleiben hingegen auf den vergleichsweise „langsamen" Festplatten.
[10] Iraci, Joe: The Relative Stabilities of Optical Disc Formats. In: Restaurator, Vol. 26 (2005), No. 2, S. 134–150.
[11] Rohde-Enslin, Stefan; und Allen, Keith: Nicht von Dauer: Kleiner Ratgeber für die Bewahrung digitaler Daten in Museen, 2. aktual. Aufl., Frankfurt a.M. 2009, S. 48.
[12] Eduardo, Pinheiro; Weber, Wolf-Dietrich; und Barroso, Luiz André: „Failure Trends in a Large Disk Drive Population." Paper-Präsentation bei: 5th USENIX Conference on File and Storage Technologies – Paper, 12./13.02.2007: 1-13 San Jose, California, 08.08.2007 <http://209.85.163.132/papers/disk_failures.pdf>.
[13] Schroeder, Bianca; und Gibson, Garth A.: „Disk Failures in the Real World: What does an MTTF of 1,000,000 Hours Mean to You?" Paper Präsentation bei: 5th USENIX Conference on File and Storage Technologies – Paper, 12./13.02.2007: San Jose, California, 26.03. 2007 <www.usenix.org/events/fast07/tech/schroeder/schroeder_html/index.html>.
[14] LTFS = Linear Tape File System bezeichnet eine hierarchische Speichermethode für Datenbänder, die es ermöglicht, die Daten in einem Filesystem abzulegen und direkt darauf zuzugreifen.
[15] <www.lto.org/About/faq.html> (abgerufen am 22.07.2013).
[16] Komprimiert 3 TB <www.lto.org/technology/index.html> (abgerufen am 22.07.2013).
[17] <www.lto.org/pdf/LTO Program LTFS Booth Presentation – NAB April 2013.pdf.
[18] Addis, wie Anm. 7, S. 6.
[19] Das U.S. National Institute of Standards and Technology (NIST) definiert „cloud computing" dabei folgendermaßen: Es handelt sich um einen Service auf Abruf, bei dem die in einem Pool bereitgestellten Ressourcen über ein Netzwerk abrufbar sind. Dabei kennen die Nutzer den Bereitstellungsort meist nicht. Die Nutzer können den Bedarf schnell ausweiten oder herunterfahren. Die Abrechnung erfolgt auf Basis der genutzten Services; vgl. Aitken, Brian; McCann, Patrick; McHugh, Andrew; und Miller, Kerry: Digital Curation and the Cloud – Final Report. Glasgow 2012, S. 4.
[20] <http://de.wikipedia.org/wiki/PRISM_%28%C3%9Cberwachungsprogramm%29> (abgerufen am 22.07.2013).
[21] Reuters: Web hosting firm says FBI took servers in raid. 22.06.2011 (www.reuters.com/article/2011/06/22/us-cybersecurity-raid-idUSTRE75L4S820110622).
[22] Bei der Berechnung der Kosten dürfen nicht alleine die Medienkosten herangezogen werden. Zusätzlich anfallende Kosten wie Infrastruktur und Pflege, Migration und Bereitstellung, Energie sowie Gebäudegröße werden für die „Total Costs of Ownership" (= TCO) berücksichtigt. Auch wenn es sich bei der Analyse um einen Report handelt, der auf den Seiten des LTO Konsortiums zu finden ist, scheinen die Berechnungen stimmig zu sein <www.lto.org/pdf/Clipper_Group-Long-Term_Storage-TCO_Analysis_of_Tape_and_Disk-May_13_2013.pdf (abgerufen am 22.07.2013)
[23] Addis, wie Anm. 7, S. 6f.
[24] Eine Migration sollte – je nach Trägersystem – etwa alle fünf Jahre stattfinden; vgl. Rohde-Enslin und Allen, wie Anm. 11, S. 52.

Initiative

„Medium Film nutzen"

Ende August 2013 hatten *Martin Luchterhandt* (Landesarchiv Berlin, Foto-Technischer Ausschuß der Archivreferentenkonferenz), *Bernhard Preuß* (Bundesamt für Bevölkerungshilfe und Katastrophenschutz) und *Wolfram Fiedler* (Herrmann und Kraemer) an interessierte Vertreter des Archiv- und Bibliothekswesens sowie anderer verwandter Bereiche, an Hersteller, Händler und Dienstleister wie Nutzer des Mikrofilms eine Einladung zu einer Ad-hoc-Veranstaltung auf dem 83. Deutschen Archivtag in Saarbrücken mit dem Titel Initiative „Medium Film erhalten" versendet, in der es hieß:

„Sehr geehrte Damen und Herren,
liebe Kolleginnen und Kollegen,
in den letzten Monaten hat es für die Nutzer von Mikrofilmen für die Langzeitarchivierung (sowohl für schwarzweiße Silberhalogenid-Mikrofilme als auch für auf dem Silberfarbbleichverfahren basierende Farbmikrofilme) immer wieder Informationen über Produktionsveränderungen, Engpässe bei der Lieferung von Filmen und Chemie sowie – teilweise drastische – Preiserhöhungen gegeben.

Das hat die betreffenden Institutionen wie auch die Dienstleister stark verunsichert und bereits zur Verzögerung von analogen Verfilmungs- und computer-output-on-microfilm-Projekten geführt. Zudem gibt es Unsicherheiten über die langfristige und stabile Fortführung von Verfilmungsprojekten und auch bzgl. der Sicherungsverfilmung von Archivalien im Rahmen des Kulturgutschutzes auf der Grundlage der Haager Konvention.

Darüber hinaus sind auch andere Bereiche von den Versorgungsproblemen bei Mikrofilmen betroffen, die weniger im Licht der Öffentlichkeit stehen, wie z.B. Bergbaubehörden oder Plan- und Konstruktionsarchive.

Da nicht angenommen werden kann, daß kurzfristig eine digitale Langzeitarchivierung die in der Haager Konvention geforderten Maßnahmen übernehmen kann, scheint es dringend an der Zeit zu sein, alle mit dieser unguten Situation zusammenhängenden Konsequenzen umfassend im Sinne einer raschen Lösung und über Deutschland hinaus zu erörtern und dazu alle beteiligten Seiten zusammenzuführen.

Um diese drängenden Fragen gemeinsam – von Archiven, Bibliotheken, Museen, Fotografen, Dienstleistern, dem Bundesamt für Bevölkerungsschutz und Katastrophenhilfe, der FA Berg bis hin zu den Vertriebseinrichtungen der Filmindustrie und den Film-Herstellern selbst – zu beraten und Wege einer Entspannung der genannten Situation zu finden, laden wir Sie zu einer kurzfristigen Zusammenkunft für eine Initiative „Medium Film erhalten" während des diesjährigen Archivtages in Saarbrücken ein. (...)

Die Veranstaltung soll nach jetzigem Stand folgende Punkte umfassen: