

Kanton St.Gallen
Departement des Innern

Amt für Kultur
Staatsarchiv St.Gallen



Archivische Dateiformate

Inhaltsverzeichnis

Einleitung	3
Formatkategorie 'Text'	3
Formatkategorie 'Bild'	3
Formatkategorie 'Audio'	3
Formatkategorie 'Video'	3
Formatkategorie 'Tabellenkalkulation'	4
Formatkategorie 'Datenbanken'	4
Formatkategorie 'Hypertext'	4
Formatkategorie 'GIS'	4
Formatkategorie 'CAD/CAM'	4
Konvertierung	5

Version vom: September 2019
Ersetzt Version vom: September 2013

© Staatsarchiv St.Gallen

Einleitung

Nur eine kleine Auswahl der vielen verschiedenen Dateiformate ist für die Archivierung tauglich. Hinsichtlich der Übernahme von digitalen Daten zur Langzeitarchivierung macht das Staatsarchiv hiermit Vorgaben, die auf dem *Katalog archivischer Dateiformate* der KOST¹ basieren und abgesehen von den Dateiformaten auch Empfehlungen bezüglich Konvertierung² enthalten. Konvertierungen sind nötig, falls Dateien nicht von Beginn an in archivischen Dateiformaten erstellt werden.

Dateiformate lassen sich in neun Formatkategorien einteilen: Text, Bild, Audio, Video, Tabellenkalkulation, Datenbanken, Hypertext, GIS und CAD/CAM. Für jede dieser Kategorien gibt es bestimmte geeignete Dateiformate, welche aufgrund verschiedener Kriterien, nach momentanem Wissensstand, für die Archivierung tauglich sind. Die Kriterien sind Offenheit, Lizenzfreiheit, Verbreitung, Funktionalität, Implementierung und Verifizierbarkeit.

Nicht für alle Formatkategorien gibt es eindeutige Vorgaben. Unter Umständen müssen der konkrete Anwendungsfall und die Aufbewahrungsdauer der Daten in den Formatentscheidungen miteinbezogen werden.

Formatkategorie 'Text'

Kurzbeschreibung

Unter Textdaten werden Daten zur Repräsentation von wenig strukturiertem Text verstanden, wie sie unter anderem von Textverarbeitungssoftware oder Texteditoren erzeugt werden (Office-Dateien, E-Mail etc.).

Vorgabe

PDF/A-1, PDF/A-2 oder OOXML (z.B. .docx)

Formatkategorie 'Bild'

Kurzbeschreibung

Bei Bilddaten ist zwischen Raster- und Vektorgrafiken zu unterscheiden. Ein einziges ideales Archivformat für Bilddaten gibt es nicht.

Vorgabe

Rastergrafiken: TIFF oder JPEG2000; bei JPEG Rücksprache mit Staatsarchiv nötig
Vektorgrafiken: Rücksprache mit Staatsarchiv nötig

Formatkategorie 'Audio'

Kurzbeschreibung

Bei Audiodaten wird zwischen unkomprimierten und komprimierten Dateien beziehungsweise Dateiformaten (z.B. MP3) unterschieden.

Vorgabe

Unkomprimierte Dateien: WAV
Komprimierte Dateien: Rücksprache mit Staatsarchiv nötig

Formatkategorie 'Video'

Kurzbeschreibung

Videoformate können nicht per se als archivtauglich oder nicht archivtauglich klassifiziert werden. Da Videodateien sehr speicherplatzintensiv sind, ist Datenkompression ein wichtiger Aspekt: Hierdurch kann eine hohe Datenreduktionsrate erreicht werden.

¹ Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen. Version 6.0 vom Juli 2019. URL: https://kost-ceco.ch/cms/kad_main_de.html [Stand: 06. September 2019].

² Umwandlung oder Erstellung von archivischen Dateiformaten.

Vorgabe

Rücksprache mit Staatsarchiv nötig

Formatkategorie 'Tabellenkalkulation'

Kurzbeschreibung

Mit Tabellenkalkulation werden numerische und alphanumerische Daten dargestellt, die mittels Funktionen verarbeitet werden können. Diese verlieren ihre Bedeutung, sobald sich die Daten nicht mehr in einem aktiven Status befinden. Daher ist die Erhaltung der Funktionen im Archiv nicht vorgesehen.

Vorgabe

ODF oder OOXML (z.B. .xlsx); bei XLS Rücksprache mit Staatsarchiv nötig

Formatkategorie 'Datenbanken'

Kurzbeschreibung

Nebst Datenbanken im eigentlichen Sinn sind hier auch Daten aus Tabellenkalkulationen gemeint, die als Datenbanken verstanden oder behandelt werden. Bei Datenbanken, die Teil eines umfassenderen Systems sind (Fachanwendungen, CMS etc.), sollten diese nicht losgelöst vom jeweiligen umfassenden System betrachtet werden.

Vorgabe

In sich abgeschlossene Datenbanken: SIARD

Datenbanken als Teil eines umfassenderen Systems: Rücksprache mit Staatsarchiv nötig

Formatkategorie 'Hypertext'

Kurzbeschreibung

Hypertext bezeichnet eine Menge von Texten, die über Verknüpfungen (Links) miteinander verbunden sind. Die Gestaltung der Texte wie auch die Verknüpfungen erfolgen mit einer Auszeichnungssprache (markup language). Die bekannteste Hypertext-Auszeichnungssprache ist HTML.

Vorgabe

WARC oder PDF/A-2

Formatkategorie 'GIS'

Kurzbeschreibung

Geoinformationssysteme bzw. Geografische Informationssysteme (GIS-Systeme) sind Informationssysteme zur Erfassung, Bearbeitung, Organisation und Darstellung räumlicher Daten. Die für diese Aufgabe notwendigen Vektordaten und raumbezogenen Metadaten werden gemeinhin in Datenbanken verwaltet.

Vorgabe

Vektordaten: INTERLIS 1

Bild- und Grafik-Rasterdaten: Rasterdaten mit World-File (je nach zugrundeliegendem Bildformat z.B. JGW)

Produkte und Karten: Vektorgrafik in PDF

Formatkategorie 'CAD/CAM'

Kurzbeschreibung

CAD (Computer-Aided Design) bezeichnet das computerunterstützte Konstruieren von unterschiedlichen Produkten wie Autos, Flugzeuge, Bauwerke. Im engeren Sinne wird darunter das Zeichnen von Konstruktions- und Architekturplänen mit Computerunterstützung verstanden. Das Produkt ist in der Regel eine 2D- oder 3D-Vektorgrafik mit Metadaten zu den einzelnen Vektorelementen

Vorgabe

CAD-Daten: DWG

Integrierte Produktion (CAX): IFC

Konstruktions- und Architekturpläne: Vektorgrafik in PDF

Konvertierung

Kurzbeschreibung

Für die Umwandlung von Dateien in archivische Dateiformate oder für die Erstellung solcher Daten können unterschiedliche Programme verwendet werden. Die Qualität dieser Programme respektive der Konvertierungsergebnisse ist unterschiedlich. Es ist darüber hinaus in gewissen Fällen sinnvoll, Dateiformate vorerst noch nicht in archivische zu konvertieren.

Vorgabe

Rücksprache mit Staatsarchiv nötig