

Themenblätter zur Archivierung digitaler Daten - Erhaltungsstrategien 3 - *Listen und 3D/VR-Daten*

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 optimistische Prognose für die dauerhafte Bewahrung von Archivgut und bestimmen dessen Archivfähigkeit. Im vorliegenden Themenblatt werden einfache funktionelle Materialtypen betrachtet (z.B. Tabellen/Listen, 3D-Daten), für die man die weit verbreiteten und gut akzeptierten formatbasierten Erhaltungsstrategien, ergänzt um eine visuelle Repräsentation, 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 funktionelle Materialtypen** (z.B. Tabellen/Listen, einfache Geodaten). Für diese Materialtypen wird die **formatbasierte Erhaltungsstrategie erweitert**. Neben der Auswahl eines nachhaltigen Datenformats zur **funktionellen Repräsentation** werden zusätzlich separate **visuelle Repräsentationen** erstellt.

2 Listen, Tabellen, strukturierte Daten

Tabellen/Listen enthalten Einträge (Zellen), die in Zeilen und ggf. in Spalten angeordnet sind und oft bestimmten Formaten (z.B. Text, Zahl, Datum, ...) unterliegen. Die Zuordnung der Einträge zu den Zeilen und Spalten ist signifikant. Einträge können einerseits Werte sein. Andererseits können (Berechnungs-) Funktionen als Einträge verwendet werden, die durch die Berechnung zu Ergebnissen führen. Normalerweise werden die Ergebnisse signifikant sein, aber auch die zugrundeliegenden Funktionen können interessant sein. Eine maschinelle Auswertung ist aufgrund der strengen Formatierung interessant, bei Nutzung von Funktionen ist sie erforderlich.

Empfohlene Formate zur Archivierung sind:

- **CSV UTF 8** (für Werte, Ergebnisse) + Strukturbeschreibung/Metadaten
- **CSV UTF 8** (für Funktionen, Formeln, Berechnungen, ...) + Strukturbeschreibung/Metadaten
- Vorschlag für CSV-Spezifikation:
 - 1. Zeile enthält Spaltennamen,
 - Trennzeichen Tab (alternativ im Datenbankbereich: |),
 - Dezimaltrenner Komma (alternativ für englischen Sprachraum: Punkt),
 - Zahlen ohne Tausender-Trennzeichen,
 - Textqualifizierer „“ (alternativ: ohne),
 - Textinhalte ohne „“, Zeilenumbruch, | bzw. Tab ... !



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

- **XML** UTF 8+ Strukturbeschreibung/Metadaten

erforderliche Strukturbeschreibung/Metadaten:

- Spaltennamen und Datentypen
- **CSV-Spezifikation:** Trennzeichen für Zeilen und Spalten, Textqualifizierer, Dezimaltrenner, Zeichensatz usw., ggf. als Datei schema.ini
- ggf. **XSD** oder DTD zu XML

Signifikante Grundfunktionalitäten:

- maschinelles Auslesen von Werten einer oder mehrerer Zellen
- Berechnung (mathematisch, bzgl. Zeichenketten, ...) von Ergebniswerten

Empfehlung zur visuellen Repräsentation:

- PDF/A
- ggf. TIFF

Allgemeine Hinweise:

- Sofern die Tabellen/Listen Berechnungen (Funktionen) enthalten, muss dringend entschieden werden, ob diese signifikant sind und neben den Werten/Ergebnissen archiviert werden müssen.
- Da sich CSV-Daten relativ einfach in Tabellenform visualisieren lassen, ist eine zusätzliche visuelle Repräsentation nicht zwingend erforderlich, kann jedoch unterstützend für das Verständnis sein.
- XML-Dateien sollen einem XML-Schema (XSD-Datei) entsprechen, welches ihre Struktur definiert. In einfachen Fällen genügt auch eine Dokumenttypdefinition (DTD). Es muss zumindest erwähnt werden, dass es sich um eine Tabelle oder Liste handelt.

3 3D- bzw. VR-Daten

[Dieser Abschnitt ist noch nicht ausreichend verifiziert, die gelb markierten Angaben sind noch unverbindlich.]

3D- bzw. VR-Daten spiegeln dreidimensionale Sachverhalte wieder. Man unterscheidet zwischen dreidimensionalen Punktwolken (typisch für Messwerte) und flächen- bzw. körperhaften Modellen (gern für die Präsentation von Ergebnissen genutzt).

3D- bzw. VR -Daten sind **einfache funktionelle Materialien**.

Sie können mit einer **erweiterten formatbasierten Erhaltungsstrategie** archiviert werden.

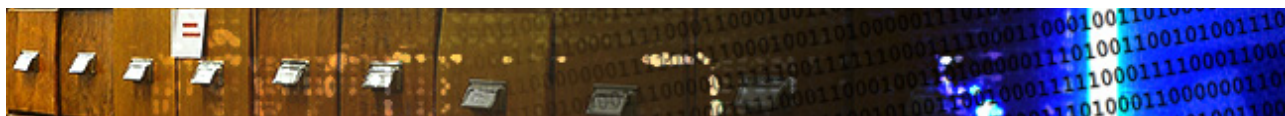
Neben der Auswahl eines nachhaltigen Datenformats zur **funktionellen Repräsentation** ist eine **Strukturbeschreibung** erforderlich und es wird eine separate **visuelle Repräsentation** dringend empfohlen.

Empfohlene Formate zur Archivierung sind:

- Zu 3D- bzw. VR- (Virtual Reality) Daten liegen nur wenige Erfahrungen vor, so dass hier kein Format empfohlen wird.

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

- **CSV** als ASCII XYZ (ggf. + RGB) + Strukturbeschreibung/Metadaten
- ggf. X3D [???, Praxis fehlt]
- ggf. OBJ [???, Praxis fehlt]
- ggf. PLY [???, Praxis fehlt]



Empfehlung zur visuellen Repräsentation:

- PDF/A (vgl. Abschnitt 3.1)
- ggf. TIFF (vgl. Abschnitt 3.2)

Strukturbeschreibung/Metadaten:

- benutztes **Messsystem** (3D-Scanner)
- ggf. Angaben zur **Auflösung** und **Genauigkeit**
- benutztes **Koordinatenreferenzsystem** (z.B. ETRS89_UTM33, RD-83_3GK5, ...)
- ggf. **CSV-Spezifikation** (vgl. Abschnitt 3.5),
zumindest Anordnung der Koordinaten (XYZ) und ggf. Farbwerte (z.B. RGB) im CSV

Signifikante Grundfunktionalitäten:

- dreidimensionale lagerichtige graphische Darstellung in einem geeigneten Viewer oder GIS
- ggf. grundlegende geometrische Operationen im dreidimensionalen Raum (analog Simple Features)

Allgemeine Hinweise:

- Das gewählte Format muss für das Archiv kuratierbar sein, d.h. es müssen adäquate Werkzeuge (z.B. 3D-Editor + Konverter + Toolbox) zur Visualisierung und Bearbeitung sowie Kompetenz zu deren Nutzung verfügbar sein.

Source/Quellen:

- Archivierungs-Ratgeber der VLA-Kommission Archäologie und Informationssysteme:
http://www.landesarhaeologen.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](https://creativecommons.org/licenses/by-sa/4.0/)). Dies ist eine [Free-Culture-Lizenz](https://creativecommons.org/licenses/by-sa/4.0/).

