

Leitfaden für die digitale Langzeitarchivierung audiovisueller Medien

verfasst von der
nestor-Arbeitsgruppe Media





Leitfaden
für die digitale
Langzeitarchivierung
audiovisueller Medien

verfasst von der
nestor-Arbeitsgruppe Media

nestor-materialien 19

**nestor - Kompetenznetzwerk Langzeitarchivierung
und Langzeitverfügbarkeit Digitaler Ressourcen
für Deutschland**

nestor - Network of Expertise in Long-Term Storage of Digital Resources
<http://www.langzeitarchivierung.de>

nestor - Partner:

- ..■ Bayerische Staatsbibliothek
- ..■ Bibliotheksservice-Zentrum Baden-Württemberg
- ..■ Bundesarchiv
- ..■ Computerspiele Museum Berlin
- ..■ Deutsche Kinemathek – Museum für Film und Fernsehen
- ..■ Deutsche Nationalbibliothek
- ..■ FernUniversität Hagen
- ..■ Georg-August-Universität Göttingen / Niedersächsische Staats- und
Universitätsbibliothek Göttingen
- ..■ GESIS Leibniz Institut für Sozialwissenschaften
- ..■ Goportis Leibniz-Bibliotheksverbund Forschungsinformation
- ..■ Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen
- ..■ Humboldt-Universität zu Berlin
- ..■ Institut für Deutsche Sprache
- ..■ Institut für Museumsforschung (Stiftung Preußischer Kulturbesitz)
- ..■ Konrad-Zuse-Zentrum für Informationstechnik Berlin
- ..■ Landesarchiv Baden-Württemberg
- ..■ Landesarchiv Nordrhein-Westfalen
- ..■ PDF/A Association
- ..■ Rechenzentrum der Universität Freiburg
- ..■ Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden

© 2016

**nestor - Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit
Digitaler Ressourcen für Deutschland**

Der Inhalt dieser Veröffentlichung darf vervielfältigt und verbreitet werden, sofern der Name des
Rechteinhabers "nestor - Kompetenznetzwerk Langzeitarchivierung" genannt wird.
Eine kommerzielle Nutzung ist nur mit Zustimmung des Rechteinhabers zulässig.

URN: [urn:nbn:de:0008-2016102107](http://nbn-resolving.de/urn:nbn:de:0008-2016102107)
<http://nbn-resolving.de/urn:nbn:de:0008-2016102107>

Kontakt zur nestor-AG Media:
Dr. Stefan Rohde-Enslin, Institut für Museumsforschung (SMB-PK)
s.rohde-enslin@smb.spk-berlin.de
Jürgen Keiper, Deutsche Kinemathek
jkeiper@deutsche-kinemathek.de

Titel

Leitfaden für die digitale Langzeitarchivierung audiovisueller Medien

verfasst von der nestor AG Media

Berlin, Oktober 2016

Projektpartner

- ▶ Akademie der Künste Berlin
- ▶ Bundesarchiv
- ▶ Deutsche Kinemathek - Museum für Film und Fernsehen
- ▶ Freie Universität Berlin, Center für Digitale Systeme (CeDiS)
- ▶ Institut für Museumsforschung (Staatliche Museen zu Berlin - Preußischer Kulturbesitz)
- ▶ Zuse Institute Berlin (ZIB)
- ▶ Landesarchiv Baden-Württemberg
- ▶ Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB)
- ▶ Technische Informationsbibliothek (TIB)

Inhaltsverzeichnis

Titel	3
Projektpartner	3
Inhaltsverzeichnis	4
Einleitung	6
Digitalisierung	10
Farbmanagement	11
Katalogisierungs- und Erschließungsstandards	12
Persistente Identifikatoren	13
Rechte	14
Digitalisierung	14
Verwaiste Werke	14
Langzeitarchivierung	14
Archivierung von Amateurfilmen	15
Definition	15
Beispiele	15
Empfehlungen für Privatanwender	16
Empfehlungen für kleinere Einrichtungen	20
Empfehlungen für mittlere Einrichtungen	27
Empfehlungen für größere Einrichtungen	35
Archivierung von Wissenschaftlichen Filmen	43
Definition	43

Beispiele	43
Empfehlungen für Privatanwender	44
Empfehlungen für kleinere Einrichtungen	50
Empfehlungen für mittlere Einrichtungen	58
Empfehlungen für größere Einrichtungen	68
Archivierung von professionellen Filmen	78
Archivierung von Fernsehmitschnitten	99
Empfehlungen für Privatanwender	99
Empfehlungen für kleinere Einrichtungen	102
Empfehlungen für mittlere Einrichtungen	106
Empfehlungen für größere Einrichtungen	110
Tools	114
Digitalisierungsempfehlungen	116
Analoge Filmformate	116
Analoge Bandformate (Video)	116
Digitale Bandformate (Video)	118
Digitale Discformate (Video)	118
Weitergehende Informationen	118
Glossar	120
Autoren	126
Impressum	127

Einleitung

Audiovisuelle Medien, also Film und Video in ihren unterschiedlichsten Ausprägungen, nehmen einen immer breiteren Raum in den Gedächtnisorganisationen ein. Dies betrifft nicht nur ihre Bedeutung, da die audiovisuellen Quellen zunehmend zum Kernbestand des kulturellen Erbes gezählt werden, sondern auch die bloße Menge an Daten, die mit ihrer Archivierung verbunden sind. Ein digitaler Film, der ohne Qualitätsverlust gespeichert wird, benötigt schnell mehrere Terabyte an Speicherplatz. Damit aber stoßen die Gedächtnisorganisationen schnell in Dimensionen von Rechenzentren vor und nicht selten sind diese auch die bessere Alternative zur hausinternen Archivierung. Doch die Speicher- menge allein ist noch nicht einmal das zentrale Problem bei der Archivierung audiovisueller Medien. Wie beim Film selbst tritt auch bei den Archivierungsbemühungen der Doppelcharakter des Films – einerseits Kunstform, andererseits Industrieprodukt – in seiner ganzen Ambivalenz zutage.

Die Notwendigkeit bei einer Archivierung, Prozesse und Workflows vom Ende her zu denken und auf den gesamten Lebenszyklus des audiovisuellen Objekts zu beziehen, ist bei audiovisuellen Objekten ungleich schwerer umzusetzen als bei anderen Medien wie Texten oder Fotos. Die Standards vom Beginn einer audiovisuellen Produktion bis zur Aufführung sind seitens der AV-Industrie definiert und in hohem Maße proprietär. Hinzukamen mit der Umstellung auf die digitale Produktion die der Computertechnik eigenen schnellen Produktzyklen: Codecs und Formate des audiovisuellen Bereichs haben eine Halbwertszeit, die deutlich unter denen benachbarter Disziplinen liegt.

Die Arbeit der AG Media ist von diesen Problemen geprägt. Die Notwendigkeit, sich den Fragen der Archivierung audiovisueller Produktionen zu stellen, war umgekehrt komplementär zu den Handlungsoptionen. Nach ihrer Gründung im Jahre 2006 befasste sich die AG auch mit der Frage der Archivierung von Computerspielen. Doch schnell war klar, dass beide Medien zwar ähnlich komplexe Fragestellungen aufwerfen, doch völlig unterschiedliche Antworten erforderten. Lösungsstrategien zur Archivierung von Computerspielen setzen etwa einen sehr viel stärkeren Fokus auf Emulation als dies bei den klassischen audiovisuellen Medien der Fall ist. Daher machte die Gründung einer AG Computerspiele Sinn.

Doch die Ratlosigkeit der AG blieb. Die Probleme der Speichermenge schienen lösbar, wenn auch nur im Rahmen von hochprofessionellen Lösungen. Die Frage des Codecs, insbesondere die Suche nach einem verlustfreien Codec, wurde lange mit dem seitens der Digital Cinema Initiative (DCI) favorisierten Codec M-JPEG2000 beantwortet. Doch der Codec war bei vielen Archiven unbeliebt, da er schwer zu handhaben und lizenzbehaftet war. Hinzu kamen negative Erfahrungen im Archivbereich, die eine starke Zurückhaltung forcierten. Auch OAIS-konforme Archivierungslösungen existierten lange Zeit nicht oder waren praktisch kaum zu finanzieren.

Vor diesem schwierigen Hintergrund suchte die AG Media auch den Kontakt zu den führenden Industrievertretern der Branche. Ziel war die Ausarbeitung eines Konzeptes, das Archivierungsbedürfnisse schon beim „Filmen“ über die Postproduktion bis hin zur Aufführung verankerte. Die Gespräche kamen allerdings zu keinem erfolgreichen Abschluss. Kritisch bleibt festzuhalten, dass eine echte Lobby seitens der Archive im Bereich der digitalen Langzeitarchivierung de facto nicht existierte. Allerdings soll dies nicht der letzte Anlauf gewesen sein.

Letztlich ist zu konstatieren, dass wichtige Entwicklungen wohl auch ihre Zeit brauchen.

Wichtige Arbeiten steuerten die FFmpeg-Gruppe und die Österreichische Mediathek mit der Entwicklung des Codecs FFV1 bei. Für die eigentliche Archivierung stehen auch erste Open-Source-Lösungen zur Verfügung und der Aufbau großer Speicher- und Archivierungslösungen ist zwar nicht einfacher, doch deutlich kostengünstiger geworden.

Daher ist dies nun ein guter Zeitpunkt, eine Bilanz zu ziehen und unsere Arbeit und unsere Empfehlungen vorzustellen. Ziel war zunächst, den aktuellen Stand der Diskussion zu präsentieren. Dies betrifft nahezu alle relevanten Bereiche. Allerdings haben wir darauf verzichtet, die Problematik großer Storage- und Archivierungslösungen hier vorzustellen. Einerseits ist dies ein Querschnittsthema, das grundsätzlich für alle Bereiche der Archivierung unabhängig vom Medientyp relevant ist, andererseits adressieren diese Fragen eine ohnehin schon hoch spezialisierte Gruppe. Festzuhalten bleibt allerdings, dass audiovisuelle Archive die Problematik von Bitfehlern adressieren müssen, da diese bei digitalem Film mit einer relativ hohen Wahrscheinlichkeit auftreten werden.

Erstmals hat sich die AG Media entschlossen, für die Empfehlungen einen grundsätzlich anderen, nämlich zielgruppenorientierten Ansatz zu wählen. Hintergrund ist die Einsicht, dass ideale Lösungen für viele Einrichtungen nicht umsetzbar sind. Dies gilt noch stärker für Privatpersonen. Dies bedeutet im Umkehrschluss nicht, dass die zielgruppenspezifischen Lösungen nun hinreichend sind. Im Gegenteil: Die Aufteilung in unterschiedliche Nutzergruppen soll deutlich machen, wo sich unterschiedliche Archivierungsszenarien positionieren, und eher das Interesse wecken, eine professionellere Variante zu wählen. Verbunden mit dieser Aufteilung sind textliche Redundanzen, die wir aber der besseren Lesbarkeit der einzelnen Abschnitte wegen in Kauf nehmen. Außen vor bleibt in diesem Text auch die Frage, ob Archivierung als Service, also die Inanspruchnahme von professionellen Dienstleistern, eine sinnvolle Alternative darstellt. Dies muss jede Institution, aber auch jeder Privatanwender für sich entscheiden. Zumindest lassen sich aus unseren Empfehlungen auch Kriterien für eine Entscheidung ableiten.

Wohl wissend, dass das Archivierungsparadigma „Vom Ende her denken!“ weitaus umfassender ist, haben wir trotzdem einen Abschnitt zur Digitalisierung ergänzt. An keiner anderen Schnittstelle als an dem Übergang von analog zu digital sind die Folgen von Fehl- oder nicht getroffenen Entscheidungen derart gravierend wie hier.

Schließlich haben wir die allgemeine Fragestellung „Archivierung audiovisueller Medien“ aufgelöst zugunsten konkreter Anwendungsszenarien: Wissenschaftlicher Film, Fernsehaufzeichnung, Amateurfilm und professioneller Film. Die Übergänge sind sicherlich fließend und manch einer mag fragen, ob dies gerechtfertigt ist. Zumindest die Anwender erhalten so allerdings genauere Beschreibungen, welche die besondere Situation reflektieren.

Erschließungsstandards und persistente Identifikatoren werden im audiovisuellen Bereich noch relativ selten verwendet. Als zentrale Metadatenelemente besitzen sie aber eine große Bedeutung für die Langzeitarchivierung. Daher haben wir den aktuellen Stand kurz zusammen getragen. Beide Bereiche sind in ihrer fachspezifischen Ausarbeitung außerhalb der „Community“ nur wenig bekannt und sollten daher zumindest kurz benannt werden. Eine breitere Nutzung ist sicher sinnvoll.

Noch ein Wort zu den Begrifflichkeiten. Die Bezeichnung „Film“ wird umgangssprachlich auch für digitale Medien verwendet. Der Materialcharakter, der damit ursprünglich beschrieben wurde, ist fast völlig verschwunden: „Wir gehen ins Kino und sehen dort einen Film“ ist sprachlicher Alltag, obgleich inhaltlich falsch. Mit dem Aufkommen der digitalen Medien bekommt „Film“ eine Doppelbedeutung, der wir auch in diesem Text Rechnung tragen. Daher verwenden wir weiterhin den Begriff „Film“ als Metapher für audiovisuelle Medien. Wo diese Begrifflichkeit zu Missverständnissen führen könnte, haben wir die korrekten, wenngleich etwas sperrigeren Begriffe benutzt.

Dieser Text ist selbst historisch und wird in einigen Jahren veraltet sein. Wir haben uns deshalb auch mit der Frage beschäftigt, wie wir eine Aktualisierung leisten können und für eine Hybridpublikation entschlossen. Einige Punkte des Textes sind nur online verfügbar (z.B. die Übersicht über Dateiformate, Codecs und Formate). Der Grund sind die kontinuierlichen Weiterentwicklungen in diesem Bereich. Ähnliches gilt auch für andere Bereiche. Alle im Text referenzierten Inhalte finden Sie auf der Homepage der nestor AG Media:

<https://wiki.dnb.de/display/NESTOR/AG+Media>

Ein immer wiederkehrendes Thema in der nestor AG Media waren auch die unterschiedlichen Adressaten dieses Leitfadens. Diese definieren sich nicht nur über unterschiedliche Interessen, sondern auch über unterschiedliche Ressourcen, die ihnen zur Verfügung stehen. Einem Privatanwender steht nun mal kein Rechenzentrum zur Verfügung und trotzdem möchte er Vorsorge treffen. Daher haben wir diesen Leitfaden grob an die vier Interessengruppen adressiert:

- ▶ Privatanwender
- ▶ kleinere Einrichtung
- ▶ mittlere Einrichtung
- ▶ größere Einrichtung

Dabei wird angenommen, dass der Privatanwender mit einfachsten, erschwinglichen technischen Möglichkeiten auskommen muss und kaum über personelle Ressourcen verfügt.

Eine kleine Einrichtung kann inhouse auf IT-Support zurückgreifen, die personellen und finanziellen Ressourcen sind aber eher knapp bemessen. Unter einer mittleren Einrichtung wird im Folgenden verstanden, dass sie zwar auf geringe personelle Ressourcen zurückgreifen kann, aber die IT-Unterstützung gut ist und finanzielle Mittel für die Langzeitarchivierung eingeplant sind.

Eine größere Einrichtung kann aufgrund der gewidmeten personellen und finanziellen Ressourcen die digitale Langzeitarchivierung professionell und nach dem aktuellen State of the Art betreiben.

Die folgenden Kapitel behandeln zunächst globale Fragestellungen, die im Kontext der digitalen Langzeitarchivierung von AV-Material relevant sind. Hierzu zählen die Digitalisierung, das oft vergessene Farbmanagement, die Erschließung und Katalogisierung sowie rechtliche Fragen und persistente Identifikatoren.

Im Anschluss hieran werden bestimmte Szenarien (Wissenschaftlicher Film, Fernsehmittschnitte etc.) für die unterschiedlichen Adressaten erörtert. Schließlich verweisen wir auf Tabellen mit technischen Spezifikationen und erörtern wichtige Grundbegriffe in einem Glossar.

Digitalisierung

Nicht alle audiovisuellen Medien liegen digital vor, sondern viele sind noch analog hergestellt worden. Werden analoge Medien digitalisiert, so empfiehlt es sich dringend, die Anforderungen der digitalen Langzeitarchivierung miteinzubeziehen. Andernfalls drohen Qualitätsverlust oder aufwändige Neudigitalisierungen. Im schlimmsten Fall steht das Ausgangsmaterial gar nicht mehr zur Verfügung und eine Neudigitalisierung ist gar nicht mehr möglich. Das Paradigma der Langzeitarchivierung, nämlich alle Prozesse „vom Ende her zu denken“, gilt also gerade auch für die Digitalisierung.

Der Fokus liegt auf LZA-relevanten Punkten wie den signifikanten Eigenschaften, technischen Parametern, der Qualität des Ausgangsmaterials und Prinzipien der Digitalisierung.

Diese Tabelle kann nur eine allgemeine Empfehlung sein. Für die konkrete Auswahl sollte immer hinterfragt werden, ob die Empfehlung dem spezifischen Quellenmaterial gerecht wird, welche Annahmen hinter bestimmten Auflösungs-, Farbraum- und Quantisierungsempfehlungen stehen und ob das richtige Verhältnis zwischen Nutzeranforderungen, Inhalt, Originalqualität und Quellenkritik auf der einen und Archivfähigkeit, Speicher- und Bandbreitenbedarf, Kosten und Migrationsszenarien auf der anderen Seite gefunden wird. So macht es beispielsweise keinen Sinn, Audiospuren von Geräten mit bekannter Grenzfrequenz von 8kHz Mono mit 96kHz oder gar 192kHz Abtastfrequenz zu digitalisieren. Es sei denn, man möchte das Material aus verschiedenen Quellen auf ein Format mit bestimmten Parametern normalisieren, weil sich z.B. der Präsentationsworkflow vereinfacht.

Andererseits kann es sinnvoll sein, bei Material, welches aus Bestandserhaltungsgründen digitalisiert werden soll, einen archäologischen Ansatz zu wählen, d.h. Parameter wie Auflösung und Quantisierung so festzulegen, dass der aktuelle Zustand nicht interpretativ gesichert wird. Korrekturen und Angleichungen an menschliche Wahrnehmung (Filterung, Kennlinienanhebung, Kontrastspreizung usw.) werden dann neben dem zu bewahrenden unbehandelten Digitalisat an einer Arbeitskopie vorgenommen.

Weiterführende Informationen finden Sie auf der Website der nestor AG Media unter <https://wiki.dnb.de/display/NESTOR/AG+Media>

Farbmanagement

Die nachfolgende Empfehlung bezieht sich primär auf die professionelle AV-Produktion, da Privatanwendern in den seltensten Fällen diese Optionen zur Verfügung stehen werden.

Dem Farbmanagement kommt unter Aspekten der Langzeitarchivierung eine besondere Bedeutung bei, da die Ausgabeformate und entsprechenden Standards von der jeweils existierenden Hardware abhängig sind. Allerdings führen die schnellen Innovationszyklen der digitalen Industrie zu einer veränderten Rolle von Standards. Sie schreiben nicht nur das technisch Mögliche fest, sondern definieren auch Innovationsanforderungen, die teilweise nur partiell, mittelfristig oder mit sehr hohem Aufwand zu erfüllen sind.

Gerade im Bereich der Farbräume wird dies zunehmend deutlich. Die neuen Standards (z.B. UHD) erweitern nicht nur die Auflösung, sondern auch den Kontrastumfang und den Farbraum (geplant ist BT.2020). Sinnvoll erscheint daher, den von dem Postproduktionspezialisten Jan Fröhlich gemachten Vorschlag aufzugreifen, sowohl szenen- als auch displayreferenzierte Farbräume zu archivieren.¹ Nur dies sichert die Zukunftsfähigkeit des Materials, aber auch die Möglichkeit der historischen Rekonstruktion.

Die AG Media wird sich in ihrer nächsten Publikation eingehender mit dieser Thematik beschäftigen. Einerseits soll die aktuelle Entwicklung noch etwas beobachtet werden, andererseits fehlt es auch schlicht an Ressourcen, um sich in dieses sehr komplexe Thema einzuarbeiten. Als allgemeine Empfehlung bleibt aber bestehen, dass die Archivierung auf nur einen Referenzfarbraum nicht hinreichend ist, sondern dieser sowohl die zeitgenössischen Abspielgeräte als auch den Produktionszusammenhang berücksichtigen sollte.

¹ Fröhlich, Jan. 2013. "Farbraum und Bildzustand im Kontext der Langzeitarchivierung". In: Klimpel Paul und Kleiper, Jürgen (Hg.), Was bleibt? Nachhaltigkeit der Kultur in der digitalen Welt. Berlin, 119ff. (http://files.dnb.de/nestor/weitere/collab_was_bleibt.pdf).

Katalogisierungs- und Erschließungsstandards

Im Unterschied zur bibliografischen oder archivarischen Erfassung entwickelten sich brauchbare Standards für den audiovisuellen Bereich erst in den letzten Jahren. Zu nennen sind hier:

- ▶ CEN TC 372 Metadata standards. Zur Dokumentation von audiovisuellen Werken. Ein junger europäischer Standard, der zunehmend Verbreitung findet.
- ▶ FIAF Cataloguing Rules. Sehr weit fortgeschrittenes Regelwerk der FIAF (International Federation of Film Archives), das aber noch nicht abgeschlossen ist.
- ▶ REM:Regelwerk Mediendokumentation 1.0. Das aktuelle Regelwerk der ARD-Archive ist das Pendant im Fernsehbereich, allerdings kaum übertragbar.

Persistente Identifikatoren

In modernen Dokumentationssystemen spielen Identifikatoren eine zentrale Rolle, v.a. wenn sie über das Web verfügbar sind. Während im Bibliotheksbereich hier die „Internationale Standardbuchnummer“ (ISBN) und die Gemeinsame Normdatei (GND) eine herausragende Rolle spielen, findet sich kein vergleichbares Pendant für den audiovisuellen Bereich.

Erste Versuche, entsprechende Identifikatoren zu etablieren, stammen von der ISO: Die „International Standard Audiovisual Number“ (ISAN) wurde 2002 eingeführt und erlaubt in einer später aktualisierten Fassung auch die Zuordnung unterschiedlicher Ausprägungen und Manifestationen eines Filmwerkes (z.B. gekürzte oder restaurierte Fassungen). Allerdings ist die filmografische Qualität der ISAN eher gering. Zudem basiert sie auf einem kostenpflichtigen Modell.

Ein Zusammenschluss von Industrie, Archiven und Standardisierungsinstitutionen gründete 2010 die Entertainment Identifier Registry (EIDR). Das Konsortium ist sehr breit aufgestellt, allerdings aufgrund der Kostenstruktur nur für Institutionen interessant. Es basiert auf modernen Standards und unterstützt ebenfalls die Identifikation unterschiedlicher Fassungen. Das System wird ebenfalls seitens der SMPTE (Society of Motion Picture & Television Engineers) zur Hinterlegung in MFX-Containern empfohlen. Aufgrund der hohen Verbreitung und der professionellen Umsetzung etablierte sich hier ein neuer Standard, der allerdings für Privatpersonen nicht geeignet ist, insbesondere da die aktive Registrierung mit hohen Kosten verbunden ist. Sowohl ISAN als auch EIDR sind zwar für Recherchen nutzbar und auch die Nummern sind sichtbar², doch müssen mehr oder weniger starke Einschränkungen in Kauf genommen werden.

Zur Identifikation von Personen steht über die GND eine frei nutzbare Datenbank zur Verfügung³. Diese wurde über das 2012 gestartete Projekt IN2N⁴ mit filmspezifischen Personendaten aus filmportal.de ergänzt.

Wikipedia bietet auch die Möglichkeit persistente, also dauerhafte Identifikatoren zu übernehmen. Allerdings ist der Vorgang selbst etwas unkomfortabel⁵. Alternativ kann man auch in Wikidata⁶ direkt recherchieren. Dort findet man auch zahlreiche weitere Identifikatoren und Links zu den Sprachversionen der Wikipedia.

Zusammenfassend bieten die Identifikatoren der GND/VIAF für die Personen sicher eine gute Ausgangsbasis. Im Bereich der Filmwerke scheint EIDR für den professionellen Bereich deutlich erfolgreicher als ISAN zu sein. Grundsätzlich stellt aber auch Wikidata aus dem Wikipedia-Universum eine freie und kostengünstige Alternative dar, die seit Jahren schon ihre Artikel mit Normdaten anreichert. Da seitens Wikipedia auch das „Linked-Data“-Modell vorangetrieben wird, ist die Verankerung der Daten im Semantic Web auch relativ gesichert.

² <https://ui.eidr.org>, <http://web.isan.org/lookup/> (mit Einschränkungen)

³ Das BSZ stellt eine Online-Suche zur Verfügung: <http://swb.bsz-bw.de>

⁴ <http://in2n.de/>

⁵ Man recherchiert im Seitenquelltext „t-wikibase“ und überträgt die sich daran anschließende URL.

⁶ <https://www.wikidata.org>

Rechte

Die Rechtesituation von audiovisuellen Materialien gestaltet sich aufgrund der zahlreichen Beteiligten in der Regel sehr komplex. Oft sind die Rechte aber überhaupt nicht zu ermitteln (verwaiste Werke). Gerade im Filmbereich ist dies sehr häufig der Fall. Ebenfalls typisch für den AV-Bereich sind Rechte, die auf unterschiedlichen kreativen Bereichen beruhen und teilweise nicht global an einen Rechteinhaber (wie z.B. die Produktionsgesellschaft) übertragen wurden. Folgende Texte setzen sich mit der Rechteproblematik unter verschiedenen Gesichtspunkten auseinander:

Digitalisierung

digiS: Servicestelle Digitalisierung, Handreichung Recht

<http://dx.doi.org/10.12752/2.0.002.2>

Paul Klimpel (Hg.), Bewegte Bilder – starres Recht? Das Filmerbe und seine rechtlichen Rahmenbedingungen. Berlin 2011

Verwaiste Werke

Paul Klimpel, Verwaiste Werke: Die Regelung kommt, die Probleme bleiben: <https://irights.info/artikel/verwaiste-werke-die-regelung-kommt-die-probleme-bleiben/12099>

Richtlinie 2012/28/EU (Verwaiste-Werke-Richtlinie)

[https://de.wikipedia.org/wiki/Richtlinie_2012/28/EU_\(Verwaiste-Werke-Richtlinie\)](https://de.wikipedia.org/wiki/Richtlinie_2012/28/EU_(Verwaiste-Werke-Richtlinie))

Langzeitarchivierung

Eric W. Steinhauer, „Wissen ohne Zukunft? Der Rechtsrahmen der digitalen Langzeitarchivierung von Netzpublikationen“ In Paul Klimpel, Jürgen Keiper (Hg.), Was bleibt? Nachhaltigkeit in der digitalen Welt. Berlin 2013. iRights Media. S. 61-80. (Auch wenn der Fokus hier auf Netzpublikationen liegt, werden grundlegende Rechtsfragen bei der digitalen Langzeitarchivierung erörtert.)

http://files.dnb.de/nestor/weitere/collab_was_bleibt.pdf

Archivierung von Amateurfilmen

Definition

Unter Amateurfilmen sind diejenigen filmischen Erzeugnisse zu verstehen, die von Privatpersonen (Laien) oder Betrieben und Institutionen, die nicht zum Rundfunk- und kinematografischen Bereich gehören, produziert wurden und i.d.R. nicht kommerziell weiterverbreitet werden. Innerhalb von Gedächtnisinstitutionen können sie u.a. aus Behördenüberlieferung, sonstigen staatlichen Institutionen (insbesondere aus dem Bereich Kunst und Kultur) sowie aus dem nichtstaatlichen Bereich (Nachlässe von Politikern, Journalisten, Künstlern und sonstigen Personen der Zeitgeschichte) stammen.

Beispiele

Videodokumentationen von Polizeieinsätzen: Hierbei handelt es sich um dienstlich von Polizeibeamten oder in deren Auftrag erstellte Videoproduktionen, etwa in Zusammenhang mit Zwischenfällen bei einer Demonstration. Entstehungszweck ist die Überwachung der öffentlichen Sicherheit und ggf. die Beweisaufnahme bei Rechtsverstößen. Für Gedächtnisinstitutionen sind diese Aufnahmen aufgrund des potenziellen zeitgeschichtlichen Dokumentationswerts interessant (z.B. Aufnahmen vom sogenannten „Schwarzen Donnerstag“ im Zusammenhang mit den Auseinandersetzungen um das Projekt „Stuttgart 21“).

Aufführungsvideos eines Staatstheaters: Sie gehören ebenfalls zur Überlieferung staatlicher Institutionen und dienen v.a. der Dokumentation der Aktivitäten der jeweiligen Institution.

Videoaufzeichnungen von Parlamentssitzungen: Diese Aufzeichnungen dienen der Dokumentation der allgemeinen Parlamentstätigkeit und als Ergänzung der stenografischen Aufzeichnungen, da auch Gestik und Mimik der beteiligten Personen festgehalten werden können. Sie besitzen einen besonders hohen historischen Dokumentationswert und werden nicht selten auch von Rundfunkanstalten verwertet.

Film- und Videoaufzeichnungen von Gerichtsverhandlungen: Auch hier besteht, je nach Bedeutung des gefilmten Prozesses, ggf. ein sehr hoher historischer Dokumentationswert und die Möglichkeit der späteren Verwertung durch Rundfunkanstalten.

Private Videoaufzeichnungen: Hierbei handelt es sich um die große Masse der Amateurfilme. In Zeiten des analogen Films verblieben diese Produktionen meist im engen familiären Kontext. Zu einer großen Verbreitung von Amateurfilmen trugen gleichermaßen Videoportale wie YouTube und die Entwicklung von Handykameras bei. Inhaltlich betrachtet wird hier das gesamte Lebensspektrum abgebildet.

Sowohl in technischer als auch in inhaltlicher Hinsicht gibt es eine große Bandbreite unterschiedlicher Ausprägungen von Amateurfilmen. Entsprechend unterschiedlich sollten diese Probleme gelöst werden, je nachdem, von wem die Filme archiviert werden. Die nachfolgenden Empfehlungen unterscheiden daher zwischen Privatanwendern, kleineren, mittleren und größeren Einrichtungen.

Empfehlungen für Privatanwender

Auswahl des zu archivierenden Materials

Wenn Amateurfilme von Privatpersonen dauerhaft gesichert werden, ist der Produzent zumeist auch der Betreuer der Filmsammlung. Der Produzent hat möglicherweise die Absicht, seine Videoaufnahmen der Nachwelt zu überlassen. Dabei handelt es sich im Regelfall zunächst um die Erben innerhalb der eigenen Familie. Es kann aber auch vorkommen, dass an den Aufnahmen eines Amateurfilmers öffentliches Interesse besteht. Hier sollte sich der Produzent überlegen, ob er seine Werke nicht bereits zu Lebzeiten und möglichst zeitnah nach ihrer Entstehung an eine öffentliche oder private Gedächtnisinstitution übergeben möchte. Dann wäre diese in einem Vertrag zur dauerhaften Erhaltung des Materials verpflichtet. Es sollte dann eine Einrichtung bevorzugt werden, die das dafür nötige Know-how in ausreichendem Umfang besitzt.

Ist es dem Produzenten jedoch wichtig, einen längeren oder lebenslangen Zugriff auf die eigenen Filmmaterialien zu haben, sind gewisse Maßnahmen zur Langzeitsicherung und Bestandserhaltung, auf die weiter unten noch näher eingegangen wird, unerlässlich. Er sollte seine Werke zumindest grob erschließen, damit sie für einen selbst, aber auch die Nachwelt, eindeutig identifizierbar bleiben. Ihm steht es frei, seine Materialien zu bewerten und solche, die er nicht für dauerhaft erhaltenswert einstuft, nach eigenem Gutdünken zu vernichten. Dies sollte sinnvollerweise dokumentiert werden. Ebenso wichtig wie die Archivwürdigkeit ist aber auch die Archivfähigkeit des zu erhaltenden Materials. Daher sollte dieses in einem geeigneten Archivierungsformat und nicht nur in einem reinen Nutzungsformat vorliegen. Liegen von den Ausgangsfilmen mehrere Versionen vor, was sich sowohl auf den Inhalt als auch auf das Dateiformat beziehen kann, sollte ein geeigneter Archivmaster ausgewählt bzw. generiert werden.

Beschreibung der vorgesehenen Nutzungsszenarien

Je mehr der Produzent von Amateurfilmen vom Interesse seiner Nachfahren oder gar vom öffentlichen Interesse an seinen Werken überzeugt ist, desto mehr sollte er in die Langzeitarchivierung investieren oder sich um eine geeignete Gedächtnisinstitution zur dauerhaften Aufbewahrung kümmern. Der Aufbau einer eigenen Infrastruktur zur digitalen Archivierung ist grundsätzlich mit sehr hohen Kosten verbunden. Im Gegensatz dazu findet eine externe Nutzung vermutlich kaum statt. In jedem Fall sollte der Produzent sich darüber im Klaren sein, welche Eigenschaften seiner Materialien als signifikant und erhaltenswert betrachtet werden können und dies dokumentieren.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren, welche die Qualität der Objekte im digitalen Langzeitarchiv sicherstellen sollen.

Empfehlung

Von Privatanwendern ist ein umfassendes Qualitätsmanagement nicht zu leisten. Trotzdem wird Privatanwendern empfohlen, im Rahmen ihrer Möglichkeiten ggf. eine Eingangskontrolle durchzuführen und die Ergebnisse von Formatmigrationen anhand festgelegter Kriterien zu prüfen.

Auch Privatanwender sollten im Rahmen der gegebenen Möglichkeiten eine Eingangskontrolle durchführen und die Ergebnisse von Formatmigrationen anhand festgelegter Kriterien prüfen.

Erhaltungsstrategien

Bitstream-Preservation

Die Erhaltung des Datenstroms ist die Basis jeder digitalen Archivierung. Dieser sollte gewährleistet werden durch:

- ▶ Verwendung geeigneter Prüfsummen
- ▶ redundante Sicherung an verschiedenen Orten
- ▶ regelmäßige Virenprüfung
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ regelmäßige Sichtprüfung

Empfehlung

CD-ROMs, DVDs usw. sind als Sicherungsmedien zu vermeiden. Auch die Sicherung auf externen Festplatten stellt allenfalls eine Grundsicherung dar. Alles darüber hinaus kann sich ein Privatanwender jedoch häufig nicht leisten. Realistisch ist meist nur der regelmäßige Austausch der Speichermedien bei räumlich getrennter Aufbewahrung. Besser wäre jedoch ein RAID-10-System, etwa ein NAS-Server. Denkbar ist auch eine Cloudlösung, wobei der Anbieter sich um die Verwaltung und Erhaltung der Speichermedien kümmert.

Realistisch für Privatanwender ist meist nur der regelmäßige Austausch der Speichermedien bei räumlich getrennter Aufbewahrung oder eine Cloudlösung.

Kriterien für einen Dienstleister bzw. eine Cloudlösung sind:

- ▶ Gewährleistung der Langzeitverfügbarkeit
- ▶ keine Nutzungsrechte für den Provider
- ▶ Schutz vor unbefugtem Zugriff

Die Inanspruchnahme eines Dienstleisters bzw. Cloudspeichers ist meist mit laufenden Kosten als monatliche Pauschale oder nach Datenaufkommen verbunden. Neben den Kosten sind auch die Nutzungsbedingungen zu beachten. Die Sicherheit kann signifikant erhöht werden, wenn ein Film zwei- oder dreimal auf unterschiedlichen Datenträgern gespeichert wird und diese an mehreren Orten abgelegt werden.

Content-Preservation

Content-Preservation meint die Sicherstellung, dass der Inhalt eines Objekts langfristig wiedergegeben und interpretiert werden kann.

Auch Bestandserhaltungsmaßnahmen kann ein Privatanwender kaum durchführen. Trotzdem sollte ein genereller Risikofaktor festgehalten werden: Wenn es für ein Objekt zunehmend schwieriger wird, aktuelle Wiedergabesoftware zu finden, ist dies ein deutlicher Hinweis darauf, dass ein Format von aktueller Software nicht mehr unterstützt wird und von Obsoleszenz bedroht ist. In diesem Fall sollten so bald wie möglich alle Objekte, die in dem betroffenen Format vorliegen, in ein neues Zielformat migriert werden.

Empfehlung

Privatanwender können eine umfassende Content-Preservation kaum leisten. Aus diesem Grund sollen hier nur minimale Anforderungen aufgeführt werden:

- ▶ Ablage der Masterdatei in einem nicht oder verlustfrei komprimierten Format
- ▶ Begrenzung auf wenige ausgewählte Archivformate durch Konvertierungs-

maßnahmen

- ▶ Identifizierung und Validierung der Dateiformate und Codecs
- ▶ Extraktion technischer Metadaten
- ▶ Festlegung einiger weniger signifikanter Eigenschaften (z.B. Seitenverhältnis, Farbraum und Auflösung)
- ▶ Dokumentation der Festlegungen und Maßnahmen
- ▶ Überprüfung der Anforderungen nach jeder Bestandserhaltungsmaßnahme

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt das ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigration zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ Nachweis von Veränderungen am Objekt in den Metadaten

Empfehlung

Die Sicherung der Authentizität wird von einem Privatanwender grundsätzlich nicht gefordert. Sie kann dann aber notwendig sein, wenn beabsichtigt wird, die eigene Amateurfilmsammlung später an eine Gedächtnisinstitution zu übergeben. Nützlich ist es in jedem Fall auch, wenn grundlegende signifikante Eigenschaften definiert werden und die Ergebnisse einer Formatmigration gegen diese Anforderungen zu evaluieren sind, z.B.:

- ▶ Beibehaltung des Seitenverhältnisses
- ▶ gleichbleibende Bildqualität
- ▶ Erhaltung des Farbraums
- ▶ Erhaltung von Untertiteln
- ▶ Erhaltung der Synchronität von Bild und Ton

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus

der Datei extrahiert werden.

- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmeteradaten: Provenienz meint zunächst die Stelle, an der das Objekt entstanden ist bzw. den Urheber. Provenienzen sind auch diejenigen Stellen, bei denen das Objekt später verwahrt oder migriert wird. Idealerweise beschreiben diese Metadaten die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.
- ▶ • Protokollmetadaten zur Erfassung der vorgenommenen Maßnahmen.

Empfehlung

Die Erfassung von Metadaten bedeutet für Privatanwender einen erheblichen Aufwand, wenn diese nicht aus einer Quelle importiert werden können. Während für Kinofilme häufig deskriptive Metadaten in Form von Filmdatenbanken wie der Internet Movie Database (IMDB) vorhanden sind, ist dies für Amateurfilme aufgrund der geringeren Verbreitung eher unwahrscheinlich.

Deshalb sollten Privatanwender sich auf grundlegende deskriptive und technische Metadaten beschränken. Im Falle einer späteren Abgabe an eine wissenschaftliche Institution sollten aber auch die Provenienzmeteradaten bekannt sein.

Die erfassten Metadaten müssen verwaltet werden. Es wird dringend davon abgeraten, die Metadaten ausschließlich in der Datei selbst zu speichern. Bei einer Formatmigration können diese Metadaten verloren gehen. Sinnvoller ist es, Objekt und Metadaten getrennt voneinander zu verwalten bzw. zu kapseln. Von der Metadatenverwaltung in Excel- oder Worddateien wird abgeraten. Besser geeignet ist frei verfügbare Datenbanksoftware für Multimediaobjekte. Hierbei ist allerdings darauf zu achten, dass die Metadaten leicht aus dem System exportiert werden können.

Privatanwender sollten sich auf grundlegende deskriptive und technische und Metadaten beschränken, im Falle einer späteren Abgabe an eine wissenschaftliche Institution sollten aber auch die Provenienzmeteradaten bekannt sein.

WICHTIG: Metadaten nicht ausschließlich in der Datei selbst speichern! Bei einer Formatmigration können diese verloren gehen. Sinnvoller: Objekt und Metadaten getrennt voneinander verwalten bzw. kapseln.

Präsentation/Zugriff

Weitverbreitete, offene Formate und Codecs sollten bevorzugt werden. Proprietäre Formate und Codecs sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung.

Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein. Nutzungsformate können entweder für eine allgemeine Präsentation der Objekte oder on demand generiert werden.

Empfehlung

Es empfiehlt sich die parallele Speicherung von Master- und Nutzungsdateien. Für die Nutzung reichen i.d.R. verlustbehaftet komprimierte Dateiformate aus. Zudem sind die Anforderungen an die Hardware geringer. Werden die Archivdateien, wie bei Privatanwendern oftmals der Fall, gleichzeitig auch als Nutzungsdateien verwendet, ist es wichtig, dass Abspielgeräte nur einen Lesezugriff bekommen. Die Präsentation über Webplattformen steht dem Amateurfilmer, der eigene Produkti-

onen als Privatanwender archiviert, neben der lokalen Nutzung und dem Versand von Datenträgern vorbehaltlich der Wahrung der Rechte Dritter frei.

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate und Codecs zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software des Herstellers gebunden und stellen eine Gefahr für die längerfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivformats auch die Kompatibilität eines Formats mit den jeweiligen Codecs. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Bei Amateurfilmen handelt es sich meistens um Consumerformate, d.h. Dateien, die durch verlustbehaftete Komprimierung vergleichsweise wenig Speicherplatz benötigen.

Aus Archivsicht sollte wenn möglich eine Normalisierung in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als Open-Source-Software zur Verfügung.

Als Archivformat bietet sich v.a. Matroska (.mkv) an. Matroska unterstützt zahlreiche Codecs, zu empfehlen ist der verlustfreie Codec FFV1 v.3.*

Empfehlung

Bei Amateurfilmen ist die Vielfalt potenzieller Produktions- und Ablieferungsformate besonders groß. Meistens handelt es sich dabei um Consumerformate, d.h. Dateien, die durch verlustbehaftete Komprimierung vergleichsweise wenig Speicherplatz benötigen und auch von Laien ohne großen Aufwand abgespielt werden können. Teilweise haben diese auch proprietären Charakter. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Content-Preservation). Es ist deshalb bei Amateurfilmen besonders zu betonen, dass die Anzahl der Dateiformate im digitalen Langzeitarchiv sich auf eine kleine Auswahl beschränken sollte.

Wenn möglich, sollte eine Normalisierung in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als Open-Source-Software zur Verfügung. Wenn Normalisierung nicht möglich ist, ist darauf zu achten, dass das Dateiformat die o.g. Anforderungen so gut wie möglich erfüllt.

Als Archivformat für Privatanwender bietet sich v.a. Matroska (*.mkv) an. Dabei handelt es sich um ein freies, robustes, wenig komplexes, plattformunabhängiges und weitverbreitetes Containerformat mit vergleichsweise geringen Dateigrößen. Die Konvertierung nach Matroska ist mit gängiger Konvertierungssoftware (z.B. Wondershare, XMedia Recode, FFmpeg) ohne Weiteres möglich.

Ebenso wichtig wie das Containerformat an sich ist ein geeigneter Codec. Matroska unterstützt zahlreiche Codecs, von denen die meisten allerdings aufgrund der verlustbehafteten Kompression nicht für die Langzeitsicherung geeignet sind. Zu empfehlen ist der verlustfreie Codec FFV1 v.3. Bei einer Konversion ist sicherzustellen, dass keine signifikanten Eigenschaften verloren gehen.

Empfehlungen für kleinere Einrichtungen

Kleinere Einrichtungen besitzen wenig Mitarbeiter, eine begrenzte IT-Infrastruktur und ein geringes Budget. Darunter sind u.U. auch Abteilungen größerer Einrichtungen zu verstehen, bei denen die Archivierung von Filmen und Videomaterialien aufs Ganze betrachtet nur marginal betrieben wird. Die Betreuer sind dann oft keine Medienfachleute, sondern Quereinsteiger (Archivare, Bibliothekare, Dokumentare usw.). Bei den Nutzern handelt es

sich i.d.R. um Personen mit historischen Forschungsinteressen, weniger um Medienfachleute.

Auswahl des zu archivierenden Materials

Amateurfilme und -videos gelangen entweder durch Sammlungstätigkeit oder eine Pflichtübernahme aufgrund eines gesetzlichen Auftrags an eine Gedächtnisinstitution. Diese prüft das angebotene Material auf Archivwürdigkeit. Wichtigstes Kriterium für die Archivwürdigkeit ist die inhaltliche Relevanz für die künftige historische Forschung.

Die Verwahrung von Amateurfilmen durch Gedächtnisinstitutionen ist urheberrechtlich geregelt. Die Modalitäten sind mit dem Produzenten bzw. der abgebenden Stelle vertraglich zu vereinbaren. Die Initiative kann sowohl vom Produzenten (Wunsch, seine Werke der Nachwelt zu erhalten) als auch von der Gedächtniseinrichtung (Feststellung öffentlichen Interesses an diesen Werken) ausgehen. Der Produzent sollte sich ggf. überlegen, ob er seine Werke nicht bereits zu Lebzeiten und möglichst zeitnah nach ihrer Entstehung an eine öffentliche oder private Gedächtnisinstitution übergeben möchte. Dann wäre diese in einem Vertrag zur dauerhaften Erhaltung des Materials verpflichtet und er wiederum von dieser Aufgabe entlastet. Es sollte dann eine Einrichtung bevorzugt werden, die das dafür nötige Know-how in ausreichendem Umfang besitzt.

Gerade für kleinere Gedächtnisinstitutionen wäre es von Vorteil, wenn der Produzent seine Werke bereits im Vorfeld der Abgabe zumindest grob erschlossen hat, sodass sie für die Nachwelt eindeutig identifizierbar bleiben. Der Institution steht es frei, die angebotenen Materialien auf Archivwürdigkeit hin zu bewerten und solche, die sie nicht für dauerhaft erhaltenswert einstuft, entweder in Absprache mit dem Produzenten zu vernichten oder bei diesem zu belassen. Dies sollte dann im Verzeichnis festgehalten werden. Ebenso wichtig wie die Archivwürdigkeit ist aber auch die Archivfähigkeit des zu erhaltenden Materials. Daher sollte dieses in einem geeigneten Archivierungsformat und nicht nur in einem reinen Nutzungsformat vorliegen. Liegen von den Ausgangsfilmen mehrere Versionen vor, was sich sowohl auf den Inhalt als auch auf das Dateiformat beziehen kann, sollte ein geeigneter Archivmaster ausgewählt bzw. generiert werden

Empfehlung

Kleinere Einrichtungen sind aufgrund ihrer beschränkten Mittel gezwungen, besonders stark zu selektieren. Redundanzen müssen vermieden und Ressourcen geschont werden. Die Archivwürdigkeit bemisst sich insbesondere auch am Unikatcharakter der zu übernehmenden Objekte. Selektionskriterien können wie folgt priorisiert werden:

- ▶ Gibt es andere/größere Institutionen, welche die digitale Langzeitarchivierung besser/sicherer gewährleisten können?
- ▶ Ist das Objekt archivwürdig, d.h. kann man annehmen, dass sich in Zukunft Menschen dafür interessieren?
- ▶ Entspricht das Objekt dem Erwerbungsprofil?
- ▶ Liegen alle benötigten Rechte vor (Langzeitarchivierung, Veränderung, Verbreitung)?
- ▶ Sind ausreichend Ressourcen vorhanden, um die digitale Langzeitarchivierung gewährleisten zu können?

Kleinere Einrichtungen sind aufgrund ihrer beschränkten Mittel gezwungen, besonders stark zu selektieren. Redundanzen müssen vermieden und Ressourcen geschont werden. Die Archivwürdigkeit bemisst sich insbesondere auch am Unikatcharakter der zu übernehmenden Objekte.

Beschreibung der vorgesehenen Nutzungsszenarien

Je mehr der Produzent von Amateurfilmen vom öffentlichen Interesse an seinen Werken überzeugt ist, desto mehr sollte er sich um eine geeignete Gedächtnisinstitution zur dauerhaften Verwahrung kümmern. Der Aufbau einer eigenen Infrastruktur zur digitalen Archivierung ist grundsätzlich mit sehr hohen Kosten verbunden. Die künftige Nutzungsfrequenz durch die (film-)historisch interessierte Öffentlichkeit ist hingegen nicht absehbar.

Potenzielle Nutzer können z.B. (Film-)Historiker, sonstige Geistes- und Sozialwissenschaftler, Journalisten oder Studenten sein. Einrichtungen mit einem kleinen Zuständigkeitsbereich werden seltener genutzt als solche mit landes- oder gar bundesweiter Zuständigkeit. In jedem Fall sollte die Einrichtung sich im Klaren darüber sein, welche Eigenschaften der übernommenen Materialien als signifikant und erhaltenswert betrachtet werden können.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren, die die Qualität der Objekte im digitalen Langzeitarchiv sicherstellen sollen. Kleinere Einrichtungen werden i.d.R. wenig Aufwand für das Qualitätsmanagement betreiben. Einige grundsätzliche Maßnahmen sind jedoch erforderlich:

- ▶ Eingangskontrolle, d.h. Prüfung der Objekte auf Lesbarkeit und inhaltliche Vollständigkeit
- ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten (falls diese bereits vorhanden sind)
- ▶ Überprüfung auf technische Fehlerfreiheit
- ▶ Überprüfung auf Standardkonformität
- ▶ Dokumentation und Evaluierung von Verfahren
- ▶ Kontrollverfahren
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigration
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards

Empfehlung

Von kleineren Einrichtungen ist ein umfassendes Qualitätsmanagement nicht zu leisten. Trotzdem wird empfohlen, im Rahmen ihrer Möglichkeiten eine Eingangskontrolle durchzuführen und die Ergebnisse von Formatmigrationen anhand festgelegter Kriterien zu prüfen.

Realistisch für Privat-anwender ist meist nur der regelmäßige Austausch der Speichermedien bei räumlich getrennter Aufbewahrung oder eine Cloudlösung.

Erhaltungsstrategien

Bitstream-Preservation

Die Erhaltung des Datenstroms ist die Basis jeder digitalen Archivierung. Dieser sollte gewährleistet werden durch:

- ▶ Verwendung geeigneter Prüfsummen
- ▶ redundante Sicherung an verschiedenen Orten
- ▶ regelmäßige Virenprüfung
- ▶ regelmäßiger Austausch von Datenträgern

- ▶ Monitoring
- ▶ Skalierbarkeit

Empfehlung

CD-ROMs, DVDs usw. sind als Sicherungsmedien zu vermeiden. Auch die Sicherung auf externen Festplatten stellt allenfalls eine Grundsicherung dar. Realistisch ist meist nur der regelmäßige Austausch der Speichermedien bei räumlich getrennter Aufbewahrung. Besser wäre jedoch ein RAID-10-System, etwa ein NAS-Server. Eine Geodiversität, d.h. die redundante Speicherung der Daten an verschiedenen Orten, ist für kleinere Einrichtungen meist nur mithilfe eines Dienstleisters bzw. einer Cloudlösung denkbar, wobei der Anbieter sich um die Verwaltung und Erhaltung der Speichermedien kümmert.

Kriterien für einen Dienstleister bzw. eine Cloudlösung:

- ▶ Gewährleistung der Langzeitverfügbarkeit
- ▶ keine Nutzungsrechte für den Provider
- ▶ Schutz vor unbefugtem Zugriff

Die Sicherheit kann signifikant erhöht werden, wenn ein Film zwei- oder dreimal auf unterschiedlichen Datenträgern gespeichert wird und diese an mehreren Orten abgelegt werden (Geodiversität).

Content-Preservation

Content-Preservation meint die Sicherstellung, dass der Inhalt eines Objekts langfristig wiedergegeben und interpretiert werden kann.

Voraussetzungen hierfür sind:

- ▶ Identifikation und Validierung von Format, Codec und Version jeder Datei
- ▶ Extraktion technischer Metadaten
- ▶ Beschaffung geeigneter Wiedergabesoftware
- ▶ Identifikation von Risikofaktoren
- ▶ Festlegung signifikanter Eigenschaften
- ▶ Bereitstellung von Kontextinformation für die Wiedergabe und Interpretation der Objekte

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Grundsätzlich gibt es hierfür die beiden Verfahren Migration und Emulation (vgl. nestor Handbuch 2.3, Kap. 8.3 und 8.4). Für audiovisuelle Objekte wird die Formatmigration empfohlen, da es nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten.

Empfehlung

Kleinere Einrichtungen können eine umfassende Content-Preservation wie oben beschrieben oft nicht leisten. Aus diesem Grund sollen hier nur minimale Anforderungen aufgeführt werden:

- ▶ Ablage der Masterdatei in einem nicht oder verlustfrei komprimierten Format
- ▶ Begrenzung auf wenige ausgewählte Archivformate durch Konvertierungs-

maßnahmen

- ▶ Identifizierung und Validierung der Dateiformate und Codecs
- ▶ Extraktion technischer Metadaten
- ▶ Festlegung einiger weniger signifikanter Eigenschaften (z.B. Seitenverhältnis, Farbraum und Auflösung)
- ▶ Dokumentation der Festlegungen und Maßnahmen
- ▶ Überprüfung der Anforderungen nach jeder Bestandserhaltungsmaßnahme

Bestandserhaltungsmaßnahmen können von kleineren Einrichtungen nicht geleistet werden. Ggf. muss über eine Abgabe an eine größere Institution nachgedacht werden. Ein genereller Risikofaktor sollte festgehalten werden: Wenn es für ein Objekt zunehmend schwieriger wird, aktuelle Wiedergabesoftware zu finden, ist dies ein deutlicher Hinweis darauf, dass ein Format von aktueller Software nicht mehr unterstützt wird und von Obsoleszenz bedroht ist. In diesem Fall sollten so bald wie möglich alle Objekte, die in dem betroffenen Format vorliegen, in ein neues Zielformat migriert werden.

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt das ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigration zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ Nachweis von Veränderungen am Objekt in den Metadaten
- ▶ Definition signifikanter Eigenschaften

Empfehlung

Die Sicherung der Authentizität muss auch für kleinere Einrichtungen ein wichtiges Anliegen sein. Grundlegende signifikante Eigenschaften müssen definiert und die Ergebnisse einer Formatmigration gegen diese Anforderungen evaluiert werden, z.B. durch:

- ▶ Beibehaltung des Seitenverhältnisses
- ▶ gleichbleibende Bildqualität
- ▶ Erhaltung des Farbraums
- ▶ Erhaltung von Untertiteln
- ▶ Erhaltung der Synchronität von Bild und Ton

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.

- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.
- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmeteradaten: Provenienz meint zunächst die Stelle, an der das Objekt entstanden ist bzw. den Urheber. Provenienzen sind auch diejenigen Stellen, bei denen das Objekt später verwahrt oder migriert wird. Idealerweise beschreiben diese Metadaten die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.
- ▶ Protokollmetadaten zur Erfassung der vorgenommenen Maßnahmen

Empfehlung

Bereits kleinere Einrichtungen sollten nach Möglichkeit Metadaten aller Kategorien erfassen. Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden. Diese Metadaten sollten möglichst vollständig oder – bei mangelnden Ressourcen – zumindest in Auswahl in ein Langzeitarchivierungssystem überführt werden, das in der Lage ist, auch administrative und Provenienzmeteradaten zu erzeugen und zu verwalten.

Bereits kleinere Einrichtungen sollten nach Möglichkeit Metadaten aller Kategorien erfassen. Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden.

Erschließung

Erschließungsarbeit ist mit verhältnismäßig hohem Aufwand verbunden, falls es hierfür keine Automatisierungsmöglichkeiten gibt. Besonders für kleinere Einrichtungen ist es von großem Vorteil, wenn die verwahrten Amateurfilme bereits vom Produzenten/Urheber bei dessen Tätigkeit oder im Anschluss daran dokumentiert wurden. Der Urheber besitzt im Gegensatz zur verwahrenden Institution oft wichtiges Detailwissen für eine (spätere) Erschließung, das sich der Mitarbeiter einer Gedächtnisinstitution möglicherweise erst erwerben muss. Deskriptive Metadaten aus Filmdatenbanken stehen bei Amateurfilmen, anders als bei Kinofilmen, nicht zur Verfügung. Kleinere Einrichtungen werden ihre Erschließungsmaßnahmen, wenn keine umfangreichen Dokumentationen der abgebenden Stelle vorliegen, auf die Gewährleistung der schnellen Auffindbarkeit beschränken. Eine Tiefenerschließung, etwa die Dokumentation von Einzelszenen, findet i.d.R. nicht statt.

Für die Verwaltung der Metadaten ist eine vom Objekt getrennte Sicherung notwendig, da diese sonst bei der Formatmigration verloren gehen können. Die Metadatenverwaltung sollte in frei verfügbarer und offener Datenbanksoftware erfolgen. Für die Ermittlung von Identifiern bieten sich ISAN, EIDR oder GND an.

Ist eine vollständige Übernahme der Metadaten nicht möglich, sollten mindestens deskriptive, rechtliche und technische Metadaten erfasst werden.

Neben den formalen Metadaten sollten auch inhaltsbeschreibende Schlagworte verwendet werden. Die Erschließungstiefe ist abhängig von den Vorgaben der Institution und den Anforderungen der Nutzer.

Wenn das Nachweissystem der Institution für die Erschließung und Verwaltung von Multimediaelementen geeignet ist, können die audiovisuellen Objekte mit dem Nachweissystem verwaltet und beschrieben werden. Ist das vorhandene Nachweissystem nicht geeignet, sollte frei verfügbare Datenbanksoftware verwendet werden, die für Multimediaobjekte geeignet ist.

Nachweissysteme sind i.d.R. nicht für die Langzeitarchivierung der Objekte selbst geeignet. Es kommt auch ein Digital Asset Management System (DAM)/Media Asset Management System (MAM) in Betracht. Ein DAM/MAM ist allerdings i.d.R. nicht für die Langzeitarchivierung geeignet und muss daher dahingehend untersucht werden.

Für den Nachweis der Objekte und die Erfassung von deskriptiven und strukturellen Metadaten sind Nachweissysteme oder DAM/MAM-Systeme jedoch gut geeignet. Ob rechtliche und technische Metadaten erfasst werden können, muss im Einzelfall geprüft werden.

Langzeitarchivierung

Für die Erfassung von Provenienzmetadaten kann überlegt werden, ob diese separat zu erfassen sind und eine Verknüpfung mit dem bestehenden Nachweissystem erfolgen kann. Besser wäre jedoch der Einsatz eines OAIS-konformen Langzeitarchivierungssystems oder die Abgabe der Objekte an eine Institution, die die Langzeitarchivierung in professionellem Umfang leisten kann.

Präsentation/Zugriff

Weitverbreitete, offene Formate und Codecs sollten bevorzugt werden. Proprietäre Formate und Codecs sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung. Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein. Nutzungsformate können entweder für eine allgemeine Präsentation der Objekte oder on demand generiert werden.

Empfehlung

Es empfiehlt sich die parallele Speicherung von Master- und Nutzungsdateien. Für die Nutzung reichen i.d.R. verlustbehaftet komprimierte Dateiformate aus. Zudem sind die Anforderungen an die Abspielhardware geringer. Werden die Archivdateien gleichzeitig auch als Nutzungsdateien verwendet, ist es wichtig, dass Abspielgeräte nur einen Lesezugriff bekommen.

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate und Codecs zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software des Herstellers gebunden und stellen eine Gefahr für die längerfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivformats auch die Kompa-

Bei Amateurfilmen handelt es sich meistens um Consumerformate, d.h. Dateien, die durch verlustbehaftete Komprimierung vergleichsweise wenig Speicherplatz benötigen.

Aus Archivsicht sollte wenn möglich eine Normalisierung in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als Open-Source-Software zur Verfügung.

Als Archivformat bietet sich v.a. Matroska (.mkv) an. Matroska unterstützt zahlreiche Codecs, zu empfehlen ist der verlustfreie Codec FFV1 v.3.*

tibilität eines Formats mit den jeweiligen Codecs. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Empfehlung

Bei Amateurfilmen ist die Vielfalt potenzieller Produktions- und Ablieferungsformate besonders groß. Meistens handelt es sich dabei um Consumerformate, d.h. Dateien, die durch verlustbehaftete Komprimierung vergleichsweise wenig Speicherplatz benötigen und auch von Laien ohne großen Aufwand abgespielt werden können. Teilweise haben diese auch proprietären Charakter. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Content-Preservation). Es ist deshalb bei Amateurfilmen besonders zu betonen, dass die Anzahl der Dateiformate im digitalen Langzeitarchiv sich auf eine kleine beschränken sollte. Wenn möglich, sollte eine Normalisierung in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als Open-Source-Software zur Verfügung. Wenn Normalisierung nicht möglich ist, ist darauf zu achten, dass das Dateiformat die o.g. Anforderungen so gut wie möglich erfüllt.

Als Archivformat bietet sich v.a. Matroska (*.mkv) an. Dabei handelt es sich um ein freies, robustes, wenig komplexes, plattformunabhängiges und weitverbreitetes Containerformat mit vergleichsweise geringen Dateigrößen. Die Konvertierung nach Matroska ist mit gängiger Konvertierungssoftware (z.B. Wondershare, XMedia Recode, FFmpeg) ohne Weiteres möglich. Ebenso wichtig wie das Containerformat an sich ist ein geeigneter Codec. Matroska unterstützt zahlreiche Codecs, von denen die meisten allerdings aufgrund der verlustbehafteten Kompression nicht für die Langzeitsicherung geeignet sind. Zu empfehlen ist der verlustfreie Codec FFV1 v.3. Bei einer Konversion ist sicherzustellen, dass keine signifikanten Eigenschaften verloren gehen.

Empfehlungen für mittlere Einrichtungen

Mittlere Einrichtungen besitzen eine etwas größere Anzahl an Mitarbeitern, eine IT-Infrastruktur und ein weitreichendes Budget. Es handelt sich um selbstständige Einrichtungen, bei denen die Archivierung von Filmen und Videomaterialien als Hauptaufgabe betrieben wird. Die Betreuer sind sowohl Medienfachleute als auch Quereinsteiger (Archivare, Bibliothekare, Dokumentare usw.). Bei den Nutzern handelt es sich i.d.R. um Personen, die sich speziell mit dieser Art von Medien beschäftigen.

Auswahl des zu archivierenden Materials

Amateurfilme und -videos gelangen entweder durch Sammlungstätigkeit oder eine Pflichtübernahme aufgrund eines gesetzlichen Auftrags an eine Gedächtnisinstitution. Diese prüft das angebotene Material auf Archivwürdigkeit. Wichtigstes Kriterium für die Archivwürdigkeit ist die inhaltliche Relevanz für die künftige historische Forschung.

Die Verwahrung von Amateurfilmen durch Gedächtnisinstitutionen ist urheberrechtlich geregelt. Die Modalitäten sind mit dem Produzenten bzw. der abgebenden Stelle vertraglich zu vereinbaren. Die Initiative kann sowohl vom Produzenten (Wunsch, seine Werke der Nachwelt zu erhalten) als auch von der Gedächtniseinrichtung (Feststellung öffentlichen Interesses an diesen Werken) ausgehen. Der Produzent sollte sich ggf. überlegen, ob er

seine Werke nicht bereits zu Lebzeiten und möglichst zeitnah nach ihrer Entstehung an eine öffentliche oder private Gedächtnisinstitution übergeben möchte. Dann wäre diese in einem Vertrag zur dauerhaften Erhaltung des Materials verpflichtet und er wiederum von dieser Aufgabe entlastet. Es sollte dann eine Einrichtung bevorzugt werden, die das dafür nötige Know-how in ausreichendem Umfang besitzt.

Auch für mittlere Gedächtnisinstitutionen wäre es von Vorteil, wenn der Produzent seine Werke bereits im Vorfeld der Abgabe zumindest grob erschlossen hat, sodass sie für die Nachwelt eindeutig identifizierbar bleiben. Der Institution steht es frei, die angebotenen Materialien auf Archivwürdigkeit hin zu bewerten und solche, die sie nicht für dauerhaft erhaltenswert einstuft, entweder in Absprache mit dem Produzenten zu vernichten oder bei diesem zu belassen. Dies sollte dann im Verzeichnis festgehalten werden.

Ebenso wichtig wie die Archivwürdigkeit ist aber auch die Archivfähigkeit des zu erhaltenden Materials. Daher sollte dieses in einem geeigneten Archivierungsformat und nicht nur in einem reinen Nutzungsformat vorliegen. Liegen von den Ausgangsfilmen mehrere Versionen vor, was sich sowohl auf den Inhalt als auch auf das Dateiformat beziehen kann, sollte ein geeigneter Archivmaster ausgewählt bzw. generiert werden.

Redundanzen sollten vermieden und Ressourcen geschont werden. Die Archivwürdigkeit bemisst sich insbesondere auch am Unikatcharakter der zu übernehmenden Objekte.

Empfehlung

Auch mittlere Einrichtungen betreiben i.d.R. keine vollständige Archivierung, sondern selektieren das ihnen angebotene Filmmaterial. Redundanzen sollten vermieden und Ressourcen geschont werden. Die Archivwürdigkeit bemisst sich insbesondere auch am Unikatcharakter der zu übernehmenden Objekte. Selektionskriterien können wie folgt priorisiert werden:

- ▶ Gibt es andere/größere Institutionen, welche die digitale Langzeitarchivierung besser/sicherer gewährleisten können?
- ▶ Ist das Objekt archivwürdig, d.h. kann man annehmen, dass sich in Zukunft Menschen dafür interessieren?
- ▶ Entspricht das Objekt dem Erwerbungsprofil?
- ▶ Liegen alle benötigten Rechte vor (Langzeitarchivierung, Veränderung, Verbreitung)?
- ▶ Sind ausreichend Ressourcen vorhanden, um die digitale Langzeitarchivierung gewährleisten zu können?

Beschreibung der vorgesehenen Nutzungsszenarien

Je mehr der Produzent von Amateurfilmen vom öffentlichen Interesse an seinen Werken überzeugt ist, desto mehr sollte er sich um eine geeignete Gedächtnisinstitution zur dauerhaften Verwahrung kümmern. Der Aufbau einer eigenen Infrastruktur zur digitalen Archivierung ist grundsätzlich mit sehr hohen Kosten verbunden. Die künftige Nutzungsfrequenz durch die (film-)historisch interessierte Öffentlichkeit ist hingegen nicht absehbar. Potenzielle Nutzer können z.B. (Film-)Historiker, sonstige Geistes- und Sozialwissenschaftler, Journalisten oder Studenten sein.

Einrichtungen mit einem mittleren Zuständigkeitsbereich werden häufiger genutzt als kleinere Einrichtungen. Die Nutzer können daher auch größere Erwartungen in Bezug auf die Anwendung von Langzeitarchivierungsstandards, etwa OAI-Konformität, erwarten. In jedem Fall sollte die Einrichtung sich im Klaren darüber sein, welche Eigenschaften der übernommenen Materialien als signifikant und erhaltenswert betrachtet werden können.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren, welche die Qualität der Objekte im digitalen Langzeitarchiv sicherstellen sollen. Folgende Maßnahmen sind grundsätzlich erforderlich:

- ▶ Eingangskontrolle, d.h. Prüfung der Objekte auf Lesbarkeit und inhaltliche Vollständigkeit
- ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten (falls diese bereits vorhanden sind)
- ▶ Überprüfung auf technische Fehlerfreiheit
- ▶ Überprüfung auf Standardkonformität
- ▶ Dokumentation und Evaluierung von Verfahren
- ▶ Kontrollverfahren
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigration
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards

Empfehlung

Mittlere Einrichtungen sollten in einem vertretbaren Maß zudem auch Ressourcen für ein Qualitätsmanagement aufwenden. Im Rahmen der eigenen Möglichkeiten sind zumindest eine Eingangskontrolle durchzuführen und die Ergebnisse von Formatmigrationen anhand festgelegter Kriterien zu prüfen. Auch eine technische Fehlerfreiheit, Standardkonformität und die Anwendung von Kontrollverfahren wird von mittleren Einrichtungen i.d.R. erwartet.

Erhaltungsstrategien

Bitstream-Preservation

Die Erhaltung des Datenstroms ist die Basis jeder digitalen Archivierung. Dieser sollte gewährleistet werden durch:

- ▶ Erhebung geeigneter Prüfsummen
- ▶ redundante Sicherung an verschiedenen Orten
- ▶ regelmäßige Virenprüfung
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Empfehlung

CD-ROMs, DVDs usw. sind als Sicherungsmedien in jedem Fall zu vermeiden. Auch die Sicherung auf externen Festplatten stellt allenfalls eine Grundsicherung dar. Mittleren Einrichtungen sollte zumindest ein RAID-10-System, etwa ein NAS-Server, zur Verfügung stehen. Die Ansprüche an mittlere Einrichtungen sind allerdings höher. Anzustreben ist in jedem Fall eine Geodiversität (Sicherung an verschiedenen Orten) durch den Aufbau eines eigenen Rechenzentrums mit externen Backup-Möglichkeiten, die Beauftragung eines externen Rechenzentrums oder die Einrichtung einer Verbundlösung mit anderen Institutionen. Die Entscheidung hierfür muss im Einzelfall getroffen werden.

CD-ROMs, DVDs usw. sind als Sicherungsmedien in jedem Fall zu vermeiden. Anzustreben ist in jedem Fall eine Geodiversität (Sicherung an verschiedenen Orten).

Kriterien für einen Dienstleister, eine Verbundlösung oder ein eigenes Rechenzentrum:

- ▶ Gewährleistung der Langzeitverfügbarkeit
- ▶ keine Nutzungsrechte für den Provider
- ▶ Schutz vor unbefugtem Zugriff
- ▶ redundante, geodiverse Speicherung
- ▶ Integritätschecks
- ▶ Viruschecks
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Die Speicherlösung sollte möglichst professionell sein und aus einer Kombination aus Festplatten- und Bandspeicher bestehen. Darüber hinaus müssen Sicherheitsbestimmungen beachtet werden, u.a.:

- ▶ Absicherung des physischen Zugangs durch Zugriffsberechtigungen und Einbruchmeldeanlage
- ▶ digitale Zugriffsberechtigungen
- ▶ Brandschutz
- ▶ Disaster Recovery

Content-Preservation

Content-Preservation meint die Sicherstellung, dass der Inhalt eines Objekts langfristig wiedergegeben und interpretiert werden kann.

Voraussetzungen hierfür sind:

- ▶ Identifikation und Validierung von Format, Codec und Version jeder Datei
- ▶ Extraktion technischer Metadaten
- ▶ Beschaffung geeigneter Wiedergabesoftware
- ▶ Identifikation von Risikofaktoren
- ▶ Festlegung signifikanter Eigenschaften
- ▶ Bereitstellung von Kontextinformation für die Wiedergabe und Interpretation der Objekte

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Grundsätzlich gibt es hierfür die beiden Verfahren Migration und Emulation (vgl. nestor Handbuch 2.3, Kap. 8.3 und 8.4). Für audiovisuelle Objekte wird die Formatmigration empfohlen, da es nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten.

Empfehlung

Mittlere Einrichtungen sollten im Rahmen ihrer Möglichkeiten in die Content-Preservation wie o.g. investieren. Minimale Anforderungen sind:

- ▶ Begrenzung auf wenige ausgewählte Archivformate durch Konvertierungsmaßnahmen
- ▶ Ablage der Masterdatei in einem nicht oder verlustfrei komprimierten Format
- ▶ Identifikation und Validierung der Dateiformate und Codecs
- ▶ Extraktion technischer Metadaten
- ▶ Festlegung der signifikanten Eigenschaften (z.B. Seitenverhältnis, Farbraum und Auflösung)

- ▶ Überprüfung der Anforderungen nach jeder Bestandserhaltungsmaßnahme
- ▶ Identifikation von Risikofaktoren
- ▶ Festlegung eines Turnus für die Überprüfung der Risikofaktoren
- ▶ Festlegung von Bestandserhaltungsmaßnahmen
- ▶ Dokumentation der Festlegungen und Maßnahmen
- ▶ Dokumentation der Verwendung von Tools für die Bestandserhaltungsmaßnahmen

Für mittlere Einrichtungen wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Objekte nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten. In Abhängigkeit von den personellen Ressourcen sollte die Anzahl der Dateiformate begrenzt werden, da gerade bei Amateurfilmen das Ausgangsformat sehr stark variiert und es sich oftmals um ein Consumerformat handelt. Es wird empfohlen, Objekte, die in anderen Dateiformaten vorliegen, nach dem Eingang der Originaldatei zu kopieren und in eines der ausgewählten Dateiformate umzuwandeln. Dafür stehen eine Reihe von Open-Source-Tools zur Verfügung. Die Originaldatei sollte unverändert aufbewahrt werden.

Das Dateiformat sollte identifiziert und validiert werden. Die Dateieindung genügt als Indikator nicht, da diese entweder falsch sein kann oder sich dahinter eine Vielzahl unterschiedlicher Codecs verbergen können, was beträchtliche Unterschiede hinsichtlich der Qualität und der signifikanten Eigenschaften bedeuten kann. Außerdem sollten technische Metadaten extrahiert werden. Das Format und der Codec eines Objekts müssen bekannt sein, da auf diesen alle Erhaltungsmaßnahmen basieren.

Signifikante Eigenschaften sollten definiert werden. Der Fokus sollte auf den für die Institution und deren Nutzer wichtigen Eigenschaften eines Objekts liegen.

Es sollten organisatorische und technische Eigenschaften definiert werden. Diese Anforderungen sollten nach einer Bestandserhaltungsmaßnahme überprüft werden.

Wichtig ist auch die Definition grundlegender Risikofaktoren, z.B.:

- ▶ Gefahr der Obsoleszenz
- ▶ das Objekt liegt in einem proprietären oder kaum verbreiteten Format vor

Tritt einer der Risikofaktoren auf, sollten alle Objekte im betroffenen Format in ein neues Zielformat migriert werden.

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt das ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigration zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ Nachweis von Veränderungen am Objekt in den Metadaten
- ▶ Definition signifikanter Eigenschaften

Empfehlung

Die Sicherung der Authentizität muss für mittlere Einrichtungen ein wichtiges Anliegen sein. Alle Versionen des Originalobjekts sollten aufbewahrt und Änderungen in den Metadaten dokumentiert werden. Grundlegende signifikante Eigenschaften

Die Sicherung der Authentizität muss für mittlere Einrichtungen ein wichtiges Anliegen sein. Alle Versionen des Originalobjekts sollten aufbewahrt und Änderungen in den Metadaten dokumentiert werden. Grundlegende signifikante Eigenschaften müssen definiert und die Ergebnisse einer Formatmigration gegen diese Anforderungen evaluiert werden.

müssen definiert und die Ergebnisse einer Formatmigration gegen diese Anforderungen evaluiert werden, z.B.:

- ▶ Beibehaltung des Seitenverhältnisses
- ▶ gleichbleibende Bildqualität
- ▶ Erhaltung des Farbraums
- ▶ Erhaltung von Untertiteln
- ▶ Erhaltung der Synchronität von Bild und Ton

Gerade bei Amateurfilmen, die oftmals nicht auf hohem technischem Qualitätsniveau entstanden sind, ist es wichtig, dass eine Formatanpassung oder -migration möglichst verlustfrei abläuft und genügend Qualitätsreserven für die Zukunft vorhanden sind.

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifizier.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.
- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmeteradaten: Provenienz meint zunächst die Stelle, an der das Objekt entstanden ist bzw. den Urheber. Provenienzen sind auch diejenigen Stellen, bei denen das Objekt später verwahrt oder migriert wird. Idealerweise beschreiben diese Metadaten die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.
- ▶ Protokollmetadaten zur Erfassung der vorgenommenen Maßnahmen.

Mittlere Einrichtungen sollten Metadaten aller Kategorien erfassen. Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden. Diese Metadaten sollten möglichst vollständig in ein Langzeitarchivierungssystem überführt werden.

Empfehlung

Mittlere Einrichtungen sollten Metadaten aller Kategorien erfassen. Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden. Diese Metadaten sollten möglichst vollständig oder – bei mangelnden Ressourcen – zumindest in Auswahl in ein Langzeitarchivierungssystem überführt werden, das in der Lage ist, auch administrative und Provenienzmeteradaten zu erzeugen und zu verwalten.

Erschließung

Erschließungsarbeit ist mit verhältnismäßig hohem Aufwand verbunden, falls es hierfür keine Automatisierungsmöglichkeiten gibt. Auch für mittlere Einrichtungen ist es von großem Vorteil, wenn die verwahrten Amateurfilme bereits vom Produzenten/Urheber bei dessen Tätigkeit oder im Anschluss daran dokumentiert wurden. Der Urheber besitzt im Gegensatz zur verwahrenden Institution oft wichtiges Detailwissen für eine (spätere) Erschließung, das sich der Mitarbeiter einer Gedächtnisinstitution möglicherweise erst erwerben muss. Deskriptive Metadaten aus Filmdatenbanken stehen bei Amateurfilmen, anders als bei Kinofilmen, nicht zur Verfügung.

Mittlere Einrichtungen sollten ihre Erschließungsmaßnahmen, auch wenn keine umfangreichen Dokumentationen der abgebenden Stelle vorliegen, nicht nur auf die Gewährleistung der schnellen Auffindbarkeit beschränken. Auch eine Tiefenerschließung, etwa die Dokumentation von Einzelszenen, sollte zumindest auf lange Sicht hin angestrebt werden. In jedem Fall ist die Vergabe inhaltsbeschreibender Schlagworte zu empfehlen. Zusätzlich zu den Erschließungsmetadaten müssen unbedingt rechtliche Metadaten erfasst werden.

Für die Verwaltung der Metadaten ist eine vom Objekt getrennte Sicherung notwendig, da diese sonst bei der Formatmigration verloren gehen können. Die Metadatenverwaltung sollte in frei verfügbarer und offener Datenbanksoftware erfolgen. Für die Generierung von Identifiern bieten sich ISAN, EIDR oder GND an.

Neben den formalen Metadaten sollten auch inhaltsbeschreibende Schlagworte verwendet werden. Die Erschließungstiefe ist abhängig von den Vorgaben der Institution und den Anforderungen der Nutzer.

Wenn das Nachweissystem der Institution für die Erschließung und Verwaltung von Multimediaelementen geeignet ist, können die audiovisuellen Objekte mit dem Nachweissystem verwaltet und beschrieben werden. Ist das vorhandene Nachweissystem nicht geeignet, sollte frei verfügbare Datenbanksoftware verwendet werden, die für Multimediaobjekte geeignet ist. Nachweissysteme sind i.d.R. nicht für die Langzeitarchivierung der Objekte selbst geeignet.

Es kommt auch ein Digital Asset Management System (DAM)/Media Asset Management System (MAM) in Betracht. Ein DAM/MAM ist allerdings i.d.R. nicht für die Langzeitarchivierung geeignet. Ob das System für die Langzeitarchivierung geeignet ist, muss im Einzelfall untersucht werden. Für den Nachweis der Objekte und die Erfassung von deskriptiven und strukturellen Metadaten sind Nachweissysteme oder DAM/MAM-Systeme jedoch gut geeignet. Ob rechtliche und technische Metadaten erfasst werden können, muss im Einzelfall geprüft werden. Zusätzlich ist auch darauf zu achten, dass die Metadaten leicht aus dem System exportiert werden können.

Langzeitarchivierung

Für die Erfassung von Provenienzenmetadaten sollte ein OAIS-konformes Langzeitarchivierungssystem eingesetzt werden. Alternativ sollte in Betracht gezogen werden, die Objekte an eine Institution abzugeben, die die Langzeitarchivierung in professionellerem Umfang leisten kann.

Präsentation/Zugriff

Weitverbreitete, offene Formate und Codecs sollten bevorzugt werden. Proprietäre Formate und Codecs sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung. Für verschiedene Nutzungsszenarien kön-

nen verschiedene Nutzungsformate erforderlich sein. Nutzungsformate können entweder für eine allgemeine Präsentation der Objekte oder on demand generiert werden.

Empfehlung

Es empfiehlt sich die parallele Speicherung von Master- und Nutzungsdateien. Für die Nutzung reichen i.d.R. verlustbehaftet komprimierte Dateiformate aus. Zudem sind die Anforderungen an die Abspielhardware geringer. Mittlere Einrichtungen sollten den Anspruch haben, Archiv- und Nutzungsdateien strikt zu trennen. Für verschiedene Nutzungsszenarien können auch verschiedene Nutzungsformate erforderlich sein.

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate und Codecs zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software des Herstellers gebunden und stellen eine Gefahr für die längerfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivformats auch die Kompatibilität eines Formats mit den jeweiligen Codecs. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Es ist bei Amateurfilmen besonders zu betonen, dass die Anzahl der Dateiformate im digitalen Langzeitarchiv sich auf eine kleine Auswahl beschränken sollte. Um den Migrationsaufwand zu verringern, wird empfohlen, gegenüber den abgebenden Stellen eine geringe Anzahl bevorzugter Dateiformate zu definieren und diese in einer Liste zu veröffentlichen.

Als Archivformat bietet sich v.a. Matroska (.mkv) an. Matroska unterstützt zahlreiche Codecs, zu empfehlen ist der verlustfreie Codec FFV1 v.3.*

Empfehlung

Bei Amateurfilmen ist die Vielfalt potenzieller Produktions- und Ablieferungsformate besonders groß. Meistens handelt es sich dabei um Consumerformate, d.h. Dateien, die durch verlustbehaftete Komprimierung vergleichsweise wenig Speicherplatz benötigen und auch von Laien ohne großen Aufwand abgespielt werden können. Teilweise haben diese auch proprietären Charakter. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Content-Preservation). Es ist deshalb bei Amateurfilmen besonders zu betonen, dass die Anzahl der Dateiformate im digitalen Langzeitarchiv sich auf eine kleine Auswahl beschränken sollte. Wenn möglich, sollte eine Normalisierung in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als Open-Source-Software zur Verfügung. Wenn Normalisierung nicht möglich ist, ist darauf zu achten, dass das Dateiformat die o.g. Anforderungen so gut wie möglich erfüllt. Um den Migrationsaufwand zu verringern, wird empfohlen, gegenüber den abgebenden Stellen eine geringe Anzahl bevorzugter Dateiformate zu definieren und diese in einer Liste zu veröffentlichen. Die Vorgehensweise insgesamt sollte in einer öffentlich verfügbaren Policy dokumentiert werden, um den Prozess für Produzenten, abgebende Stellen und Nutzer transparent zu machen.

Als Archivformat bietet sich v.a. Matroska (*.mkv) an. Dabei handelt es sich um ein freies, robustes, wenig komplexes, plattformunabhängiges und weitverbreitetes Containerformat mit vergleichsweise geringen Dateigrößen. Die Konvertierung nach Matroska ist mit gängiger Konvertierungssoftware (z.B. Wondershare, XMedia Recode, FFmpeg) ohne Weiteres möglich. Ebenso wichtig wie das Containerformat an sich ist ein geeigneter Codec. Matroska unterstützt zahlreiche Codecs, von denen die meisten allerdings aufgrund der verlustbehafteten Kompression nicht für die Langzeitsicherung geeignet sind. Zu empfehlen ist der verlustfreie Codec FFV1 v.3. Bei einer Konversion ist sicherzustellen, dass keine signifikanten Eigenschaften verloren gehen.

Empfehlungen für größere Einrichtungen

Größere Einrichtungen besitzen eine hohe Anzahl an fachlich gut ausgebildeten Mitarbeitern, eine sehr gute IT-Infrastruktur und ein langfristig gesichertes Budget, das den Einsatz neuester Techniken („State of the Art“) bzw. den Einkauf erstklassiger Dienstleistungen ermöglicht. Es handelt sich um selbstständige Einrichtungen, bei denen die Archivierung von Filmen und Videomaterialien als Hauptaufgabe betrieben wird. Die Betreuer sind fast ausschließlich Medienfachleute und seltener Quereinsteiger. Bei den Nutzern handelt es sich i.d.R. um Personen, die sich speziell mit dieser Art von Medien beschäftigen.

Auswahl des zu archivierenden Materials

Amateurfilme und -videos gelangen entweder durch Sammlungstätigkeit oder eine Pflichtübernahme aufgrund eines gesetzlichen Auftrags an eine Gedächtnisinstitution. Diese prüft das angebotene Material auf Archivwürdigkeit. Wichtigstes Kriterium für die Archivwürdigkeit ist die inhaltliche Relevanz für die künftige historische Forschung.

Die Verwahrung von Amateurfilmen durch Gedächtnisinstitutionen ist urheberrechtlich geregelt. Die Modalitäten sind mit dem Produzenten bzw. der abgebenden Stelle vertraglich zu vereinbaren. Die Initiative kann sowohl vom Produzenten (Wunsch, seine Werke der Nachwelt zu erhalten) als auch von der Gedächtniseinrichtung (Feststellung öffentlichen Interesses an diesen Werken) ausgehen.

Der Produzent sollte sich ggf. überlegen, ob er seine Werke nicht bereits zu Lebzeiten und möglichst zeitnah nach ihrer Entstehung an eine öffentliche oder private Gedächtnisinstitution übergeben möchte. Dann wäre diese in einem Vertrag zur dauerhaften Erhaltung des Materials verpflichtet und er wiederum von dieser Aufgabe entlastet. Es sollte dann eine Einrichtung bevorzugt werden, die das dafür nötige Know-how in ausreichendem Umfang besitzt. Auch für größere Gedächtnisinstitutionen wäre es von Vorteil, wenn der Produzent seine Werke bereits im Vorfeld der Abgabe zumindest grob erschlossen hat, sodass sie für die Nachwelt eindeutig identifizierbar bleiben. Der Institution steht es frei, die angebotenen Materialien auf Archivwürdigkeit hin zu bewerten und solche, die sie nicht für dauerhaft erhaltenswert einstuft, entweder in Absprache mit dem Produzenten zu vernichten oder bei diesem zu belassen. Dies sollte dann im Verzeichnis festgehalten werden. Ebenso wichtig wie die Archivwürdigkeit ist aber auch die Archivfähigkeit des zu erhaltenden Materials. Daher sollte dieses in einem geeigneten Archivierungsformat und nicht nur in einem reinen Nutzungsformat vorliegen. Liegen von den Ausgangsfilmen mehrere Versionen vor, was sich sowohl auf den Inhalt als auch auf das Dateiformat beziehen kann, sollte ein geeigneter Archivmaster ausgewählt bzw. generiert werden.

Empfehlung

Auch größere Einrichtungen betreiben i.d.R. keine vollständige Archivierung, sondern selektieren das ihnen angebotene Filmmaterial. Redundanzen sollten vermieden und Ressourcen geschont werden. Die Archivwürdigkeit misst sich insbesondere auch am Unikatcharakter der zu übernehmenden Objekte. Selektionskriterien können wie folgt priorisiert werden:

- ▶ Ist das Objekt archivwürdig, d.h. kann man annehmen, dass sich in Zukunft Menschen dafür interessieren?
- ▶ Entspricht das Objekt dem Erwerbungsprofil?
- ▶ Liegen alle benötigten Rechte vor (Langzeitarchivierung, Veränderung, Verbreitung)?

Beschreibung der vorgesehenen Nutzungsszenarien

Je mehr der Produzent von Amateurfilmen vom öffentlichen Interesse an seinen Werken überzeugt ist, desto mehr sollte er sich um eine geeignete Gedächtnisinstitution zur dauerhaften Verwahrung kümmern. Der Aufbau einer eigenen Infrastruktur zur digitalen Archivierung ist grundsätzlich mit sehr hohen Kosten verbunden. Größere Einrichtungen sind finanziell am besten ausgestattet, sollten aber trotzdem bedenken, dass die künftige Nutzungsfrequenz durch die (film-)historisch interessierte Öffentlichkeit nicht letztgültig absehbar ist. Potenzielle Nutzer können z.B. (Film-)Historiker, sonstige Geistes- und Sozialwissenschaftler, Journalisten oder Studenten sein. Einrichtungen mit einem großen Zuständigkeitsbereich werden häufiger genutzt als mittlere oder kleinere Einrichtungen.

Die Nutzer können daher auch größere Erwartungen in Bezug auf die Anwendung von Langzeitarchivierungsstandards, etwa OAIS-Konformität, erwarten. In jedem Fall sollte sich die Einrichtung im Klaren darüber sein, welche Eigenschaften der übernommenen Materialien als signifikant und erhaltenswert betrachtet werden können.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren, die die Qualität der Objekte im digitalen Langzeitarchiv sicherstellen sollen. Folgende Maßnahmen sind grundsätzlich erforderlich:

- ▶ Eingangskontrolle, d.h. Prüfung der Objekte auf Lesbarkeit und inhaltliche Vollständigkeit
- ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten (falls diese bereits vorhanden sind)
- ▶ Überprüfung auf technische Fehlerfreiheit
- ▶ Überprüfung auf Standardkonformität
- ▶ Dokumentation und Evaluierung von Verfahren
- ▶ Kontrollverfahren
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigration
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards
- ▶ Beachtung der Kriterien für die vertrauenswürdige Langzeitarchivierung (nestor-Siegel nach DIN 31644)

Größere Einrichtungen sollten in jedem Fall auch angemessene Ressourcen für ein Qualitätsmanagement aufwenden.

Empfehlung

Größere Einrichtungen sollten in jedem Fall auch angemessene Ressourcen für ein Qualitätsmanagement aufwenden. Im Rahmen der eigenen Möglichkeiten sollten sie alle o.g. Maßnahmen durchführen. Auch eine technische Fehlerfreiheit, Standardkonformität und die Anwendung von Kontrollverfahren wird von größeren Einrichtungen i.d.R. erwartet.

Erhaltungsstrategien

Bitstream-Preservation

Die Erhaltung des Datenstroms ist die Basis jeder digitalen Archivierung. Dieser sollte gewährleistet werden durch:

- ▶ Erhebung geeigneter Prüfsummen
- ▶ redundante Sicherung an verschiedenen Orten (Geodiversität)
- ▶ regelmäßige Virenprüfung
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Empfehlung

Größere Einrichtungen sollten den Anspruch haben, ihre Bitstream-Preservation eigenständig und professionell durchzuführen und nicht von externen Dienstleistern abhängig zu sein. Dabei sind folgende Anforderungen zu erfüllen:

- ▶ Gewährleistung der Langzeitverfügbarkeit
- ▶ Schutz vor unbefugtem Zugriff
- ▶ redundante, geodiverse Speicherung
- ▶ Integritätschecks
- ▶ Viruschecks
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Größere Einrichtungen sollten ihre Bitstream-Preservation eigenständig und professionell durchführen. Die Speicherlösung sollte professionell sein und aus einer Kombination aus Festplatten- und Bandspeicher bestehen. Werden dennoch Dienstleister hinzugezogen, sollte mit diesen ein Service-Level-Agreement (SLA) vereinbart werden.

Die Speicherlösung sollte möglichst professionell sein und aus einer Kombination aus Festplatten- und Bandspeicher bestehen. Darüber hinaus müssen Sicherheitsbestimmungen beachtet werden, u.a.:

- ▶ Absicherung des physischen Zugangs durch Zugriffsberechtigungen und Einbruchmeldeanlage
- ▶ digitale Zugriffsberechtigungen
- ▶ Brandschutz
- ▶ Disaster Recovery

Werden dennoch Dienstleister hinzugezogen, sollte mit diesen ein Service-Level-Agreement (SLA) vereinbart werden. Der Umfang der Dienstleistung muss genau evaluiert werden, um sicherzustellen, dass alle Anforderungen an die Bitstream-Preservation erfüllt werden.

Content-Preservation

Content-Preservation meint die Sicherstellung, dass der Inhalt eines Objekts auch nach vielen Jahren wiedergegeben und interpretiert werden kann.

Voraussetzungen hierfür sind:

- ▶ Identifikation und Validierung von Format, Codec und Version jeder Datei
- ▶ Extraktion technischer Metadaten
- ▶ Beschaffung geeigneter Wiedergabesoftware
- ▶ Identifikation von Risikofaktoren
- ▶ Festlegung signifikanter Eigenschaften
- ▶ Bereitstellung von Kontextinformation für die Wiedergabe und Interpretation der Objekte

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Grundsätzlich gibt es hierfür die beiden Verfahren Migration und Emulation (vgl. nestor Handbuch 2.3, Kap. 8.3 und 8.4). Für audiovisuelle Objekte wird die Formatmigration empfohlen, da es nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten.

Empfehlung

Größere Einrichtungen sollten großzügig in die Content-Preservation wie o.g. investieren und eine klare Erhaltungsstrategie erarbeiten, die folgende Punkte abdecken sollte:

- ▶ Begrenzung auf wenige ausgewählte Archivformate durch Konvertierungsmaßnahmen
- ▶ Ablage der Masterdatei in einem nicht oder verlustfrei komprimierten Format
- ▶ Identifikation und Validierung der Dateiformate und Codecs
- ▶ Extraktion technischer Metadaten
- ▶ Festlegung der signifikanten Eigenschaften (z.B. Seitenverhältnis, Farbraum und Auflösung)
- ▶ Überprüfung der Anforderungen nach jeder Bestandserhaltungsmaßnahme
- ▶ Identifikation von Risikofaktoren
- ▶ Festlegung eines Turnus für die Überprüfung der Risikofaktoren
- ▶ Festlegung von Erhaltungsmaßnahmen für Bestandsgruppen
- ▶ Dokumentation der Festlegungen und Maßnahmen
- ▶ Definierung geeigneter Tools für die Erhaltungsmaßnahmen
- ▶ Erstellung von Policies nach dem SCAPE Policy Framework

Für größere Einrichtungen wird die Formatmigration empfohlen, da für die meisten audiovisuellen Objekte die ursprüngliche Nutzungsumgebung nicht erhalten werden muss. Objekte, die in anderen Dateiformaten vorliegen, sollten in eines der ausgewählten Dateiformate umgewandelt werden. Die Originaldatei sollte unverändert aufbewahrt werden.

Für größere Einrichtungen wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Objekte nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten. In Abhängigkeit von den personellen Ressourcen sollte die Anzahl der Dateiformate begrenzt werden, da gerade bei Amateurfilmen das Ausgangsformat sehr stark variiert und es sich oftmals um ein Consumerformat handelt. Es wird empfohlen, Objekte, die in anderen Dateiformaten vorliegen, nach dem Eingang der Originaldatei zu kopieren und in eines der ausgewählten Dateiformate umzuwandeln. Dafür stehen eine Reihe von Open-Source-Tools zur Verfügung. Die Originaldatei sollte unverändert aufbewahrt werden. Das Dateiformat sollte identifiziert und validiert werden.

Die Dateieindung genügt als Indikator nicht, da diese entweder falsch sein kann oder sich dahinter eine Vielzahl unterschiedlicher Codecs verbergen können, was beträchtliche Unterschiede hinsichtlich der Qualität und der signifikanten Eigenschaften bedeuten kann. Außerdem sollten technische Metadaten extrahiert werden. Das Format und der Codec eines Objekts müssen bekannt sein, da auf diesen alle Erhaltungsmaßnahmen basieren.

Signifikante Eigenschaften sollten definiert werden. Der Fokus sollte auf den für die Institution und deren Nutzer wichtigen Eigenschaften eines Objekts liegen. Es sollten organisatorische und technische Eigenschaften definiert werden. Diese Anforderungen sollten nach einer Bestandserhaltungsmaßnahme überprüft werden.

Wichtig ist auch die Definition grundlegender Risikofaktoren, z.B.:

- ▶ Gefahr der Obsoleszenz
- ▶ das Objekt liegt in einem proprietären oder kaum verbreiteten Format vor

Tritt einer der Risikofaktoren auf, sollten alle Objekte im betroffenen Format in ein neues Zielformat migriert werden.

Authentizität

Authentizität bedeutet der Nachweis, dass ein Objekt das ist, was zu sein es vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigration zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ Nachweis von Veränderungen am Objekt in den Metadaten
- ▶ Definition signifikanter Eigenschaften

Empfehlung

Die Sicherung der Authentizität muss für größere Einrichtungen ein wichtiges Anliegen sein. Alle Versionen des Originalobjekts sollten aufbewahrt und Änderungen in den Metadaten dokumentiert werden. Grundlegende signifikante Eigenschaften müssen definiert und die Ergebnisse einer Formatmigration gegen diese Anforderungen evaluiert werden, z.B.:

- ▶ Beibehaltung des Seitenverhältnisses
- ▶ gleichbleibende Bildqualität
- ▶ Erhaltung des Farbraums
- ▶ Erhaltung von Untertiteln
- ▶ Erhaltung der Synchronität von Bild und Ton

Gerade bei Amateurfilmen, die oftmals nicht auf hohem technischem Qualitätsniveau entstanden sind, ist es wichtig, dass eine Formatanpassung oder -migration möglichst verlustfrei abläuft und genügend Qualitätsreserven für die Zukunft vorhanden sind.

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.
- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren

Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.

- ▶ Provenienzmetadaten: Provenienz meint zunächst die Stelle, an der das Objekt entstanden ist bzw. den Urheber. Provenienzen sind auch diejenigen Stellen, bei denen das Objekt später verwahrt oder migriert wird. Idealerweise beschreiben diese Metadaten die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.
- ▶ Protokollmetadaten zur Erfassung der vorgenommenen Maßnahmen.

Erschließungsarbeit ist mit verhältnismäßig hohem Aufwand verbunden, insbesondere dann, wenn es hierfür keine Automatisierungsmöglichkeiten gibt. Auch für größere Einrichtungen ist es von großem Vorteil, wenn die verwahrten Amateurfilme bereits vom Produzenten/Urheber bei dessen Tätigkeit oder im Anschluss daran dokumentiert wurden.

Der Urheber besitzt im Gegensatz zur verwahrenden Institution oft wichtiges Detailwissen für eine (spätere) Erschließung, das sich der Mitarbeiter einer Gedächtnisinstitution möglicherweise erst erwerben muss. Deskriptive Metadaten aus Filmdatenbanken stehen bei Amateurfilmen, anders als bei Kinofilmen, nicht zur Verfügung. Größere Institutionen sollten ihre Erschließungsmaßnahmen, auch wenn keine umfangreichen Dokumentationen der abgebenden Stelle vorliegen, nicht nur auf die Gewährleistung der schnellen Auffindbarkeit beschränken. Auch eine Tiefenerschließung, etwa die Dokumentation von Einzelszenen, sollte baldmöglichst angestrebt werden. In jedem Fall ist die Vergabe inhaltsbeschreibender Schlagworte zu empfehlen. Zusätzlich zu den Erschließungsmetadaten müssen unbedingt rechtliche Metadaten erfasst werden.

Damit die Metadaten bei der Formatmigration nicht verloren gehen, sollten sie vom Objekt getrennt gesichert werden. Größere Einrichtungen sollten Metadaten aller Kategorien erfassen. Das genutzte Nachweis-system sollte neben deskriptiven auch technische, rechtliche und strukturelle Metadaten erfassen. Das Langzeitarchivierungssystem sollte in der Lage sein, auch administrative und Provenienzmetadaten zu erzeugen und zu verwalten.

Empfehlung

Für die Verwaltung der Metadaten ist eine vom Objekt getrennte Sicherung notwendig, da diese sonst bei der Formatmigration verloren gehen können. Die Metadatenverwaltung sollte in frei verfügbarer und offener Datenbanksoftware erfolgen. Für die Generierung von Identifiern bieten sich ISAN, EIDR, GND oder Wikipedia an.

Größere Einrichtungen sollten Metadaten aller Kategorien erfassen. Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dieses sollte so ausgestattet sein, dass neben deskriptiven Metadaten auch technische, rechtliche und strukturelle Metadaten erfasst werden können. Diese Metadaten sollten möglichst vollständig in ein Langzeitarchivierungssystem überführt werden, das in der Lage ist, auch administrative und Provenienzmetadaten zu erzeugen und zu verwalten.

Erschließung

Neben den formalen Metadaten sollten auch inhaltsbeschreibende Schlagworte verwendet werden. Die Erschließungstiefe ist abhängig von den Vorgaben der Institution und den Anforderungen der Nutzer, sollte aber bei größeren Einrichtungen dementsprechend ausgeprägt sein.

Wenn das Nachweissystem der Institution für die Erschließung und Verwaltung von Multimediaelementen geeignet ist, können die audiovisuellen Objekte mit dem Nachweissystem verwaltet und beschrieben werden. Ist das vorhandene Nachweissystem nicht geeignet, sollte frei verfügbare Datenbanksoftware verwendet werden, die für Multime-

diaobjekte geeignet ist. Nachweissysteme sind i.d.R. nicht für die Langzeitarchivierung der Objekte selbst geeignet.

Es kommt auch ein Digital Asset Management System (DAM)/Media Asset Management System (MAM) in Betracht. Ein DAM/MAM ist allerdings i.d.R. nicht für die Langzeitarchivierung geeignet. Ob das System für die Langzeitarchivierung geeignet ist, muss im Einzelfall untersucht werden.

Für den Nachweis der Objekte und die Erfassung von deskriptiven und strukturellen Metadaten sind Nachweissysteme oder DAM/MAM-Systeme jedoch gut geeignet. Größere Einrichtungen sollten nach Möglichkeit ein System verwenden, in dem auch rechtliche und technische Metadaten erfasst werden können, was vor der Beschaffung genau zu prüfen ist. Zusätzlich ist auch darauf zu achten, dass die Metadaten leicht aus dem System exportiert werden können.

Langzeitarchivierung

Größere Einrichtungen werden möglicherweise des Öfteren angefragt, Objekte und Bestände zu übernehmen, deren Langzeitarchivierung für kleinere oder mittlere Einrichtungen nicht möglich ist. Daher muss ein OAIS-konformes Langzeitarchivierungssystem eingesetzt werden, das in der Lage ist, alle erforderlichen Metadaten zu verwalten.

Präsentation/Zugriff

Weitverbreitete, offene Formate und Codecs sollten bevorzugt werden. Proprietäre Formate und Codecs sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung. Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein. Nutzungsformate können entweder für eine allgemeine Präsentation der Objekte oder on demand generiert werden.

Empfehlung

Es empfiehlt sich die parallele Speicherung von Master- und Nutzungsdateien. Für die Nutzung reichen i.d.R. verlustbehaftet komprimierte Dateiformate aus. Zudem sind die Anforderungen an die Abspielhardware geringer. Größere Einrichtungen sollten den Anspruch haben, Archiv- und Nutzungsdateien strikt zu trennen. Für verschiedene Nutzungsszenarien können auch verschiedene Nutzungsformate erforderlich sein.

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate und Codecs zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software des Herstellers gebunden und stellen eine Gefahr für die längerfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivformats auch die Kompatibilität eines Formats mit den jeweiligen Codecs. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Es empfiehlt sich die parallele Speicherung von Master- und Nutzungsdateien. Für die Nutzung reichen i.d.R. verlustbehaftet komprimierte Dateiformate aus, auch sind die Anforderungen an die Abspielhardware geringer. Archiv- und Nutzungsdateien sollten strikt getrennt werden. Für verschiedene Nutzungsszenarien können auch verschiedene Formate erforderlich sein.

Es wird eine Normalisierung in ein offenes und standardisiertes Dateiformat empfohlen, wofür verschiedene Konvertierungsprogramme frei zur Verfügung stehen. Um den Migrationsaufwand zu verringern, sollte gegenüber den abgebenden Stellen eine geringe Anzahl bevorzugter Dateiformate definiert und in einer Liste veröffentlicht werden.

Als Archivformat bietet sich v.a. Matroska (.mkv) an. Dabei handelt es sich um ein freies, robustes, wenig komplexes, plattformunabhängiges Containerformat mit vergleichsweise geringen Dateigrößen. Ebenso wichtig wie das Containerformat ist ein geeigneter Codec. Zu empfehlen ist der verlustfreie Codec FFV1 v.3.*

Empfehlung

Bei Amateurfilmen ist die Vielfalt potenzieller Produktions- und Ablieferungsformate besonders groß. Meistens handelt es sich dabei um Consumerformate, d.h. Dateien, die durch verlustbehaftete Komprimierung vergleichsweise wenig Speicherplatz benötigen und auch von Laien ohne großen Aufwand abgespielt werden können. Teilweise haben diese auch proprietären Charakter. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Content-Preservation). Es ist deshalb bei Amateurfilmen besonders zu betonen, dass die Anzahl der Dateiformate im digitalen Langzeitarchiv sich auf eine kleine Auswahl beschränken sollte.

Wenn möglich, sollte eine Normalisierung in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als Open-Source-Software zur Verfügung. Wenn Normalisierung nicht möglich ist, ist darauf zu achten, dass das Dateiformat die o.g. Anforderungen so gut wie möglich erfüllt.

Um den Migrationsaufwand zu verringern, wird empfohlen, gegenüber den abgebenden Stellen eine geringe Anzahl bevorzugter Dateiformate zu definieren und diese in einer Liste zu veröffentlichen. Die Vorgehensweise insgesamt sollte in einer öffentlich verfügbaren Policy dokumentiert werden, um den Prozess für Produzenten, abgebende Stellen und Nutzer transparent zu machen.

Als Archivformat bietet sich v.a. Matroska (*.mkv) an. Dabei handelt es sich um ein freies, robustes, wenig komplexes, plattformunabhängiges und weitverbreitetes Containerformat mit vergleichsweise geringen Dateigrößen. Die Konvertierung nach Matroska ist mit gängiger Konvertierungssoftware (z.B. Wondershare, XMedia Recode, FFmpeg) ohne Weiteres möglich. Ebenso wichtig wie das Containerformat an sich ist ein geeigneter Codec. Matroska unterstützt zahlreiche Codecs, von denen die meisten allerdings aufgrund der verlustbehafteten Kompression nicht für die Langzeitsicherung geeignet sind. Zu empfehlen ist der verlustfreie Codec FFV1 v.3. Bei einer Konversion ist sicherzustellen, dass keine signifikanten Eigenschaften verloren gehen.

Archivierung von Wissenschaftlichen Filmen

Definition

Unter Wissenschaftlichem Film verstehen wir die Nutzung von zu wissenschaftlichen Zwecken hergestellten Produktionen. Wesentliches Merkmal des Wissenschaftlichen Films ist die Erstellung des Materials zu wissenschaftlichen Zwecken in Forschung und Lehre.

Prinzipiell kann jedes audiovisuelle Werk wissenschaftlich genutzt werden.

Zum Wissenschaftlichen Film zählen wir u.a. Vorlesungsaufzeichnungen, Konferenzaufzeichnungen, Filmaufnahmen von Experimenten und Dokumentationsfilme, die sich inhaltlich und konzeptionell voneinander unterscheiden. Sie sind nicht auf bestimmte Themengebiete festgelegt.

Die Zielgruppe des Wissenschaftlichen Films ist ein wissenschaftliches Fachpublikum (Wissenschaftler, Studierende, Tagungsteilnehmer). Zweck des Films ist nicht Unterhaltung, sondern die Verwendung in Forschung und Lehre.

Beispiele

Vorlesungs- und Konferenzaufzeichnungen

Zu unterscheiden ist die bloße Aufzeichnung eines Vortrags von Vortragsaufzeichnungen, die multimedial, z.B. mit Präsentationsfolien, angereichert wurden. Hier sind oft die Präsentationsfolien und der Referent gleichzeitig im Bild zu sehen. Spezialisierte Aufnahme-Software in Form von Open-Source- oder kommerziellen Lösungen unterstützen den Produktionsprozess dieser Vortragsaufzeichnungen.

Grenzbereiche wie interaktive Flashfilme von Webinaren werden nicht dazugezählt. Multimediale Vermittlungsformen wie Webinare sind ebenfalls nicht Teil dieser Betrachtung.

Beispiele für die Aufzeichnung einer Vorlesung:

- ▶ DOI: <http://dx.doi.org/10.3203/IWF/G-119>
- ▶ DOI: <http://dx.doi.org/10.5446/14759>

Beispiel für die Aufzeichnung einer Vorlesung mit im Film eingebetteten Präsentationsfolien:

- ▶ DOI: <http://dx.doi.org/10.5446/13208>

Letztere sind heute aufgrund des größeren Mehrwerts für die Zielgruppe und spezialisierter Aufnahmesoftware zunehmend verbreiteter.

Experimentaufzeichnungen

Oft beinhalten Experimentaufzeichnungen neben der Durchführung des Experimentes eine Beschreibung des Versuchsaufbaus, der Methode und des Ergebnisses. Manchmal sind sie zusätzlich mit Texten oder Bildern angereichert.

Beispiel für eine Experimentaufzeichnung:

- ▶ DOI: <http://dx.doi.org/10.3203/IWF/C-14881>

Dokumentarische Filme

Der dokumentarische Film kann Ereignisse von wissenschaftlichem Interesse festhalten. Die wissenschaftlichen Verfahren sollen zugänglich gemacht und die Erkenntnisse vermittelt werden.

Beispiel für einen Dokumentationsfilm mit Kommentar:

- ▶ DOI: <http://dx.doi.org/10.3203/IWF/C-1790>

Beispiel für einen Dokumentationsfilm ohne Kommentar:

- ▶ DOI: <http://dx.doi.org/10.3203/IWF/E-2896>

Empfehlungen für Privatanwender

Auswahl des zu archivierenden Materials

Selektionskriterien bilden die Entscheidungsgrundlage für die Frage, ob ein Objekt langfristig gesichert werden soll. Hierbei spielen neben rechtlichen auch inhaltliche Kriterien eine Rolle. Besondere Bedeutung kommt der Frage nach den Ressourcen zu.

Privatanwender sollten bewusst selektieren, z.B. nach thematischen Schwerpunkten oder der Relevanz eines Objekts für eine spezifische Fragestellung. Auch größtmögliche Vollständigkeit einer Sammlung kann eine Rolle spielen.

Empfehlung

Privatanwender sollten in Abhängigkeit von ihren persönlichen Anforderungen bewusst selektieren, z.B. nach thematischen Schwerpunkten oder der Relevanz eines Objekts für eine spezifische Fragestellung. Auch größtmögliche Vollständigkeit einer Sammlung kann eine Rolle spielen.

Beschreibung der vorgesehenen Nutzungsszenarien

Empfehlung

Vorgesehene Nutzerzielgruppen sind für Privatanwender nicht relevant, es sollte aber die zukünftige Nutzung der Objekte berücksichtigt werden.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren und Prozesse, die die Qualität der Objekte sowie der Prozesse im digitalen Langzeitarchiv sicherstellen sollen.

Darunter fallen z.B.:

- ▶ Übernahmevereinbarung (ein Dokument, das organisatorische, rechtliche und technische Punkte bei der Übernahme berücksichtigt)
- ▶ Eingangskontrolle der Objekte (Überprüfung der Objekte beim Eingang in das digitale Langzeitarchiv)
 - ▶ Prüfung auf Lesbarkeit und inhaltliche Vollständigkeit
 - ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten
 - ▶ technische Fehlerfreiheit
 - ▶ Standardkonformität
- ▶ Dokumentation von Prozessen (Genaue Beschreibung von Prozessen im digitalen Langzeitarchiv: Was wird wann warum wie gemacht? Z.B. schriftlich oder in einer Notation für Prozessdokumentation wie BPMN 2.0.)
- ▶ Evaluierung von Prozessen (Dokumentierte Prozesse können nach Kriterien wie Praxistauglichkeit, Zeitaufwand, Ressourcenaufwand, Überschneidungen mit anderen Prozessen usw. untersucht werden. Die Analyse ist Voraussetzung für die Optimierung der Prozesse.)
- ▶ Kontrollverfahren (technische und organisatorische Verfahren für die Kontrolle von definierten Regeln und Strukturen oder zur Evaluierung von Ergebnissen nach Änderungen an Objekten)
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigrationen
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards
- ▶ Beachtung der Kriterien für die vertrauenswürdige Langzeitarchivierung (nestor-Siegel nach DIN 31644[1]⁷, auch Data Seal of Approval⁸)

Empfehlung

Von Privatanwendern ist ein umfassendes Qualitätsmanagement nicht zu leisten. Trotzdem wird Privatanwendern empfohlen, im Rahmen ihrer Möglichkeiten eine Eingangskontrolle durchzuführen und die Ergebnisse von Formatmigrationen anhand festgelegter Kriterien zu prüfen.

Erhaltungsstrategien

Erhaltungsstrategien stellen sicher, dass das Objekt als Datenstrom erhalten und nutzbar bleibt. Dabei sichern Verfahren für die Bitstream-Preservation den Datenstrom und Verfahren für die Content-Preservation die Les- und Interpretierbarkeit des Objekts.

Bitstream-Preservation

Digitale Datenträger sind nur von begrenzter Haltbarkeit. Sie sind nicht nur von Verschleiß betroffen, auch Hardwareobsoleszenz ist problematisch. Aus diesem Grund ist ein Speicherkonzept essenziell für die Erhaltung des Bitstroms der digitalen Objekte.

⁷ http://files.d-nb.de/nestor/materialien/nestor_mat_17.pdf

⁸ <http://datasealofapproval.org/en/information/guidelines/>

Grundlegende Anforderungen an das Speicherkonzept sind:

- ▶ redundante, geografisch verteilte Speicherlösung
- ▶ Integritätschecks, z.B. mit Prüfsummen
- ▶ Viruscheck
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Die Sicherung des Bitstroms eines Objekts ist ein stetig wiederkehrender Prozess und Basis jeder Erhaltungsstrategie.

Statt des Betriebes eines entsprechenden Speichers können Privatanwender auch mehrere Kopien eines Objekts auf mehreren Festplatten, geografisch getrennt, redundant speichern. Die Speichermedien müssen regelmäßig ersetzt werden. Eine Alternative ist die Speicherung in einer Cloud, wobei der Anbieter sich um die Verwaltung und Erhaltung der Speichermedien kümmert. Wenn die Kosten hierfür zu hoch sind, sollte in Betracht gezogen werden die Objekte an eine Gedächtnisinstitution abzugeben.

Empfehlung

Von Privatanwendern ist kaum zu erwarten, einen entsprechenden Speicher zu betreiben. Deswegen werden an dieser Stelle Alternativen empfohlen.

Sinnvoll ist die geografisch getrennte, redundante Speicherung mehrerer Kopien eines Objekts, z.B. eine Arbeitskopie und zwei Backups auf mehreren Festplatten, und regelmäßiges Ersetzen der Speichermedien. Ein Viruscheck sollte beim Eingang eines Objekts durchgeführt werden.

Eine Alternative ist die Speicherung der Objekte in einer Cloud, wobei der Anbieter sich um die Verwaltung und Erhaltung der Speichermedien kümmert. Cloudspeicher ist meist mit laufenden Kosten als monatliche Pauschale oder nach Datenaufkommen verbunden. Neben den Kosten sind auch die Nutzungsbedingungen zu beachten: Werden dem Anbieter mit dem Upload Nutzungsrechte an den Objekten eingeräumt? Wie schützt der Anbieter die Objekte vor unbefugtem Zugriff? In welchem Land stehen die Server?

Spezialisierte Dienstleister sind eine weitere Möglichkeit. Zu beachten ist hierbei, ob der Dienstleister lediglich die Speicherung der Objekte anbietet oder darüber hinaus tatsächlich die digitale Langzeitarchivierung der Objekte gewährleistet. Einige große Informations- und Gedächtnisinstitutionen in Deutschland konzipieren derzeit Dienstleistungsangebote, es gibt aber auch kommerzielle Anbieter. Sinnvoll ist die Wahl von zwei verschiedenen Speicherkonzepten.

Wenn der Verwaltungsaufwand der Festplatten und die Kosten für den Cloudspeicher und den Dienstleister zu hoch sind, sollte in Betracht gezogen werden, ob die Objekte an eine Gedächtnisinstitution abgegeben werden können, die die adäquate Langzeitarchivierung leisten kann.

Content-Preservation

Content-Preservation soll sicherstellen, dass der Inhalt eines Objekts auch nach Jahren wiedergegeben und interpretiert werden kann.

Dafür ist es erforderlich:

- ▶ Das Dateiformat, die -version und den Codec zu identifizieren und zu validieren. Bei AV-Medien wird diese Erkenntnis durch den Codec erschwert, der in dem Containerformat enthalten ist.

- ▶ technische Metadaten zu extrahieren (siehe Tabelle mit Tools, S. 110 ff.)
- ▶ Informationen über geeignete Wiedergabesoftware zu sammeln
- ▶ Risikofaktoren zu identifizieren
- ▶ signifikante Eigenschaften festzulegen
- ▶ Kontextinformationen bereitzustellen, die die Wiedergabe des Objekts und die Interpretation des Inhalts ermöglichen

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Dafür gibt es zwei Verfahren: Formatmigration und Emulation⁹.

Empfehlung

Privatanwender können umfassende Bestandserhaltungsmaßnahmen wie oben beschrieben kaum leisten. Aus diesem Grund sollen hier nur minimale Anforderungen aufgeführt werden.

Es wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Objekte nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten. Eine Ausnahme stellt Videokunst dar, die im Wissenschaftlichen Film aber nicht als solche, sondern höchstens als Repräsentation vorkommt.

Gleichzeitig sollte die Anzahl der Archivformate nach Möglichkeit auf ein bis einige wenige ausgewählte Dateiformate begrenzt werden. Es wird empfohlen, Objekte, die in anderen Dateiformaten vorliegen, in eines der ausgewählten Dateiformate umzuwandeln. Dafür stehen eine Reihe von Open-Source- und kommerzieller Software zur Verfügung.

Das Dateiformat sollte identifiziert und validiert werden. Es genügt nicht, ein Objekt über die Dateiendung zu identifizieren, da diese auch irreführend sein kann. Außerdem sollten technische Metadaten extrahiert werden. Das Dateiformat eines Objekts und der Codec müssen bekannt sein, da alle Erhaltungsmaßnahmen hierauf basieren.

Signifikante Eigenschaften sollten in geringem Umfang festgelegt werden. Der Fokus sollte auf den Eigenschaften eines Objektes liegen, die dem Privatanwender wichtig sind (z.B. das Seitenverhältnis, der Farbraum, die Auflösung). Diese Anforderungen sollten nach einer Bestandserhaltungsmaßnahme überprüft werden.

Risikomanagement kann von Privatanwendern kaum geleistet werden, trotzdem soll ein Risikofaktor hier festgehalten werden: Wenn es für ein Objekt zunehmend schwieriger wird, aktuelle Wiedergabesoftware zu finden, ist dies ein deutlicher Hinweis darauf, dass das Format von aktueller Software nicht mehr unterstützt wird und von Obsoleszenz bedroht ist. In diesem Fall sollten so bald wie möglich alle Objekte, die in dem betroffenen Format vorliegen, in ein neues Zielformat migriert werden.

Die Anzahl der Archivformate sollte nach Möglichkeit auf ein bis einige wenige ausgewählte Dateiformate begrenzt werden. Außerdem sollten technische Metadaten extrahiert werden. Signifikante Eigenschaften sollten in geringem Umfang festgelegt werden. Der Fokus sollte auf den Eigenschaften eines Objekts liegen, die dem Privatanwender wichtig sind.

⁹ Siehe http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf Kapitel 8.3 und 8.4.

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigrationen zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Von einem Privatanwender wird die Sicherung der Authentizität nicht gefordert. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ die Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ der Nachweis von Veränderungen am Objekt in den Metadaten

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier, Nutzerdaten, Rollen- und Rechtevergabe.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, den Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.
- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmeteradaten: Diese Metadaten beschreiben die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für die Authentizität.

Die Erfassung von Metadaten bedeutet für Privatanwender einen erheblichen Aufwand, wenn diese nicht aus einer Quelle importiert werden können. Deshalb sollten sich Privatanwender auf grundlegende deskriptive und technische Metadaten beschränken. Es wird dringend davon abgeraten, die Metadaten ausschließlich in der Datei selbst zu speichern. Bei einer Formatmigration können diese Metadaten verloren gehen.

Empfehlung

Die Erfassung von Metadaten bedeutet für Privatanwender einen erheblichen Aufwand, wenn diese nicht aus einer Quelle importiert werden können. Während für Kinofilme häufig deskriptive Metadaten in Form von Filmdatenbanken wie der Internet Movie Database (IMDB) vorhanden sind, ist dies für den Wissenschaftlichen Film aufgrund der geringeren Popularität eher unwahrscheinlich. Deshalb sollten sich Privatanwender auf grundlegende deskriptive und technische Metadaten beschränken.

Die erfassten Metadaten müssen verwaltet werden. Es wird dringend davon abgeraten, die Metadaten ausschließlich in der Datei selbst zu speichern. Bei einer Formatmigration können diese Metadaten verloren gehen. Sinnvoller ist es, Objekt und Metadaten getrennt voneinander zu verwalten. Von der Metadatenverwaltung in Excel- oder Worddateien wird abgeraten. Besser geeignet ist frei verfügbare Datenbanksoftware für Multimediaobjekte (siehe Tabelle mit Tools, S. 110 ff).

Präsentation/Zugriff

Weitverbreitete, offene Formate und Codecs sollten bevorzugt werden; proprietäre Formate und Codecs sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung.

Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein.

Empfehlung

Bei Privatanwendern ist das Nutzungsformat oft identisch mit dem Archivformat. In diesem Fall muss das Archivformat für die Nutzung geeignet sein. Alternativ kann eine Konvertierung in ein geeigneteres Format in Betracht gezogen werden.

Bei Privatanwendern kommen in erster Linie folgende Nutzungsszenarien vor:

- ▶ lokale Nutzung
- ▶ Versand von Datenträgern
- ▶ Onlinezugriff

Dateiformate und Codecs

Für das Archivformat und den verwendeten Codec sind offene, standardisierte und weitverbreitete Formate zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software eines Herstellers gebunden und stellen eine potenzielle Gefahr für die langfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivformats auch die Kompatibilität von Format und Codec. Eine Zusammenstellung finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Darüber hinaus muss basierend auf den Anforderungen der vorgesehenen Nutzerzielgruppen auch eine Entscheidung über die Formate für Nutzungskopien getroffen werden.

Empfehlung

Die große Vielfalt potenzieller Ablieferungsformate und Codecs stellt Privatanwender vor eine große Herausforderung. Je höher die Anzahl der Dateiformate und Codecs, desto größer ist der Aufwand für die Bestandserhaltung (Content-Preservation). Es ist deshalb sinnvoll, die Anzahl der Dateiformate und Codecs im digitalen Langzeitarchiv auf eine kleine Auswahl zu beschränken.

Wenn möglich, sollte eine Normalisierung (Vereinheitlichung) in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren. Für die Normalisierung stehen verschiedene Konvertierungsprogramme als proprietäre oder Open-Source-Software zur Verfügung.

Wenn Normalisierung nicht möglich ist, ist darauf zu achten, dass das Dateiformat die o.g. Anforderungen so gut wie möglich erfüllt.

Die große Vielfalt potenzieller Ablieferungsformate und Codecs stellt Privatanwender vor eine große Herausforderung. Wenn möglich, sollte eine Normalisierung (Vereinheitlichung) in ein offenes und standardisiertes Dateiformat erfolgen, um den Verwaltungsaufwand für verschiedene Dateiformate und Wiedergabesoftware zu minimieren.

Soll für die Nutzung auf den Archivmaster zugegriffen werden, ist unbedingt zu beachten, dass die Wiedergabesoftware den Archivmaster nicht (versehentlich) manipulieren kann. Dies lässt sich z.B. durch read-only-Zugriff auf den Archivmaster erreichen.

Als Archivformat bietet sich v.a. Matroska (*.mkv) an. Dabei handelt es sich um ein freies, robustes, wenig komplexes, plattformunabhängiges und weitverbreitetes Containerformat mit vergleichsweise geringen Dateigrößen. Die Konvertierung nach Matroska ist mit gängiger Konvertierungssoftware (z.B. Wondershare, XMedia Recode, FFmpeg) ohne Weiteres möglich. Ebenso wichtig wie das Containerformat an sich ist ein geeigneter Codec. Matroska unterstützt zahlreiche Codecs, von denen die meisten allerdings aufgrund der verlustbehafteten Kompression nicht für die Langzeitsicherung geeignet sind. Zu empfehlen ist der verlustfreie Codec FFV1 v.3. Bei einer Konversion ist sicherzustellen, dass keine signifikanten Eigenschaften verloren gehen.

Empfehlungen für kleinere Einrichtungen

Auswahl des zu archivierenden Materials

Selektionskriterien bilden die Entscheidungsgrundlage für die Frage, ob ein Objekt langfristig gesichert werden soll. Hierbei spielen neben rechtlichen auch inhaltliche Kriterien eine Rolle. Besondere Bedeutung kommt der Frage nach den Ressourcen zu.

Empfehlung

Kleine Einrichtungen sind aufgrund ihrer beschränkten Mittel häufig gezwungen, stark zu selektieren. Dabei sollte die Frage berücksichtigt werden, ob Objekte auch anderen Institutionen überlassen/angeboten werden können, um die eigene Belastung zu reduzieren.

Selektionskriterien können wie folgt priorisiert werden:

- ▶ Gibt es andere Institutionen, die die digitale Langzeitarchivierung des Objekts besser/sicherer gewährleisten können? Ist das Objekt archivwürdig?
- ▶ Entspricht das Objekt dem Sammlungsprofil bzw. dem Archivauftrag?
- ▶ Liegen alle benötigten Rechte vor (Langzeitarchivierung, Veränderung, Verbreitung)?
- ▶ Sind ausreichend Ressourcen vorhanden, um die digitale Langzeitarchivierung gewährleisten zu können?

Beschreibung der vorgesehenen Nutzungsszenarien

Die vorgesehenen Zielgruppen sind die aktuellen und künftigen Nutzergruppen, die mit dem digitalen Langzeitarchiv oder daran angeschlossenen Nutzungsplattformen interagieren. Unter Nutzerzielgruppen sind an dieser Stelle nicht nur die klassischen Nutzer zu verstehen. Vorgesehene Nutzerzielgruppen sind auch externe Nutzer wie Produzenten, die Objekte abliefern, und interne Nutzer wie die Mitarbeiter der Institution. Die vorgesehenen Nutzerzielgruppen entsprechen den klassischen, bekannten Zielgruppen der Institution und können unterschiedliche Anforderungen haben, die gelegentlich bereits analysiert worden sind.

Zusätzlich kommen als vorgesehene Nutzerzielgruppen hinzu:

- ▶ die Datenproduzenten
- ▶ die mit der digitalen Langzeitarchivierung befassten Mitarbeiter

Die verschiedenen vorgesehenen Nutzerzielgruppen haben bestimmte Anforderungen. Daraus ergeben sich Nutzungsszenarien und signifikante Eigenschaften, die Einfluss auf Erhaltungsmaßnahmen haben.

Empfehlung

Die Nutzerzielgruppen sollten der Institution bekannt sein. Sofern die personellen Ressourcen ausreichen, sollte eine Analyse der Nutzerzielgruppen durchgeführt werden, die neben den klassischen Zielgruppen der Institution auch die Datenproduzenten und die Mitarbeiter berücksichtigt.

Die Nutzerzielgruppen sollten der Institution bekannt sein. Sofern die personellen Ressourcen ausreichen, sollte eine Analyse der Nutzerzielgruppen durchgeführt werden, die neben den klassischen Zielgruppen der Institution auch die Datenproduzenten und die Mitarbeiter berücksichtigt.

Die Analyse der Nutzerzielgruppen sollte folgende Punkte beinhalten:

- ▶ die Beschreibung und Abgrenzung der verschiedenen Nutzerzielgruppen
- ▶ die Beschreibung der Anforderungen der Nutzerzielgruppen
- ▶ in den Nutzerzielgruppen verbreitete Nutzerformate

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren und Prozesse, die die Qualität der Objekte sowie der Prozesse im digitalen Langzeitarchiv sicherstellen sollen.

Darunter fallen z.B.:

- ▶ Übernahmevereinbarung (ein Dokument, das organisatorische, rechtliche und technische Punkte bei der Übernahme berücksichtigt)
- ▶ Eingangskontrolle der Objekte (Überprüfung der Objekte beim Eingang in das digitale Langzeitarchiv)
 - ▶ Prüfung auf Lesbarkeit und inhaltliche Vollständigkeit
 - ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten
 - ▶ technische Fehlerfreiheit
 - ▶ Standardkonformität
- ▶ Dokumentation von Prozessen (Genauere Beschreibung von Prozessen im digitalen Langzeitarchiv: Was wird wann warum wie gemacht? Z.B. schriftlich oder in einer Notation für Prozessdokumentation wie BPMN 2.0.)
- ▶ Evaluierung von Prozessen (Dokumentierte Prozesse können nach Kriterien wie Praxistauglichkeit, Zeitaufwand, Ressourcenaufwand, Überschneidungen mit anderen Prozessen usw. untersucht werden. Die Analyse ist Voraussetzung für die Optimierung der Prozesse.)
- ▶ Kontrollverfahren (technische und organisatorische Verfahren für die Kontrolle von definierten Regeln und Strukturen oder zur Evaluierung von Ergebnissen nach Änderungen an Objekten)
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigrationen
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards
- ▶ Beachtung der Kriterien für die vertrauenswürdige Langzeitarchivierung (nestor-Siegel nach DIN 31644¹⁰, auch Data Seal of Approval¹¹)

¹⁰ http://files.d-nb.de/nestor/materialien/nestor_mat_17.pdf

¹¹ <http://datasealofapproval.org/en/information/guidelines/>

Empfehlung

Kleine Institutionen sollten im Rahmen ihrer personellen und finanziellen Ressourcen Maßnahmen für das Qualitätsmanagement implementieren.

Es wird empfohlen, eine Eingangskontrolle durchzuführen und Kontrollverfahren einzurichten. Darüber hinaus sollten Standards genutzt werden.

Die Eingangskontrolle sollte eine (stichprobenartige) Sichtprüfung des Materials und die Prüfung der Übereinstimmung mit den deskriptiven Metadaten beinhalten. Zusätzlich sollten die Objekte identifiziert werden (siehe Tabelle mit Tools, S. 110 ff).

Die Kontrollverfahren können teilautomatisiert werden. Die Prüfung auf Lesbarkeit und Vollständigkeit nach Formatmigrationen sollte (stichprobenartig) manuell erfolgen. Die Prüfung auf Regelkonformität von Transferpaketen z.B. kann automatisiert erfolgen, indem ein entsprechendes Prüfverfahren bei der Übernahme implementiert wird.

Die Nutzung von Standards bezieht sich auf das Standard-Referenzmodell OAIS, auf Metadatenstandards (z.B. PREMIS, Dublin Core, METS) sowie auf Dateiformatstandards. Darüber hinaus können auch standardisierte Prozesse und einheitliche Paketstrukturen darunter verstanden werden. Einheitlich strukturierte Daten und definierte Prozesse erleichtern nicht nur die digitale Langzeitarchivierung, sondern jegliche Art von Datenverwaltung.

Erhaltungsstrategien

Erhaltungsstrategien stellen sicher, dass das Objekt als Datenstrom erhalten und nutzbar bleibt. Dabei sichern Verfahren für die Bitstream-Preservation den Datenstrom und Verfahren für die Content-Preservation die Les- und Interpretierbarkeit des Objekts.

Bitstream-Preservation

Digitale Datenträger sind nur von begrenzter Haltbarkeit. Sie sind nicht nur von Verschleiß betroffen, auch Hardwareobsoleszenz ist problematisch. Aus diesem Grund ist ein Speicherkonzept essenziell für die Erhaltung des Bitstroms der digitalen Objekte.

Grundlegende Anforderungen an das Speicherkonzept sind:

- ▶ redundante, geografisch verteilte Speicherlösung
- ▶ Integritätschecks, z.B. mit Prüfsummen
- ▶ Viruscheck
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Die Sicherung des Bitstreams eines Objekts ist ein stetig wiederkehrender Prozess und Basis jeder Erhaltungsstrategie.

Empfehlung

Die Institution sollte prüfen, ob die Möglichkeit besteht, allein oder im Verbund mit einem Rechenzentrum zu kooperieren, das die professionelle Bitstream-Preservation übernimmt. Ein Service-Level-Agreement sollte abgeschlossen werden.

Besteht diese Möglichkeit nicht oder ist die Nutzung eines Dienstleisters nicht finanzierbar, so muss die Institution die Bitstream-Preservation selbst übernehmen.

Die institutionsinterne Bitstream-Preservation sollte mindestens die folgenden Anforderungen erfüllen:

- ▶ redundante Speicherung
- ▶ Integritätschecks
- ▶ Viruschecks
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring

Dafür kommt der Einsatz eines RAID-Systems mit einer Speicherverwaltungssoftware in Betracht. Das RAID-System sollte nach Möglichkeit regelmäßig gespiegelt werden.

Die Institution sollte prüfen, ob die Möglichkeit besteht, allein oder im Verbund mit einem Rechenzentrum zu kooperieren, das die professionelle Bitstream-Preservation übernimmt. Ein Service-Level-Agreement sollte abgeschlossen werden.

Content-Preservation

Content-Preservation soll sicherstellen, dass der Inhalt eines Objekts auch nach Jahren wiedergegeben und interpretiert werden kann.

Dafür ist es erforderlich:

- ▶ das Dateiformat, die -version und den Codec zu identifizieren und zu validieren. Bei AV-Medien wird diese Erkenntnis durch den Codec erschwert, der in dem Containerformat enthalten ist.
- ▶ technische Metadaten zu extrahieren (siehe Tabelle mit Tools, S. 110 ff.)
- ▶ Informationen über geeignete Wiedergabesoftware zu sammeln
- ▶ Risikofaktoren zu identifizieren
- ▶ signifikante Eigenschaften festzulegen
- ▶ Kontextinformationen bereitzustellen, die die Wiedergabe des Objekts und die Interpretation des Inhalts ermöglichen

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Dafür gibt es zwei Verfahren: Formatmigration und Emulation.¹²

¹² Siehe http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf, Kapitel 8.3 und 8.4.

Es wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Objekte die ursprüngliche Nutzungsumgebung nicht erhalten werden muss. Eine Ausnahme stellt Videokunst dar, die im Wissenschaftlichen Film aber höchstens als Repräsentation vorkommt. Dafür steht eine Reihe von Open-Source- und kommerzieller Software zur Verfügung.

Empfehlung

Kleine Institutionen können eine umfassende Content-Preservation kaum leisten. Aus diesem Grund sollen hier nur minimale Anforderungen aufgeführt werden.

Es wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Objekte nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten. Eine Ausnahme stellt Videokunst dar, die im Wissenschaftlichen Film aber nicht als solche, sondern höchstens als Repräsentation vorkommt. Dafür stehen eine Reihe von Open-Source- Tools und kommerzielle Software zur Verfügung.

Das Dateiformat sollte identifiziert und validiert werden. Es genügt nicht, ein Objekt über die Dateiendung zu identifizieren, da diese auch irreführend sein kann. Außerdem sollten technische Metadaten extrahiert werden. Das Dateiformat und der Codec eines Objekts müssen bekannt sein, da alle Erhaltungsmaßnahmen hierauf basieren.

Signifikante Eigenschaften sollten in geringem Umfang festgelegt werden. Der Fokus sollte auf den für die Institution und deren Nutzer wichtigen Eigenschaften eines Objektes liegen. Diese können sein:

- ▶ das Seitenverhältnis
- ▶ der Farbraum
- ▶ die Auflösung
- ▶ bestimmte Zusatzfunktionen wie Untertitel, Mehrsprachigkeit usw.

Diese Anforderungen sollten nach einer Bestandserhaltungsmaßnahme überprüft werden.

Grundlegende Risikofaktoren sollten definiert werden, z.B.:

- ▶ das Format oder der Codec ist von Obsoleszenz bedroht
- ▶ das Objekt liegt in einem proprietären oder kaum verbreiteten Format oder Codec vor

Tritt einer der Risikofaktoren auf, sollten alle Objekte im betroffenen Format in ein neues Zielformat migriert werden.

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigrationen zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ die Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ der Nachweis von Veränderungen am Objekt in den Metadaten

Empfehlung

Aufgrund der begrenzten Ressourcen kann es für kleinere Institutionen schwierig sein, alle Versionen eines Objekts aufzubewahren und somit die vollständige Nachweiskette zu erhalten, denn die Speicherkapazität spielt bei audiovisuellen Objekten nach wie vor eine Rolle. Reichen die Ressourcen nicht aus, sollte wenigstens das Original und die jeweils aktuelle Version des Archivformats aufbewahrt werden.

Aufgrund der begrenzten Ressourcen kann es für kleinere Institutionen schwierig sein, alle Versionen eines Objekts aufzubewahren und somit die vollständige Nachweiskette zu erhalten, denn die Speicherkapazität spielt bei audiovisuellen Objekten nach wie vor eine Rolle. Reichen die Ressourcen nicht aus, sollte wenigstens das Original und die jeweils aktuelle Version des Archivformats aufbewahrt werden.

Der vollständige Nachweis über Veränderungen am Objekt muss in den Metadaten erfolgen. Die Protokollierung von Änderungen muss technisch und organisatorisch sichergestellt sein.

Diese beinhaltet z.B.:

- ▶ Art und Zeitpunkt der Änderung
- ▶ Name des Benutzers, der die Änderung veranlasst hat
- ▶ Name des Benutzers oder des Programms, der/das die Änderung durchgeführt hat
- ▶ Grund der Änderung

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier, Nutzerdaten, Rollen- und Rechtevergabe.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, den Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.
- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmetadaten: Diese Metadaten beschreiben die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.

Auch kleinere Institutionen sollten nach Möglichkeit Metadaten aller Kategorien erfassen. Ist das nicht möglich, sollten mindestens deskriptive, rechtliche und technische Metadaten erfasst werden.

Empfehlung

Auch kleinere Institutionen sollten nach Möglichkeit Metadaten aller Kategorien erfassen.

Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden. Diese Metadaten sollten vollständig oder in Auswahl in das Langzeitarchivierungssystem überführt werden, das in der Lage ist, administrative und Provenienzmetadaten zu erzeugen und zu verwalten.

Ist das nicht möglich, sollten mindestens deskriptive, rechtliche und technische Metadaten erfasst werden.

Erschließung

Neben den formalen Metadaten sollten auch inhaltsbeschreibende Schlagworte verwendet werden. Die Erschließungstiefe ist abhängig von den Vorgaben der Institution und den Anforderungen der Nutzer.

Wenn das Nachweissystem der Institution für die Erschließung und Verwaltung von Multimediaelementen geeignet ist, können die audiovisuellen Objekte mit dem Nachweissystem verwaltet und beschrieben werden. Ist das vorhandene Nachweissystem nicht geeignet, sollte frei verfügbare Datenbanksoftware verwendet werden, die für Multimediaobjekte geeignet ist (siehe Tabelle mit Tools, S. 110ff.).

Nachweissysteme sind i.d.R. nicht für die Langzeitarchivierung geeignet.

Es kommt auch ein Digital Asset Management System (DAM)/Media Asset Management System (MAM) in Betracht. Ein DAM/MAM ist allerdings i.d.R. nicht für die Langzeitarchivierung geeignet. Ob das System für die Langzeitarchivierung geeignet ist, muss im Einzelfall untersucht werden.

Für den Nachweis der Objekte und die Erfassung von deskriptiven und strukturellen Metadaten sind Nachweissysteme oder DAM/MAM-Systeme jedoch gut geeignet. Ob rechtliche und technische Metadaten erfasst werden können, muss im Einzelfall geprüft werden.

Provenienz- und Erhaltungsmetadaten

Für die Erfassung von Provenienzmetadaten kann überlegt werden, ob die Provenienzmetadaten separat zu erfassen sind und eine Verknüpfung mit dem bestehenden Nachweissystem erfolgen kann. Empfehlenswert ist eine METS/MODS-Datei, die die relevanten Metadaten enthält. Besser wäre jedoch der Einsatz eines OAIS-konformen Langzeitarchivierungssystems, das die Protokollierung LZA-relevanter Metadaten automatisiert gewährleistet, oder die Abgabe der Objekte an eine Institution, die die Langzeitarchivierung leisten kann.

Präsentation/Zugriff

Das Format der Nutzungskopien sollte die Anforderungen der vorgesehenen Nutzerzielgruppen berücksichtigen.

Weitverbreitete, offene Formate sollten bevorzugt werden; proprietäre Formate sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung. Im Idealfall soll die Nutzungskopie ohne die Installation zusätzlicher Software lesbar sein. Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein.

Empfehlung

In kleineren Institutionen ist das Nutzungsformat oft identisch mit dem Archivmaster. In diesem Fall muss das Archivformat für die Nutzung geeignet sein. Alternativ müssen Nutzungskopien in einem geeigneten Format erzeugt werden.

Reichen die technischen Ressourcen nicht aus, um Nutzungskopien vorzuhalten, können sie auch auf Nachfrage erzeugt werden, wenn die Nachfrage gering ist. Stark nachgefragte Objekte sollten in einer Nutzungskopie vorgehalten werden. Das Nutzungsformat sollte den Bedürfnissen des Nutzers entsprechen.

In kleinen Institutionen können alle Nutzungsszenarien vorkommen, wenn die Ressourcen für eine Webpräsentation zur Verfügung stehen:

- ▶ lokale Nutzung
- ▶ der Versand von Datenträgern
- ▶ Onlinezugriff

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software eines Herstellers gebunden und stellen eine potenzielle Gefahr für die langfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivmasters auch die Kompatibilität eines Formats und Codecs. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Darüber hinaus muss basierend auf den Anforderungen der vorgesehenen Nutzerzielgruppen auch eine Entscheidung über die Formate für Nutzungskopien getroffen werden.

Empfehlung

Die große Vielfalt potenzieller Ablieferungsformate stellt die Institutionen vor eine große Herausforderung. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Preservation Planning, Preservation Watch, Preservation Actions, Risikomanagement usw.). Es ist deshalb für kleine Institutionen sinnvoll, die Anzahl der Dateiformate im digitalen Langzeitarchiv auf eine Auswahl zu beschränken. Das kann durch folgende Verfahren geschehen:

- ▶ die Veröffentlichung einer Liste bevorzugter Dateiformate
- ▶ durch Normalisierung im digitalen Langzeitarchiv
- ▶ oder durch eine Kombination beider Verfahren

In kleineren Institutionen ist das Nutzungsformat oft identisch mit dem Archivmaster. In diesem Fall muss das Archivformat für die Nutzung geeignet sein. Alternativ müssen Nutzungskopien in einem geeigneten Format erzeugt werden. Reichen die technischen Ressourcen nicht aus, um Nutzungskopien vorzuhalten, können sie auch auf Nachfrage erzeugt werden, wenn die Nachfrage gering ist.

Die Rahmenbedingungen und das Vorgehen bei der Normalisierung sollten in einer öffentlich verfügbaren Policy dokumentiert werden, um den Prozess für Produzenten und Nutzer transparent zu machen.

Diesem pragmatischen Ansatz steht die Anforderung der Gewährleistung der Authentizität gegenüber. Auch ein kleines digitales Langzeitarchiv muss nachweisen können, dass ein Objekt ist, was es auch zu sein scheint. Diese Anforderung wird u.a. durch den Erhalt des Originals und die Dokumentation von Veränderungen in den Metadaten gewährleistet. Kleinere Institutionen verfügen meist nicht über die Kapazitäten, um eine große Anzahl von Dateiformaten zu verwalten. Die Normalisierung eingehender Objekte reduziert die Arbeitsbelastung erheblich.

Als Format wird Matroska (.mkv), ffv1-codiert empfohlen, Audio: Wave Linear PCM 48 kHz. Für die Archivmaster sollten möglichst ein oder wenige Formate festgelegt werden, die sich im Idealfall auch als Nutzungskopien eignen. Für die Konvertierung stehen verschiedene Konvertierungsprogramme zur Verfügung. Gleichzeitig sollte die Anzahl verschiedener Formate bereits bei der Anlieferung begrenzt werden.*

Formatempfehlung

- ▶ Matroska (*.mkv), ffv1-codiert, Audio: Wave Linear PCM 48 kHz

Die empfohlene Kombination aus Dateiformat und Codec kann mit FFmpeg erzeugt werden. Weitere Dateiformate werden unter <https://wiki.dnb.de/display/NESTOR/AG+Media> behandelt.

Wenn möglich, sollten ein oder wenige Formate für die Archivmaster festgelegt werden, die sich im Idealfall auch als Nutzungskopien eignen. Für die Konvertierung stehen verschiedene Konvertierungsprogramme als proprietäre oder Open-Source-Software zur Verfügung.

Gleichzeitig sollten kleinere Institutionen versuchen, die Anzahl verschiedener Formate bereits bei der Anlieferung zu begrenzen, z.B. durch die Veröffentlichung einer Liste bevorzugter Dateiformate.

Soll für die Nutzung auf den Archivmaster zugegriffen werden, ist unbedingt zu beachten, dass die Wiedergabesoftware den Archivmaster nicht (versehentlich) manipulieren kann. Dies lässt sich z.B. durch read-only-Zugriff auf den Archivmaster erreichen.

Empfehlungen für mittlere Einrichtungen

Auswahl des zu archivierenden Materials

Selektionskriterien bilden die Entscheidungsgrundlage für die Frage, ob ein Objekt langfristig gesichert werden soll. Hierbei spielen neben rechtlichen auch inhaltliche Kriterien eine Rolle. Besondere Bedeutung kommt der Frage nach den Ressourcen zu.

Empfehlung

Mittlere Institutionen sollten anhand ihres Sammlungsprofils und ihres Mandats selektieren:

- ▶ Ist das Objekt archivwürdig?
- ▶ Entspricht das Objekt dem Sammlungsprofil bzw. dem Archivauftrag?
- ▶ Liegen alle benötigten Rechte vor (Langzeitarchivierung, Veränderung, Verbreitung)?

Es kann auch überlegt werden, ob die Objekte an eine Institution abgegeben werden, die die Langzeitarchivierung als Dienstleistung gewährleisten kann.

Beschreibung der vorgesehenen Nutzungsszenarien

Die vorgesehenen Nutzerzielgruppen sind die aktuellen und künftigen Nutzergruppen, die mit dem digitalen Langzeitarchiv oder daran angeschlossenen Nutzungsplattformen interagieren. Unter Nutzerzielgruppen sind an dieser Stelle nicht nur die klassischen Nutzer zu verstehen. Vorgesehene Nutzerzielgruppen sind auch externe Nutzer wie Produzenten, die Objekte abliefern, und interne Nutzer wie die Mitarbeiter der Institution.

Verschiedene vorgesehene Nutzerzielgruppen können unterschiedliche Anforderungen haben. Die vorgesehenen Nutzerzielgruppen entsprechen den klassischen Zielgruppen der Institution. Diese sind bekannt und deren Anforderungen sind bereits analysiert worden. Zusätzlich kommen als vorgesehene Nutzerzielgruppen hinzu:

- ▶ die Datenproduzenten
- ▶ die mit der digitalen Langzeitarchivierung befassten Mitarbeiter

Die verschiedenen vorgesehene Nutzerzielgruppen haben bestimmte Anforderungen. Daraus ergeben sich Nutzungsszenarien und signifikante Eigenschaften, die Einfluss auf die Erhaltungsmaßnahmen nehmen.

Empfehlung

Eine Analyse der Nutzerzielgruppen sollte durchgeführt werden, die neben den klassischen Zielgruppen der Institution auch die Datenproduzenten und die Mitarbeiter berücksichtigt.

Die Analyse der Nutzerzielgruppen sollte folgende Punkte beinhalten:

- ▶ die Beschreibung und Abgrenzung der verschiedenen Nutzerzielgruppen
- ▶ die Beschreibung der Anforderungen der Nutzerzielgruppen
- ▶ in den Nutzerzielgruppen verbreitete Dateiformate

Die Nutzerzielgruppen sollten in wiederkehrenden Abständen erneut analysiert werden, da sich die Anforderungen im Lauf der Zeit ändern können.

Eine Analyse der Nutzerzielgruppen sollte durchgeführt werden, die neben den klassischen Zielgruppen der Institution auch die Datenproduzenten und die Mitarbeiter berücksichtigt. Die Nutzerzielgruppen sollten in wiederkehrenden Abständen erneut analysiert werden, da sich die Anforderungen im Lauf der Zeit ändern können.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren und Prozesse, die die Qualität der Objekte sowie der Prozesse im digitalen Langzeitarchiv sicherstellen sollen. Darunter fallen z.B.:

- ▶ Übernahmevereinbarung (ein Dokument, das organisatorische, rechtliche und technische Punkte bei der Übernahme berücksichtigt)
- ▶ Eingangskontrolle der Objekte (Überprüfung der Objekte beim Eingang in das digitale Langzeitarchiv)
 - ▶ Prüfung auf Lesbarkeit und inhaltliche Vollständigkeit
 - ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten
 - ▶ technische Fehlerfreiheit
 - ▶ Standardkonformität

Maßnahmen für das Qualitätsmanagement müssen implementiert werden. Es wird empfohlen, eine Eingangskontrolle durchzuführen und Kontrollverfahren einzurichten. Darüber hinaus sollten Standards genutzt und die Kriterien für die vertrauenswürdige Langzeitarchivierung beachtet werden. Prozesse sollten mindestens grundlegend dokumentiert werden. Die Produzenten sollten die Objekte möglichst mit einer Prüfsumme an das Archiv abgeben. Diese muss nach der Annahme im Archiv geprüft werden. Darüber hinaus sollte sich die Institution an den Kriterien für die vertrauenswürdige Langzeitarchivierung orientieren. Selbst wenn keine Zertifizierung angestrebt wird, sind die dort beschriebenen Kriterien hilfreich, um die Qualität der digitalen Langzeitarchivierung sicherzustellen.

- ▶ Dokumentation von Prozessen (Genauere Beschreibung von Prozessen im digitalen Langzeitarchiv: Was wird wann warum wie gemacht? Z.B. schriftlich oder in einer Notation für Prozessdokumentation wie BPMN 2.0.)
- ▶ Evaluierung von Prozessen (Dokumentierte Prozesse können nach Kriterien wie Praxistauglichkeit, Zeitaufwand, Ressourcenaufwand, Überschneidungen mit anderen Prozessen usw. untersucht werden. Die Analyse ist Voraussetzung für die Optimierung der Prozesse.)
- ▶ Kontrollverfahren (technische und organisatorische Verfahren für die Kontrolle von definierten Regeln und Strukturen oder zur Evaluierung von Ergebnissen nach Änderungen an Objekten)
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigrationen
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards
- ▶ Beachtung der Kriterien für die vertrauenswürdige Langzeitarchivierung (nestor-Siegel nach DIN 31644¹³, auch Data Seal of Approval¹⁴)

Empfehlung

Maßnahmen für das Qualitätsmanagement müssen implementiert werden.

Es wird empfohlen, eine Eingangskontrolle durchzuführen und Kontrollverfahren einzurichten. Darüber hinaus sollten Standards genutzt und die Kriterien für die vertrauenswürdige Langzeitarchivierung beachtet werden. Prozesse sollten mindestens grundlegend dokumentiert werden.

Die Eingangskontrolle sollte eine (stichprobenartige) Sichtprüfung des Materials und die Prüfung der Übereinstimmung mit den deskriptiven Metadaten beinhalten. Zusätzlich sollten die Objekte identifiziert werden (siehe Tabelle mit Tools, S. 110 ff.).

Wenn möglich, sollten die Produzenten die Objekte mit einer Prüfsumme an das Archiv abgeben. Diese Prüfsumme muss nach der Annahme im Archiv geprüft werden. Stimmen die Prüfsummen nicht überein, muss das Objekt abgewiesen und vom Produzenten erneut abgegeben werden, bis ein erfolgreicher Transfer mit übereinstimmenden Prüfsummen erfolgt ist.

Die Kontrollverfahren können teilautomatisiert werden. Die Prüfung auf Lesbarkeit und Vollständigkeit nach Formatmigrationen sollte (stichprobenartig) manuell erfolgen. Darüber hinaus muss nach einer Bestandserhaltungsmaßnahme geprüft werden, ob die signifikanten Eigenschaften erhalten geblieben sind.

Die Prüfung auf Regelkonformität von Transferpaketen kann automatisiert erfolgen, indem ein entsprechendes Prüfverfahren bei der Übernahme implementiert wird.

Die Nutzung von Standards bezieht sich auf das Standard-Referenzmodell OAIS, auf Metadatenstandards (z.B. PREMIS, Dublin Core, METS) sowie Dateiformatstandards. Darüber hinaus können darunter auch standardisierte Prozesse und einheitliche Paketstrukturen verstanden werden. Das Langzeitarchivierungssystem muss in der Lage sein, regelkonforme Archivpakete zu erstellen.

¹³ http://files.d-nb.de/nestor/materialien/nestor_mat_17.pdf

¹⁴ <http://datasealofapproval.org/en/information/guidelines/>

Einheitlich strukturierte Daten und definierte Prozesse erleichtern nicht nur die digitale Langzeitarchivierung, sondern jegliche Art von Datenverwaltung.

Darüber hinaus sollte die Institution sich an den Kriterien für die vertrauenswürdige Langzeitarchivierung orientieren. Selbst wenn keine Zertifizierung angestrebt wird, sind die dort beschriebenen Kriterien hilfreich, um die Qualität der digitalen Langzeitarchivierung sicherzustellen.

Erhaltungsstrategien

Erhaltungsstrategien stellen sicher, dass das Objekt als Datenstrom erhalten und nutzbar bleibt. Dabei sichern Verfahren für die Bitstream-Preservation den Datenstrom und Verfahren für die Content-Preservation die Les- und Interpretierbarkeit des Objekts.

Bitstream-Preservation

Digitale Datenträger sind nur von begrenzter Haltbarkeit. Sie sind nicht nur von Verschleiß betroffen, auch Hardwareobsoleszenz ist problematisch. Aus diesem Grund ist ein Speicherkonzept essenziell für die Erhaltung des Bitstroms der digitalen Objekte.

Grundlegende Anforderungen an das Speicherkonzept sind:

- ▶ redundante, geografisch verteilte Speicherlösung
- ▶ Integritätschecks, z.B. mit Prüfsummen
- ▶ Viruscheck
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Die Sicherung des Bitstroms eines Objekts ist ein stetig wiederkehrender Prozess und Basis jeder Erhaltungsstrategie.

Empfehlung

Mittlere Institutionen können ein eigenes Speichersystem betreiben, wenn die technischen Ressourcen vorhanden sind. Es kann aber auch sinnvoll sein, im Verbund mit anderen Institutionen ein Rechenzentrum zu betreiben oder einen Dienstleister mit der Bitstream-Preservation zu beauftragen. Diese Entscheidung muss im Einzelfall getroffen werden.

Die institutionsinterne Bitstream-Preservation sollte die folgenden Anforderungen erfüllen:

- ▶ redundante, räumlich getrennte Speicherung auf verschiedenen Speichermedien
- ▶ Integritätschecks
- ▶ Viruschecks
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Mittlere Institutionen können ein eigenes Speichersystem betreiben, wenn die technischen Ressourcen vorhanden sind. Man kann im Einzelfall aber auch entscheiden, im Verbund mit anderen Institutionen ein Rechenzentrum zu betreiben oder einen Dienstleister mit der Bitstream-Preservation zu beauftragen.

Es sollte eine professionelle Speicherlösung gewählt werden. Es wird zu einer Kombination aus Festplatten- und Bandspeicher geraten. Darüber hinaus müssen Sicherheitsbestimmungen beachtet werden. Darunter fallen:

- ▶ physischer Zugang
- ▶ Zugriffsberechtigungen
- ▶ Einbruchmeldeanlage
- ▶ digitale Zugriffsberechtigungen
- ▶ Brandschutz
- ▶ Disaster Recovery

Mit einem Dienstleister sollte ein Service-Level-Agreement (SLA) vereinbart werden. Der Umfang der Dienstleistung muss genau evaluiert werden, um sicherzustellen, dass alle Anforderungen an die Bitstream-Preservation erfüllt werden.

Mittlere Institutionen sollten eine Erhaltungsstrategie mit folgenden Inhalten ausarbeiten: Formatidentifikation und -validierung und daraus folgend das Bilden von homogenen Bestandsgruppen; Risikofaktoren identifizieren; Turnus für die Überprüfung der Risikofaktoren sowie Erhaltungsmaßnahmen für Bestandsgruppen festlegen; geeignete Tools für die Erhaltungsmaßnahmen definieren. Die Originaldatei sollte unverändert aufbewahrt, signifikante Eigenschaften sollten definiert werden. Der Fokus sollte auf den für die Institution und deren Nutzer wichtigen Eigenschaften eines Objekts liegen.

Content-Preservation

Content-Preservation soll sicherstellen, dass der Inhalt eines Objekts auch nach Jahren wiedergegeben und interpretiert werden kann.

Dafür ist es erforderlich:

- ▶ das Dateiformat, die -version und den Codec zu identifizieren und zu validieren. Bei AV-Medien wird diese Erkenntnis durch den Codec erschwert, der in dem Containerformat enthalten ist
- ▶ technische Metadaten zu extrahieren (siehe Tabelle mit Tools, S. 110ff.)
- ▶ Informationen über geeignete Wiedergabesoftware zu sammeln
- ▶ Risikofaktoren zu identifizieren
- ▶ signifikante Eigenschaften festzulegen
- ▶ Kontextinformationen bereitzustellen, die die Wiedergabe des Objekts und die Interpretation des Inhalts ermöglichen

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Dafür gibt es zwei Verfahren: Formatmigration und Emulation.¹⁵

Empfehlung

Mittlere Institutionen sollten eine Erhaltungsstrategie ausarbeiten.

Die Erhaltungsstrategie sollte folgende Punkte abdecken:

- ▶ Formatidentifikation und -validierung. Daraus folgend das Bilden von homogenen Bestandsgruppen.
- ▶ Risikofaktoren identifizieren
- ▶ Turnus für die Überprüfung der Risikofaktoren festlegen
- ▶ Erhaltungsmaßnahmen für Bestandsgruppen festlegen
- ▶ geeignete Tools für die Erhaltungsmaßnahmen definieren

Es wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Ob

¹⁵ Siehe http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf, Kapitel 8.3 und 8.4.

jekte nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten. Eine Ausnahme stellt Videokunst dar, die im Wissenschaftlichen Film aber nicht als solche, sondern höchstens als Repräsentation vorkommt.

In Abhängigkeit von den personellen Ressourcen sollte die Anzahl der Archivformate im Archiv nach Möglichkeit auf ein bis einige wenige ausgewählte Dateiformate begrenzt werden. Es wird empfohlen, Objekte, die in anderen Dateiformaten vorliegen, nach dem Eingang der Originaldatei zu kopieren und in eines der ausgewählten Dateiformate umzuwandeln. Dafür stehen eine Reihe von Open-Source- und kommerzieller Software zur Verfügung.

Die Originaldatei sollte unverändert aufbewahrt werden.

Das Dateiformat sollte identifiziert und validiert werden. Es genügt nicht, ein Objekt über die Dateiendung zu identifizieren, da diese auch irreführend sein kann. Außerdem sollten technische Metadaten extrahiert werden. Das Dateiformat und der Codec eines Objekts müssen bekannt sein, da alle Erhaltungsmaßnahmen hierauf basieren.

Signifikante Eigenschaften sollten definiert werden. Der Fokus sollte auf den für die Institution und deren Nutzer wichtigen Eigenschaften eines Objektes liegen.

Signifikante Eigenschaften hängen u.a. auch davon ab, welchen Zweck die Aufnahme hat. Ist der Zweck der Aufnahme z.B. die Dokumentation eines Ereignisses, eines Experiments usw., so ergibt sich daraus, was besonders erhaltenswert ist. Davon wiederum hängen ggf. weitere technische signifikante Eigenschaften ab.

Diese Anforderungen sollten nach einer Bestandserhaltungsmaßnahme überprüft werden.

Beispiele für technische signifikante Eigenschaften:

Eigenschaft	Komponente	Definition der Eigenschaft
Videokanäle	video	Anzahl der unterschiedlichen Videokanäle im digitalen Objekt
Audiokanäle	audio	Anzahl der unterschiedlichen Audiokanäle im digitalen Objekt
Länge	video	Anzahl der Frames in der Aufnahme
Breite	frame	Breite des Einzelbildes
Höhe	frame	Höhe des Einzelbildes
Bittiefe	pixel	Anzahl der Farbinformationen pro Pixel
Farbmodell	pixel	verwendetes Farbmodell bei der Aufnahme
Farbraum	pixel	verwendeter Farbraum bei der Aufnahme
Seitenverhältnis	pixel	Seitenverhältnis der Pixel aspect ratio of the pixels making up the image
Bildrate	video	Anzahl der Frames in einem definierten Zeitintervall
Interlace	frame	falls es sich um interlaced video handelt
Metadaten		zum Objekt gehörende Metadaten

(Quelle: Archivemata Wiki¹⁶)

Grundlegende Risikofaktoren sollten definiert werden, z.B.:

- ▶ das Format oder der Codec ist von Obsoleszenz bedroht
- ▶ das Objekt liegt in einem proprietären oder kaum verbreiteten Format vor

Tritt einer der Risikofaktoren auf, sollten alle Objekte im betroffenen Format in ein neues Zielformat migriert werden.

¹⁶ https://wiki.archivemata.org/Significant_characteristics_of_video_files

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigrationen zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ die Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ der Nachweis von Veränderungen am Objekt in den Metadaten

Empfehlung

Mittlere Institutionen sollten das Originalobjekt sowie alle Versionen des Originalobjekts aufbewahren und Änderungen in den Metadaten dokumentieren. Die Protokollierung von Änderungen muss technisch und organisatorisch sichergestellt sein. Diese beinhaltet z.B.:

- ▶ Art und Zeitpunkt der Änderung
- ▶ Name des Benutzers, der die Änderung veranlasst hat
- ▶ Name des Benutzers oder des Programms, der/das die Änderung durchgeführt hat
- ▶ Grund der Änderung

Mittlere Institutionen sollten das Originalobjekt sowie alle Versionen des Originalobjekts aufbewahren und Änderungen in den Metadaten dokumentieren. Die Protokollierung von Änderungen muss technisch und organisatorisch sichergestellt sein.

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier, Nutzerdaten, Rollen- und Rechtevergabe.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, den Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.
- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmetadaten: Diese Metadaten beschreiben die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.

Mittlere Institutionen sollten Metadaten aller Kategorien erfassen. Die Erschließung der Objekte erfolgt für gewöhnlich in einem Nachweissystem, in dem neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden können.

Empfehlung

Mittlere Institutionen sollten Metadaten aller Kategorien erfassen.

Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden. Diese Metadaten sollten vollständig oder in Auswahl in das Langzeitarchivierungssystem überführt werden, das in der Lage ist, administrative und Provenienzmetadaten zu erzeugen und zu verwalten.

Erschließung

Neben den formalen Metadaten sollten auch inhaltsbeschreibende Schlagworte verwendet werden. Die Erschließungstiefe ist abhängig von den Vorgaben der Institution und den Anforderungen der Nutzer.

Wenn das Nachweissystem der Institution für die Erschließung und Verwaltung von Multimediaelementen geeignet ist, können die audiovisuellen Objekte mit dem Nachweissystem verwaltet und beschrieben werden. Ist das vorhandene Nachweissystem nicht geeignet, sollte frei verfügbare Datenbanksoftware verwendet werden, die für Multimediaobjekte geeignet ist. Nachweissysteme sind i.d.R. nicht für die Langzeitarchivierung geeignet.

Es kommt auch ein Digital Asset Management System (DAM)/Media Asset Management System (MAM) in Betracht. Ein DAM/MAM ist allerdings i.d.R. nicht für die Langzeitarchivierung geeignet. Ob das System für die Langzeitarchivierung geeignet ist, muss im Einzelfall untersucht werden.

Für den Nachweis der Objekte und die Erfassung von deskriptiven und strukturellen Metadaten sind Nachweissysteme oder DAM/MAM-Systeme jedoch gut geeignet. Ob rechtliche und technische Metadaten erfasst werden können, muss im Einzelfall geprüft werden.

Provenienz- und Erhaltungsmetadaten

Für die Erfassung von Provenienzmetadaten muss ein OAIS-konformes Langzeitarchivierungssystem eingesetzt werden. Alternativ sollte in Betracht gezogen werden, die Objekte an eine Institution abzugeben, die die Langzeitarchivierung leisten kann. Bereits vorhandene Metadaten, z.B. deskriptive und rechtliche Metadaten, können aus dem bestehenden Nachweissystem vollständig oder in Auswahl in das digitale Langzeitarchivierungssystem über eine Schnittstelle importiert werden. Im Langzeitarchivierungssystem sollten alle relevanten Metadaten in einer METS/MODS-Datei gespeichert werden, um eine schnellere Verarbeitung in Fällen der Formatmigration oder Metadaten-Updates zu gewährleisten.

Präsentation/Zugriff

Das Format der Nutzungskopien sollte die Anforderungen der vorgesehenen Nutzerzielgruppen berücksichtigen. Weitverbreitete, offene Formate sollten bevorzugt werden; proprietäre Formate sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung. Im Idealfall soll die Nutzungskopie ohne die Installation zusätzlicher Software lesbar sein.

Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein.

Empfehlung

Mittlere Institutionen sollten Nutzungskopien erzeugen und für die Nutzung vorhalten. Das Format der Nutzungskopie sollte den Anforderungen der Nutzerzielgruppen entsprechen.

Es muss beachtet werden, wie die Objekte für die Nutzung zur Verfügung gestellt werden dürfen. Folgende Nutzungsszenarien sind denkbar:

- ▶ lokale Nutzung
- ▶ der Versand von Datenträgern
- ▶ Onlinezugriff mit und ohne Registrierung

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software eines Herstellers gebunden und stellen eine potenzielle Gefahr für die langfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivmasters auch die Kompatibilität eines Formats und Codecs. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Darüber hinaus muss basierend auf den Anforderungen der vorgesehenen Nutzerzielgruppen auch eine Entscheidung über die Formate für Nutzungskopien getroffen werden.

Empfehlung

Die große Vielfalt potenzieller Ablieferungsformate stellt die Institutionen vor eine große Herausforderung. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Preservation Planning, Preservation Watch, Preservation Actions, Risikomanagement usw.).

Es ist deshalb sinnvoll, die Anzahl der Dateiformate im digitalen Langzeitarchiv möglichst gering zu halten, sofern die Institution Einfluss darauf nehmen kann. Dies kann durch folgende Verfahren geschehen:

- ▶ die Veröffentlichung einer Liste bevorzugter Dateiformate
- ▶ das Erzeugen einer neuen Version durch Formatmigration nach Eingang der Objekte. Die Originaldateien müssen dabei unverändert aufbewahrt werden.
- ▶ eine Kombination beider Verfahren

Die Vorgehensweise sollte in einer öffentlich verfügbaren Policy dokumentiert werden, um den Prozess für Produzenten und Nutzer transparent zu machen.

Als Format empfiehlt sich Matroska (.mkv) PUID: fmt/569, ffv1-codiert, Audio: Wave Linear PCM 48 kHz. Die empfohlene Kombination aus Dateiformat und Codec kann mit FFmpeg erzeugt werden. Weitere Dateiformate sind in der Übersicht über „Dateiformate, Codecs und Bildformate“, unter <https://wiki.dnb.de/display/NESTOR/AG+Media> aufgeführt.

Die große Vielfalt potenzieller Ablieferungsformate ist eine große Herausforderung. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung. Die Anzahl der Dateiformate im digitalen Langzeitarchiv sollte daher möglichst gering gehalten werden, sofern darauf Einfluss genommen werden kann. Die Vorgehensweise sollte hinsichtlich der Transparenz in einer öffentlich verfügbaren Policy dokumentiert werden. Als Format empfiehlt sich Matroska (.mkv) PUID: fmt/569, ffv1-codiert, Audio: Wave Linear PCM 48 kHz.

Empfehlungen für größere Einrichtungen

Auswahl des zu archivierenden Materials

Selektionskriterien bilden die Entscheidungsgrundlage für die Frage, ob ein Objekt langfristig gesichert werden soll. Hierbei spielen neben rechtlichen auch inhaltliche Kriterien eine Rolle. Besondere Bedeutung kommt der Frage nach den Ressourcen zu.

Größere Institutionen sollten anhand ihres Sammlungsprofils und ihres Mandats selektieren, ob das Objekt archivwürdig ist, ob es dem Sammlungsprofil bzw. dem Archivauftrag entspricht, und ob alle benötigten Rechte vorliegen.

Empfehlung

Größere Institutionen sollten anhand ihres Sammlungsprofils und ihres Mandats selektieren:

- ▶ Ist das Objekt archivwürdig?
- ▶ Entspricht das Objekt dem Sammlungsprofil bzw. dem Archivauftrag?
- ▶ Liegen alle benötigten Rechte vor (Langzeitarchivierung, Veränderung, Verbreitung)?

Beschreibung der Nutzungsszenarien

Die vorgesehenen Nutzerzielgruppen sind die aktuellen und künftigen Nutzergruppen, die mit dem digitalen Langzeitarchiv oder daran angeschlossenen Nutzungsplattformen interagieren. Unter Nutzerzielgruppen sind an dieser Stelle nicht nur die klassischen Nutzer zu verstehen. Vorgesehene Nutzerzielgruppen sind auch externe Nutzer wie Produzenten, die Objekte abliefern, und interne Nutzer wie die Mitarbeiter der Institution.

Verschiedene vorgesehene Nutzerzielgruppen können unterschiedliche Anforderungen haben. Die vorgesehenen Nutzerzielgruppen entsprechen den klassischen Zielgruppen der Institution. Diese sind bekannt und deren Anforderungen gelegentlich bereits analysiert worden.

Zusätzlich kommen als vorgesehene Nutzerzielgruppen hinzu:

- ▶ die Datenproduzenten
- ▶ die mit der digitalen Langzeitarchivierung befassten Mitarbeiter

Die verschiedenen vorgesehene Nutzerzielgruppen haben bestimmte Anforderungen. Daraus ergeben sich Nutzungsszenarien und signifikante Eigenschaften, die Einfluss auf die Erhaltungsmaßnahmen nehmen.

Es sollte eine Analyse der Nutzerzielgruppen durchgeführt werden, die neben den klassischen Zielgruppen der Institution auch die Datenproduzenten und die Mitarbeiter berücksichtigt. Die Nutzerzielgruppen sollten in wiederkehrenden Abständen erneut analysiert werden, da sich die Anforderungen im Lauf der Zeit ändern können.

Empfehlung

Die Nutzerzielgruppen müssen der Institution bekannt sein.

Eine Analyse der Nutzerzielgruppen sollte durchgeführt werden, die neben den klassischen Zielgruppen der Institution auch die Datenproduzenten und die Mitarbeiter berücksichtigt.

Die Analyse der Nutzerzielgruppen sollte folgende Punkte beinhalten:

- ▶ die Beschreibung und Abgrenzung der verschiedenen Nutzerzielgruppen
- ▶ die Beschreibung der Anforderungen der Nutzerzielgruppen
- ▶ in den Nutzerzielgruppen verbreitete Nutzerformate

Die Nutzerzielgruppen sollten in wiederkehrenden Abständen erneut analysiert werden, da sich die Anforderungen im Lauf der Zeit ändern können.

Qualitätsmanagement

Qualitätsmanagement bezeichnet Verfahren und Prozesse, die die Qualität der Objekte sowie der Prozesse im digitalen Langzeitarchiv sicherstellen sollen.

Darunter fallen z.B.:

- ▶ Übernahmevereinbarung (ein Dokument, das organisatorische, rechtliche und technische Punkte bei der Übernahme berücksichtigt)
- ▶ Eingangskontrolle der Objekte (Überprüfung der Objekte beim Eingang in das digitale Langzeitarchiv)
 - ▶ Prüfung auf Lesbarkeit und inhaltliche Vollständigkeit
 - ▶ Prüfung auf inhaltliche Übereinstimmung mit den deskriptiven Metadaten
 - ▶ technische Fehlerfreiheit
 - ▶ Standardkonformität
- ▶ Dokumentation von Prozessen (Genauere Beschreibung von Prozessen im digitalen Langzeitarchiv: Was wird wann warum wie gemacht? Z.B. schriftlich oder in einer Notation für Prozessdokumentation wie BPMN 2.0.)
- ▶ Evaluierung von Prozessen (Dokumentierte Prozesse können nach Kriterien wie Praxistauglichkeit, Zeitaufwand, Ressourcenaufwand, Überschneidungen mit anderen Prozessen usw. untersucht werden. Die Analyse ist Voraussetzung für die Optimierung der Prozesse.)
- ▶ Kontrollverfahren (technische und organisatorische Verfahren für die Kontrolle von definierten Regeln und Strukturen oder zur Evaluierung von Ergebnissen nach Änderungen an Objekten)
 - ▶ Regelkonformität von Transfer- und Archivpaketen
 - ▶ Qualitätsprüfung nach Formatmigrationen
 - ▶ Lesbarkeit und Vollständigkeit nach Änderungen am Objekt
- ▶ Nutzung von Standards
- ▶ Beachtung der Kriterien für die vertrauenswürdige Langzeitarchivierung (nestor-Siegel nach DIN 31644¹⁷, auch Data Seal of Approval¹⁸)

Empfehlung

Maßnahmen für das Qualitätsmanagement müssen implementiert werden. Es wird empfohlen, eine Eingangskontrolle durchzuführen und Kontrollverfahren einzurichten. Darüber hinaus sollten Standards genutzt und die Kriterien für die vertrauenswürdige Langzeitarchivierung beachtet werden. Prozesse sollten dokumentiert und evaluiert, die Dokumentation sollte regelmäßig auf Aktualität geprüft werden.

Maßnahmen für das Qualitätsmanagement müssen implementiert werden. Prozesse sollten dokumentiert und evaluiert werden. Die Dokumentation sollte regelmäßig auf Aktualität geprüft werden.

¹⁷ http://files.d-nb.de/nestor/materialien/nestor_mat_17.pdf

¹⁸ <http://datasealofapproval.org/en/information/guidelines/>

Die Kontrollverfahren können teilautomatisiert werden. Die Prüfung auf Lesbarkeit und Vollständigkeit nach Formatmigrationen sollte manuell erfolgen. Zudem muss nach Bestandserhaltungsmaßnahmen geprüft werden, ob die signifikanten Eigenschaften erhalten geblieben sind. Darüber hinaus sollte die Institution den Kriterien für die vertrauenswürdige Langzeitarchivierung folgen und eine Zertifizierung des digitalen Langzeitarchivs anstreben.

Die Eingangskontrolle sollte eine (stichprobenartige) Sichtprüfung des Materials und die Prüfung der Übereinstimmung mit den deskriptiven Metadaten beinhalten. Zusätzlich sollten die Objekte identifiziert werden (siehe Tabelle mit Tools, S. 110 ff.).

Wenn möglich, sollten die Produzenten die Objekte mit einer Prüfsumme an das Archiv abgeben. Diese Prüfsumme muss nach der Annahme im Archiv geprüft werden. Stimmen die Prüfsummen nicht überein, muss das Objekt abgewiesen und vom Produzenten erneut abgegeben werden, bis ein erfolgreicher Transfer mit übereinstimmenden Prüfsummen erfolgt ist.

Die Kontrollverfahren können teilautomatisiert werden. Die Prüfung auf Lesbarkeit und Vollständigkeit nach Formatmigrationen sollte (stichprobenartig) manuell erfolgen.

Darüber hinaus muss nach Bestandserhaltungsmaßnahmen geprüft werden, ob die signifikanten Eigenschaften erhalten geblieben sind.

Die Prüfung auf Regelkonformität von Transferpaketen kann automatisiert erfolgen, indem ein entsprechendes Prüfverfahren bei der Übernahme implementiert wird.

Die Nutzung von Standards bezieht sich auf das Standard-Referenzmodell OAIS, auf Metadatenstandards (z.B. PREMIS, Dublin Core, METS) sowie Dateiformatstandards. Darüber hinaus können darunter auch standardisierte Prozesse und einheitliche Paketstrukturen verstanden werden. Das Langzeitarchivierungssystem muss in der Lage sein, regelkonforme Archivpakete zu erstellen.

Einheitlich strukturierte Daten und definierte Prozesse erleichtern nicht nur die digitale Langzeitarchivierung, sondern jegliche Art von Datenverwaltung.

Darüber hinaus sollte die Institution den Kriterien für die vertrauenswürdige Langzeitarchivierung folgen und eine Zertifizierung des digitalen Langzeitarchivs anstreben. Zusätzlich zur öffentlichen Sichtbarkeit, dass die digitale Langzeitarchivierung definierten Qualitätsansprüchen genügt, erhält die Institution so auch Gelegenheit, ihre Prozesse, Verfahren und Strukturen von externen Experten überprüfen zu lassen und zu optimieren.

Erhaltungsstrategien

Erhaltungsstrategien stellen sicher, dass das Objekt als Datenstrom erhalten und nutzbar bleibt. Dabei sichern Verfahren für die Bitstream-Preservation den Datenstrom und Verfahren für die Content-Preservation die Les- und Interpretierbarkeit des Objekts.

Bitstream-Preservation

Digitale Datenträger sind nur von begrenzter Haltbarkeit. Sie sind nicht nur von Verschleiß betroffen, auch Hardwareobsoleszenz ist problematisch. Aus diesem Grund ist ein Speicherkonzept essenziell für die Erhaltung des Bitstroms der digitalen Objekte.

Grundlegende Anforderungen an das Speicherkonzept sind:

- ▶ redundante, geografisch verteilte Speicherlösung
- ▶ Integritätschecks, z.B. mit Prüfsummen
- ▶ Viruscheck
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Die Sicherung des Bitstroms eines Objekts ist ein stetig wiederkehrender Prozess und Basis jeder Erhaltungsstrategie.

Empfehlung

Größere Institutionen können allein oder im Verbund ein Rechenzentrum betreiben oder einen Dienstleister mit der Bitstream-Preservation beauftragen. Diese Entscheidung muss im Einzelfall getroffen werden.

Die institutionsinterne Bitstream-Preservation sollte die folgenden Anforderungen erfüllen:

- ▶ redundante, räumlich getrennte Speicherung auf verschiedenen Speichermedien
- ▶ Integritätschecks mit Prüfsummen
- ▶ Viruschecks
- ▶ Datenschutzkonzept
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring
- ▶ Skalierbarkeit

Prüfsummen können für ein gesamtes Objekt oder mit einem Tool, z.B. FFmpeg, auf Framebasis erstellt werden.

Es sollte eine professionelle Speicherlösung gewählt werden. Es wird zu einer Kombination aus Festplatten- und Bandspeicher geraten.

Darüber hinaus müssen Sicherheitsbestimmungen beachtet werden. Darunter fallen:

- ▶ physischer Zugang
 - ▶ Zugriffsberechtigungen
 - ▶ Einbruchmeldeanlage
- ▶ digitale Zugriffsberechtigungen
- ▶ Brandschutz
- ▶ Disaster Recovery

Mit einem Dienstleister sollte ein Service-Level-Agreement (SLA) vereinbart werden. Der Umfang der Dienstleistung muss genau evaluiert werden, um sicherzustellen, dass alle Anforderungen an die Bitstream-Preservation erfüllt werden.

Größere Institutionen können allein oder im Verbund ein Rechenzentrum betreiben oder einen Dienstleister mit der Bitstream-Preservation beauftragen. Diese Entscheidung muss im Einzelfall getroffen werden.

Content-Preservation

Content-Preservation soll sicherstellen, dass der Inhalt eines Objekts auch nach Jahren wiedergegeben und interpretiert werden kann.

Dafür ist es erforderlich:

- ▶ das Dateiformat, die -version und den Codec zu identifizieren und zu validieren. Bei AV-Medien wird diese Erkenntnis durch den Codec erschwert, der in dem Containerformat enthalten ist
- ▶ technische Metadaten zu extrahieren (siehe Tabelle mit Tools, S. 110ff.)
- ▶ Informationen über geeignete Wiedergabesoftware zu sammeln
- ▶ Risikofaktoren zu identifizieren
- ▶ signifikante Eigenschaften festzulegen
- ▶ Kontextinformationen bereitzustellen, die die Wiedergabe des Objekts und die Interpretation des Inhalts ermöglichen

Auf Basis dieser Informationen können Bestandserhaltungsmaßnahmen konzipiert und durchgeführt werden. Dafür gibt es zwei Verfahren: Formatmigration und Emulation.¹⁹

Große Institutionen sollten eine Erhaltungsstrategie ausarbeiten. Diese sollte folgende Punkte abdecken: Formatidentifikation und -validierung und daraus folgend das Bilden von homogenen Bestandsgruppen; Risikofaktoren identifizieren; Turnus für Überprüfung der Risikofaktoren festlegen; Erhaltungsmaßnahmen für Bestandsgruppen festlegen; geeignete Tools für die Erhaltungsmaßnahmen definieren; Policies nach dem SCAPE Policy Framework erstellen.

Empfehlung

Große Institutionen sollten eine Erhaltungsstrategie ausarbeiten. Die Erhaltungsstrategie sollte folgende Punkte abdecken:

- ▶ Formatidentifikation und -validierung. Daraus folgend das Bilden von homogenen Bestandsgruppen.
- ▶ Risikofaktoren identifizieren
- ▶ Turnus für Überprüfung der Risikofaktoren festlegen
- ▶ Erhaltungsmaßnahmen für Bestandsgruppen festlegen
- ▶ geeignete Tools für die Erhaltungsmaßnahmen definieren
- ▶ Policies nach dem SCAPE Policy Framework erstellen²⁰

Es wird die Formatmigration empfohlen, da es für die meisten audiovisuellen Objekte nicht erforderlich ist, die ursprüngliche Nutzungsumgebung zu erhalten. Eine Ausnahme stellt Videokunst dar, die im Wissenschaftlichen Film aber nicht als solche vorkommt.

Das Dateiformat sollte identifiziert und validiert werden. Es genügt nicht, ein Objekt über die Dateiendung zu identifizieren, da diese auch irreführend sein kann. Außerdem sollten technische Metadaten extrahiert werden.

Das Dateiformat und der Codec eines Objekts müssen bekannt sein, da alle Erhaltungsmaßnahmen hierauf basieren. Proprietäre Dateiformate sollten, wenn es möglich ist, nach dem Eingang der Originaldatei kopiert werden. Diese Kopie sollte in ein offenes, standardisiertes Archivformat konvertiert werden. Die Originaldatei soll unverändert erhalten bleiben.

¹⁹ Siehe http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf, Kapitel 8.3 und 8.4.

²⁰ <http://wiki.opf-labs.org/display/SP/SCAPE+Policy+Framework>

Signifikante Eigenschaften sollten definiert werden. Der Fokus sollte auf den für die Institution und deren Nutzer wichtigen Eigenschaften eines Objektes liegen.

Signifikante Eigenschaften hängen u.a. auch davon ab, welchen Zweck die Aufnahme hat. Ist der Zweck der Aufnahme z.B. die Dokumentation eines Ereignisses, eines Experiments usw., so ergibt sich daraus, was besonders erhaltenswert ist. Davon wiederum hängen ggf. weitere technische signifikante Eigenschaften ab.

Diese Anforderungen sollten nach einer Bestandserhaltungsmaßnahme überprüft werden.

Beispiele für technische signifikante Eigenschaften:

Eigenschaft	Komponente	Definition der Eigenschaft
Videokanäle	video	Anzahl der unterschiedlichen Videokanäle im digitalen Objekt
Audiokanäle	audio	Anzahl der unterschiedlichen Audiokanäle im digitalen Objekt
Länge	video	Anzahl der Frames in der Aufnahme
Breite	frame	Breite des Einzelbildes
Höhe	frame	Höhe des Einzelbildes
Bittiefe	pixel	Anzahl der Farbinformationen pro Pixel
Farbmodell	pixel	verwendetes Farbmodell bei der Aufnahme
Farbraum	pixel	verwendeter Farbraum bei der Aufnahme
Seitenverhältnis	pixel	Seitenverhältnis der Pixel
Bildrate	video	Anzahl der Frames in einem definierten Zeitintervall
Interlace	frame	falls es sich um interlaced video handelt
Metadaten		zum Objekt gehörende Metadaten

(Quelle: Archivematica Wiki²¹)

²¹ https://wiki.archivematica.org/Significant_characteristics_of_video_files

Grundlegende Risikofaktoren sollten definiert werden, z.B.:

- ▶ das Format oder der Codec ist von Obsoleszenz bedroht
- ▶ das Objekt liegt in einem proprietären oder kaum verbreiteten Format oder Codec vor

Tritt einer der Risikofaktoren auf, sollten alle Objekte im betroffenen Format in ein neues Zielformat migriert werden.

Authentizität

Die Sicherung der Authentizität bedeutet, nachweisen zu können, dass ein Objekt ist, was es zu sein vorgibt. Da Bestandserhaltungsmaßnahmen wie Formatmigrationen zwangsläufig eine Veränderung eines Objekts verursachen, ist die Sicherung der Authentizität nicht ganz unproblematisch. Folgende Maßnahmen sind zur Sicherung der Authentizität etabliert:

- ▶ die Aufbewahrung des Originals und aller Bearbeitungen des Objekts (Versionierung)
- ▶ der Nachweis von Veränderungen am Objekt in den Metadaten

Größere Institutionen sollten das Originalobjekt sowie alle Versionen des Originalobjekts aufbewahren und Änderungen in den Metadaten dokumentieren. Die Protokollierung von Änderungen muss technisch und organisatorisch sichergestellt sein.

Empfehlung

Größere Institutionen sollten das Originalobjekt sowie alle Versionen des Originalobjekts aufbewahren und Änderungen in den Metadaten dokumentieren. Die Protokollierung von Änderungen muss technisch und organisatorisch sichergestellt sein. Diese beinhaltet z.B.:

- ▶ Art und Zeitpunkt der Änderung
- ▶ Name des Benutzers, der die Änderung veranlasst hat
- ▶ Name des Benutzers oder des Programms, der/das die Änderung durchgeführt hat
- ▶ Grund der Änderung

Metadaten

Metadaten beschreiben ein Objekt. Es wird zwischen folgenden Metadatenkategorien unterschieden:

- ▶ deskriptive Metadaten: Es handelt sich um klassische Erschließungsmetadaten. Diese beschreiben ein Objekt und gewährleisten, dass es wiederauffindbar ist. Erschließungsmetadaten unterteilen sich in formale Metadaten wie Titel, Urheber, Erscheinungsjahr usw. und inhaltsbeschreibende Metadaten wie Schlagworte, ein Abstract, ein Transkript, Sequenzbeschreibungen usw.
- ▶ administrative Metadaten: Administrative Metadaten sind für die Verwaltung der Objekte im Langzeitarchiv und den Betrieb des Systems erforderlich, z.B. systeminterne Identifier, Nutzerdaten, Rollen- und Rechtevergabe.
- ▶ rechtliche Metadaten: Rechtliche Metadaten beschreiben den Nutzerzugriff auf die Objekte. Sie legen auch fest, ob die Objekte bearbeitet und wie sie bereitgestellt werden dürfen.
- ▶ technische Metadaten: Technische Metadaten beschreiben z.B. Dateiformat und -version, den Codec, Laufzeit, Dateigröße, Bitrate usw. Technische Metadaten können aus der Datei extrahiert werden.

- ▶ strukturelle Metadaten: Strukturelle Metadaten beschreiben die Zusammensetzung eines Objekts und Zusammenhänge mit anderen, z.B. wenn ein Objekt aus mehreren Dateien in einer festgelegten Reihenfolge besteht oder wenn Objekte Teil einer Serie sind.
- ▶ Provenienzmeteradaten: Diese Metadaten beschreiben die Objektgeschichte, um zu dokumentieren, welcher Nutzer wann mit welchen Tools welche Bearbeitung vorgenommen hat. Diese Metadaten sind von großer Wichtigkeit für den Nachweis der Authentizität.

Empfehlung

Größere Institutionen müssen Metadaten aller Kategorien erfassen.

Die Erschließung der Objekte wird für gewöhnlich in einem Nachweissystem erfolgen. Dort können neben deskriptiven Metadaten abhängig vom System auch technische, rechtliche und strukturelle Metadaten erfasst werden. Diese Metadaten sollten vollständig oder in Auswahl in das Langzeitarchivierungssystem überführt werden, das in der Lage ist, administrative und Provenienzmeteradaten zu erzeugen und zu verwalten.

Größere Institutionen müssen Metadaten aller Kategorien erfassen. Diese sollten vollständig oder in Auswahl in das Langzeitarchivierungssystem überführt werden, das in der Lage ist, administrative und Provenienzmeteradaten zu erzeugen und zu verwalten.

Erschließung

Neben den formalen Metadaten sollten auch inhaltsbeschreibende Schlagworte verwendet werden. Die Erschließungstiefe ist abhängig von den Vorgaben der Institution und den Anforderungen der Nutzer.

Wenn das Nachweissystem der Institution für die Erschließung und Verwaltung von Multimediaelementen geeignet ist, können die audiovisuellen Objekte mit dem Nachweissystem verwaltet und beschrieben werden. Ist das vorhandene Nachweissystem nicht geeignet, sollte frei verfügbare Datenbanksoftware verwendet werden, die für Multimediaobjekte geeignet ist.

Nachweissysteme sind in der Regel nicht für die Langzeitarchivierung geeignet.

Es kommt auch ein Digital Asset Management System (DAM)/Media Asset Management System (MAM) in Betracht. Ein DAM/MAM ist allerdings i.d.R. nicht für die Langzeitarchivierung geeignet. Ob das System für die Langzeitarchivierung geeignet ist, muss im Einzelfall untersucht werden.

Für den Nachweis der Objekte und die Erfassung von deskriptiven und strukturellen Metadaten sind Nachweissysteme oder DAM/MAM-Systeme jedoch gut geeignet. Ob rechtliche und technische Metadaten erfasst werden können, muss im Einzelfall geprüft werden.

Provenienz- und Erhaltungsmeteradaten

Es muss ein OAIS-konformes Langzeitarchivierungssystem eingesetzt werden, das in der Lage ist, alle erforderlichen Metadaten zu verwalten. Bereits vorhandene Metadaten, z.B. deskriptive und rechtliche Metadaten, können aus dem bestehenden Nachweissystem vollständig oder in Auswahl in das digitale Langzeitarchivierungssystem über eine Schnittstelle importiert werden. Im Langzeitarchivierungssystem sollten alle relevanten Metadaten in einer METS/MODS-Datei gespeichert werden, um eine schnellere Verarbeitung in Fällen der Formatmigration oder Metadaten-Updates zu gewährleisten.

Präsentation/Zugriff

Das Format der Nutzungskopien sollte die Anforderungen der vorgesehenen Nutzerzielgruppen berücksichtigen.

Weitverbreitete, offene Formate sollten bevorzugt werden; proprietäre Formate sind häufig an eine spezielle Wiedergabesoftware gebunden, binden an einen Hersteller und begrenzen die Nutzung. Im Idealfall soll die Nutzungskopie ohne die Installation zusätzlicher Software lesbar sein.

Für verschiedene Nutzungsszenarien können verschiedene Nutzungsformate erforderlich sein.

Größere Institutionen sollten Nutzungskopien erzeugen und für die Nutzung vorhalten. Es muss beachtet werden, wie die Objekte für die Nutzung zur Verfügung gestellt werden dürfen. Außerdem kann es erforderlich sein, dass der Zugriff auf angemeldete Benutzer oder nur bestimmte Nutzergruppen beschränkt werden muss.

Empfehlung

Größere Institutionen sollten Nutzungskopien erzeugen und für die Nutzung vorhalten. Das Format der Nutzungskopie sollte den Anforderungen der Nutzerzielgruppen entsprechen.

Es muss beachtet werden, wie die Objekte für die Nutzung zur Verfügung gestellt werden dürfen. Folgende Nutzungsszenarien sind denkbar:

- ▶ lokale Nutzung
- ▶ der Versand von Datenträgern
- ▶ Onlinezugriff mit und ohne Registrierung

Für die lokale Nutzung muss ein Format gewählt werden, das in der betreffenden Umgebung wiedergegeben werden kann.

Beim Versand von Datenträgern ist die Größe des Nutzungsformats ggf. durch die Speicherkapazität begrenzt. In diesem Fall kann eine Kompression erforderlich sein.

Den Webzugriff muss die Institution technisch bewältigen können: Da die Dateigröße die Übertragungsdauer beeinflusst, ist für den Webzugriff die Kompression des Objekts erforderlich. Außerdem kann es erforderlich sein, dass der Zugriff auf angemeldete Benutzer oder nur bestimmte Nutzergruppen beschränkt werden muss.

Dateiformate und Codecs

Für den Archivmaster sind offene, standardisierte und weitverbreitete Formate zu bevorzugen, die nicht oder verlustfrei komprimieren, um einem späteren Qualitätsverlust vorzubeugen. Proprietäre Dateiformate sind oft an eine spezifische Software eines Herstellers gebunden und stellen eine potenzielle Gefahr für die langfristige Verfügbarkeit dar, da der Code nicht öffentlich zugänglich ist und Formatmigrationen so erschwert oder sogar unmöglich werden. Zu beachten ist bei der Wahl des Archivmasters auch die Kompatibilität eines Formats und Codecs.

Darüber hinaus muss basierend auf den Anforderungen der vorgesehenen Nutzerzielgruppen auch eine Entscheidung über die Formate für Nutzungskopien getroffen werden. Eine Zusammenstellung empfohlener Formate und Codecs finden Sie auf den Seiten der nestor AG media unter <https://wiki.dnb.de/display/NESTOR/AG+Media>

Empfehlung

Die große Vielfalt potenzieller Ablieferungsformate stellt die Institutionen vor eine große Herausforderung. Je höher die Anzahl der Dateiformate, desto größer ist der Aufwand für die Bestandserhaltung (Preservation Planning, Preservation Watch, Preservation Actions, Risikomanagement usw.).

Es ist deshalb sinnvoll, die Anzahl der Dateiformate im digitalen Langzeitarchiv möglichst gering zu halten, sofern die Institution Einfluss darauf nehmen kann. Dies kann durch folgende Verfahren geschehen:

- ▶ die Veröffentlichung einer Liste bevorzugter Dateiformate
- ▶ das Erzeugen einer neuen Version durch Formatmigration nach dem Eingang proprietärer Objekte. Die Originaldateien müssen dabei unverändert aufbewahrt werden.
- ▶ eine Kombination beider Verfahren

Die Vorgehensweise sollte in einer öffentlich verfügbaren Policy dokumentiert werden, um den Prozess für Produzenten und Nutzer transparent zu machen.

Als Format empfiehlt sich Matroska (.mkv), ffv1-codiert, Audio: Wave Linear PCM 48 kHz. Die empfohlene Kombination aus Dateiformat und Codec kann mit FFmpeg erzeugt werden.

Die große Vielfalt potenzieller Ablieferungsformate stellt die Institutionen vor eine große Herausforderung. Es ist deshalb sinnvoll, die Anzahl der Dateiformate im digitalen Langzeitarchiv möglichst gering zu halten. Die Vorgehensweise sollte in einer öffentlich verfügbaren Policy dokumentiert werden, um den Prozess für Produzenten und Nutzer transparent zu machen.

Archivierung von professionellen Filmen

Im Gegensatz zu den anderen Medientypen dieser Handreichung werden die Empfehlungen für die Archivierung von professionellen Filmen nicht in die einzelnen Gruppen der Privatanwender und der verschiedenen Institutionstypen unterteilt. Die digitale Langzeitarchivierung von professionellen Filmen erfordert hohe personelle und technische Ressourcen, die i.d.R. durch private Personen oder kleinere Institutionen nur sehr schwer zu leisten sind. Selbst große Institutionen, die mit der Archivierung von professionellem Film betraut sind, liegen in der Entwicklung einer vertrauenswürdigen Langzeitarchivierung hinter Bibliotheken und Archiven weit zurück.

Das Kapitel „Archivierung von professionellen Filmen“ orientiert sich an DIN 31644 „Information und Dokumentation – Kriterien für vertrauenswürdige digitale Langzeitarchivierung“ und nimmt Bezug auf die Besonderheiten und Herausforderungen für professionelle AV-Materialien. Dabei werden vorwiegend jene Kriterien der DIN 31644, die durch diesen Medientyp beeinflusst werden, in den Vordergrund gerückt. In manchen Abschnitten wie z.B. über den Umgang mit deskriptiven Metadaten können dabei Empfehlungen ausgesprochen werden, in anderen Fällen gelingt lediglich eine Bestandsaufnahme der Herausforderungen. Der Text stellt deshalb gleichsam einen Anfangspunkt für noch zu entwickelnde Lösungen dar.

Die Differenzierung zwischen Amateurfilmen und professionellen Filmen wird in dieser Handreichung nicht technisch oder qualitativ unterschieden, sondern strukturell. Professioneller Film zeichnet sich durch eine Produktionsgesellschaft oder -firma aus, die die Herstellung, also insbesondere Finanzierung, Dreharbeiten und Postproduktion koordiniert. Bekannte Produktionsgesellschaften in Deutschland sind z.B. Constantin Film AG oder Bavaria Film GmbH.

Für kleinere Filmproduktionen werden jedoch auch kleinere Produktionsgesellschaften gegründet, an denen Regisseure und andere Filmschaffende beteiligt sein können oder diese innehaben. Ein Beispiel ist Harun Farocki GbR (Berlin). An Fernsehproduktionen können Produktionsgesellschaften beteiligt sein, oft übernehmen auch Fernsehanstalten diese Rolle selbst. Nicht selten sind Koproduktionen verschiedener Produktionsfirmen für Film- oder Fernsehproduktionen.

Auswahl des zu archivierenden Materials

Für die Auswahl des zu archivierenden Materials bilden in der digitalen – genauso wie in der analogen – Archivierung zunächst sogenannte Selektionskriterien die Grundlage. Selektionskriterien können inhaltlich, formal, rechtlich oder qualitativer Art sein.²² So können z.B. ein spezifisches Sammlungsprofil, ein gesetzlicher Auftrag oder ein Depositvertrag zwischen Archiv und Produzent rechtliche Selektionskriterien sein. Ein Archiv bestimmt damit, welche Objekte es für archivwürdig hält.

Daneben bilden auch formale Selektionskriterien eine Rolle. Diese beinhalten, mit welchen digitalen Formaten ein Archiv umgehen kann. Solche digitalen Formate werden auch als archivfähig beschrieben. Durch eine Übernahmeregelung bzw. einen Erhaltungsplan legt ein Archiv fest, welche Formate es als archivfähig definiert.

²² Keitel, Christian und Schoger, Astrid. 2013. Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Beuth-Kommentar. Berlin, 47-48..

Professionelle Filmproduktionen zeichnen sich jedoch mehr als andere Medientypen durch komplexere Produktionsprozesse und verschiedene Distributionsformate und -fassungen aus. Beispielsweise durch Verleihfassungen für den nationalen oder internationalen Markt, veränderte Schnittfassungen wie beispielsweise ein Director's Cut oder durch die Veröffentlichung für bestimmte Zielgruppen wie z.B. eine Kinodistribution oder eine erwerbzbare DVD.

Durch die unterschiedlichen Varianten, aber auch durch die verschiedenen Displayformate in den jeweiligen Anwendungskontexten entstehen große Datenmengen. Soll neben autorisierten Masterfassungen auch noch das Drehmaterial archiviert werden – ein Drehverhältnis für professionelle Film- und Fernsehproduktionen reicht von 1:6 bis 1:100 – potenzieren sich entsprechend Aufwand und notwendige Ressourcen.

All diese Umstände stellen spezielle Herausforderungen bei der Auswahl des zu archivierenden Materials dar:

1. Ein Archiv muss dafür Sorge tragen, dass die sorgfältige Erschließung der professionellen Filme, ihrer unterschiedlichen Varianten und Manifestationen gewährleistet wird, um Informationsobjekte genau zu bestimmen. Erst dadurch kann eine Auswahl hinreichend erfolgen, die Authentizität gewahrt, und gleichzeitig eine mögliche redundante Archivierung in verschiedenen Institutionen vermieden werden. Als derzeitige dafür voraussetzende Standards für die Erschließung filmografischer Daten sind CEN TC 372 Metadata Standards und die FIAF Cataloguing Rules zu nennen.
2. Ein Archiv sollte seine Sammlungspolitik und seine Selektionskriterien genau bestimmen und im optimalen Fall transparent veröffentlichen. Diese sollten auch in Abhängigkeit der verfügbaren Ressourcen betrachtet werden. So wäre z.B. möglich, dass eine Institution, deren Zielgruppe durch die Vermittlung von Filmen bestimmt wird, sich lediglich auf die Archivierung von autorisierten Manifestationen beschränkt, die für die Veröffentlichung bestimmt sind. Die dLZA des Drehmaterials könnte z.B. die Produktionsgesellschaft selbst übernehmen oder an eine andere Institution abgeben. Dabei ist eine redundante Speicherung in verschiedenen Institutionen zu vermeiden. Mehr als in der analogen Archivierung ist in der dLZA eine vorausschauende Planung von der Produktion bis zur Fertigstellung eines Filmprojekts notwendig. Sowohl für das Archiv, welches die Ressourcen für die Archivierung bereitstellen muss, als auch für die Produzenten, die sich auf die Anforderungen und Bedingungen des Archivs einstellen müssen, können dabei Übernahmevereinbarungen bzw. -richtlinien hilfreich sein, die rechtliche, formale und technische Absprachen zwischen Archiv und Produzenten definieren.
3. Abhängig der jeweiligen unterschiedlichen Produktions-, aber auch Digitalisierungsverfahren sieht sich ein Archiv mit unterschiedlichsten Dateiformaten und Codecs konfrontiert. Infolgedessen muss ein Archiv deshalb definieren, welche Formate archivfähig sind.²³

Beschreibung der vorgesehenen Nutzungsszenarien

Die erläuterte Auswahl von Archivmaterialien hängt mit den Nutzungsszenarien zusammen. Durch eine entsprechende Auswahl, beispielsweise durch eine entsprechende Sammlungspolitik oder andere Selektionskriterien, ergeben sich zunächst bestimmte Nutzungsszenarien. Darunter werden Anfragen einer oder mehrerer bestimmter Gruppen

²³ Zur Wahl von geeigneten Archivformaten siehe <https://wiki.dnb.de/display/NESTOR/AG+Media>.

an das Archiv verstanden, z.B. können digitale Filme an Kinos für Vorführungen verliehen werden. Auch der Filmwissenschaftler, der einen Filmausschnitt in seine Präsentation einarbeiten will, kann ein Nutzungsszenario darstellen.

Nutzungsszenarien von digitalen AV-Materialien zeichnen sich jedoch durch unterschiedliche Anforderungen und eine Bandbreite an Akteuren aus. Das Kino wird Wert auf eine hohe Auflösung legen, der Filmwissenschaftler eher auf eine geringe Dateigröße. Das Archiv bzw. der Regisseur legt Wert darauf, dass sich innerhalb der verschiedenen Nutzungsszenarios die Authentizität des Werks nicht verfälscht; beispielsweise durch eine Farbverschiebung infolge einer Formatmigration.

Die erste Herausforderung für die Archive liegt darin, die jeweiligen Nutzungsszenarien zu beschreiben, um den Bedürfnissen und Anforderungen der Nutzer entsprechend gerecht werden zu können. Eine geeignete Methode dafür ist eine Zielgruppenanalyse anhand einer Befragung oder auch eine Evaluierung der bisherigen Nutzungsszenarien. Dadurch können die häufigsten und relevanten Szenarien definiert bzw. eingegrenzt werden. In Anbetracht der Vielfalt von technischen Systemen und Akteuren wird ein Archiv jedoch nicht alle Anforderungen und Spezifikationen für die Nutzung von AV-Materialien abdecken können.

Die zweite Herausforderung für das Archiv beinhaltet die Definition der sogenannten signifikanten Eigenschaften. Darunter versteht man Attribute eines Informationsobjekts, welche für die Authentizität als relevant erachtet werden.²⁴ Signifikante Eigenschaften für alle Nutzungsszenarien können z.B. Farbigkeit, Aufnahmegeschwindigkeit oder das Seitenverhältnis sein – Eigenschaften, die sich also keinesfalls ändern sollten. Eine signifikante Eigenschaft für ein spezifisches Nutzungsszenario, beispielsweise die Projektion eines Films im Kino, kann z.B. der Raumklang, also ein Mehrkanalton sein.

Drittens muss das Archiv auf Änderungen von Anforderungen entsprechend reagieren. Beispielsweise ändern sich durch entsprechende Displayentwicklungen wie z.B. UHD oder die digitale Projektion von AV-Materialien auch deren Dateiformate und Codecs. Das Archiv muss infolgedessen notwendige Nutzungsformate auch anpassen können.

Die Nutzergruppen sollten klar beschrieben und abgegrenzt werden, die signifikanten Eigenschaften definiert, die spezifischen Formate für die Nutzung festgelegt und die Nutzerbedürfnisse kontinuierlich analysiert werden.

Empfehlung

- ▶ Beschreibung und klare Abgrenzung der Nutzergruppen
- ▶ Definition der signifikanten Eigenschaften
- ▶ Festlegung von spezifischen Formaten für die Nutzung
- ▶ kontinuierliche Analyse der Bedürfnisse der Nutzer

Qualitätsmanagement

Aufgrund der leichten Manipulierbarkeit von digitalen Objekten sind Qualitätssicherung und Validierung wesentliche Arbeitsschritte in der dLZA, die v.a. in der Übernahme, aber auch in den anderen funktionalen Bereichen des OAIS-Modells vorausgesetzt werden.²⁵ In der Übernahme sind diese Verfahren ein besonders kritischer Prozess, da sie maßgeblich alle weiteren Prozesse der dLZA beeinflussen.

²⁴ Zur Erläuterung der signifikanten Eigenschaften vgl. Abschnitt „Authentizität“, S. 80.

²⁵ Sie kommen z.B. auch bei Erhaltungsmaßnahmen (also in der Bitstream- und Content-Preservation), in der Datenverwaltung und im Archivspeicher zum Tragen.

Die Qualitätssicherung stellt durch die Validierung der Persistenzinformationen sicher, dass das Inhaltsinformationsobjekt sich nicht undokumentiert verändert. Verfahren der Qualitätssicherung sind erstens Dateiintegritätsprüfungen, z.B. durch zyklische Redundanzprüfungen (CRC-Verfahren) oder durch Prüfsummen, die Transferprozesse dokumentieren, und zweitens die Verfahren der Formatvalidierungen und Formatspezifizierungen. Nach den Vorgaben der DIN 31645²⁶ lassen sich Validierungen demnach in zwei verschiedene Klassen von Vergleichszielen unterscheiden:

Einerseits wird das zu validierende Objekt anhand seines „Eltern-Objekts“ überprüft, indem die Prüfsummen des digitalen Objekts vor und nach dem Transferprozess miteinander verglichen werden. Diese erste Klasse kann gleichgesetzt werden mit der Definition der Überprüfung der Dateiintegrität.

Andererseits wird nach inhaltlichen und formalen Vorgaben überprüft, ob das Format einer Datei auch seiner dokumentierten Spezifikation entspricht, bzw. inwieweit es davon abweicht. Dafür sind bestimmte Validierungs- und Verifizierungsschlüssel erforderlich²⁷, anhand derer die Authentizität des digitalen Objekts überprüft wird. Es wird damit sichergestellt, dass eine Information auch tatsächlich das ist, was sie vorgibt zu sein. Diese zweite Klasse kann mit den Begriffen Formatvalidierung und Formatspezifizierung gleichgesetzt werden.

Das Archiv muss dabei zunächst diese beiden Klassen von Validierungsprozessen (also Dateiintegritätsprüfung und Formatspezifizierung und -validierung) konkret definieren, abgrenzen und benennen.²⁸ Schließlich muss das Archiv den Rahmen und die Bedingungen dieser Prozesse, d.h. ihren Erfüllungsgrad bestimmen.

Abhängig von den Richtlinien der Institution und dem Erfüllungsgrad der Qualitätssicherung und Validierung werden bei der Übernahme spezifische Dateiformate oder definierte Übergabepakete (SIP) unter Berücksichtigung der Dateiintegrität übernommen. Zur Veranschaulichung lassen sich zwei verschiedene Extrema nennen:²⁹ Im optimalen Fall liegt einerseits ein definiertes Übergabepaket vor, das andererseits Dateiformate beinhaltet, die seitens des Archivs als Archivformate akzeptiert sind. Folglich lassen sich in der Qualitätssicherung und der Validierung des Dateiformats bestimmte – anschließend erläuterte – Prozesse anwenden, und es kann darüber hinaus auf eine Migration der Dateiformate verzichtet werden. Im ungünstigsten Fall übernimmt das Archiv Dateiformate, die sich nicht als Archivformate eignen und außerdem nicht als Übergabepaket definiert sind. Dementsprechend lassen sich in der Qualitätssicherung und Validierung bestimmte Maßnahmen zum einen nicht durchführen, zum anderen ist ein Migrationsprozess in ein akzeptiertes Archivformat obligatorisch.

In der Praxis werden Qualitätssicherung und Validierungsprozesse durch Software unterstützt. In vielen Fällen greift dabei eine Software auf gepflegte Registraturen von Dateiformaten zurück, in welchen Formate beschrieben und charakterisiert sind. An dieser Stelle sollen drei weitverbreitete, softwareunterstützte Registraturen vorgestellt werden, die von verschiedenen Institutionen bzw. Initiativen derzeit weiterentwickelt und gepflegt werden.

26 Deutsches Institut für Normung e. V. 2011. „Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive“, Nr. 31645.

27 nestor-Arbeitsgruppe Digitale Bestandserhaltung (2012). „Leitfaden zur digitalen Bestandserhaltung“.

28 Deutsches Institut für Normung e. V. (2011). „Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive“, Nr. 31645.

29 Keitel, Christian, 2010. „Archive.“ In Neuroth, Oßwald, Scheffel, and Strathmann, S., Huth, K., nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, 36.

Erstens die Initiative PRONOM³⁰ der britischen Behörde The National Archives. Das Online-Informationssystem PRONOM stellt dabei eine Formatregistratur bereit, die sich mithilfe der Software DROID (Digital Record Object Identification)³¹, ebenfalls entwickelt von The National Archives, abfragen lässt. PRONOM stellt zudem ein API (application programming interface) bereit, also eine Schnittstelle, die ermöglicht, dass PRONOM neben DROID auch mit anderer Software kommunizieren kann. Beispielsweise nutzt die Software FIDO³² mittels dieser Schnittstelle das Verzeichnis des PRONOM. Ein weiteres bedeutendes Merkmal seitens PRONOM ist die Bereitstellung von Persistent Unique Identifiers (PUIs) für die jeweiligen Formate.

Als zweite Initiative einer softwareunterstützten Registratur ist JHOVE (JSTOR/Harvard Object Validation Environment)³³ zu nennen. Sie wurde durch die gemeinnützige Organisation Journal Storage (JSTOR) und die Harvard University Library entwickelt und unterstützt – neben anderen Funktionen – die Formaterkennung und -validierung von digitalen Objekten.

Als dritte Initiative ist die Koordinationsstelle für die dauerhafte Archivierung von elektronischen Unterlagen (KOST) zu nennen, die ebenfalls einen Validator zur Verfügung stellt.³⁴

Für die Formaterkennung und -validierung von AV-Objekten lassen sich in den drei Initiativen jedoch Unterschiede erkennen: Der Validator der KOST unterstützt derzeit nur spezifische Text- und Bildformate und Formate für die Archivierung von Datenbanken. Die Software JHOVE schließt bestimmte AV-Materialien ebenso von der Validierung und Formaterkennung aus. Lediglich PRONOM beinhaltet Unterstützung für AV-Materialien, dabei vornehmlich jedoch für digitale Videoformate, während sich die im Kapitel „Digitaler Film“ angeführten Formate des DCP, DCDM oder IMF nicht durch PRONOM validieren lassen.

Einen sehr vielversprechenden Ausblick, insbesondere für die Validierung von AV-Materialien, bietet das im Januar 2014 gestartete EU-Projekt „Preservation formats for culture information/e-archives (PREFORMA)“.³⁵ Die Hauptziele des mit 2,8 Millionen Euro ausgestatteten Projekts sind die Entwicklung und Bereitstellung von Open-Source-Software für die Formatvalidierung und Formatstandardisierung von spezifischen Dateiformaten der Medientypen Text, Bilder und AV-Medien. Für die Umsetzung des Arbeitspakets AV-Materialien wurde die Firma MediaArea beauftragt, welche für die Containerformate MKV und Ogg sowie die Codecs FFV1 und Dirac v2.2.3 den Validator MediaConch³⁶ entwickeln wird.

Erhaltungsstrategien

Die Erhaltung digitaler Daten ist das Kerngeschäft des digitalen Langzeitarchivs. Maßnahmen für die Erhaltung digitaler Daten sollten nicht ad hoc erfolgen, sondern durch eine langfristige Planung sichergestellt werden.³⁷ Die Beschreibung von Erhaltungsstrategien

30 The National Archives. 2015. „The technical registry PRONOM.“ Zugriff: 25. Juli 2015. <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>.

31 „DROID“. Zugriff: 25. Juli 2015. <http://sourceforge.net/projects/droid/>.

32 „Format Identification for Digital Objects (FIDO) – Open Preservation Foundation“. Zugriff: 25. Juli 2015. <http://openpreservation.org/technology/products/fido/>.

33 „JHOVE – JSTOR/Harvard Object Validation Environment“ 2012. Zugriff: 25. Juli 2015. <http://jhove.sourceforge.net/>.

34 Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen und KOST. Zugriff: 25. Juli 2015. http://kost-ceco.ch/cms/index.php?kost_val_de.

35 Vgl. ausführlich dazu: Preforma Projekt. <http://www.preforma-project.eu/>

36 Der Entwicklungsstand ist verfügbar unter „MediaArea 2015a“. „Sustainability Factors.“ Zugriff: 25. Juli 2015. <https://mediaarea.net/MediaConch/>.

<http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>.

37 Keitel, Christian und Schoger, Astrid. 2013. Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Beuth-Kommentar. Berlin: Beuth, K11.

wird auch als Preservation Strategies bzw. Policies oder auch als Erhaltungsplan bezeichnet. Ein Erhaltungsplan lässt sich in zwei Bereiche gliedern: die Bitstream- und die Content-Preservation.

Bitstream-Preservation

Dabei behandelt die Bitstream-Preservation den reinen physischen Erhalt der Daten. Maßgebliche Verfahren für die Bitstream-Preservation sind eine redundante Datenhaltung, also eine Herstellung von Sicherheitskopien und eine Diversität eingesetzter Speichertechnik. Seitens der Technologie, d.h. also der Entwicklung von Speichermedien und Speichertechnik, stellt dies keine große Herausforderung mehr dar. Aktuelle technologische Lösungen sind z.B. Festplattenverbundsysteme (RAID = Redundant Array of Independent Disks) oder Bandbibliotheken (Tape Libraries). Unter Berücksichtigung der Notwendigkeit einer ständigen Aktualisierung solcher Systeme kann diese Frage prinzipiell als gelöst betrachtet werden.³⁸ Die Frage der Bitstream-Preservation ist v.a. eine ökonomische, die abhängig von verfügbaren Mitteln und Ressourcen beantwortet wird.

Grundlegende Anforderungen an die Bitstream-Preservation sind:

- ▶ Gewährleistung der Langzeitverfügbarkeit
- ▶ Schutz vor unbefugtem Zugriff durch ein Rechtemanagement
- ▶ redundante, geodiverse Speicherung
- ▶ Integritätschecks
- ▶ Viruschecks
- ▶ regelmäßiger Austausch von Datenträgern
- ▶ Monitoring und Skalierbarkeit
- ▶ Absicherung des physischen Zugangs durch Zugriffsberechtigungen und Einbruchmeldeanlage
- ▶ Brandschutz
- ▶ Disaster Recovery

Content-Preservation

Zum anderen benötigt die dLZA jedoch tragfähige Konzepte und Verfahren für die Erhaltung der logischen und konzeptuellen digitalen Objekte. Bedingt durch die Veränderung von Interpretationen der Daten wie z.B. Software, aber auch durch Veränderungen der Hardwareumgebung können Dateiformate und/oder Software obsolet werden. D.h., es sind Strategien notwendig, um die Informationen auch zukünftig verstehbar zu halten. Diese Strategien werden unter dem Begriff **Content-Preservation** zusammengefasst.

Dabei haben sich maßgeblich zwei verschiedene Konzepte etabliert, die beide unabhängig voneinander oder auch kombiniert zum Tragen kommen können: Migration und Emulation. Eine Definition für **Migration** ist „die rechtzeitige und genau dokumentierte, hinreichend ausgetestete Überführung von Daten aus einem nicht mehr beherrschbaren Format in ein anderes, das [...] eine Zukunftsperspektive hat.“³⁹ Zu differenzieren ist die Migration in vier verschiedene Arten von Migrationsprozessen. Diese sind: Refreshment, Replication, Repackaging und Transformation.⁴⁰

38 Straube, Arnim und Beinert, Tobias. 2013. „Aktuelle Herausforderungen der digitalen Langzeitarchivierung.“ In: Klimpel, Paul und Keiper, Jürgen Was bleibt? Nachhaltigkeit der Kultur in der digitalen Welt. Berlin: iRights Media, 36

39 Althenhöner, Reinhard und Sabine Schrimpf. 2014. „Langzeitarchivierung.“ In: Kuhlen, Rainer Semar, Wolfgang und Strauch, Dietmar (Hg.), Grundlagen der praktischen Information und Dokumentation: Handbuch zur Einführung in die Informationswissenschaft und -praxis. Völlig neu gefasste Ausg., Berlin: De Gruyter, 529-540.

40 Consultative Committee for Space Data Systems 2011, 5-4-5-5. OAIS Referenz. Zugriff: 22. März 2015. <http://public.ccsds.org/publications/archive/650x0m2.pdf>.

Dabei zählen Refreshment (der Austausch von gleichwertigen Datenträgern) und Replication (ein Austausch von Datenträgern, der zudem eine Anpassung der Speicherinfrastruktur vornimmt) genau genommen zu den Methoden der Bitstream-Preservation, da sie den Datenstrom nicht verändern. Repackaging und Transformation hingegen überführen die Dateien in eine andere Struktur bzw. ein anderes Format und wandeln den Bitstream um. Nur diese beiden Migrationsverfahren werden im Begriff Formatmigration zusammengefasst und sind ein Teil der Content-Preservation.

Bei der Formatmigration entsteht eine neue Repräsentation des digitalen Objekts. Es handelt sich eben nicht mehr um eine bitgenaue Kopie, sondern um eine Veränderung des Bitstreams. Der Anspruch, den verstehbaren Inhalt des Informationsobjekts, also dessen Authentizität, nicht zu verändern, kommt dabei aber zur Geltung⁴¹; es werden schließlich nur die Schreib- und Lesart des Objekts geändert, nicht aber der daraus extrahierte Inhalt.

Die zweite Erhaltungsstrategie für digitale Daten bildet die **Emulation**. Das Konzept der Emulation sieht keine Formatmigration vor, die digitalen Objekte liegen in der Langzeitarchivierung deshalb unverändert, bitgenau vor. Zweifelsohne wird damit die Authentizität der digitalen Objekte vollständig erhalten.

Um jedoch der Obsoleszenz von veralteten Dateiformaten, der Veränderung von Hardwareumgebungen oder Betriebssystemen vorzubeugen, werden stattdessen andere Vorkehrungen getroffen. Die Emulation bildet die originale Umgebung der digitalen Objekte nach. Die originalen Anwendungsprogramme oder Betriebssysteme werden innerhalb einer weiteren, moderneren Hardware- und Softwareumgebung mit Emulatoren ausgeführt, um das Informationsobjekt zu interpretieren. Die Erforschung und Anwendung von Emulationsstrategien kommt derzeit in der Archivierung von Software, insbesondere von Computerspielen, zum Tragen – hauptsächlich weil Formatmigration sich aufgrund der Komplexität dieser digitalen Objekte ausschließt.

Für die dLZA von AV-Materialien sind derzeit noch keine der beiden Erhaltungsstrategien entschieden. Auf fachlicher Ebene überwiegen jedoch die Debatten über die Formatmigration als zukünftige Strategie. Dabei ist anzumerken, dass sich innerhalb der Archivierung von AV-Archivalien – insbesondere durch die Digitalisierung analoger Materialien – bereits ein Bewusstsein für die Problematik geschärft hat: Es wurde von Beginn an auf archivfähige Formate gesetzt, deshalb ist bislang auch noch keine im größeren Umfang stattfindende zwingende Formatmigration bekannt.⁴²

Authentizität

Ein wesentliches Kriterium für die digitale Langzeitarchivierung ist der Erhalt der Authentizität des Informationsobjekts in den jeweiligen Repräsentationen sowie Bearbeitungsstufen. Dabei lässt sich der Erhalt bzw. die Sicherstellung der Authentizität in drei verschiedene Teilaspekte in der dLZA untergliedern. Diese sind Aufnahme, Erhaltungsmaßnahmen und schließlich Nutzung von Informationsobjekten.⁴³

Zunächst muss bei der Aufnahme, also dem Übernahmeprozess in das Archiv sichergestellt werden, dass das Informationsobjekt authentisch ist. Damit wird sichergestellt, dass ein digitales Objekt auch dem entspricht, was es inhaltlich vorgibt zu sein.

41 Keitel, Christian und Schoger, Astrid. 2013. Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Beuth-Kommentar. Berlin: Beuth, 14.

42 Straube, Arnim und Beinert, Tobias. 2013. „Aktuelle Herausforderungen der digitalen Langzeitarchivierung.“ In: Klimpel, Paul und Keiper, Jürgen (Hg.), Was bleibt?, 37.

43 Keitel, Christian und Schoger, Astrid. 2013. Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Beuth-Kommentar. Berlin: Beuth, K17-K19.

Wie im Abschnitt „Erhaltungsstrategien“ beschrieben, können durch notwendige Formatmigrationen jedoch z.B. neue Repräsentationen des Informationsobjekts entstehen. Zwangsläufig wird dadurch das Objekt verändert. Um die Authentizität nach der Formatmigration sicherzustellen, bieten sich folgende Maßnahmen an:

- ▶ die Aufbewahrung des Originals und aller folgenden Bearbeitungen des Objekts (Versionierung der Repräsentationen)
- ▶ der Nachweis von Veränderungen am Objekt in den Metadaten

der dritte und letzte Teilaspekt bei der Sicherstellung der Authentizität betrifft die Nutzung von digitalen Objekten. Für die jeweiligen Nutzungsszenarien werden die verschiedenen Archivpakete zu Nutzungspaketen transformiert. Dies erfordert für AV-Materialien i.d.R. auch unterschiedliche Transkodierungen für z.B. unterschiedliche Displays. Auch hierbei muss sichergestellt werden, dass die Authentizität in den jeweiligen Nutzungsrepräsentationen gewährleistet ist.

Signifikante Eigenschaften

Ein Konzept, welches als Sicherstellung der Authentizität für alle drei Teilaspekte gilt, ist die Definition von signifikanten Eigenschaften. Das digitale Archiv identifiziert, welche Eigenschaften der digitalen Objekte als wichtig erachtet werden und damit sowohl für mögliche Erhaltungsmaßnahmen, aber auch Nutzungsszenarien voraussetzend sind. Dabei steht der Umfang der zu bewahrenden Eigenschaften einerseits in Zusammenhang mit dem Aufwand, der für die digitale Langzeitarchivierung notwendig wird, und andererseits mit den Bedürfnissen der jeweiligen Nutzungsszenarien.⁴⁴

Dabei lassen sich die signifikanten Eigenschaften digitaler AV-Materialien in folgenden Gruppen mit unterschiedlichen Attributen unterscheiden:

- ▶ signifikante Eigenschaften für einzelne Pixel
 - ▶ Pixelseitenverhältnis, Farbraum, Farbmodell, Bittiefe, Subsampling
- ▶ signifikante Eigenschaften für einen Frame (Einzelbild)
 - ▶ Auflösung horizontal und vertikal, Bildseitenverhältnis, Bildkaschierung (Pillarboxed, Letterboxed), Zeilensprungverfahren
- ▶ signifikante Eigenschaften für die Bildfolge (Video)
 - ▶ Bildrate (Aufnahmegeschwindigkeit und Abspielgeschwindigkeit), Laufzeit, Kanäle
- ▶ signifikante Eigenschaften für den Ton (Audio)
 - ▶ Kanäle, Abtastrate, Bittiefe, Rauschunterdrückung, Raumklang
- ▶ signifikante Eigenschaften für Text (Untertitel, Zwischentitel o.a. Titel)
 - ▶ Schriftart, Schriftgröße, Einblend- und Ausblendzeiten, Schriftfarbe
- ▶ signifikante Eigenschaften für Dimensionen (2D/3D/VR)
- ▶ signifikante Eigenschaften für Kompositionen bzw. Abspielfolgen (Timecode und Elemente)

Für AV-Materialien bestimmen die signifikanten Eigenschaften auch maßgeblich Archiv- und Nutzungsformate, da diese in unmittelbarer Abhängigkeit zueinander stehen. So lassen sich z.B. Untertitel oder auch bestimmte Farbräume nur mit bestimmten Formaten abbilden.

Metadaten

Eine hohe Relevanz von Standardisierungen kommt innerhalb der dLZA insbesondere bei den Metadaten zum Tragen. Für die Dokumentation und Archivierung digitaler AV-Materi-

⁴⁴ Keitel, Christian und Schoger, Astrid. 2013. Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Beuth-Kommentar. Berlin: Beuth, K 13.

alien lassen sich zunächst Metadaten hinsichtlich ihrer Funktion kategorisieren:

- ▶ deskriptive Metadaten (descriptive) – beschreibende filmografische Daten
- ▶ technische Metadaten (technical) – technische und formatspezifische Attribute von AV-Dateien
- ▶ strukturelle Metadaten (structural) – Daten über die Zusammensetzung von digitalen Objekten
- ▶ Erhaltungsmetadaten (preservation) – Metadaten für die langfristige Interpretation
- ▶ administrative Metadaten (administrative) – Metadaten für die Verwaltungsinformation
- ▶ rechtliche Metadaten (rights) – Metadaten über Rechteinhaber

Für jede Funktion existieren verschiedene Metadatenstandards, die Grenzen sind dabei jedoch nicht immer klar zu ziehen, mehrere Standards können auch mehrere Funktionen beinhalten. Einen universellen Metadatenstandard für die dLZA von AV-Materialien gibt es nicht, die Herausforderung liegt vielmehr darin, für den jeweiligen Anforderungskontext die geeigneten Metadatenstandards festzulegen und dabei die wesentlichen Funktionen abzudecken.

Deskriptive Metadaten

Unter deskriptiven Metadaten versteht man die bibliografische, identifizierende Beschreibung eines Objekts, dazu zählen der Autor oder der Titel eines Werks. Die Funktionen deskriptiver Metadaten lassen sich anhand der wichtigsten theoretischen Grundlage innerhalb der bibliografischen Dokumentation nachvollziehen: Das FRBR-Modell (Functional Requirements for Bibliographic Records) postuliert vier Anforderungen an deskriptive Metadaten. Sie müssen es ermöglichen, Objekte durch eine Suchanfrage aufzufinden, ein spezifisches Informationsobjekt zu identifizieren und zu filtern, um letztlich Zugang zum Objekt zu erhalten.⁴⁵

Ein weitverbreitetes Metadatenchema – das Dublin Core Metadata Element Set (DCMES), welches für diese Anforderungen entwickelt wurde – stammt von der Dublin Core Metadata Initiative (DCMI). Auf einer Konferenz in Dublin in den 1990er Jahren einigten sich die Teilnehmer auf 15 Elemente, die alle inhaltlichen Kernelemente (core elements) eines Dokuments abdecken sollen. Das Ergebnis wurde schließlich als ISO Norm 15836 zertifiziert;⁴⁶ dabei lag das Hauptaugenmerk der Konzeption auf Retrieval, d.h. auf der stichwortbasierten Abfrage innerhalb von Webressourcen.

Für vielschichtigere Dokumente bzw. Objekte erweist sich dieser – wenn auch von der Idee her richtige – universelle Ansatz jedoch als unzureichend. So wird bei der Katalogisierung häufig eine höhere Erschließungstiefe angestrebt als durch die Beschreibung der in den core elements erfassten Kernelementen möglich ist. Ein komplexes Objekt wie Film lässt sich nicht nur auf einen Urheber (creator) reduzieren, neben dem Regisseur kommen auch die Verantwortlichen für Kamera, Schnitt u.v.m. zum Tragen.

Standardisierte, deskriptive Metadatenschemata von Filmen und audiovisuellen Medien sind dagegen unterrepräsentiert. In den verschiedenen Institutionen, die mit der Archivierung und der Distribution von AV-Materialien betraut sind, haben sich – v.a. mangels eines geeigneten Datenmodells – individuelle Erschließungskonzepte entwickelt, die sich international sowie national folglich unterscheiden.

45 „Find materials, identify an entity, select an entity, acquire and obtain access.“
Zugriff: 9. April 2015 http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf.

46 „Dublin Core Metadata Element Set, Version 1.1“.

Zugriff: 3. Mai 2015 <http://dublincore.org/documents/dces/>.

Nicht unbegründet forderte die Politik seitens des Europäischen Rates im November 2005 die EU-Mitgliedsstaaten auf, „die europaweite Standardisierung und die Interoperabilität filmografischer Datenbanken sowie deren Verfügbarkeit für die Öffentlichkeit [zu] verbessern [...]“.⁴⁷ Die Umsetzung dieser politischen Forderung lässt sich durch die Beauftragung des Europäischen Komitees für Normung (CEN, Comité Européen de Normalisation) nachvollziehen: Das Mandat beinhaltet die Entwicklung eines „umfassende[n] Metadatenvokabular[s] sowie eine[r] umfassende[n] und einheitliche[n] Terminologie zur Beschreibung kinematographischer Werke“ und die Definition von Datenelementen, „die als vorrangig zur Unterscheidung und Identifikation einzelner kinematographischer Werke angesehen werden.“⁴⁸

Zwischen 2009 und 2010 wurde durch das CEN der Standard CEN TC 372 schließlich veröffentlicht und damit ein erster Schritt zur Standardisierung von deskriptiven filmografischen Daten getan. Der Standard untergliedert sich in die Normen DIN EN 15744:2009 (Identifikation von Filmen – Mindestsatz von Metadaten für kinematographische Werke) und DIN EN 15907:2010 (Identifikation von Filmen – Verbesserung der Interoperabilität von Metadaten – Elementsätze und Strukturen). Ein wesentlicher Teil dieser Normen beinhaltet die Eingliederung der anfangs vorgestellten Funktionen deskriptiver Metadaten nach dem FRBR-Modell, angepasst an die Bedürfnisse und Bedingungen zur Beschreibung von AV-Medien. Die schematische Darstellung (Abb. 1: CEN TC 372) zeigt die Differenzierungen von Entitäten und ihre jeweiligen Assoziationen analog zum FRBR-Schema.

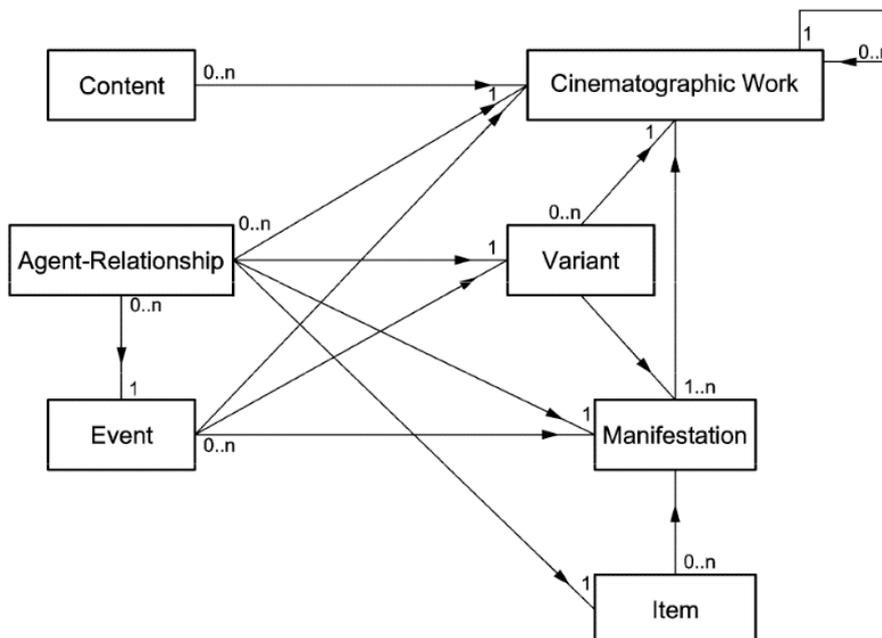


Abb. 1: CEN TC 372. Entitäten und ihre Assoziationen

47 Europäisches Komitee für Normung 2010. Identifikation von Filmen – Verbesserung der Interoperabilität von Metadaten – Elementsätze und Strukturen, 15907:2010. Zugriff: 3. Mai 2015. http://filmstandards.org/media/EN_15907_Deutsch.pdf.

48 Europäisches Komitee für Normung 2010. Identifikation von Filmen – Verbesserung der Interoperabilität von Metadaten – Elementsätze und Strukturen, 15907:2010. Zugriff: 3. Mai 2015. http://filmstandards.org/media/EN_15907_Deutsch.pdf. 5.

Anschaulicher lässt sich der Vorteil dieses Datenmodells an einem Beispiel erläutern: Das kinematographische Werk (Cinematographic Work) METROPOLIS (FRITZ LANG, DE 1925/26) wird durch mehrere Varianten (Variant), so z.B. durch abweichende Schnittfassungen für die jeweilige länderspezifische Distribution, repräsentiert. Die jeweiligen Varianten können jedoch durch eine Restaurierung zudem eine neue Ausprägung (Manifestation) erhalten, die sich wiederum in spezifische physische Objekte (Items) unterteilt. In Verbindung mit den Entitäten Events und Agent-Relationship lässt sich deshalb in der deskriptiven Beschreibung folgende hypothetische Identifizierung herstellen:

Beispiel

Das Werk METROPOLIS (FRITZ LANG, DE 1925/26) in seiner Variante (Zensurfassung, DE, 1926) restauriert durch Hans Mustermann (Agent-Relationship) wurde als Distributionsfassung (HANS MUSTERMANN, DE 2012) (Manifestation) am 26.10.2012 uraufgeführt (Event) in Form eines Digital Cinema Package (Item).

Diese differenzierte Beschreibung erreicht damit eine hohe Genauigkeit bei der Identifizierung von AV-Materialien. Deshalb ist CEN TC 372, wenngleich auch noch ein sehr junger, der derzeit vielversprechendste europäische Standard für die Erhebung deskriptiver, filmografischer Daten. Beispiele für die Umsetzung des Standards sind z.B. am British Film Institute (BFI), im EU-Projekt EFG 1914 (2013-2014) oder auch im Bestandskatalogprojekt des Kinematheksverbunds (2013-2015) zu verzeichnen. Auch das größte deutsche Filmarchiv, das Bundesarchiv, hat CEN TC 372 in der neuen Datenbankapplikation implementiert.

Kritisch zu hinterfragen bleibt die Praxis, bei diesen unterschiedlichen Implementierungen von CEN TC 372 auf Teile der Standardisierung zu verzichten bzw. diese zu modifizieren. Das angepasste CEN TC 372-Datenmodell des BFI, des EYE-Film Instituts oder auch der Deutschen Kinemathek verzichtet z.B. auf die Differenzierung in Varianten. Eine wenn auch geringfügige Modifikation untergräbt die Idee von Standardisierungen, also deren Vorteile der Interoperabilität und der zukünftigen Nachhaltigkeit. Zusätzlich fehlen weitestgehend Implementierungen in verschiedene Datenbankmanagementsysteme, weshalb der Handlungsspielraum bei der Auswahl von Datenbanken eingeschränkt ist.

Technische Metadaten

Neben der deskriptiven Ebene lassen sich digitale Objekte durch technische Metadaten erfassen, welche einerseits qualitative Aussagen beinhalten, andererseits auch die Provenienz und den Herstellungsprozess dokumentieren. Im Gegensatz zu deskriptiven müssen technische Metadaten nicht zwingend erschlossen werden, da eine Bereitstellung durch Dateispezifikationen erfolgt.

Typisch für technische Metadaten ist deshalb, dass sie sich aus digitalen Objekten extrahieren lassen. Bei AV-Materialien lassen sich Elemente wie z.B. das Dateiformat, die Auflösung eines Videos, dessen Codierung oder auch der Farbraum aus den digitalen Objekten herauslesen. Diese Metadaten liegen in der Datei bereits vor (embedded Metadata). Eine aufwändige Erschließung erübrigt sich deshalb ganz im Gegensatz zur Erfassung deskriptiver Metadaten; die Herausforderung liegt vielmehr darin, geeignete automatisierbare Methoden zu benutzen und in die bestehenden Systeme und Workflows zu integrieren. Softwares für die Extraktion technischer Metadaten von AV-Objekten sind z.B. ffprobe⁴⁹,

⁴⁹ „ffprobe Documentation.“ 2015. Zugriff: 8. Mai 2015.
<http://www.ffmpeg.org/ffprobe.html>.

MedialInfo⁵⁰ und VideoSpec⁵¹. Der Support und die Entwicklung der Software sind eingestellt worden und deshalb nur bedingt zu empfehlen (vgl. „VideoSpec by houdini Software – Analyze your video“ 2015).

Erhaltungsmetadaten

In der dLZA werden neben der deskriptiven und technischen Erfassung von digitalen Objekten Erhaltungsmetadaten differenziert. Für die digitale Archivierung kommen Langzeiterhaltungsmetadaten eine besonders hohe Bedeutung zu, so liegt in ihnen erst die Grundvoraussetzung für die langfristige Erhaltung digitaler Objekte: „Langzeitarchivierungsmetadaten sind – vereinfacht ausgedrückt – strukturierte Informationen über digitale Objekte, ihre Kontexte, ihre Beziehungen und Verknüpfungen, welche die Prozesse der digitalen Langzeitarchivierung ermöglichen, unterstützen oder dokumentieren.“⁵²

Erhaltungsmetadaten gewährleisten somit, dass der langfristige Verbleib, die Lesbarkeit und Verwendbarkeit von digitalen AV-Materialien im Archiv gesichert sind. Sie verknüpfen ausgesuchte Metadaten aus allen Funktionen miteinander, geben Information über die zukünftige Erreichbarkeit der Datei und initiieren und dokumentieren notwendige Maßnahmen für den weiteren Erhalt.

Die Bereitstellung von Metadaten-Schemata ist jedoch überschaubar, es lassen sich maßgeblich zwei etablierte Ansätze verfolgen: PREMIS und LMER.

Neben einem Datenmodell stellt PREMIS zudem ein Data Dictionary bereit, welches die Implementierung von Langzeiterhaltungsmetadaten in einem Metadaten-Schema ermöglicht. Darüber hinaus können mit PREMIS Erhaltungsmaßnahmen protokolliert werden. Das Schema greift dabei bewusst lediglich auf eine Schnittmenge von Metadaten zurück, welche ein digitales Langzeitarchiv insgesamt erhebt.⁵³ Abb. 2 veranschaulicht, dass sich PREMIS im Zentrum befindet und nur eine Teilmenge (dunkelgrau) von technischen, deskriptiven sowie rechtlichen und administrativen Metadaten, welche zu den Erhaltungsmetadaten (grau) zählen, vereinnahmt. Die fehlenden Metadaten, welche für eine digitale Archivierung als wichtig erachtet werden, stehen nicht im Fokus des Datenmodells PREMIS und müssen demnach aus anderen Metadatenstandards bereitgestellt werden.

50 MediaArea 2015b. „MedialInfo.“ Zugriff: 8. Mai 2015.

<https://mediarea.net/de/MedialInfo>.

51 „VideoSpec by houdini Software – Analyze your video.“ 2015. Zugriff: 8. Mai 2015.

<http://videospec.free.fr/english/>.

52 Brandt, Olaf 2010, „PREMIS-Metadatenstandards im Bereich der digitalen LZA.“ In: Neuroth, Oßwald, Scheffel and Strathmann, S., Huth, K., Nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, 128-132.

53 Caplan, Priscilla. 2009, „PREMIS verstehen.“ Zugriff: 10. Mai 2015.

http://www.loc.gov/standards/premis/understanding_premis_german.pdf. 5.

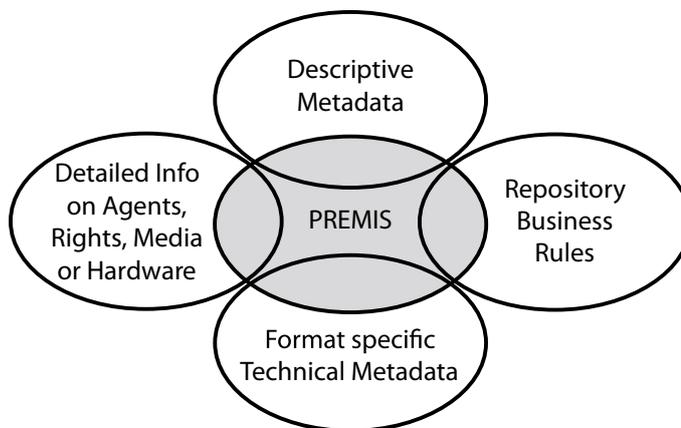


Abb. 2: Schnittmenge von Metadatenfunktionen innerhalb PREMIS

Die dabei erhobenen Daten werden im PREMIS-Modell zusätzlich klassifiziert (Abb. 3), es sind fünf Entities (Entitäten) definiert: Objects, Rights, Agents, Events und Intellectual Entities werden durch das PREMIS Data Dictionary als semantische Einheiten (semantic units) näher beschrieben.⁵⁴

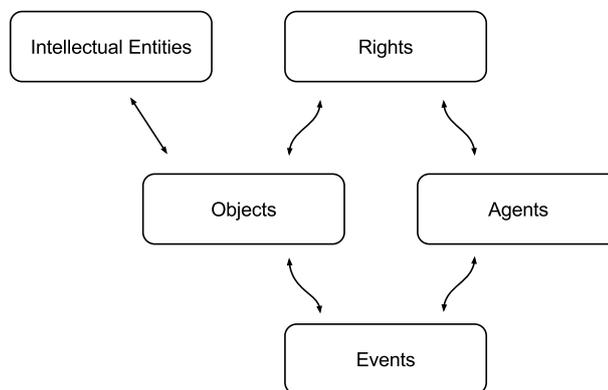


Abb. 3: PREMIS Datenmodell

Der Vollständigkeit halber sei an dieser Stelle das zweite Schema für Erhaltungsmetadaten LMER (Langzeitarchivierungsmetadaten für elektronische Ressourcen) erwähnt, entwickelt von der Deutschen Nationalbibliothek (DNB).⁵⁵ Aufgrund einer anhaltenden Weiterentwicklung sowie der Bereitstellung eines geeigneten Datenmodells seitens PREMIS ist erstgenanntes Schema gegenüber LMER zu bevorzugen.

⁵⁴ Caplan, Priscilla. 2009, „PREMIS verstehen.“ Zugriff: 10. Mai 2015.

http://www.loc.gov/standards/premis/understanding_premis_german.pdf. 7.

⁵⁵ DeutscheNationalbibliothek.2014.„DeutscheNationalbibliothek-Langzeitarchivierung.“ Zugriff: 10. Mai 2015. http://www.dnb.de/DE/Standardisierung/LMER/lmer_node.html.

Rechtliche und administrative Metadaten

Administrative und rechtliche Metadaten sind für die Organisation, Verwaltung und Nutzung von Objekten im digitalen Archiv notwendig. Beispielsweise können somit bestimmte vertragsbedingte Nutzungen formuliert werden, etwa nicht kommerzieller Verleih oder auch nur interne Nutzung. Rechtliche/administrative Metadaten sollten systematisch erschlossen werden, damit im Verlauf der Langzeitarchivierung auch nur vereinbarte Maßnahmen bzw. Nutzungen durchgeführt werden. So können z.B. eine redundante Speicherung oder auch Formatmigrationen vertraglich festgesetzt sein. Ansätze für Standardisierungen von administrativen Metadaten lassen sich z.B. bei PREMIS „Rights“ nachweisen.

Strukturelle Metadaten

Die breitgefächerte Funktionalität und die Anforderungen von Metadaten sowie die Vielfalt von existierenden Metadatenstandards wurden dargestellt. Dabei entsteht der nachvollziehbare Wunsch, die relevanten Metadaten und ihre Standards zu strukturieren. Der de facto – da gewissermaßen alternativlose – Standard für die Bereitstellung von strukturellen Metadaten in der dLZA ist der Metadata Encoding and Transmission Standard (METS). Entwickelt von der Library of Congress reichen die Anfänge des METS bis 2001 zurück; eine Weiterentwicklung findet aktuell statt. Eine umfassende Dokumentation des Standards sowie seines XML-Schemas wird auf der METS-Webseite bereitgestellt.⁵⁶

Die Entwicklung von METS erfolgte, um folgende Ziele zu erreichen:

- ▶ Kodierung von Erschließungs-, Verwaltungs- und Strukturangaben
- ▶ Beschreibung der komplexen Bezüge zwischen Metadatenkategorien
- ▶ Austausch von digitalen Objekten zwischen Sammlungen
- ▶ Verknüpfung von Metadatenstandards
- ▶ Flexibilität zwischen Metadatenstandards

Präsentation/Zugriff

Das Langzeitarchiv muss sicherstellen, dass berechtigte Nutzer einen Zugang zu den Informationsobjekten sowie seinen Repräsentationen erhalten. Dabei ist es zunächst unerheblich, ob eine Nutzung z.B. über einen Fernzugriff wie z.B. Webseiten oder VoD-Plattformen gewährleistet wird oder die Objekte nur vor Ort genutzt werden können. Gerade durch die für Filme zu klärenden Rechte sind Online-Nutzungen i.d.R. nur eingeschränkt möglich.

Für die Präsentation und den Zugriff sind weiterhin für die entsprechenden Nutzungsszenarien die entsprechenden Formate zu definieren. Grundsätzlich gelten dafür dieselben Bedingungen wie für Archivformate. So sollten auch beim Zugriff Formate zum Einsatz kommen, die offen, weitverbreitet und robust sind.⁵⁷

Dateiformate und Codecs

Für die Archivierung von digitalen Objekten sind zunächst die Formate zu definieren, welche im Archiv behandelt werden sollen. In einer Übernahmeregulierung werden diese archivwürdigen Formate formuliert und damit die Voraussetzung für ihre Überprüfung geschaffen. Gegebenenfalls müssen Formate migriert werden, wenn sie den Erhaltungsrichtlinien des Archivs widersprechen. Technisch gesehen muss deshalb das Archiv in der Lage sein, Formate zu erkennen, Archivformate zu definieren und diese schließlich zu validieren.

⁵⁶ Library of Congress. 2014. „METS Metadata Encoding & Transmission Standard.“ Zugriff: 10. Mai 2015. <http://www.loc.gov/standards/mets/mets-schemadocs.html>.

⁵⁷ Siehe hierzu auch <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Für die dLZA lassen sich verschiedene Kriterien für archivwürdige Dateiformate definieren, die auf eine Vielzahl von Autoren und Publikationen zurückgehen, die diese Faktoren beschreiben. Dabei lassen sich Abstraktionen zusammenfassen, die in verschiedene Kriterien differenziert werden können⁵⁸: Maßgeblich für ein geeignetes Archivformat ist demnach erstens die **Offenheit** eines Formats, d.h. inwieweit Lizenzierungen davon abhängig sind und ob eine Norm vorliegt. Zweitens wie weit ein Dateiformat **verbreitet** ist; dies ist ein Indiz für die Bereitstellung von unterschiedlicher Software, die unabhängig mit dem Format umgehen kann. Als drittes Kriterium kann differenziert werden, wie **komplex** ein Dateiformat aufgebaut ist. Komplexität erschwert zum einen die Interpretation und zum anderen die Nutzung von Dateiformaten. Vorhandene **Schutzmechanismen** für Dateiformate sind ein Ausschlusskriterium für die dLZA, nur die exakte und nachhaltige Kenntnis über Verschlüsselungsmechanismen kann zur Akzeptanz solcher Formate führen.⁵⁹ Als letzte Kriterien sind **Robustheit** eines Dateiformats, d.h. die Fehleranfälligkeit für die Veränderung einzelner Bits für die Interpretierbarkeit des digitalen Objekts, und die **Abhängigkeiten** von spezieller Hard- oder Software ausschlaggebend.

Die aufgeführten Kriterien werden überwiegend gleichsetzend auch von der Library of Congress (LoC) verwendet, die eine ausführliche Dokumentation von Dateiformaten hinsichtlich der Kriterien auf ihrer Webseite bereitstellt.⁶⁰ Die Dokumentation wird beständig aktualisiert und bietet eine gute Übersicht über jeweilige Dateiformate anhand ihrer definierten Kriterien.

Für AV-Materialien gilt hinsichtlich dieser Kriterien jedoch zu hinterfragen, inwieweit diese für ein vertrauenswürdige Archivformat gelten können. Anders als digitale Text- oder Bildformate sind digitale AV-Formate schwieriger und vielfältiger zu definieren. Oft wird dieser Aspekt in der Forschung der dLZA von AV-Objekten nicht hinreichend beleuchtet.

Vergleicht man zunächst analoge und digitale AV-Materialien, lassen sich zwei Kategorien differenzieren. Zum einen die Kategorie Film, zum anderen die Kategorie Video. In der digitalen Welt ist diese Trennung zwar nicht so eindeutig und belastbar wie im analogen Bereich. In der Projektion handelt es sich auch im digitalen Kino um Videobilder. Dennoch ist es sinnvoll, die Unterteilung zu treffen, da sich digitale AV-Objekte durch eine hohe Anzahl von unterschiedlichen Dateiformaten, Kompressionsmöglichkeiten und Standardisierungen auszeichnen. Die Kategorisierung hilft dabei, unterschiedliche Formatgruppen zu definieren und dabei auch unterschiedliche Konzepte für Archivformate auseinanderzuhalten. Die folgende Einteilung wird dabei die Strukturierung der vorhandenen digitalen Objekte ermöglichen.

Digitaler Film

Digitaler Film zeichnet sich darin aus, dass er für die Projektion in digitalen Kinos bestimmt ist. Merkmale von digitalem Film sind u.a. eine hohe Bild- und Tonauflösung, ein großer Farbraum, Mehrkanal-Ton und eine hohe Bildfrequenz – und auch nicht selten eine Verschlüsselung. Unter einer Verschlüsselung wird ein Digital Rights Management (DRM)-System verstanden. Ein DRM-System verhindert das unerlaubte Kopieren von digitalen Objekten, des Weiteren kann es den wiederholten Zugriff einschränken. Digitaler Film wird zudem professionell produziert, i.d.R. sind dabei Produktionsgesellschaften oder eine hin-

58 Ludwig, Jens. 2010, „Auswahlkriterien“. In: Neuroth, Oßwald, Scheffel, and Strathmann, S., Huth, K., nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, 147-148.

59 Ludwig, Jens. 2010, „Auswahlkriterien“. In: Neuroth, Oßwald, Scheffel, and Strathmann, S., Huth, K., nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, 147-148.

60 Vgl. Library of Congress. 2015a. „Sustainability Factors.“ Zugriff: 17. Juli 2015. <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>.

reichende Finanzierung involviert. Kennzeichnend für digitalen Film ist, dass sich das Informationsobjekt durch Bildsequenzen auszeichnet, die aus jeweils einzelnen Bilddateien bestehen. Dieser Umstand ist herstellungsbedingt, denn digitale Filmkameras generieren Einzelbilder, somit kommen auch in der Projektion des D-Cinema (digitales Kino) Einzelbildsequenzen zum Einsatz. Bei einer typischen Bildfrequenz von 24 Bildern pro Sekunde werden 84.000 Einzelbilder pro Stunde Filmdauer projiziert. Die Audioinformation liegt bei digitalem Film separiert in eigenen Dateien bzw. Dateiformaten vor, dabei können die Audioinformationen verschiedene Synchronfassungen beinhalten.

Ein Konzept für die Speicherung von digitalem Film liegt darin, die separierten Bildsequenzen und Audiodateien in einem Paket zu definieren. Zum einen, um den Transport zu gewährleisten, zum anderen auch für die Distribution, Präsentation und Archivierung. Standardisierungen von Paketen für digitalen Film werden durch die Digital Cinema Initiatives (DCI), einem Dachverband, der sich aus Akteuren der US-amerikanischen Filmindustrie zusammensetzt, bereitgestellt.⁶¹ Die gleichnamige Spezifikation DCI beinhaltet zwei definierte Pakete: das Digital Cinema Package (DCP) und das Digital Cinema Distribution Master (DCDM). Das Hauptziel des DCI-Dachverbandes liegt in der Standardisierung digitaler AV-Objekte für die Distribution in den digitalen Kinos. Wesentliche technische Merkmale der Spezifizierung des DCI-Workflows sind die Komprimierung von Bilddateien durch den JPEG 2000-Codec⁶² und die Verwendung eines MXF-Containers.⁶³

Wenngleich diese Standardisierungen erfolgreich sind und DCPs als Standard für die Projektion von digitalem Film gewissermaßen vollständig etabliert sind, so ist die Frage, wie sich DCI-Pakete für die Archivierung eignen, umstritten.⁶⁴ Zum einen ist bei DCPs nicht auszuschließen, dass eine Verschlüsselung vorliegt, zum anderen verwenden DCPs auch eine verlustbehaftete Komprimierung der Bildinformationen. Das DCDM erfüllt per Definition zwar eine verlustfreie Komprimierung, dient jedoch lediglich zur Erstellung eines DCPs innerhalb des DCI-Distribution-Workflows. Keinesfalls kann also das DCDM als Masterformat eines abgeschlossenen Postproduktionsprozesses angesehen werden.⁶⁵ D.h. die Pakete DCP und DCDM eignen sich nicht zwingend für die Herstellung von anderen Formaten jenseits des DCI-Workflows.

Vergleicht man nun DCP und DCDM mit den Anforderungen, die seitens des OAIS-Modells an ein AIP gestellt werden, so lassen sich die beiden Konzepte von Paketen schlichtweg nicht vergleichen. Weder Metadaten noch Erschließungsinformationen werden im DCI-Workflow OAIS-konform bereitgestellt. Dieser Umstand ist auch nicht verwunderlich, für eine OAIS-konforme Übernahme wurde der DCI-Workflow nicht konzipiert. Die für Archive zunächst verführerische Möglichkeit, diese Pakete als Archivformate zu etablieren, sollte deshalb kritisch hinterfragt werden. In der Deutschen Kinemathek werden derzeit beide Formate des DCI-Workflows übernommen und als Archivformate akzeptiert. Dem widerspricht gegenwärtig auch nicht die LoC, die den DCI-Workflow als ein „middle-state format for archiving“ beschreibt.⁶⁶

61 Zur ausführlichen Beschreibung siehe die Webseite der DCI Digital Cinema Initiatives LLC 2015 sowie zur Spezifikation des DCI Digital Cinema Initiatives LLC [2005] 2012. <http://www.dcmovies.com/>.

62 „JPEG – JPEG 2000.“ 2015. Zugriff: 6. Juli 2015. <http://www.jpeg.org/jpeg2000/>.

63 Ferreira, Pedro. 2002. „MXF - the Material eXchange Format.“

64 Fornaro, Peter und Gubler, David. 2014. „DCP/A: Discussion of an Archival Digital Cinema Package for AV-Media.“

65 Digital Cinema Initiatives LLC. (2005) 2012. „Digital Cinema System Specification v.1.2.“ http://dcmovies.com/specification/DCI_DCSS_v12_with_errata_2012-1010.pdf, 20.

66 Library of Congress. 2011. „Digital Cinema Initiative Distribution Package (DCP), Version 1.0.“ Zugriff: 5. Juli 2015. <http://www.digitalpreservation.gov/formats/fdd/fdd000200.shtml>.

Eine andere Initiative, ein dezidiertes Archivformat für digitalen Film zu etablieren, wurde durch das Enhanced-Digital-Cinema-Projekt (EDCINE) unternommen. Das in den Jahren 2006 bis 2009 mit 16 Millionen Euro großzügig ausgestattete Projekt entwickelte für die Archivreiche die beiden Pakete Master Archive Package (MAP) und Intermediate Access Package (IAP). Beide Pakete beinhalten Spezifikationen des DCI-Distribution-Workflows. Auch hier finden die JPEG2000-Kodierung und MXF-Container Anwendung. Neben der Entwicklung von MAP und IAP ist der zweite Meilenstein des Projekts die Bereitstellung einer Softwarelösung, welche in der Lage ist, die Pakete zu erstellen und beinhaltete Objekte darzustellen. Die Software Curator Archive Suite wird federführend vom Fraunhofer-Institut für Integrierte Schaltungen (IIS) entwickelt und vertrieben.⁶⁷ Es lässt sich fünf Jahre nach Projektende feststellen, dass sowohl Archivpakete als auch Softwarelösungen nicht breit etabliert sind.

In der Deutschen Kinemathek konnte die Software durch das Projekt EFG 1914 getestet werden, als Nachteile erwiesen sich in der Anwendung die zeitaufwändigen Kompressions- und Dekompressionsmethoden, die mangelnde Standardisierung der Pakete, die Abhängigkeit vom Entwickler und dem Fraunhofer-IIS sowie die vergleichsweise hohen Anschaffungskosten. Zudem stellt die Software keine ausreichende Hardwarebeschleunigung seitens geeigneter Videokarten bereit. Als entscheidendes Defizit dient der Umstand, dass weder MAP noch IAP konform mit der OAIS-Definition von AIPs sind. Über die Weiterentwicklung stellt weder EDCINE noch das Fraunhofer-IIS Informationen bereit, die Projektwebseite ist derzeit offline.

Letztlich wird v.a. seit 2014 ein dritter Standard für die Definition von Archivformaten für digitalen Film immer stärker diskutiert. Dabei handelt es sich um ein Paket, das sich wiederum in die Reihe von Akronymen von Paketen einreicht: das sogenannte Interoperable Master Format (IMF). Der Standard wurde erstmalig 2012 durch die Society of Motion Picture & Television Engineers (SMPTE) publiziert und wird seitdem kontinuierlich weiterentwickelt. Genauso wie bei den anderen Standards durch DCI und EDCINE ist auch hier die Industrie beteiligt. Neben der SMPTE arbeiten wiederum die führenden Major US-Studios sowie v.a. die Firma DVS an der Entwicklung des Standards. Dieser Umstand ist mit einer gewissen Skepsis zu sehen, da seitens der Industrie stets wirtschaftliche Gründe eine Rolle spielen. Die wirtschaftliche Abhängigkeit, sei es von Hardware oder Software oder auch Lizenzmodellen, ist für die dLZA stets zu hinterfragen, da eine Nachhaltigkeit nicht gewährleistet werden kann.

Auch der Kostenfaktor ist natürlich immanent. Dies erklärt, dass das IMF zunächst als Business-to-Business-Format konzipiert ist und damit v.a. Unabhängigkeit von Dienstleistern im Postproduction-Workflow von digitalem Film gewährleisten soll. So ist denkbar, dass ein Dienstleister den Schnitt eines Films verantwortet und daraufhin Special Effects (VFX) oder auch Lichtbestimmung (Grading) von anderen Dienstleistern übernommen werden. Im Gegensatz zum EDCINE-Projekt wird hier jedoch eine komplette Standardisierung seitens SMPTE vorausgesetzt, dabei orientiert sich IMF auch an DCI-Standards.⁶⁸

Zusammenfassend lässt sich für digitalen Film darstellen, dass derzeit kein etabliertes Archivformat vorhanden ist. Anhand der einleitend vorgestellten Kriterien treten bei den vorgestellten Archivformaten Schwächen auf, entweder sind die Formate nicht weitreichend verbreitet oder es kommen Schutzmechanismen zum Tragen, darüber hinaus werden die Formate und Pakete v.a. durch einzelne Hersteller oder Gruppen der Industrie bestimmt. Auch wenn Bestrebungen für die Etablierung eines archivwürdigen Standards

67 Fraunhofer IIS. 2011. „CURATOR Archive Suite.“ Zugriff: 2. August 2015.
<http://www.iis.fraunhofer.de/de/pr/2011/curator.html>.

68 Rhode & Schwarz DVS GmbH 2015. „Webinar: IMF – Interoperable Master Format: Entwicklung eines Standards“. Zugriff: 5. Juli 2015.
<https://www.youtube.com/watch?v=RVBVH065IS8&feature=youtu.be>.

deutlich werden, lässt sich derzeit eine Empfehlung unter Berücksichtigung einer verantwortungsvollen dLZA nicht aussprechen.

Vergleicht man die dLZA von digitalem Film mit der von Textdokumenten, wird offensichtlich, dass ein PDF/A-Dokument für digitalen Film derzeit nicht existiert. Die PDF/A-Entwicklung kämpft zwar noch mit Problemen der Validierung, die Etablierung eines einheitlichen Dateiformats ist aber gewährleistet. Ein „Archival Digital Cinema Package“, welches gleichzeitig OASIS-konform ist und zudem funktionale Anforderungen für digitalen Film besitzt, ist demnach noch ein Desiderat.

Eine Alternative zu den dargestellten Ansätzen, digitalen Film in standardisierten Paketen zu archivieren, bietet derzeit lediglich die Speicherung auf nativen, unkomprimierten Formatdefinitionen für Bildsequenzen wie Digital Picture Exchange (DPX) oder Tagged Image File Format (TIFF). Archive, die den o.g. Paketen gegenüber skeptisch eingestellt sind, orientieren sich deshalb an diesen Formaten. Die Nachteile dabei liegen in den vergleichsweise hohen Datenmengen sowie in den fehlenden Möglichkeiten für eine spezifische Referenzierung der unterschiedlichen Objekttypen der intellektuellen Einheit. Untertitel, Tonspuren und andere Objekttypen des Films können demnach nur separiert gespeichert werden, ohne dass der Bezug zueinander eindeutig definiert ist.

Digitales Video

Als zweite Kategorie neben dem digitalen Film lassen sich wie vorgeschlagen AV-Materialien als digitales Video definieren. Digitales Video zeichnet sich zunächst dadurch aus, dass es nicht primär für die Projektion im D-Cinema-Bereich konzipiert ist, dagegen findet es in TV-Produktionen, in der Medienkunst und natürlich im Bereich von privaten Aufzeichnungen Verwendung. I.d.R. liegt digitales Video in einem singulären Dateiformat vor, welches Video- und Audioinformation zusammen speichert; ein dafür weitverbreiteter, etablierter Begriff ist das Containerformat.

Kennzeichen für digitales Video sind niedrigere Auflösungen, einfachere Kopierschutzmechanismen sowie kostengünstigere Workflows verglichen mit den Anforderungen des digitalen Films. Außerdem lässt sich eine große Vielfalt von unterschiedlichen Dateiformaten und Kompressionsstandards nachweisen: der sogenannte Format-Zoo. Gewissermaßen kehren sich bei digitalem Video im Vergleich zu digitalem Film die Gegebenheiten um. Das Problem liegt nicht darin, dass sich nur wenige Initiativen für die Etablierung eines hinreichend vertrauenswürdigen Archivformates aufzeigen lassen, sondern in der Herausforderung, innerhalb der vorhandenen Masse der unterschiedlichen Formate das geeignete Format für die vertrauensvolle dLZA zu bestimmen.

Der Ansatz, ein Archivformat für digitales Video zu definieren, ist im Gegensatz zu digitalem Film jedoch nicht vorwiegend durch die Industrie und einzelne Firmen bestimmt. Gleichwohl werden auch bei digitalem Video Dateiformate von der Industrie durch Microsoft Corporation, Apple Inc. oder Avid Technology Inc. bereitgestellt. Verschiedene Initiativen seitens der Informationswissenschaft, die dem Open-Source-Gedanken folgen, haben dazu beigetragen, Dateiformate und Codecs zu entwickeln, die den Kerngedanken der dLZA folgen. Außerdem werden Technologien und Algorithmen erarbeitet, um jene Dateiformate und Codecs durch eine Transkodierung und Formatwandlung zu migrieren. Zu nennen ist die in diesem Zusammenhang sicherlich bedeutendste Entwickler-Community, FFmpeg.

An diesem eigenständigen FFmpeg-Projekt haben dabei auch andere Open-Source-Entwickler, das Netzwerk nestor und einzelne Gedächtnisinstitutionen Anteil, indem sie diesen Entwicklungsprozess fördern und in Teilen auch finanziell unterstützen. SLUB und die Deutsche Kinemathek planen die Entwicklung von FFV1 mitzutragen (vgl. nestor Media AG, Beschluss vom 12. Juni 2015).

Für digitales Video können die Kriterien für eine vertrauenswürdige dLZA, die einleitend formuliert wurden, hinreichend in Betracht gezogen werden, um den Format-Zoo zu beurteilen. So sind Kriterien wie Robustheit, Offenheit, Schutzmechanismen, Komplexität, Verbreitung und Standardisierung zu bewerten und zu differenzieren. In der derzeitigen Praxis von Filmarchiven lassen sich dabei grundsätzlich zwei unterschiedliche Ansätze unterscheiden:

Zum einen existiert in Archiven, die mit der Digitalisierung und Archivierung des audiovisuellen Erbes betraut sind, der Anspruch, den größtmöglichen Informationsgehalt durch möglichst native Objekte zu erhalten. Dieser Umstand könnte auf die Tradition von analogen (Kopier-)Verfahren in einem Archiv zurückzuführen sein, welche gleichwohl immer mit einem Informationsverlust verbunden sind. Eine Veränderung von Informationsobjekten durch technologische Prozesse ist dem entgegengesetzt. Infolgedessen werden seitens der Filmarchive Technologien, die auf der Migration durch Transkodierung von Objekten beruhen, sehr skeptisch beurteilt. Es werden Verfahren der Transkodierung vermieden und bevorzugt nicht-komprimierte bzw. native Dateien für die Archivierung verwendet.⁶⁹ Vorzugsweise wird deshalb digitales Video bspw. in Codecs wie v210 oder YUY digitalisiert und archiviert. Dabei müssen auch Nachteile in Kauf genommen werden: Einerseits wird das Kriterium der Offenheit vernachlässigt, so kommt z.B. auch deshalb die Gruppe von proprietären Formaten (d.h. Formate, die durch Hersteller lizenziert und nicht veröffentlicht bzw. quelloffen sind) in verschiedenen Institutionen zum Einsatz.⁷⁰ Auch die Deutsche Kinemathek akzeptiert derzeit proprietäre Formate, wenn sie nativ erstellt sind. Andererseits zieht dieses Konzept der nativen Formate eine große Bandbreite von unterschiedlichen Dateiformaten und Dateikodierungen nach sich, die archiviert werden müssen.

Im zweiten Ansatz für die Behandlung von digitalem Video existieren Konzepte, welche die o.g. Maximen bei der Formatwahl stärker berücksichtigen und dabei bestrebt sind, spezifische Dateiformate zu bestimmen und davon abweichende Formate zu migrieren. Vergleichbar sind der Einsatz von TIFF- und JPEG-Formaten bei der dLZA von Bildern und das PDF/A-Format im Bereich der Textdokumente.

Im Wesentlichen stehen derzeit zwei zu beachtende Dateiformate zur Diskussion: Das erste Konzept geht auf die im Abschnitt zum digitalen Film erwähnte DCI- bzw. EDCINE-Konzeption zurück. Infolgedessen wird vorgeschlagen, digitales Video ebenso in einem MXF-Containerformat durch eine JPEG-2000-Kodierung abzuspeichern. Filmarchive, die diesem Konzept folgen, sind u.a. EYE Film Institute Netherlands, La Cinémathèque Royale de Belgique oder auch Danish Film Institute/Det Danske Filminstitut. Sowohl für das Containerformat MXF als auch den Codec JPEG 2000 können die Kriterien, die für ein Dateiformat der dLZA definiert sind, zunächst als ausreichend beschrieben werden. Dateiformat und Codec sind offene, robuste Standards, verbreitet und unabhängig von Kopierschutzmechanismen. Der Codec ist zudem lossless, d.h. verlustfrei spezifiziert. Als Problem von JPEG 2000 ist jedoch die Implementierung in Software kennzeichnend. Der Umgang mit diesem Format setzt – genauso wie bei digitalem Film – spezifische Softwarelösungen voraus. Wie bei digitalem Film kommt darüber hinaus auch das Problem der Performance der JPEG-2000-Kodierung durch spezifische Hard- und Software zum Tragen.

Neben diesem Konzept für ein digitales Video-Archivformat, welches als JPEG 2000 codiert in einem MXF-Dateiformat vorliegt, lässt sich derzeit ein zweiter Ansatz nachweisen, der aktuell immer größere Beachtung und Unterstützung findet. Die österreichische Media-

69 Memoriav. [2015] 2015. „Memoriav Empfehlungen, Digitale Archivierung von Film und Video: Grundlagen und Orientierung“. Zugriff: 5. Juli 2015. http://memoriav.ch/wp-content/uploads/2015/04/Empfehlungen_Digitale_Archivierung_Version1.0.pdf.

70 Murray, Kate. 2015. „All in the (Apple ProRes 422 Video Codec) Family | The Signal: Digital Preservation.“ Zugriff: 21. März 2015. <http://blogs.loc.gov/digitalpreservation/2015/01/all-in-the-apple-prores-422-video-codec-family/>.

thek verfolgt das Konzept des Codecs FFV1 bereits seit 2013 und hat dabei den Entscheidungsprozess, Implementierung sowie Vor- und Nachteile dokumentiert, vgl. Bubestinger 2015. Darüber hinaus findet aktuell eine Debatte durch die nestor Media AG statt, die plant, FFV1 und MKV als zukünftiges Archivformat für digitales Video zu etablieren. Das Konzept sieht vor, digitale Video- und Audiodateien in einem Matroska-Containerformat (MKV)⁷¹ zu speichern und durch den Videocodec FFV1 zu codieren.⁷² Der verwendete Codec FFV1 geht auf den Entwickler Michael Niedermayer zurück, der Codec ist seit 2003 Teil der freien Codec-Sammlung des FFmpeg-Projekts (libavcodec). FFV1 und MKV zeichnen sich primär durch den Open-Source-Gedanken von offenen Spezifikationen aus; Nachhaltigkeit und Weiterentwicklung sind deshalb unabhängig von Lizenzierungen oder einem einzelnen Entwickler gegeben. Die Softwareunterstützung wird durch die Software des FFmpeg-Projekts gewährleistet und bietet dadurch die Möglichkeit zur Implementierung. Vorteile gegenüber dem Codec JPEG 2000 liegen in der besseren Performance⁷³, die durch effizientere Algorithmen gewährleistet ist und die eine verlustfreie Komprimierung ermöglicht. Darüber hinaus ist ein Alleinstellungsmerkmal des FFV1, dass durch Checksummenbildung von einzelnen Bildern der Videodatei die Überprüfung der Dateiintegrität ermöglicht wird. Die Nachteile sind dagegen die fehlende Standardisierung von FFV1, die geringe Verbreitung und derzeit eingeschränkte Implementierung in verschiedene Videoapplikationen⁷⁴ sowie eine Fehleranfälligkeit in der Transkodierung.⁷⁵

Zusammenfassend zeigt sich, dass für die dLZA von digitalem Video zwar verschiedene Ansätze für Archivformate vorhanden sind, ein hinreichendes Konzept, welches für eine Empfehlung notwendig ist, jedoch nicht ausgemacht werden kann. Auf der einen Seite stehen die Archive mit einer abwartenden, zögerlichen Haltung, ein Format zu bestimmen. Die Konsequenz davon sind multiple, dabei jedoch auch native Formate.

Auf der anderen Seite steht der Ansatz, multiple Formate zu vermeiden und konsequent in ein bestimmtes Format zu migrieren. Es ist festzustellen, dass zwischen dem Konzept JPEG 2000 und FFV1 die Vorteile des FFV1-Formates überwiegen. Vielversprechend beim FFV1-Konzept sind die vergleichbar höhere Performance und die Möglichkeit für eine bildgenaue Dateiintegrität von digitalen AV-Objekten. Einen weiteren Vorteil bietet die Unterstützung seitens des nestor-Netzwerkes und einzelner Gedächtnisinstitutionen, die sich derzeit diesem Format verschrieben haben. FFV1 ist demnach ein vielversprechender Ansatz, der sich als Archivformat jedoch zukünftig erst beweisen muss.

Weitere Literatur

Europäisches Parlament und Rat. 2005. „EMPFEHLUNG DES EUROPÄISCHEN PARLAMENTES UND DES RATES vom 16. November 2005 zum Filmerbe und zur Wettbewerbsfähigkeit der einschlägigen Industriezweige.“ Amtsblatt der Europäischen Union (L 323/57). <http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32005H0865&from=DE>. Zugriff: 9. April 2015.

71 Matroska. 2013. „Specifications | Matroska.“ Zugriff: 21. Juli 2015.

<http://www.matroska.org/technical/specs/index.html>.

72 Niedermayer, Michael. 2013. „FFV1 Video Codec Specification.“ Zugriff: 21. Juli 2015.

<http://www.ffmpeg.org/%7Emichael/ffv1.html>.

73 Romeyke, Andreas. 2015. „Kulturreste – Was von uns übrig bleibt...: Verlustfreie Videocodecs in der Langzeitarchivierung – Ein Vergleich.“ Zugriff: 21. Juli 2015.

<http://kulturreste.blogspot.de/2014/12/verlustfreie-videocodecs-in-der.html>.

74 Im Testsystem der Deutschen Kinemathek konnte sowohl mit Avid Media Composer 6 als auch Final Cut 7 der Codec nicht verarbeitet werden.

75 Bubestinger, Peter. 2015. „Comparing video codecs and containers for archives.“ Zugriff: 21. Juli 2015. http://download.das-werkstatt.com/pb/mthk/info/video/comparison_video_codecs_containers.html#codec_tests.

Korb, Nikola und Andreas Roth. 2008. „Persistent Identifier: ...eindeutige Bezeichner für digitale Inhalte...“ Zugriff: 25. Juli 2015. <http://www.persistent-identifier.de/>.

Lange, Andreas. 2013. „Die Gaming-Community als Pionier der digitalen Bewahrung.“ In: Klimpel, Paul und Keiper, Jürgen, Was bleibt? Nachhaltigkeit der Kultur in der digitalen Welt. Berlin: iRights Media

———. 2015b. „Metadata Object Description Schema: MODS (Library of Congress).“ Zugriff: 11. Mai 2015. <http://www.loc.gov/standards/mods/>.

nestor-Arbeitsgruppe OAIS-Übersetzung / Terminologie. 2012. „Referenzmodell für ein Offenes Archiv-Informationssystem: Deutsche Übersetzung.“

Prestocentre. 2014. „Webinar: Creation of Master Archive Package (MAP).“ Zugriff: 5. Juli 2015. <https://www.youtube.com/watch?v=5oMUhw-1GZM#t=1842>.

Archivierung von Fernsehmitschnitten

Empfehlungen für Privatanwender

Auswahl des zu archivierenden Materials

In der dLZA wird mit der Auswahl die begründete Entscheidung getroffen, warum etwas archiviert werden sollte. I.d.R. werden all diejenigen Materialien archiviert, die von kulturellem Wert sind. Viele Gedächtnisorganisationen haben einen gesetzlichen Auftrag, durch den sie verpflichtet sind, bestimmte Materialien dauerhaft zu sichern.

Für den Privatanwender gibt es diese Pflichten nicht, er wird oft auch Materialien sichern, die nicht einzigartig sind oder einen hohen kulturellen Wert haben. I.d.R. hat er ein privates Interesse, etwas langfristig verfügbar zu halten. Sei es sein eigenes Filmarchiv, Teile seines Lebenswerkes, oder vielleicht nur eine Materialsammlung seiner Hobbies.

Auch wenn Privatanwender bei der Auswahl Dritten gegenüber nicht verpflichtet sind, so sollten sie dennoch Notizen hinterlegen, warum sie etwas archiviert haben. Dies ermöglicht es, noch zu Lebzeiten oder in der Beurteilung des Nachlasses festzustellen, ob eine Nutzung (noch) sinnvoll ist oder nicht.

Die Festlegung, warum etwas archiviert werden soll, wird „Archivwürdigkeit“ genannt. Ob das Material auch „archivfähig“ ist, sprich: in einem Format vorliegt, welches guten Wissens und Gewissens als Archivformat genutzt werden kann, wird in den Abschnitten Erhaltungsstrategien und Formate näher beleuchtet.

Beschreibung der vorgesehenen Nutzungsszenarien

Die Archivierung und insbesondere die Langzeitarchivierung sind mit hohen Kosten verbunden. Um sicherzustellen, dass das ausgewählte Material auch langfristig nicht nur gesichert, sondern auch zur Nutzung verfügbar gehalten wird, ist es wichtig, im Auge zu behalten, welche Nutzer des Archivs infrage kommen und welche Anforderungen sich daraus ergeben.

Für den Privatanwender ist es nicht notwendig, diese Nutzerzielgruppen zu definieren, da er und sein Umfeld i.d.R. die einzigen Nutzer des Archivs sein werden.

Auch ein Privatanwender sollte sich Gedanken machen, welche Aspekte des Fernsehmitschnittes für ihn wichtig sind. Soll der Zweikanalton erhalten bleiben, weil man die Sendung später noch einmal in Englisch schauen möchte? Spielt die Audioqualität eine besondere Rolle, da es sich um ein Konzert aus dem Gewandhaus handelt? Oder ist die Bildqualität wichtig, weil man nur dann die Vortragsfolien lesen kann?

Auch ist es sinnvoll, sich Gedanken darüber zu machen, wie man später auf seine Fernsehmitschnitte zugreifen will. Wie findet man diese wieder? Wie möchte man das Archiv nutzen? Es kann nicht schaden, sich auch hierüber ein paar Notizen zu machen.

Auch ein Privatanwender sollte sich Gedanken machen, welche Aspekte des Fernsehmitschnittes für ihn wichtig sind. Soll der Zweikanalton erhalten bleiben, weil man die Sendung später noch einmal in Englisch schauen möchte? Oder ist die Bildqualität wichtig, weil man nur dann die Vortragsfolien lesen kann? Auch sollte man sich Gedanken darüber machen, wie man später auf seine Fernsehmitschnitte zugreifen will.

Qualitätsmanagement

Unter Qualitätsmanagement wird verstanden, dass das zu archivierende Material in einer Form übernommen wird, die technisch fehlerfrei ist, voll den Spezifikationen entspricht und auch inhaltlich mit den Metadaten übereinstimmt. Der Privatanwender wird i.d.R. keinen Aufwand für das Qualitätsmanagement betreiben und zumindest eine grobe Sichtprüfung des Fernsehmittschnittes vornehmen.

Erhaltungsstrategien

Bitstream-Preservation

Basis der Erhaltungsstrategien ist immer die Bitstream-Preservation, d.h. die bitgenaue digitale Kopie des zu archivierenden Materials. Um festzustellen, ob die zu sichernden Materialien korrekt gespeichert sind, bietet es sich an, Prüfsummen zu erstellen (z.B. Sha256), mit denen man feststellen kann, ob der Inhalt verändert ist (Integritätssicherung). Diese sollten direkt nach der Erstellung des Materials (hier: Aufzeichnung des Fernsehmittschnittes) berechnet und nach jedem Kopiervorgang geprüft werden. Es gibt verschiedene Technologien, um Datenströme effizient und möglichst sicher zu speichern.

Nach aktuellem Stand der Technik ist für den Privatanwender nur die Speicherung von Mehrfach-Kopien sinnvoll. Als einfachste, automatische Lösung bieten sich hier Raid-10-Systeme, z.B. in Form von NAS-Boxen, an:

- ▶ Raid-5-Systeme sind nicht zu empfehlen, da bei der Wiederherstellung von ausgefallenen Platten die restliche Hardware übermäßig beansprucht wird.
- ▶ Eine Sicherung auf DVDs und CD-ROMs wird ebenfalls nicht empfohlen, da handelsübliche Medien schnell altern und Endkunden-Hardware keine der speziellen Rohlinge für die Archivierung brennen kann.
- ▶ Im Notfall tut es auch das händische Sichern des Archivs auf mehrere USB-Festplatten oder (verschlüsselt) in Clouds.

Content-Preservation

Um das zu archivierende Material nicht nur zu sichern, sondern auch langfristig nutzbar zu halten, im Falle der Fernsehmittschnitte also abzuspielen und anzuschauen, muss die Darstellbarkeit des Inhalts gesichert werden (Content-Preservation).

Dabei sollte frühzeitig bekannt sein, welche Eigenschaften des Materials über die Zeit unbedingt erhalten bleiben müssen, da diese signifikant für die Nutzung sind. Dies kann in dem einen Fall der hochwertige Mehrkanalton sein, im zweiten die Bildqualität und im dritten die Farbtreue.

Bei den Erhaltungsstrategien unterscheidet man grob zwischen Emulation und Formatmigration. Im ersten Fall erstellt man eine Emulation einer bestimmten Umgebung, die Hardware, Betriebssystem und installierte Software umfasst, die zum Archivierungszeitpunkt in der Lage ist, die gesicherten Materialien wiederzugeben. Sollte ein Videoformat eines Fernsehmittschnittes nicht mehr von aktueller Software unterstützt werden, so kann man dieses über die emulierte Umgebung noch abspielen. Im zweiten Fall versucht man, die zu archivierenden Materialien in einem besonders für die Archivierung geeigneten Format zu speichern. Sollte dann ein Dateiformat veralten, kann rechtzeitig und unter der Beachtung der zu bewahrenden, signifikanten Eigenschaften eine Konvertierung in ein anderes Dateiformat vorgenommen werden.

Für den Privatanwender empfiehlt sich der zweite Ansatz. Eine Empfehlung für archivfähige Formate finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Nach aktuellem Stand der Technik ist für den Privatanwender nur die Speicherung von Mehrfach-Kopien sinnvoll. Als einfachste, automatische Lösung bieten sich hier Raid-10-Systeme, z.B. in Form von NAS-Boxen, an. Im Notfall tut es auch das händische Sichern des Archivs auf mehrere USB-Festplatten oder (verschlüsselt) in Clouds.

Authentizität

Die Sicherung der Authentizität ist i.d.R. keine Anforderung, die an einen Privatanwender gestellt wird. Damit ist gemeint, dass nachgewiesen wird, dass ein gesichertes Objekt auch das ist, was es vorgibt zu sein. Dies ist z.B. bei elektronischen Urkunden wichtig, wo man auch bei zwischenzeitlich veralteten elektronischen Signaturen nachweisen möchte, dass die Urkunde nicht verändert wurde.

Metadaten

Unter Erschließung wird zum einen die Erfassung von filmografischen Metadaten verstanden (Katalogisierung), zum anderen die Inhaltsererschließung.

Die filmografischen Metadaten umfassen i.d.R. die Urheber des Werkes, das Erscheinungsdatum, Herausgeber und Verleger, sonstige Mitwirkende und Besonderheiten wie Aufnahmezeitpunkt, Sender usw.⁷⁶

Die inhaltliche Erschließung von Filmmitschnitten ermöglicht es, in einer größeren Sammlung von Werken nach bestimmten Szenen, gesprochenen Worten, Stimmungen usw. zu suchen. Für den Privatanwender ist die Erschließung mit hohem Aufwand verbunden. Dennoch erleichtert eine leichtgewichtige, standardisierte Metadaten-Datei die spätere Nutzung. Viele Mediendatenbanken erlauben eine komfortable Eingabe. Als ein Beispiel seien die freien Softwarepakete Tellico (<<http://tellico-project.org/>>), DataCrow (<<http://www.datacrow.net/>>) und GCStar (<<http://www.gcstar.org/>>) genannt.

Präsentation/Zugriff

Für die Nutzung der archivierten Medien wird oft auf Vorschauen zurückgegriffen, die das archivierte Material in reduzierter Auflösung, speziell kodiert oder in Sequenzen unterteilt repräsentieren. Dies macht insbesondere dann Sinn, wenn das Archiv über Webplattformen erreichbar sein soll und die Kosten für die Datenübertragung eine große Rolle spielen.

Dateiformate und Codecs

Für den Privatanwender ist die Trennung in Präsentationsformate und Archivformate i.d.R. nicht sinnvoll. Wenn der Zugriff (Access) auf die archivierten Daten (Master) direkt erfolgen soll, ist unbedingt darauf zu achten, dass die Abspielprogramme das Archiv nicht kompromittieren können (z.B. durch read-only-Zugriff).

Im Folgenden werden Anforderungen an ein Archivformat (Dateiformat und Codec) aufgeführt:

- ▶ Standardisierung und Offenheit des Format (notfalls offen spezifiziert)
- ▶ Robustheit (einzelne Bitfehler sind in Wirkung begrenzt)
- ▶ hohe Verbreitung
- ▶ niedrige Komplexität (benötigte Rechenzeit)
- ▶ Selbstbeschreibung (Bin ich ein Matroska-File?)
- ▶ Lizenz- und Patentfreiheit
- ▶ Erkennbarkeit/Validierbarkeit des Formates (Gibt es ein Prüftool?)
- ▶ Verschlüsselungsfreiheit
- ▶ Unabhängigkeit von spezieller Hard-/Software

Für den Privatanwender ist die Erschließung mit hohem Aufwand verbunden. Dennoch erleichtert eine leichtgewichtige, standardisierte Metadaten-Datei die spätere Nutzung. Viele Mediendatenbanken erlauben eine komfortable Eingabe.

⁷⁶ Vgl. hierzu den Metadatenstandard EN 15744 (http://filmstandards.org/fsc/index.php/EN_15744).

Fernsehmitschnitte werden i.d.R. via DVB-T, DVB-S oder DVB-C empfangen. Die Ausstrahlung über die digitalen Fernsehkanäle nach Digital Video Broadcasting (DVB) 3 resultiert oft in Dateien (MPEG-TS), die verschiedene Datenströme für Video, Audio, Teletext, MHP und EPG enthalten. Die Videodaten sind MPEG-2 oder MPEG-4 (in H.264) kodiert, die Audiodaten in MP2 oder AC-3. Da diese MPEG-TS-Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer (<<http://www.mplayerhq.hu/>>) oder FFmpeg (<<https://www.ffmpeg.org/>>) in Matroska-Container (Endung .mkv) zu konvertieren.

Dabei können die Codecs für die Video- und Audiodatenströme von DVB ohne Rekodierung weiterverwendet werden, da o.g. Codecs standardisiert sind und als weitverbreitet angesehen werden können. Dabei sollte man im Hinterkopf behalten, dass diese Codecs verlustbehaftet und für den Videoschnitt eher ungeeignet sind.

Matroska (<<http://www.matroska.org/>>) als Container bietet sich an, weil es sich um ein freies Containerformat für Videodateien handelt, welches Synchronität gewährleistet und eine gute Verbreitung besitzt.

Empfehlungen für kleinere Einrichtungen

Auswahl des zu archivierenden Materials

In der dLZA wird mit der Auswahl die begründete Entscheidung getroffen, warum etwas archiviert werden sollte. I.d.R. werden all diejenigen Materialien archiviert, die von kulturellem Wert sind. Viele Gedächtnisorganisationen haben ein Mandat (z.B. einen gesetzlichen Auftrag), durch den sie verpflichtet sind, bestimmte Materialien dauerhaft zu sichern.

Kleinere Einrichtungen müssen oft abwägen, wie sie ihre sehr knappen Ressourcen einsetzen. Um diese manchmal auch schmerzhaften Entscheidungen transparent zu machen, sollten sie anhand der erwarteten Nutzergruppen (siehe nächster Abschnitt) einen objektiven Kriterienkatalog anlegen.

Mit diesem Kriterienkatalog können sie Dritten gegenüber argumentieren, warum sie etwas archivieren und wieso dies die entstehenden Kosten rechtfertigt. Auch ermöglicht dies festzustellen, ob man und wie man eine notwendige Langzeitarchivierung auslagert, z.B. wenn eine ordentliche Betreuung des Archivs nach aktuellem Stand der Technik oder aufgrund der Qualifikationen der Mitarbeiter nicht möglich ist.

Diese Festlegung, warum etwas archiviert werden soll, wird „Archivwürdigkeit“ genannt. Ob das Material auch „archivfähig“ ist, sprich: in einem Format vorliegt, welches guten Wissens und Gewissens als Archivformat genutzt werden kann, wird in den Abschnitten „Erhaltungsstrategien“, S. 99, und „Datenformate und Codecs“, S. 101, näher beleuchtet.

Beschreibung der vorgesehenen Nutzungsszenarien

Die Archivierung und insbesondere die Langzeitarchivierung sind mit hohen Kosten verbunden. Um sicherzustellen, dass das ausgewählte Material auch langfristig nicht nur gesichert, sondern auch zur Nutzung verfügbar gehalten wird, ist es wichtig, im Auge zu behalten, welche Nutzer des Archivs denn infrage kommen und welche Anforderungen sich daraus ergeben. Wenn man die Anforderungen der verschiedenen Nutzer gruppiert und die für diese jeweils relevanten Aspekte festhält, erreicht man zwei Dinge:

Erstens lernt man seine Nutzergruppen kennen und kann seine Arbeit priorisieren. Beides erleichtert die Zielgruppenanalyse, d.h. die stetige Überprüfung, ob man als Einrichtung auf die Anforderungen seiner Nutzer eingestellt ist und ggf. Korrekturen vornehmen sollte.

Kleinere Einrichtungen müssen abwägen, wie sie ihre Ressourcen einsetzen. Um der Transparenz willen sollten sie anhand der erwarteten Nutzergruppen einen objektiven Kriterienkatalog anlegen. Wenn man die Anforderungen der verschiedenen Nutzer gruppiert und die für diese jeweils relevanten Aspekte festhält, lernt man seine Nutzergruppen kennen und kann seine Arbeit priorisieren. Zudem ergeben sich daraus die signifikanten Eigenschaften, d.h. welche Aspekte des Fernsehmitschnitts im Laufe des Archivlebens in seinen grundlegenden Eigenschaften bewahrt werden müssen.

Zweitens ergeben sich daraus die signifikanten Eigenschaften, d.h. welche Aspekte des Fernsehmitschnittes im Laufe des Archivlebens in seinen grundlegenden Eigenschaften bewahrt werden müssen. Soll der Zweikanalton erhalten bleiben, weil Nutzer die Sendung später auf Englisch und Deutsch schauen wollen? Spielt die Audioqualität eine besondere Rolle, da es sich um ein Konzertmitschnitt aus dem örtlichen Konzerthaus handelt und die Einrichtung Musikstudenten als Nutzer hat? Oder ist die Bildqualität wichtig, weil man nur dann die Vortragsfolien lesen kann?

Aus der Analyse der Nutzerzielgruppen ergibt sich in Teilen auch die Frage des Zugriffs auf das Archiv. Wie und wem (einem Bereitstellungssystem oder dem Kunden direkt) werden die Daten bereitgestellt? Müssen Nutzungsrechte berücksichtigt werden und wo werden diese geprüft?

Qualitätsmanagement

Unter Qualitätsmanagement wird verstanden, dass das zu archivierende Material in einer Form übernommen wird, die technisch fehlerfrei ist, voll den Spezifikationen entspricht und auch inhaltlich mit den Metadaten übereinstimmt.

Kleinere Einrichtungen werden in der Regel kaum Aufwand für das Qualitätsmanagement betreiben können. Sie sollten zumindest eine Sichtprüfung des Fernsehmitschnittes vornehmen und sich nicht nur auf die Ausgabe von Standardsoftware wie Videoschnittprogrammen verlassen.

Erhaltungsstrategien

Bitstream-Preservation

Basis der Erhaltungsstrategien ist immer die Bitstream-Preservation, d.h. die bitgenaue digitale Kopie des zu archivierenden Materials. Um festzustellen, ob die zu sichernden Materialien korrekt gespeichert sind, bietet es sich an, Prüfsummen zu erstellen (z.B. Sha256), mit denen sich feststellen lässt, ob der Inhalt verändert ist (Integritätssicherung). Diese sollten direkt nach der Erstellung des Materials (hier: Aufzeichnung des Fernsehmitschnittes) berechnet und nach jedem Kopiervorgang geprüft werden.

Es gibt verschiedene Technologien, um Datenströme effizient und möglichst sicher zu speichern. Nach aktuellem Stand der Technik ist für kleine Einrichtungen nur die Speicherung von Mehrfach-Kopien sinnvoll. Als einfachste, automatische Lösung bieten sich hier Raid-10-Systeme, z.B. in Form von NAS-Boxen, an die regelmäßig auf Magnetbänder abgezogen werden.

Raid-5-Systeme sind nicht zu empfehlen, da bei der Wiederherstellung von ausgefallenen Platten die restliche Hardware übermäßig beansprucht wird.

Eine Sicherung auf DVDs und CD-ROMs ist generell nicht empfehlenswert, da handelsübliche Medien schnell altern. In Einzelfällen können Spezialaufwerke mit speziellen Rohlingen für DVDs zur Archivierung genutzt werden.

Wenn man die Wahl hat, sollte man auf Bänder mit LTO-Bandlaufwerken sichern. Dies ist günstiger und bei ordentlicher Lagerung die haltbarste Variante.

Content-Preservation

Um aber das zu archivierende Material nicht nur zu sichern, sondern auch langfristig nutzbar zu halten, im Falle der Fernsehmitschnitte also abzuspielen und anzuschauen, muss die Darstellbarkeit des Inhalts gesichert werden (Content-Preservation). Dabei sollte früh-

Es gibt verschiedene Technologien, um Datenströme effizient und möglichst sicher zu speichern. Nach aktuellem Stand der Technik ist für kleine Einrichtungen nur die Speicherung von Mehrfach-Kopien sinnvoll. Als einfachste, automatische Lösung bieten sich hier Raid-10-Systeme, z.B. in Form von NAS-Boxen an. Wenn man die Wahl hat, sollte man auf Bänder mit LTO-Bandlaufwerken sichern. Dies ist günstiger und bei ordentlicher Lagerung die haltbarste Variante.

zeitig bekannt sein, welche Eigenschaften des Materials über die Zeit unbedingt erhalten bleiben müssen, d.h. welche Eigenschaften signifikant für die Nutzung sind. Dies kann in dem einen Fall der hochwertige Mehrkanalton sein, im zweiten die Bildqualität und im dritten die Farbtreue.

Bei den Erhaltungsstrategien unterscheidet man grob zwischen Emulation und Formatmigration. Im ersten Fall erstellt man eine Emulation einer bestimmten Umgebung, die Hardware, Betriebssystem und installierte Software umfasst und die zum Archivierungszeitpunkt in der Lage ist, die gesicherten Materialien wiederzugeben. Sollte ein Videoformat eines Fernsehmitschnittes nicht mehr von aktueller Software unterstützt werden, so kann man diesen über die emulierte Umgebung noch abspielen. Im zweiten Fall versucht man die zu archivierenden Materialien in einem besonders für die Archivierung geeigneten Format zu speichern. Sollte dann ein Dateiformat veralten, kann rechtzeitig und unter Beachtung der zu bewahrenden, signifikanten Eigenschaften eine Konvertierung in ein anderes Dateiformat vorgenommen werden.

Für kleinere Einrichtungen empfiehlt sich meist der zweite Ansatz. Eine Empfehlung für archivfähige Formate finden Sie unter <https://wiki.dnb.de/display/NESTOR/AG+Media>.

Authentizität

Unter Authentizität wird verstanden, dass nachgewiesen wird, dass ein gesichertes Objekt auch das ist, was es vorgibt zu sein. Dies ist z.B. bei elektronischen Urkunden wichtig, wo man auch bei zwischenzeitlich veralteten elektronischen Signaturen nachweisen möchte, dass die Urkunde nicht verändert wurde.

Für Fernsehmitschnitte ist in den Metadaten mitzuführen, welche Sendung wann aufgenommen wurde. Ebenso sind Prüfsummen über die Fernsehmitschnitte zu erstellen. Während der Archivierung sind Änderungen an den gespeicherten Archivdaten (AIP) zu protokollieren und ggf. auszuweisen. Für ein digitales Langzeitarchiv empfiehlt sich, immer die Originaldaten aus dem Ingest im Archiv zu behalten und auf diese mitzuverweisen. Dies gilt auch für den Fall, dass eine Formatmigration stattgefunden hat.

Durch technische und organisatorische Maßnahmen ist sicherzustellen, dass eine Überprüfung der Protokollierung möglich ist.

Metadaten

Unter Erschließung wird zum einen die Erfassung von filmografischen Metadaten verstanden (Katalogisierung), zum anderen die Inhaltserschließung. Die filmografischen Metadaten umfassen i.d.R. die Urheber des Werkes, das Erscheinungsjahr, Herausgeber und Verleger, sonstige Mitwirkende und Besonderheiten.⁷⁷

Die inhaltliche Erschließung von Filmmitschnitten ermöglicht es, in einer größeren Sammlung von Werken nach bestimmten Szenen, gesprochenen Worten, Stimmungen usw. zu suchen. Die Basiskatalogisierung kann dabei selbstständig oder, wenn möglich, aus Beständen Dritter übernommen werden.

Um eine schnellere Verarbeitung in Fällen der Formatmigration oder Metadaten-Updates zu gewährleisten, sollten die filmografischen Metadaten separat zum Mitschnitt (und nicht primär innerhalb des Datenformats des Mitschnitts) in einem allgemein akzeptierten Metadatenstandard gespeichert werden. I.d.R. liegt man mit METS/MODS richtig, welches u.a. als Standardformat im DFG-Viewer eingesetzt wird.

⁷⁷ Vgl. hierzu den Metadatenstandard EN 15744 (http://filmstandards.org/fsc/index.php/EN_15744).

Für Fernsehmitschnitte ist in den Metadaten mitzuführen, welche Sendung wann aufgenommen wurde. Ebenso sind Prüfsummen über die Fernsehmitschnitte zu erstellen. Während der Archivierung sind Änderungen an den gespeicherten Archivdaten (AIP) zu protokollieren und ggf. auszuweisen.

Präsentation/Zugriff

Für die Nutzung der archivierten Medien wird oft auf Vorschauen zurückgegriffen, die das archivierte Material in reduzierter Auflösung, speziell kodiert oder in Sequenzen unterteilt repräsentieren. Dies macht insbesondere dann Sinn, wenn das Archiv über Webplattformen erreichbar sein soll und die Kosten für die Datenübertragung eine große Rolle spielen.

Kleinere Einrichtungen sollten sich bemühen, diese Trennung in Präsentationsformate und Archivformate vorzunehmen.

Wo dies nicht sinnvoll umgesetzt werden kann, sollte der Kompromiss darin bestehen, den im folgenden Abschnitt beschriebenen Anforderungen an ein Archivformat den Vorrang einzuräumen und aus o.g. Gründen erforderliche Qualitätsabstriche abhängig von den definierten Nutzeranforderungen zu bewerten. Der Zugriff auf die archivierten Daten muss so abgesichert sein, dass die Abspielprogramme bzw. Webzugriffe das Archiv nicht kompromittieren können (z.B. durch read-only-Zugriff und Proxies).

Dateiformate und Codecs

Im Folgenden werden Anforderungen an ein Archivformat (Dateiformat und Codec) aufgeführt:

- ▶ Standardisierung und Offenheit des Formats (notfalls offen spezifiziert)
- ▶ Robustheit (einzelne Bitfehler sind in ihrer Wirkung begrenzt)
- ▶ hohe Verbreitung
- ▶ niedrige Komplexität (benötigte Rechenzeit)
- ▶ Selbstbeschreibung (Bin ich ein Matroska-File?)
- ▶ Lizenz- und Patentfreiheit
- ▶ Erkennbarkeit/Validierbarkeit des Formates (Gibt es ein Prüftool?)
- ▶ Verschlüsselungsfreiheit
- ▶ Unabhängigkeit von spezieller Hard-/Software

Fernsehmitschnitte werden i.d.R. via DVB-T, DVB-S oder DVB-C empfangen. Die Ausstrahlung über die digitalen Fernsehkanäle nach Digital Video Broadcasting (DVB) resultiert oft in Dateien (MPEG-TS), die verschiedene Datenströme für Video, Audio, Teletext, MHP und EPG enthalten. Die Videodaten sind MPEG-2 oder MPEG-4 (in H.264) kodiert, die Audiodaten in MP2 oder AC-3. Da diese MPEG-TS-Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer (<<http://www.mplayerhq.hu/>>) oder FFmpeg (<<https://www.ffmpeg.org/>>) in Matroska-Container (Endung .mkv) zu konvertieren.

Dabei können die Codecs für die Video- und Audiodatenströme von DVB ohne Rekodierung weiterverwendet werden, da o.g. Codecs standardisiert sind und als weitverbreitet angesehen werden können. Dabei sollte man im Hinterkopf behalten, dass diese Codecs verlustbehaftet und für den Videoschnitt eher ungeeignet sind.

Matroska (<<http://www.matroska.org/>>) als Container bietet sich an, weil es sich um ein freies Containerformat für Videodateien handelt, welches Synchronität gewährleistet und eine gute Verbreitung besitzt.

Kleinere Einrichtungen sollten sich bemühen, diese Trennung in Präsentationsformate und Archivformate vorzunehmen. Wo dies nicht sinnvoll umgesetzt werden kann, sollte der Kompromiss darin bestehen, den Anforderungen an ein Archivformat den Vorrang einzuräumen. Der Zugriff auf die archivierten Daten muss so abgesichert sein, dass die Abspielprogramme bzw. Webzugriffe das Archiv nicht kompromittieren können. Da diese Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer oder FFmpeg in Matroska-Container zu konvertieren. Dabei können die Codecs für die Video- und Audiodatenströme von DVB ohne Rekodierung weiterverwendet werden.

Empfehlungen für mittlere Einrichtungen

Auswahl des zu archivierenden Materials

In der dLZA wird mit der Auswahl die begründete Entscheidung getroffen, warum etwas archiviert werden sollte. I.d.R. werden all diejenigen Materialien archiviert, die von kulturellem Wert sind. Viele Gedächtnisorganisationen haben ein Mandat (z.B. einen gesetzlichen Auftrag), durch das sie verpflichtet sind, bestimmte Materialien dauerhaft zu sichern.

Auch mittlere Einrichtungen müssen abwägen, wie sie ihre Ressourcen einsetzen. Um Entscheidungen darüber treffen zu können, ob sie inhouse archivieren, externe Dienstleister beauftragen und in welchem Umfang sie ihre Mittel für die digitale Langzeitarchivierung einsetzen können, sollten sie anhand der erwarteten Nutzergruppen (siehe nächster Abschnitt) einen objektiven Kriterienkatalog anlegen.

Mit diesem Kriterienkatalog können sie Dritten gegenüber argumentieren, warum sie etwas archivieren und wieso dies die entstehenden Kosten rechtfertigt. Auch ermöglicht dies festzustellen, ob man und wie man eine notwendige Langzeitarchivierung auslagert, z.B. wenn eine ordentliche Betreuung des Archivs nach aktuellem Stand der Technik oder aufgrund der Qualifikationen der Mitarbeiter nicht möglich ist.

Diese Festlegung, warum etwas archiviert werden soll, wird „Archivwürdigkeit“ genannt. Ob das Material auch „archivfähig“ ist, sprich: in einem Format vorliegt, welches guten Wissens und Gewissens als Archivformat genutzt werden kann, wird in den Abschnitten „Erhaltungsstrategien“, S. 107, und unter <https://wiki.dnb.de/display/NESTOR/AG+Media>., näher beleuchtet.

Beschreibung der vorgesehenen Nutzungsszenarien

Die Archivierung und insbesondere die Langzeitarchivierung sind mit hohen Kosten verbunden. Um sicherzustellen, dass das ausgewählte Material auch langfristig nicht nur gesichert, sondern auch zur Nutzung verfügbar gehalten wird, ist es wichtig, im Auge zu behalten, welche Nutzer des Archivs infrage kommen und welche Anforderungen sich daraus ergeben.

Wenn man die Anforderungen der verschiedenen Nutzer gruppiert und die für diese jeweils relevanten Aspekte festhält, erreicht man zwei Dinge:

Erstens lernt man seine Nutzergruppen kennen und kann seine Arbeit priorisieren. Beides erleichtert die Zielgruppenanalyse, d.h. die stetige Überprüfung, ob man als Einrichtung auf die Anforderungen seiner Nutzer eingestellt ist und ggf. Korrekturen vornehmen sollte.

Zweitens ergeben sich daraus die signifikanten Eigenschaften, d.h. welche Aspekte des Fernsehmitschnittes im Laufe des Archivlebens in seinen grundlegenden Eigenschaften bewahrt werden müssen: Soll der Zweikanalton erhalten bleiben, weil Nutzer die Sendung später in Englisch und Deutsch schauen wollen? Spielt die Audioqualität eine besondere Rolle, da es sich um ein Konzertmitschnitt aus dem örtlichen Konzerthaus handelt und die Einrichtung Musikstudenten als Nutzer hat? Oder ist die Bildqualität wichtig, weil man nur dann die Vortragsfolien lesen kann?

Aus der Analyse der Nutzerzielgruppen ergibt sich in Teilen auch die Frage des Zugriffs auf das Archiv. Wie und wem (einem Bereitstellungssystem oder dem Kunden direkt) werden die Daten bereitgestellt? Müssen Nutzungsrechte berücksichtigt werden und wo werden diese geprüft?

Qualitätsmanagement

Unter Qualitätsmanagement wird verstanden, dass das zu archivierende Material in einer Form übernommen wird, die technisch fehlerfrei ist, voll den Spezifikationen entspricht und auch inhaltlich mit den Metadaten übereinstimmt. Mittlere Einrichtungen sollten sich gerade wegen der knappen Personalressourcen frühzeitig Gedanken über eine möglichst weitgehende Automatisierung machen. Händisches Qualitätsmanagement, z.B. Sichtprüfung, sollte auf das erforderliche Mindestmaß beschränkt bleiben. Umso wichtiger ist es für diese Einrichtungen, auf Einhaltung von Standards zu achten und auf eine weitgehende Normierung ihres zu archivierenden Materials hinzuwirken.

Erhaltungsstrategien

Bitstream-Preservation

Basis der Erhaltungsstrategien ist immer die Bitstream-Preservation, d.h. die bitgenaue digitale Kopie des zu archivierenden Materials. Um festzustellen, ob die zu sichernden Materialien korrekt gespeichert sind, bietet es sich an, Prüfsummen zu erstellen (z.B. Sha256), mit denen sich feststellen lässt, ob der Inhalt verändert ist (Integritätssicherung). Diese sollten direkt nach der Erstellung des Materials (hier: Aufzeichnung des Fernsehmitschnittes) berechnet und nach jedem Kopiervorgang geprüft werden.

Es gibt verschiedene Technologien, um Datenströme effizient und möglichst sicher zu speichern. Nach aktuellem Stand der Technik ist für mittlere Einrichtungen nur die Speicherung von Mehrfachkopien sinnvoll. Aus Kostengründen sollte man Kopien über Bandbibliotheken auf Bänder mit LTO-Bandlaufwerken sichern. Dies ist günstiger als die Speicherung auf Festplatten und bei ordentlicher Lagerung die haltbarste Variante. Eine Sicherung auf DVDs und CD-ROMs ist generell nicht empfehlenswert, da handelsübliche Medien schnell altern. In Einzelfällen können Speziallaufwerke mit speziellen Rohlingen für DVDs zur Archivierung genutzt werden.

Content-Preservation

Um das zu archivierende Material nicht nur zu sichern, sondern auch langfristig nutzbar zu halten, im Falle der Fernsehmitschnitte also abzuspielen und anzuschauen, muss die Darstellbarkeit des Inhalts gesichert werden (Content-Preservation). Dabei sollte frühzeitig bekannt sein, welche Eigenschaften des Materials über die Zeit unbedingt erhalten bleiben müssen, d.h. welche Eigenschaften signifikant für die Nutzung sind. Dies kann in dem einen Fall der hochwertigen Mehrkanalton sein, im zweiten die Bildqualität und im dritten die Farbtreue.

Bei den Erhaltungsstrategien unterscheidet man grob zwischen Emulation und Formatmigration. Im ersten Fall erstellt man eine Emulation einer bestimmten Umgebung, die Hardware, Betriebssystem und installierte Software umfasst und die zum Archivierungszeitpunkt in der Lage ist, die gesicherten Materialien wiederzugeben. Sollte ein VideofORMAT eines Fernsehmitschnittes nicht mehr von aktueller Software unterstützt werden, so kann man diesen über die emulierte Umgebung noch abspielen. Im zweiten Fall versucht man die zu archivierenden Materialien in einem besonders für die Archivierung geeigneten Format zu speichern. Sollte dann ein Dateiformat veralten, kann rechtzeitig und unter der Beachtung der zu bewahrenden, signifikanten Eigenschaften eine Konvertierung in ein anderes Dateiformat vorgenommen werden. Für mittlere Einrichtungen empfiehlt sich der zweite Ansatz.

Um den Migrationsaufwand in der Zukunft gering zu halten, sollte man bei Aufnahme des Videomaterials bereits eine Normierung in ein geeignetes Archivformat vornehmen. Eine Empfehlung für archivfähige Formate wird unter <https://wiki.dnb.de/display/NESTOR/AG+Media> gegeben.

Mittlere Einrichtungen sollten sich frühzeitig Gedanken über eine möglichst weitgehende Automatisierung machen. Händisches Qualitätsmanagement sollte auf ein Mindestmaß beschränkt bleiben. Umso wichtiger ist es, Standards einzuhalten und auf eine weitgehende Normierung des zu archivierenden Materials hinzuwirken.

Nach aktuellem Stand der Technik ist für mittlere Einrichtungen nur die Speicherung von Mehrfachkopien sinnvoll. Aus Kostengründen sollte man Kopien auf Bänder mit LTO-Bandlaufwerken sichern. Eine Sicherung auf DVDs und CD-ROMs ist generell nicht empfehlenswert, da handelsübliche Medien schnell altern. In Einzelfällen können Speziallaufwerke mit speziellen Rohlingen für DVDs zur Archivierung genutzt werden.

Um den Migrationsaufwand in der Zukunft gering zu halten, sollte man bei Aufnahme des Videomaterials bereits eine Normierung in ein geeignetes Archivformat vornehmen.

Für Fernsehmitschnitte ist in den Metadaten mitzuführen, welche Sendung wann aufgenommen wurde. Über die Mitschnitte sind Prüfsummen zu erstellen. Während der Archivierung sind Änderungen an den gespeicherten AIP zu protokollieren und ggf. auszuweisen.

Die Basiskatalogisierung kann dabei selbstständig oder aus Beständen Dritter übernommen werden. Die filmografischen Metadaten sollten separat in einem allgemein akzeptierten Metadatenstandard gespeichert werden.

Das Präsentationssystem sollte vom Archiv entkoppelt werden, wodurch die Aufwände für Wartungsarbeiten minimiert werden und Lastspitzen nicht auf das Archiv durchschlagen. Der Zugriff auf die archivierten Daten muss so abgesichert sein, dass Abspielprogramme / Webzugriffe das Archiv nicht stören können.

Authentizität

Unter Authentizität wird verstanden, dass nachgewiesen wird, dass ein gesichertes Objekt auch das ist, was es vorgibt zu sein. Dies ist z.B. bei elektronischen Urkunden wichtig, wo man auch bei zwischenzeitlich veralteten elektronischen Signaturen nachweisen möchte, dass die Urkunde nicht verändert wurde. Für Fernsehmitschnitte ist in den Metadaten mitzuführen, welche Sendung wann aufgenommen wurde. Ebenso sind Prüfsummen über die Fernsehmitschnitte zu erstellen. Während der Archivierung sind Änderungen an den gespeicherten Archivdaten (AIP) zu protokollieren und ggf. auszuweisen.

Für ein digitales Langzeitarchiv empfiehlt es sich, immer die Originaldaten aus dem Ingest im Archiv zu behalten und auf diese mitzuverweisen. Dies gilt auch für den Fall, dass eine Formatmigration stattgefunden hat. Durch technische und organisatorische Maßnahmen ist sicherzustellen, dass eine Überprüfung der Protokollierung möglich ist.

Metadaten

Unter Erschließung wird zum einen die Erfassung von filmografischen Metadaten verstanden (Katalogisierung), zum anderen die Inhaltserschließung. Die filmografischen Metadaten umfassen i.d.R. die Urheber des Werkes, das Erscheinungsjahr, Herausgeber und Verleger, sonstige Mitwirkende und Besonderheiten sowie die Kodierung der Rechte und Rechteinhaber.⁷⁸ Die inhaltliche Erschließung von Filmmitschnitten ermöglicht es, in einer größeren Sammlung von Werken nach bestimmten Szenen, gesprochenen Worten, Stimmungen usw. zu suchen. Die Basiskatalogisierung kann dabei selbstständig oder, wenn möglich, aus Beständen Dritter übernommen werden. Um eine schnellere Verarbeitung in Fällen der Formatmigration oder Metadaten-Updates zu gewährleisten, sollten die filmografischen Metadaten separat zum Mitschnitt (und nicht primär innerhalb des Datenformats des Mitschnitts) in einem allgemein akzeptierten Metadatenstandard gespeichert werden. I.d.R. liegt man mit METS/MODS richtig, welches u.a. als Standardformat im DFG-Viewer eingesetzt wird.

Präsentation/Zugriff

Für die Nutzung der archivierten Medien wird oft auf Vorschauen zurückgegriffen, die das archivierte Material in reduzierter Auflösung, speziell kodiert oder in Sequenzen unterteilt repräsentieren. Dies ergibt insbesondere dann Sinn, wenn das Archiv über Webplattformen erreichbar sein soll und die Kosten für die Datenübertragung eine große Rolle spielen. Mittlere Einrichtungen sollten sich bemühen, das eigentliche Präsentationssystem vom Langzeitarchiv zu entkoppeln. Die Vorteile liegen darin, dass Aufwände für Wartungsarbeiten minimiert werden und Lastspitzen nicht auf das Archiv durchschlagen.

Wo dies nicht sinnvoll umgesetzt werden kann, sollte der Kompromiss darin bestehen, den im folgenden Abschnitt „Dateiformate und Codecs“ beschriebenen Anforderungen an ein Archivformat den Vorrang einzuräumen und aus o.g. Gründen erforderliche Qualitätsabstriche abhängig von den definierten Nutzeranforderungen zu bewerten. Der Zugriff auf die archivierten Daten muss so abgesichert sein, dass die Abspielprogramme bzw. Webzugriffe das Archiv nicht kompromittieren können (z.B. durch read-only-Zugriff und Proxies).

⁷⁸ Vgl. hierzu den Metadatenstandard EN 15744 (http://filmstandards.org/fsc/index.php/EN_15744).

Dateiformate und Codecs

Im Folgenden werden Anforderungen an ein Archivformat (Dateiformat und Codec) aufgeführt:

- ▶ Standardisierung und Offenheit des Formats (notfalls offen spezifiziert)
- ▶ Robustheit (einzelne Bitfehler sind in ihrer Wirkung begrenzt)
- ▶ hohe Verbreitung
- ▶ niedrige Komplexität (benötigte Rechenzeit)
- ▶ Selbstbeschreibung (Bin ich ein Matroska-File?)
- ▶ Lizenz- und Patentfreiheit
- ▶ Erkennbarkeit/Validierbarkeit des Formates (Gibt es ein Prüftool?)
- ▶ Verschlüsselungsfreiheit
- ▶ Unabhängigkeit von spezieller Hard-/Software

Fernsehmitschnitte werden i.d.R. via DVB-T, DVB-S oder DVB-C empfangen. Die Ausstrahlung über die digitalen Fernsehkanäle nach Digital Video Broadcasting (DVB) 3 resultiert oft in Dateien (MPEG-TS), die verschiedene Datenströme für Video, Audio, Teletext, MHP und EPG enthalten. Die Videodaten sind MPEG-2 oder MPEG-4 (in H.264) kodiert, die Audiodaten in MP2 oder AC-3. Da diese MPEG-TS-Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer (<<http://www.mplayerhq.hu/>>) oder FFmpeg (<<https://www.ffmpeg.org/>>) in Matroska-Container (Endung .mkv) zu konvertieren.

Dabei können die Codecs für die Video- und Audiodatenströme von DVB ohne Rekodierung weiterverwendet werden, da o.g. Codecs standardisiert sind und als weitverbreitet angesehen werden können. Dabei sollte man im Hinterkopf behalten, dass diese Codecs verlustbehaftet und für den Videoschnitt eher ungeeignet sind.

Matroska (<<http://www.matroska.org/>>) als Container bietet sich an, weil es sich um ein freies Containerformat für Videodateien handelt, welches Synchronität gewährleistet und eine gute Verbreitung besitzt.

Da diese MPEG-TS-Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer oder FFmpeg in Matroska-Container zu konvertieren. Dabei können die Codecs für die Video- und Audiodatenströme von DVB ohne Rekodierung weiterverwendet werden. Allerdings sind diese Codecs verlustbehaftet und für den Videoschnitt eher ungeeignet. Matroska bietet sich als Container an, weil es sich um ein freies Containerformat für Videodateien handelt, das Synchronität gewährleistet und eine gute Verbreitung besitzt.

Große Einrichtungen sollten die Anforderungen an digitale Langzeitarchive, wie sie sich in den DIN-Normen DIN 31645 „Leitfaden zur Übernahme in digitale Langzeitarchive“ und DIN 31644 „Kriterien für vertrauenswürdige digitale Langzeitarchive“ niederschlagen, berücksichtigen und anhand der erwarteten Nutzergruppen einen objektiven Kriterienkatalog anlegen. Mit diesem kann Dritten gegenüber argumentiert werden, warum etwas archiviert wird und warum die entstehenden Kosten und Aufwände gerechtfertigt sind.

Wenn man die Anforderungen der verschiedenen Nutzer gruppiert und die jeweils relevanten Aspekte festhält, lernt man seine Nutzergruppen kennen und kann seine Arbeit priorisieren. Zudem ergeben sich daraus die signifikanten Eigenschaften, d.h. welche Aspekte des Fernsehmitschnittes im Laufe des Archivlebens in seinen grundlegenden Eigenschaften bewahrt werden müssen.

Empfehlungen für größere Einrichtungen

Auswahl des zu archivierenden Materials

In der dLZA wird mit der Auswahl die begründete Entscheidung getroffen, warum etwas archiviert werden sollte. I.d.R. werden all diejenigen Materialien archiviert, die von kulturellem Wert sind. Viele Gedächtnisorganisationen haben ein Mandat (z.B. einen gesetzlichen Auftrag), durch das sie verpflichtet sind, bestimmte Materialien dauerhaft zu sichern.

Große Einrichtungen sollten die Anforderungen an digitale Langzeitarchive, wie sie sich in den DIN Normen DIN 31645 „Leitfaden zur Übernahme in digitale Langzeitarchive“⁷⁹ und DIN 31644 „Kriterien für vertrauenswürdige digitale Langzeitarchive“⁸⁰ niederschlagen, berücksichtigen und daher anhand der erwarteten Nutzergruppen⁸¹ einen objektiven Kriterienkatalog anlegen. Mit diesem Kriterienkatalog können sie Dritten gegenüber argumentieren, warum sie etwas archivieren und wieso dies die entstehenden Kosten und Aufwände rechtfertigt. Auch ermöglicht dies festzustellen, ob man und wie man eine notwendige Langzeitarchivierung auslagert, z.B. wenn eine ordentliche Betreuung des Archivs nach aktuellem Stand der Technik oder aufgrund der Qualifikationen der Mitarbeiter nicht möglich ist.

Diese Festlegung, warum etwas archiviert werden soll, wird „Archivwürdigkeit“ genannt. Ob das Material auch „archivfähig“ ist, sprich: in einem Format vorliegt, welches guten Wissens und Gewissens als Archivformat genutzt werden kann, wird in den Abschnitten „Erhaltungsstrategien“ und „Dateiformate und Codecs“, siehe S. 107 und S. 110, näher beleuchtet.

Beschreibung der vorgesehenen Nutzungsszenarien

Die Archivierung und insbesondere die Langzeitarchivierung sind mit hohen Kosten verbunden. Um sicherzustellen, dass das ausgewählte Material auch langfristig nicht nur gesichert, sondern auch zur Nutzung verfügbar gehalten wird, ist es wichtig, im Auge zu behalten, welche Nutzer des Archivs infrage kommen und welche Anforderungen sich daraus ergeben.

Wenn man die Anforderungen der verschiedenen Nutzer gruppiert und die für diese jeweils relevanten Aspekte festhält, erreicht man zwei Dinge. Erstens lernt man seine Nutzergruppen kennen und kann seine Arbeit priorisieren. Beides erleichtert die Zielgruppenanalyse, d.h. die stetige Überprüfung, ob man als Einrichtung auf die Anforderungen seiner Nutzer eingestellt ist und ggf. Korrekturen vornehmen sollte.

Zweitens ergeben sich daraus die signifikanten Eigenschaften, d.h. welche Aspekte des Fernsehmitschnittes im Laufe des Archivlebens in seinen grundlegenden Eigenschaften bewahrt werden müssen: Soll der Zweikanalton erhalten bleiben, weil Nutzer die Sendung später in Englisch und Deutsch schauen wollen? Spielt die Audioqualität eine besondere Rolle, da es sich um ein Konzertmitschnitt aus dem örtlichen Konzerthaus handelt und die Einrichtung Musikstudenten als Nutzer hat? Oder ist die Bildqualität wichtig, weil man nur dann die Vortragsfolien lesen kann?

⁷⁹ <http://www.nabd.din.de/cmd?JSESSIONID=4A8BSBBA23TB19QRARKT0L8H.2&level=tpl-art-detailansicht&committeeid=54738855&artid=145158117&languageid=de&bcrumblevel=1&subcommitteeid=112656549>

⁸⁰ <http://www.nabd.din.de/cmd?JSESSIONID=4A8BSBBA23TB19QRARKT0L8H.2&level=tpl-art-detailansicht&committeeid=54738855&artid=147058907&languageid=de&bcrumblevel=1&subcommitteeid=112656173>

⁸¹ s. nächster Abschnitt „Auswahl der vorgesehenen Nutzergruppen“

Aus der Analyse der Nutzerzielgruppen ergibt sich in Teilen auch die Frage des Zugriffs auf das Archiv. Wie und wem (einem Bereitstellungssystem oder dem Kunden direkt) werden die Daten bereitgestellt? Müssen Nutzungsrechte berücksichtigt werden und wo werden diese geprüft? Ebenso sollte schriftlich fixiert werden, was unter „Original“ verstanden wird und warum.

Qualitätsmanagement

Unter Qualitätsmanagement wird verstanden, dass das zu archivierende Material in einer Form übernommen wird, die technisch fehlerfrei ist, voll den Spezifikationen entspricht und auch inhaltlich mit den Metadaten übereinstimmt. Große Einrichtungen sollten sich frühzeitig Gedanken über eine möglichst weitgehende Automatisierung machen. Händisches Qualitätsmanagement, z.B. Sichtprüfung, sollte auf das erforderliche Mindestmaß beschränkt bleiben.

Auf Einhaltung von Standards zu achten, Werkzeuge zur Validierung voranzutreiben und auf eine weitgehende Normierung des zu archivierenden Materials hinzuarbeiten sollten zu den Aufgaben des LZA-Teams gehören und organisatorisch vorbereitet sein.

Erhaltungsstrategien

Bitstream-Preservation

Basis der Erhaltungsstrategien ist immer die Bitstream-Preservation, d.h. die bitgenaue digitale Kopie des zu archivierenden Materials. Um festzustellen, ob die zu sichernden Materialien korrekt gespeichert sind, bietet es sich an, Prüfsummen zu erstellen (z.B. Sha256), mit denen sich feststellen lässt, ob der Inhalt verändert ist (Integritätssicherung). Diese sollten direkt nach der Erstellung des Materials (hier: Aufzeichnung des Fernsehmittschnittes) berechnet und nach jedem Kopiervorgang geprüft werden. Die Prüfsummen sollten bei Video auf Frame-Basis, bei Audio in festgelegten Blöcken berechnet werden. Dies erleichtert nicht nur die Überprüfung von Fehlern bei Kopiervorgängen, sondern erlaubt bei Verwendung von verlustfreien Codecs auch den automatisierten Vergleich mit dem Original.

Es gibt verschiedene Technologien, um Datenströme effizient und möglichst sicher zu speichern. Nach aktuellem Stand der Technik ist für große Einrichtungen die Speicherung von Mehrfachkopien an mehreren Standorten verpflichtend. Aus Kostengründen sollte man Kopien über Bandbibliotheken auf Bänder mit LTO-Bandlaufwerken sichern. Dies ist günstiger als die Speicherung auf Festplatten und bei ordentlicher Lagerung die haltbarste Variante.

Sinnvoll ist es, auf den Speicherpfaden zu den jeweiligen Kopien unterschiedliche Technologien zu benutzen, um Treiber- und Hardwareprobleme als single-point-of-failure auszuschließen (Sprich, wenn identische Hardware verwendet wird, können diese auch zeitnah mit gleichen Problemen ausfallen).

Content-Preservation

Um das zu archivierende Material nicht nur zu sichern, sondern auch langfristig nutzbar zu halten, im Falle der Fernsehmittschnitte also abzuspielen und anzuschauen, muss die Darstellbarkeit des Inhalts gesichert werden (Content-Preservation). Dabei sollte frühzeitig bekannt sein, welche Eigenschaften des Materials über die Zeit unbedingt erhalten bleiben müssen, d.h. welche Eigenschaften signifikant für die Nutzung sind. Dies kann in dem einen Fall der hochwertige Mehrkanalton sein, im zweiten die Bildqualität und im dritten die Farbtreue. Bei den Erhaltungsstrategien unterscheidet man grob zwischen Emulation und Formatmigration.

Große Einrichtungen sollten sich frühzeitig Gedanken über eine möglichst weitgehende Automatisierung machen. Händisches Qualitätsmanagement sollte auf das erforderliche Mindestmaß beschränkt bleiben. Auf Einhaltung von Standards zu achten, Werkzeuge zur Validierung voranzutreiben und auf eine weitgehende Normierung des zu archivierenden Materials hinzuarbeiten sollten zu den Aufgaben des LZA-Teams gehören und organisatorisch vorbereitet sein.

Nach aktuellem Stand der Technik ist für große Einrichtungen die Speicherung von Mehrfachkopien an mehreren Standorten verpflichtend. Aus Kostengründen sollte man Kopien auf Bänder mit LTO-Bandlaufwerken sichern. Dies ist günstiger als die Speicherung auf Festplatten und bei ordentlicher Lagerung die haltbarste Variante. Sinnvoll ist es, auf den Speicherpfaden unterschiedliche Technologien zu benutzen, um Treiber- und Hardwareprobleme als single-point-of-failure auszuschließen.

Für große Einrichtungen empfiehlt sich bisher v.a. die Speicherung der zu archivierenden Materialien in einem besonders für die Archivierung geeigneten Format. Um den Migrationsaufwand in der Zukunft gering zu halten, sollte man bei Aufnahme des Videomaterials bereits eine Normierung in ein geeignetes Archivformat vornehmen.

Die Basiskatalogisierung kann selbstständig oder aus Beständen Dritter übernommen werden. Um eine schnellere Verarbeitung in Fällen der Formatmigration oder Metadaten-Updates zu gewährleisten, sollten die filmografischen Metadaten separat zum Mitschnitt (und nicht primär innerhalb des Datenformats des Mitschnitts) in einem allgemein akzeptierten Metadatenstandard gespeichert werden. I.d.R. liegt man mit METS/MODS richtig, welches u.a. als Standardformat im DFG-Viewer eingesetzt wird.

Im ersten Fall erstellt man eine Emulation einer bestimmten Umgebung, die Hardware, Betriebssystem und installierte Software umfasst und die zum Archivierungszeitpunkt in der Lage ist, die gesicherten Materialien wiederzugeben. Sollte ein Videoformat eines Fernsehmitschnittes nicht mehr von aktueller Software unterstützt werden, so kann man diesen über die emulierte Umgebung noch abspielen.

Im zweiten Fall versucht man die zu archivierenden Materialien in einem besonders für die Archivierung geeigneten Format zu speichern. Sollte dann ein Dateiformat veralten, kann rechtzeitig und unter Beachtung der zu bewahrenden, signifikanten Eigenschaften eine Konvertierung in ein anderes Dateiformat vorgenommen werden.

Für große Einrichtungen empfiehlt sich bisher v.a. der zweite Ansatz. Um den Migrationsaufwand in der Zukunft gering zu halten, sollte man bei Aufnahme des Videomaterials bereits eine Normierung in ein geeignetes Archivformat vornehmen. Eine Empfehlung für archivfähige Formate wird im Abschnitt „Dateiformate und Codecs“, S. 110, gegeben.

Authentizität

Unter Authentizität wird verstanden, dass nachgewiesen wird, dass ein gesichertes Objekt auch das ist, was es vorgibt zu sein. Dies ist z.B. wichtig bei elektronischen Urkunden, wo man auch bei zwischenzeitlich veralteten elektronischen Signaturen nachweisen möchte, dass die Urkunde nicht verändert wurde. Für Fernsehmitschnitte ist in den Metadaten mitzuführen, welche Sendung wann aufgenommen wurde. Ebenso sind Prüfsummen über die Fernsehmitschnitte auf Frame- bzw. Block-Basis zu erstellen. Während der Archivierung sind Änderungen an den gespeicherten Archivdaten (AIP-Pakete) zu protokollieren und ggf. auszuweisen.

Für ein digitales Langzeitarchiv empfiehlt es sich, immer die Originaldaten aus dem Ingest im Archiv zu behalten und auf diese mit zu verweisen. Dies gilt auch für den Fall, dass eine Formatmigration stattgefunden hat. Durch technische und organisatorische Maßnahmen ist sicherzustellen, dass eine Überprüfung der Protokollierung möglich ist.

Metadaten

Unter Erschließung wird zum einen die Erfassung von filmografischen Metadaten verstanden (Katalogisierung), zum anderen die Inhaltserschließung. Die filmografischen Metadaten umfassen in der Regel die Urheber des Werkes, das Erscheinungsjahr, Herausgeber und Verleger, sonstige Mitwirkende und Besonderheiten sowie die Kodierung der Rechte und Rechteinhaber.

Die inhaltliche Erschließung von Filmmitschnitten ermöglicht es, in einer größeren Sammlung von Werken nach bestimmten Szenen, gesprochenen Worten, Stimmungen usw. zu suchen.

Die Basiskatalogisierung kann dabei selbstständig oder, wenn möglich, aus Beständen Dritter übernommen werden.

Um eine schnellere Verarbeitung in Fällen der Formatmigration oder Metadaten-Updates zu gewährleisten, sollten die filmografischen Metadaten separat zum Mitschnitt (und nicht primär innerhalb des Datenformats des Mitschnitts) in einem allgemein akzeptierten Metadatenstandard gespeichert werden. I.d.R. liegt man mit METS/MODS richtig, welches u.a. als Standardformat im DFG-Viewer eingesetzt wird.

Präsentation/Zugriff

Für die Nutzung der archivierten Medien wird oft auf Vorschauen zurückgegriffen, die das archivierte Material in reduzierter Auflösung, speziell kodiert oder in Sequenzen unterteilt, repräsentieren. Dies macht insbesondere dann Sinn, wenn das Archiv über Webplattformen erreichbar sein soll und die Kosten für die Datenübertragung eine große Rolle spielen.

Größere Einrichtungen sollten sich bemühen, das eigentliche Präsentationssystem vom Langzeitarchiv zu entkoppeln. Die Vorteile liegen darin, dass Aufwände für Wartungsarbeiten minimiert werden und Lastspitzen nicht auf das Archiv durchschlagen.

Dateiformate und Codecs

Im Folgenden werden Anforderungen an ein Archivformat (Dateiformat und Codec) aufgeführt:

- ▶ Standardisierung und Offenheit des Formats (notfalls offen spezifiziert)
- ▶ Robustheit (einzelne Bitfehler sind in ihrer Wirkung begrenzt)
- ▶ hohe Verbreitung
- ▶ niedrige Komplexität (benötigte Rechenzeit)
- ▶ Selbstbeschreibung (Bin ich ein Matroska-File?)
- ▶ Lizenz- und Patentfreiheit
- ▶ Erkennbarkeit/Validierbarkeit des Formates (Gibt es ein Prüftool?)
- ▶ Verschlüsselungsfreiheit
- ▶ Unabhängigkeit von spezieller Hard-/Software

Fernsehmitschnitte werden i.d.R. via DVB-T, DVB-S oder DVB-C empfangen. Die Ausstrahlung über die digitalen Fernsehkanäle nach Digital Video Broadcasting (DVB) 3 resultiert oft in Dateien (MPEG-TS), die verschiedene Datenströme für Video, Audio, Teletext, MHP und EPG enthalten. Die Videodaten sind MPEG-2 oder MPEG-4 (in H.264) kodiert, die Audiodaten in MP2 oder AC-3. Da diese MPEG-TS-Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer (<<http://www.mplayerhq.hu/>>) oder FFmpeg (<<https://www.ffmpeg.org/>>) in Matroska-Container (Endung .mkv) zu konvertieren.

Dabei können die Codecs für die Video- und Audiodatenströme von DVB ohne Rekodierung weiterverwendet werden, da o.g. Codecs standardisiert sind und als weitverbreitet angesehen werden können. Dabei sollte man im Hinterkopf behalten, dass diese Codecs verlustbehaftet und für den Videoschnitt eher ungeeignet sind.

Matroska (<<http://www.matroska.org/>>) als Container bietet sich an, weil es sich um ein freies Containerformat für Videodateien handelt, welches Synchronität gewährleistet und eine gute Verbreitung besitzt.

Größere Einrichtungen sollten das eigentliche Präsentationssystem vom Langzeitarchiv entkoppeln. Dadurch können Aufwände für Wartungsarbeiten minimiert werden und Lastspitzen nicht auf das Archiv durchschlagen.

Da MPEG-TS-Daten oftmals mehrere Audio- und Videodaten von unterschiedlichen Sendern enthalten, bietet es sich an, diese mithilfe von Werkzeugen wie mplayer oder FFmpeg in Matroska-Container zu konvertieren. Dabei können die Codecs für die Video- und Audio-datenströme von DVB ohne Rekodierung weiterverwendet werden. Dabei sollte man im Hinterkopf behalten, dass diese Codecs verlustbehaftet und für den Videoschnitt eher ungeeignet sind.

Matroska als Container bietet sich an, weil es sich um ein freies Containerformat für Videodateien handelt, welches Synchronität gewährleistet und eine gute Verbreitung besitzt.

Tools

Diese Übersicht gibt einen Überblick über Tools, die von den AG-Mitgliedern genutzt werden. Der Fokus liegt hierbei auf Tools, die für die digitale Langzeitarchivierung audiovisueller Medien genutzt werden; darüber hinaus ist auch Software verzeichnet, die generell in der digitalen Bestandserhaltung zum Einsatz kommt, z.B. digitale Langzeitarchivierungssysteme.

Name	Funktion	Anwendungsfall	Weiterentwicklung
Archi- vemata	Digitales Langzeitarchivierungssystem		x
Bagit	Ingest	„Strukturieren von SIPs mit Prüfsummen“	x
DIMAG	Digitales Langzeitarchivierungssystem	Entwicklungspartnerschaftsmodell (Zielgruppe: öffentlich-rechtliche Archive)	x
DROID	Ingest Re-Ingest	Dateiformatidentifikation	x
FFmpeg	Migration	Migration und Qualitätskontrolle der Migration (anhand von Checksums pro Frame siehe: http://dericed.com/papers/reconsidering-the-checksum-for-audiovisual-preservation/)	x
IngestList	Ingest Re-Ingest	Formatidentifikation, Validation (implementiert: DROID/PRONOM, JHOVE, MediaInfo)	x
mediainfo	Ingest Re-Ingest	Extraktion technischer Metadaten	x
media- conch	Ingest	Validator Video FFV1/Matroska	x
Rosetta	Digitales Langzeitarchivierungssystem		x

Ein ausführliches Verzeichnis, das objektartübergreifend Tools für die digitale Langzeitarchivierung verzeichnet, ist unter COPTR (http://coptr.digipres.org/Main_Page) zu finden.

Lizenz	Dokumentation	Bezugsquelle	Nutzung durch AG Mitglieder
x	x	https://www.archivematica.org/en/	ZIB
x	x	https://github.com/LibraryOfCongress/bagger/releases/tag/v2.5-RC1	SLUB
x	x	http://www.landesarchiv-bw.de	LABW
x	x	http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/ https://github.com/digital-preservation/droid	LABW,TIB, SLUB
x	x	http://ffmpeg.org/	LABW,TIB, SDK, SLUB
x	x	http://sourceforge.net/projects/ingestlist/	LABW
x	x	https://mediarea.net/de/MediaInfo	LABW,TIB, SDK, SLUB
	x	https://mediarea.net/MediaConch/	SLUB
	x	Ex Libris	"TIB SLUB"

Digitalisierungsempfehlungen

Analoge Filmformate

Material-format	Materi-alart	Auflösung	Framerate	Interlacing	Image Format
8mm	Negativ	2k	24,25	p	DPX, DNG, TIFF
8mm	Positiv	1080, 2K	24,25	p	DPX, DNG, TIFF
16mm	Negativ	2k-4k	24,25	p	DPX, DNG, TIFF
16mm	Positiv	2K	24,25	p	DPX, DNG, TIFF
35mm	Negativ	4K-8K	24	p	DPX, DNG, TIFF
35mm	Positiv	2K-4K	24	p	DPX, DNG, TIFF

Analoge Bandformate (Video)

Material-format	Materi-alart	Auflösung	Framerate	Interlacing	Container
Beta SP	Video	576	25	i	MOV, AVI, MXF, MKV
video 8 / Hi 8	Video	576	25	i	MOV, AVI, MXF, MKV
DV/mini DV	Video	576	25	i	MOV, AVI, MXF, MKV
VHS, S-VHS.VHS-C	Video	576	25	i	MOV, AVI, MXF, MKV
DVCPRO	Video	576	25	i	MOV, AVI, MXF, MKV
Umatic	Video	576	25	i	MOV, AVI, MXF, MKV
1" A,B,C	Video	576	25	i	MOV, AVI, MXF, MKV
VCR/Video 2000	Video	576	25	i	MOV, AVI, MXF, MKV

Quantisierung/ Bit	Farbraum	Subsampling	lin/log
10-16	RGB	444	log
10-16	RGB	444	log
10-16	RGB	444	log
10-16	RGB	444	log
10-16	RGB	444	log
10-16	RGB	444	log

Codec	Quantisierung/ Bit	Farbraum	Subsampling	lin/log
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin
FFV1, v210, UYVY,YUY2	8	YUV = Y'CBCR	422	lin

Digitale Bandformate (Video)

Material-format	Materialart	Auflösung	Framerate	Interlacing	Container
HDCAM SR	Video	1080	24,25	i,p	MOV, AVI, MXF, MKV
digital Betacam	Video	576	25	i	MOV, AVI, MXF, MKV
HDCAM SR +	Video	1080	24,25	i	MOV, AVI, MXF, MKV
IMX	Video	576	25	i	MOV, AVI, MXF, MKV
HDCAM	Video	1080,72	24,25	i,p	MOV, AVI, MXF, MKV

Digitale Discformate (Video)

Material-format	Anwendung	Auflösung	DRM-System	Framerate	Interlacing
BluRay	Display-format	1080	Möglich	24,25	i,p
DVD Video	Display-format	576	Möglich	25	i

Weitergehende Informationen

Vogel, Andreas (Hg.), Digitalisierungsfibel: Leitfaden für audiovisuelle Archive.
 Siehe auch <http://indi.langzeitarchivierung.de/allg/detail.php?show=4437>

Codec	Quantisierung/Bit	Farbraum	Farbraum/Grading	Subsampling	lin/log
FFV1, v210, UYVY,YUY2	10	YUV, RGB	Rec. 709, RGB	422,444	lin
FFV1, v210, UYVY,YUY2	10	YUV	Rec. 601	422	lin
FFV1, v210, UYVY,YUY2	10	YUV, RGB	Rec. 709, RGB	422,444	lin
FFV1, v210, UYVY,YUY2	8	YUV	Rec. 709	422	lin
FFV1, v210, UYVY,YUY2	8	YUV	Rec. 709	422	lin

Container	Codec	Quantisierung/Bit	Farbraum	Subsampling	lin/log
MOV, AVI, MXF, MKV	MPEG2, H264	8	YUV	420	lin
	MPEG2	8	YUV	420	lin

Glossar

Adlib	Software für Sammlungs- und Informationsmanagement in Museen und Archiven
Agent-Relationship	Beziehung zwischen Agenten in Datenmodell CEN TC 372
AIP	Archival Information Package/Archivinformationspaket
API	application programming interface
Archivwürdigkeit	Begründung, warum ein Objekt langzeitarchiviert werden sollte. Dabei spielt die Bewertung des kulturellen Nutzens eine wichtige Rolle.
Archivfähigkeit	Technische Feststellung, ob ein zu archivierendes Objekt in einer Form vorliegt, die es ermöglicht, das Objekt langzeit zu archivieren.
AV	audiovisuell
Avid	Avid Technology Inc. Ist eine US-amerikanische Firma, die Software für die professionelle Medienproduktion bereitstellt.
Avid Media Composer 6	ein Produkt der Avid Technology Inc, siehe Avid
Backup	Archivierung (Informationstechnik)
basic certification	Zertifizierungslevel für die Überprüfung der dLZA in Institutionen. (Selbstevaluierung)
BFI	British Institute of Film
Bitstream	eine zusammenhängende, logische Folge von binären Ziffern auf bestimmten Datenträgern bzw. Systemen
Bitstream-Preservation	physischer Erhalt digitaler Daten
CCSDS	Consultative Committee for Space Data Systems
CDWA	Categories for the Description of Works of Art
CEN	Comité Européen de Normalisation
CEN TC 372	Standard zur Erschließung deskriptiver Metadaten
Checksummen	Prüfsummen
Codec	Verfahren zur (De-)Kodierung digitaler Dateien

Content-Information	Inhaltsinformation
Content-Preservation	Erhaltung der logischen und konzeptuellen digitalen Objekte
core element	Kernelemente von Metadaten, vgl. DCMES
CRC	cyclic redundancy check, zyklische Redundanzprüfung
creator	Urheber
CRL	Center for Research Libraries
Curator Archive Suite	Software der Firma Fraunhofer IIS
Data Dictionary	Datenkatalog von PREMIS
Data Management	Datenverwaltung
DCDM	Digital Cinema Distribution Master
DCI	Digital Cinema Initiatives, Dachverband der US-amerikanischen Filmindustrie, Spezifikation
D-Cinema	Digital cinema = digitales Kino
DCMES	Dublin Core Metadata Element Set, auch reduziert verwendet unter dem Begriff Dublin Core
DCMI	Dublin Core Metadata Initiative
DCP	Digital Cinema Package
DFG	Deutsche Forschungsgemeinschaft
digital born	digitale Materialien ohne analoge Vorlage
Digital Policies	Richtlinien für die Erhaltung digitaler Objekte, Synonym für Digital Preservation Policies
DIN	Deutsches Institut für Normung e.V.
DIP	Dissemination Information Package/ Auslieferungsinformationpaket
dLZA	digitale Langzeitarchivierung
DNB	Deutsche Nationalbibliothek
DPX	Digital Picture Exchange

DRM	Digital Rights Management
DROID	Digital Record Object Identification
DSA	Data Seal of Approval
DSM	Digital Source Master
Digital Video Broadcasting	Übertragungsstandards für digitales Fernsehen, < http://de.wikipedia.org/wiki/Digital_Video_Broadcasting >
DVD	Digitale Versatile Disc
DVS	Rohde & Schwarz DVS GmbH ist ein Hersteller für Soft- und Hardware für die professionelle D-Cinema-Produktion
EAD	Encoded Archival Description
EBUCore	Metadaten schemata für technische Metadaten von AV-Objekten
EDCINE	Enhanced Digital CINema, EU-Forschungsprojekt für digitale Archivierung
embedded Metadata	in der Datei selbst vorliegende Metadaten
Emulation	Nachbildung und Archivierung der originalen Umgebung (Soft- und Hardware) einer digitalen Datei
entity	Entität im PREMIS Datenmodell sowie im CEN TC 372 Datenmodell
EPG	electronic program guide, „elektronische Programmzeitschrift“
event	Entität im PREMIS Datenmodell sowie im CEN TC 372 Datenmodell
EXIF	Exchangeable Image File Format
FFmpeg	Programm zur Transkodierung und Formatidentifikation von AV Dateien
FFV1	verlustfreier intra-frame Videocodec
FIAF	Fédération Internationale des Archives du Film
Framework	Regelwerk
Fraunhofer-IIS	Fraunhofer-Institut für Integrierte Schaltungen
FRBR	Functional Requirements for Bibliographic Records
FTP	File Transfer Protocol, Netzwerkprotokoll zur Übernahme digitaler Daten
GOP	Group of Pictures

Grading	Lichtbestimmung
HD Cam (SR)	Digitales Videoformat mit einer bandbasierten Speicherung
IAP	Intermediate Access Package
IMF	Interoperable Master Format
Ingest	Übernahme (digitaler) Daten in ein Archiv
IPTC	Standard zur Speicherung von Metadaten in Bilddateien
ISO	International Organization for Standardization
IT	Informationstechnik
items	physische Objekte
JHOVE	JSTOR/Harvard Object Validation Environment
JPEG 2000	ISO-Standard für die verlustbehaftete und verlustfreie Kompression von Bilddateien
JSTOR	Journal STOrage
KOST	Koordinationsstelle für die dauerhafte Archivierung von elektronischen Unterlagen
libavcodec	Codec-Sammlung
LoC	Library of Congress
lossless	verlustfrei
LTFS	Linear Tape File Section
LTO	Linear Tape Open, auf linearer Magnetbandtechnologie beruhendes Speicherverfahren
manifestation	neue Ausprägung, z.B. nach Restaurierung
MAP	Master Archive Package
mapping, (Data-)	Mapping bedeutet die Zuordnung von Informationen zwischen Metadatenschemata
Matroska	Videocontainer, < http://www.matroska.org/ >

MD5	Message-Digest Algorithm 5, kryptografische Hashfunktion zur Bildung von Prüfsummen, um die Authentizität einer archivierten Datei zu gewährleisten
Media-Asset-Management-System	Datenverwaltungssystem insbesondere für Medienobjekte wie AV-Dateien
Metadaten	strukturierte Daten über Daten
METS	Metadata Encoding and Transmission Standard
MHP	
Migration	Überführung einer digitalen Datei in ein archivierungsfähiges Format
MKV	Matroska-Containerformat
MODS	Metadata Objects Description Schema
MPEG	
MPEG-TS	
MP-2	
MXF	Material eXchange Format
NAS	Network Attached Storage
OAIS	Open Archival Information System
OCLC	Online Computer Library Center, Inc.
ÖNB	Österreichische Nationalbibliothek
PB	Petabyte
PBCore	Metadatenschemata für technische Metadaten von AV-Objekten
PCM	Puls-Code-Modulation, langzeittauglicher Audiocodec, der u.a. im unkomprimierten Wave-Format Verwendung findet und auch in den Matroska-Container implementiert werden kann
PDF/A	Portable Document Format/A = Etabliertes Format für die dLZA von Textdokumenten, standardisiert und offen
PUID	Persistent Unique Identifier
Postproduction	Arbeitsschritte für die Nachbearbeitung in Film- und Fernsehproduktionen
PREFORMA	Preservation Formats for culture information/e-archives

PREMIS	Preservation Metadata: Implementation Strategies
Preservation description information	Erhaltungsmetadaten
Preservation Strategies	Erhaltungsstrategien in der dLZA, vgl. Migration oder Emulation
QS	Qualitätssicherung
RAID	Redundant Array of Independent Disks, Methode zur Steigerung des Durchsatzes oder Ausfallsicherheit von Festplattenverbänden, < http://de.wikipedia.org/wiki/RAID >
Refreshment	Austausch von gleichwertigen Datenträgern
Repackaging	eine Methode der Content-Preservation durch eine strukturelle oder inhaltliche Veränderung von Informationspaketen
Replication	Austausch von Datenträgern, Anpassung der Speicherinfrastruktur
retrieval	stichwortbasierte Abfrage
RLG	Research Libraries Group
SIP	Submission Information Package/Übergabeinformationspaket
SLUB	Sächsische Landesbibliothek, Staats- und Universitätsbibliothek Dresden
SMPTE	Society of Motion Picture & Television Engineers
Storage	allgemeiner Begriff für Speichermöglichkeiten in der IT, darüber hinaus eine Funktionseinheit im OAIS Modell
Tape Library	Bandbibliothek
TDR	Trusted Digital Repositories (ISO 16363)
TIB	Technische Informationsbibliothek Hannover
TIFF	Tagged Image File Format. Archivformat für die dLZA von Bilddateien
Transformation	eine Methode der Content-Preservation durch z.B. Formattransformation
UUID	Universally Unique Identifier
v210	Unkomprimierter Videocodec
YUV	Farbmodell im analogen Farbfernsehen
ZIB	Zuse Institute Berlin, Konrad-Zuse-Zentrum für Informationstechnik Berlin

Autoren

Volkmar Ernst, Deutsche Kinemathek – Museum für Film und Fernsehen (Digitalisierung; Archivierung von professionellen Filmen; Tools; Übersicht über Dateiformate, Codecs und Formate)

Jürgen Keiper, Deutsche Kinemathek – Museum für Film und Fernsehen (Farbmanagement; Rechte; Katalogisierungs- und Erschließungsstandards für audiovisuelle Medien; Identifikatoren im audiovisuellen Bereich)

Johannes Renz, Landesarchiv Baden-Württemberg (Archivierung von Amateurfilmen)

Andreas Romeyke, Saxon State- and University Library Dresden (Archivierung von Fernsehmitschnitten; Übersicht über Dateiformate, Codecs und Formate; Tools)

Thomas Bähr und Franziska Schwab, Technische Informationsbibliothek Hannover (Archivierung von Wissenschaftlichen Filmen, Codecs und Formate)

Impressum

Diese Publikation wurde gemeinsam erarbeitet von der nestor AG Media.

Dr. Sebastian Barteleit

Thomas Bähr

Volkmar Ernst

Tim Hasler

Jürgen Keiper

Wolfram Lippert

Johannes Renz

Dr. Stefan Rohde-Enslin

Andreas Romeyke

Franziska Schwab

Uta Simmons

Redaktion: Tim Hasler, Jürgen Keiper

Lektorat des Manuskripts: Astrid Treusch

Layout: Anne Bauknecht

Stand: Oktober 2016

