



Themenblätter zur Archivierung digitaler Daten - Erhaltungsstrategien 2 - *Texte, Bilder, Zeichnungen, Audio/Video*

Subject/Thema (Schlagworte):

Archivierung digitaler Daten, Erhaltungsplanung, archivfähige Datenformate

Description/Beschreibung (Abstract):

Erhaltungsstrategien beschreiben, wie das Archivmaterial am besten erhalten werden kann. Sie ermöglichen eine Prognose für die dauerhafte Bewahrung von Archivgut und definieren dessen Archivfähigkeit. Im vorliegenden Themenblatt werden visuelle Materialtypen betrachtet (z.B. Texte, Bilder, Zeichnungen, auch Audio/Video), für die man die weit verbreiteten und gut akzeptierten formatbasierten Erhaltungsstrategien verwenden kann.

1 Einführung

Erhaltungsstrategien lassen sich abstrakt gliedern in Erhaltungsstrategien für einfache visuelle und für funktionelle Materialien, für systematische Informationen und für komplexe Materialien (s. dazu Themenblatt Erhaltungsstrategien 1 – Grundlagen). Dieses Themenblatt betrachtet **einfache visuelle Materialtypen**, für die man die weit verbreiteten und gut akzeptierten **formatbasierten Erhaltungsstrategien** verwenden kann. Dabei werden die entsprechenden Dateiformate so gewählt, dass sie den Kriterien einer nachhaltigen Nutzbarkeit gut entsprechen. Bei der Erhaltung stehen visuelle Kriterien im Vordergrund. Es wird eine **visuelle Repräsentation** gespeichert.

2 Texte

Texte sind maschinengeschriebene Berichte und Informationen in Textform. Formatierungen, Abbildungen und Zeichnungen sind erlaubt (formatierte Texte). Sie werden typischerweise mit einem Texteditor oder einem Office-Schreibprogramm wie z.B. MS-Word, Open-Office-Writer erstellt. Neben den formatierten Texten gibt es einfache Texte, die nur aus Schriftzeichen bestehen, ohne Formatierungen, Abbildungen usw.

Empfohlene Formate zur Archivierung sind:

- **PDF/A** (ISO 19005-1, Level A = Accessible empfohlen, Level B möglich) für formatierte Texte.
- **TXT** (UTF-8 codiert) für einfache Texte,

Allgemeine Hinweise:

- zu PDF/A:
 - PDF/A ist für die Archivierung optimiert (keine externen Referenzen, keine Verschlüsselung, ausreichende Farbdefinition, keine Skripte, keine Audios oder Videos eingebettet, aber Schriftfonts müssen enthalten sein, ...).
 - „Normale“ PDFs werden in der Regel nicht als archivfähig angesehen.



- Der bevorzugte Level A enthält strukturelle Informationen zum Text (Gliederung, Absätze, ...) und unterstützt damit eine weitere Bearbeitung. Level A lässt sich in der Regel nur aus dem Erfassungs-Original (z.B. MS-Office oder Libre-Office) erzeugen.
- Level B bietet dagegen (nur) eine visuelle Reproduzierbarkeit des Original-Dokuments.
- Als Zeichenkodierung für reine Texte (das betrifft TXT ebenso wie CSV, HTML, XML, GML usw.) sollte vorzugsweise UTF-8 benutzt werden, weil die Kompatibilität mit Viewern und Editoren gegenwärtig am größten ist und dabei eine korrekte Darstellung von Sonderzeichen gewährleistet wird. Im Internet-Bereich ist UTF-8 der vorherrschende Standard. Unter Windows muss beim Speichern von TXT-Dateien UTF-8 extra eingestellt werden (Standard ist ANSI, das für die Archivierung schlecht geeignet ist).

Darüber hinaus kann akzeptiert werden:

- ODT (ODF-Standard, Libre-/Open Office)
- DOCX (OOXML-Standard, MS-Office ab Version 2013)
- HTML (Version 3.2, 4.01).

Außerdem können CSV-Formate und einfache XML-Dialekte akzeptiert werden (sofern sie nicht als Listen, Tabellen oder strukturierte Daten zu betrachten sind → s. dazu separates Themenblatt).

3 Bilder

Bei (Raster-) Bildern werden möglichst einfache und weit verbreitete Formate angestrebt, RAW-Formate bieten sich aufgrund ihrer exklusiven Spezifik nicht zur Archivierung an. Für entzerrte, georeferenzierte Bilder gelten aufgrund der Georeferenzierung andere Vorgaben.

Empfohlene Formate zur Archivierung sind:

- **TIFF unkomprimiert** (TIFF 6.0 Baseline-Format)

Darüber hinaus kann akzeptiert werden:

- **JPEG** (wenn Komprimierungsverluste akzeptabel sind, als Kamera-Original)
- **TIFF G4-komprimiert** (Gruppe 4, verlustfrei nach ITU/CCITT), nur für Schwarz-Weiß-Bilder
- **TIFF LZW-komprimiert**

Allgemeine Hinweise:

- Das Archivformat sollte möglichst direkt aus dem Original erzeugt werden, da jede Umformatierung (bei verlustbehafteter Komprimierung ggf. sogar jede Speicherung) Abweichungen vom Original erzeugt.
- Es sollen keine Mehrseitenbilder, keine Ebenen und keine integrierten Vorschaubilder verwendet werden.
- Der Wertebereich soll 24...48 Bit bei Farbbildern bzw. 8...16 Bit bei Grauwertbildern betragen.
- Die Auflösung muss entsprechend dem Inhalt gewählt sein. Als Faustregel kann gelten, dass die kleinste zu erkennende Struktur über etwa drei Pixel reichen soll.



4 Zeichnungen

Zeichnungen sind Vektorgrafiken. Sofern sie nicht in formatierten Texten enthalten sind, werden sie als separate Dateien gespeichert. Die zumeist üblichen proprietären Formate sind zur Archivierung nicht geeignet. Für Geodaten (CAD/GIS, ggf. auch Pläne und Kartierungen) gelten aufgrund der Georeferenzierung andere Vorgaben.

Empfohlene Formate zur Archivierung sind:

- **SVG** (Scalable Vector Graphics, Version 1.0, 1.1, 1.2)

Darüber hinaus kann akzeptiert werden (sofern die Empfehlung nicht umsetzbar):

- **PDF/A** (Zeichnung innerhalb eines „formatierten Textes“)
- **ggf. DXF** (Zeichnung als „CAD“)

Allgemeine Hinweise:

- Als Zeichnung (Vektorgrafik) eignen sich Grafiken, die (auch) mit einem Stift ausgeführt werden können. Typische Zeichnungselemente sind Punkte, Linienzüge und durch geschlossene Linienzüge (Polygone) gebildete Flächen.
- Besondere Eigenschaften einer Zeichnung sind die objektbezogene Bearbeitung und die verlustfreie Skalierbarkeit. Bei einer Konvertierung in ein Rasterbild gehen diese Eigenschaften verloren.

5 Audio- und Videodateien

Audio- und Videodateien (Ton- und Filmaufnahmen) sind mit visuellen Materialien vergleichbar. Ihr Zweck ist, vom Menschen wahrgenommen zu werden (hören, sehen). Die maschinelle Verarbeitung (Funktionalität) ist dabei nur Mittel zum Zweck. Das Ziel der Archivierung ist es, beim Nutzer die bei der Aufnahme beabsichtigte Wahrnehmung zu ermöglichen. Die dazu erforderliche Qualität ist sorgfältig abzustimmen (es ist nicht unbedingt das technische Maximum erforderlich).

Empfohlene Formate zur Archivierung sind:

- Die verfügbaren Audio- und Video-Datenformate eignen sich nur bedingt zur Archivierung, daher wird hier diesbezüglich keine allgemeine Empfehlung ausgesprochen.

Zur Archivierung kann akzeptiert werden:

- **WAVE** (Audio) mit linearer PCM
- **MP3** (Audio) verlustbehaftet komprimiert
- **MPEG-4** (Video) ohne DRM

Allgemeine Hinweise:

- Das verlustbehaftete MP3 sollte nur archiviert werden, wenn es originär so generiert wurde.
- Eine Archivierung von Audio/Video sollte nur erfolgen, wenn Tools zur aktiven Bearbeitung und entsprechende Kompetenzen verfügbar sind.



6 Übersicht

Materialtyp	Bearbeitung	Software	Bearbeitungs- und Aufbewahrungs-Formate	Archiv-Format
Texte	unformatiert editieren	Win-Editor	TXT (UTF-8-codiert)	TXT
	Textverarbeitung (formatiert editieren)	MS Office Word (>2013), LibreOffice bzw. Open Office Writer,	DOCX (OOXML), ODT (ODF), ggf. PDF/A	PDF/A
	PDF/A-Konverter	Adobe Acrobat	PDF/A	
Bilder	Aufnahme	Kamera, ggf. Scanner	RAW, DNG, ggf. JPEG	baseline TIFF
	Entwickeln	Kamera-spezifische Software	TIFF, ggf. JPEG	
	Bearbeitung	Adobe Photoshop, GIMP	TIFF, ggf. JPEG	
Zeichnungen	Erfassung, Bearbeitung	MS Office Word, LibreOffice bzw. Open Office Draw	(DOCX,) ODG (ODF), SVG	SVG
Audio/Video	Aufnahme, Schnitt	diverse	diverse	WAVE, MP3, MPEG-4

Source/Quellen:

- Archivierungs-Ratgeber der VLA-Kommission Archäologie und Informationssysteme:
http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Archaeologie-Informationssysteme/Dokumente_AIS_Archivierung/Ratgeber_Archivierung_V1.0.pdf
- ADS, Preferred and Accepted File Formats:
<http://archaeologydataservice.ac.uk/advice/FileFormatTable>
- DANS, preferred formats:
<https://dans.knaw.nl/en/Deposit/information-about-depositing-data/DANSpreferredformatsUK.pdf>
- IANUS-Forschungsdatenzentrum, IT-Empfehlungen, Dateiformate:
<http://www.ianus-fdz.de/it-empfehlungen/dateiformate>
- Schweizer Bundesarchiv, Archivtaugliche Formate:
https://www.bar.admin.ch/dam/bar/de/dokumente/konzepte_und_weisungen/archivtaugliche_dateiformate.1.pdf.download.pdf/archivtaugliche_dateiformate.pdf
- The National Archives, File Formats:
<https://www.nationalarchives.gov.uk/information-management/manage-information/selection-and-transfer/digital-records-transfer/file-formats-transfer>



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International Lizenz (CC-BY-SA). Dies ist eine [Free-Culture-Lizenz](https://creativecommons.org/licenses/by-sa/4.0/).

