

Archivfähige Dateiformate und ihre Spezifikationen

Archivfähige Dateiformate (Kostenordnung für Leistungen des Hessischen Landesarchivs, § 4) sind Dateiformate, die sich aufgrund ihrer Verbreitung, der langfristigen Verfügbarkeit und ihres offenen Standards zur langfristigen Erhaltung von Informationen eignen. Im Folgenden werden die Dateiformate und Spezifikationen aufgelistet, die bei der Aussonderung unterschiedlicher Arten von Daten zu verwenden sind. Für Unterlagen, die hiervon nicht abgedeckt werden, muss im Bedarfsfall ein geeignetes Dateiformat durch das Hessische Landesarchiv festgelegt werden.

Reiner Text

TXT für Daten, die direkt unter Verwendung einer Zeichenkodierung in Text umgesetzt werden.

Spezifikation	Bevorzugt	Weitere
Zeichensatz	UTF-8	UTF-16, Normenfamilie ISO 8859, Windows 1252
Zeilenumbrüche	CRLF	

Gestalteter Text

PDF für Text, der Daten seitenbasiert und für den Ausdruck optimiert darstellt.

Spezifikation	Bevorzugt	Weitere
Standard	PDF/A-1a	PDF/A-1b, PDF/A-2a, PDF/A-2b, PDF/A-2u
Farbräume	sRGB	Adobe RGB, ProPhoto RGB, ECI RGB oder CMYK
Zeichensatz	UTF-8	UTF-16, Normenfamilie ISO 8859, Windows 1252

Rasterbilder

TIF für Bilder, die zur Darstellung ihres Inhaltes ein Pixelraster verwenden.

Spezifikation	Bevorzugt	Weitere
Komprimierung	Keine	
Farbräume	sRGB	Adobe RGB, ProPhoto RGB, ECI RGB oder CMYK

Vektorbilder

SVG für Bilder, die zur Darstellung ihres Inhaltes eine auf Vektoren basierte Beschreibung verwenden. Geeignete Formatspezifikationen werden in Kooperation mit der abgebenden Stelle entwickelt.

Akustische Informationen

WAV für auditive Daten, d. h. Audioaufzeichnungen.

Spezifikation	Bevorzugt	Weitere
Samplingtiefe	32 Bit	Alle anderen Spezifikationen sind entsprechend dem Originalformat zu übernehmen
Samplingrate	44100 Hz	
Kanäle	Stereo	

Bewegte Bilder (mit und ohne Ton)

MPEG für Videoaufnahmen bzw. Abfolgen von Bildern.

Spezifikation	Bevorzugt	Weitere
Videocodec	mpeg2video	Alle anderen Spezifikationen sind entsprechend dem Originalformat zu übernehmen.
Videobitrate	mind. 5000 kbps	
Bildwiederholungsfrequenz	mind. 25 Bilder pro Sekunde	
Audiocodec	AC3	
Audiokanäle	Stereo	
Audiosamplingtiefe / -rate	16 Bit / 48000 Hz	

Strukturierte Informationen

CSV oder **XML** für strukturierte Informationen, i. d. R. aus Datenbanken und Tabellen. Hierzu gehören auch Metadaten (z. B. aus einem DMS). Zusätzlich zum Dateiformat muss hier die Struktur des Datenexports abgestimmt werden.

Spezifikationen	Bevorzugt	Weitere
CSV-Feldtrennzeichen	Semikolon (;)	
CSV-Textbegrenzungszeichen	Anführungszeichen (") bei jedem Inhalt (Text oder Zahl)	Bei Dateien mit ausschließlich numerischen Feldern ist kein Textbegrenzungszeichen notwendig
CSV-Datensatztrenner	CRLF	
CSV-Kopfdatensatz	Erste Zeile	
CSV-Zeichensatz	UTF-8	UTF-16, Normenfamilie ISO 8859, Windows 1252
CSV-Zeilenumbruch in Feldinhalten	Zeilenumbrüche (Linefeed bzw. Carriage Return Linefeed) innerhalb von Feldinhalten sind bei vorhandenem Textbegrenzungszeichen	

	zulässig.	
CSV- Textbegrenzungszeichen in Feldinhalten	Wird das Textbegrenzungs- zeichen, d.h. das Anführungs- zeichen, innerhalb eines Feldinhaltes verwendet, muss es durch ein vorangestelltes Anführungszeichen maskiert werden. D.h. ein als Feldinhalt vorkommendes Anführungs- zeichen ist gemäß des CSV- Spezifikation http://www.ietf.org/rfc/rfc4180.txt zu verdoppeln.	
CSV-Feldbeschreibungen	Nach Absprache zwischen Archiv und abgebender Stelle.	
XML	Standardausprägungen wie xDOMEA.	Abstimmung eines Schemas (wenn kein Standard existiert).