

STANDARDS FÜR DIE ARCHIVIERUNG DIGITALER UNTERLAGEN

— *Archivtaugliche Dateiformate*



Verantwortliche Stelle: Schweizerisches Bundesarchiv
Abteilung Informationsüberlieferung
Datum: April 2020
Version: 2020/04, ersetzt die Version 2018/04

Änderungskontrolle

Version	Beschreibung, Bemerkung
06/2007	Erste Version Juli 2007
01/2014	Erweiterung mit MPEG-4 (Video), Gesamtüberarbeitung
04/2018	Erweiterung mit PDF/A 2, JPEG2000, XML/XSD, SIARD 2.1, FFV1, TIFF+EWF.XML, INTERLIS, Gesamtüberarbeitung
04/2020	Erweiterung mit PDF/A-2b

Inhaltsverzeichnis

1	Einleitung.....	3
1.1	Zusammenfassung der wichtigsten Neuerungen.....	3
2	Anwendungsbereiche.....	4
2.1	Text.....	4
2.2	Bild.....	4
2.3	Audio.....	5
2.4	Video.....	5
2.5	Tabellen und Datenbanken.....	5
2.6	Geodaten.....	5
2.7	Sonstiges.....	6
2.7.1	XML/XSD.....	6
2.7.2	Webarchivierung.....	6
2.7.3	Weitere Anwendungsbereiche.....	6
3	Archivtaugliche Dateiformate.....	7
	TEXT – unstrukturierte "Nur-Text"-Daten.....	7
	PDF/A – Portable Document Format / Archive.....	9
	CSV –Text-Datei mit Trennzeichen.....	11
	SIARD – Software Independent Archiving of Relational Databases.....	13
	TIFF – Tagged Image File Format.....	14
	JPEG2000 - Joint Photographic Experts Group 2000.....	15
	WAVE – Audioformat von Microsoft.....	16
	FFV1 Video Codec.....	17
	MPEG-4 – Videoformat.....	18
	XML/XSD - eXtensible Markup Language.....	19
	TIFF+EWF.XML – Tagged Image File Format und Extended World File.....	20
	INTERLIS.....	21

1 EINLEITUNG

Unterlagen, welche mit Hilfe von Informations- und Kommunikationstechnologien empfangen oder erstellt wurden, werden im Folgenden als „digitale Unterlagen“ und ihre Aufbewahrung auf unbestimmte Zeit im Bundesarchiv als „digitale Archivierung“ bezeichnet.

Die anbietepflichtige Stelle sorgt dafür, dass die Unterlagen so aufbereitet sind, dass sie ohne zusätzlichen Aufwand im Hinblick auf ihre Archivwürdigkeit bewertet und gegebenenfalls archiviert werden können.¹

Die für die Ablieferung verwendeten Formate und Materialien müssen archivtauglich sein und den in Kapitel 2 und 3 aufgeführten Merkmalen entsprechen. Das Bundesarchiv setzt die *Dateiformate* fest, die als archivtauglich gelten und somit den strengen Anforderungen genügen, die eine langfristige Verstehbarkeit garantieren.

Das Bundesarchiv favorisiert die Einschränkung der Anzahl Formate gegenüber der Vielfalt. Eine knappe, überschaubare und vorsichtig ausgewählte Menge an Formaten stellt die langfristige Verstehbarkeit in deutlich höherem Masse sicher als eine schwer kontrollierbare und aufwendig zu pflegende grosse Menge an Formaten, die oft von den gerade aktuellen Versionen der Computer-Applikationen abhängig ist.

Dieses Dokument gibt eine Übersicht der zurzeit akzeptierten Formate. Auf der [Webseite des Bundesarchivs](#) ist ein Ausblick auf Formate zu finden, welche aktuell als Kandidaten für die Aufnahme in die vorliegende Liste der archivtauglichen Formate gelten.

1.1 Zusammenfassung der wichtigsten Neuerungen

Neue Formate:

- PDF/A-2b

¹ VBGA, Art. 5

2 ANWENDUNGSBEREICHE

Zurzeit existieren in Bundesarchiv folgende Standards für die Archivierung digitaler Unterlagen:

Anwendungsbereich	archivtaugliche Formate	Bemerkungen
Text (unstrukturiert)	"Nur Text" ("plain text")	UTF-8 UTF-16 ISO-8859-1 ISO 8859-15 US-ASCII
„Office“ Dokumente	PDF/A	entspricht PDF 1.4 (PDF/A-1) bzw. PDF 1.7 (PDF/A-2) mit Einschränkungen
Tabellen	CSV	Text-Datei mit Trennzeichen, Encoding wie bei Text (unstrukturiert)
Relationale Datenbanken	SIARD	
Rasterbilder	TIFF JPEG2000	
Audio	WAVE	
Video	MPEG-4 FFV1	
Strukturierte (Text-)Daten	XML	UTF-8, UTF-16
Geodaten	TIFF + EWF.XML INTERLIS	

Im Folgenden werden die Formate kurz erläutert. Die ausführlichen Bundesarchiv-Standards mit Hinweisen und Einschränkungen sind im nächsten Kapitel zu finden.

2.1 Text

Bei Textdaten ist mehrheitlich der (textuelle) Inhalt relevant. Wenn hier von Textinhalten gesprochen wird, werden nicht nur klassische Textdokumente gemeint, sondern auch Präsentationen und Webseiten. Diese Inhalte werden gerne auch durch Formatierungen und Bilder ergänzt. Als Format für die Archivierung von solchen Textinhalten eignet sich das Format PDF/A.

Neben diesen „Office“-Textdokumenten gibt es auch noch unstrukturierten Text. Dieser enthält keine eingebetteten oder sichtbaren Anweisungen zur Darstellung (Fettdruck, eingerückt, farbig, etc.) oder Informationen zur Struktur (Titel, Abschnitt, Unterabschnitt, Inhaltsverzeichnis, etc.). Es hat sich gezeigt, dass gerade diese in ihrer Art sehr einfachen Dateien sich als am längsten haltbar und verstehbar erweisen. Beispiele sind einfache E-Mail-Meldungen ("Nur-Text", ohne Anhänge), Logfiles oder Kurzbeschreibungen (README.TXT). Zur Archivierung dieser Inhalte eignet sich das Format „Nur Text“ („Plain Text“).

2.2 Bild

Bei den Bildformaten wird zwischen Rasterbildern und Vektorbildern unterschieden. Rasterbilder haben eine definierte Grösse, und sind im Wesentlichen eine Ansammlung von einzelnen Pixeln. Vektorbilder hingegen

speichern die Bilddaten in Form von Vektoren, und sind daher beliebig skalierbar. Vektorbilder eignen sich besonders für Diagramme und ähnliches, für Fotografien sind sie gänzlich ungeeignet. Da die Verbreitung von Vektordaten in der Bundesverwaltung bisher eingeschränkt ist, entstand noch kein Bedarf diese in einem entsprechenden Format zu archivieren. Falls bei einer Abliefernden Stelle der Bedarf für die Archivierung von Vektordaten besteht, wird darum gebeten das Bundesarchiv zu kontaktieren um das weitere Vorgehen abzustimmen.

Für Rasterbilder bietet das Bundesarchiv zwei akzeptierte Formate an: TIFF und JPEG2000. TIFF ist das ältere und weiter verbreitete Format. JPEG2000 ist ein jüngerer Format, und adressiert einige der Schwächen von TIFF. So ist es mit JPEG2000 beispielsweise möglich, hochqualitative verlustbehaftete Kompromierung einzusetzen. Diese Verwendung wird vom Bundesarchiv zwar nur beschränkt empfohlen, kann in manchen Anwendungsfällen aber durchaus sinnvoll sein. Zudem ermöglicht JPEG2000 wesentlich bessere Kompromierung bei Bildern mit „wenig Inhalt“, sprich beispielsweise Architekturplänen, die hauptsächlich aus wenigen Linien auf einer grossen Freifläche von Hintergrund bestehen.

2.3 Audio

Um Audiodaten in hoher Qualität erhalten zu können, akzeptiert das BAR derzeit nur WAVE als Format.

2.4 Video

Die Archivierung von Videodaten kann sehr schnell zu enormen Datenmengen führen. Es ist daher vor der Archivierung abzuwägen, welche Videoinhalte hochqualitativ erhalten werden müssen.

Für Videos, die in hoher Qualität archiviert werden sollen, steht das Format FFV1 zur Verfügung. Dieses Format wurde speziell für die Archivierung entwickelt. Für optimale Ergebnisse empfiehlt das Bundesarchiv, die FFV1-Datei direkt aus den Original-(Roh-)Videodaten zu generieren.

Für grössere Mengen von Videos, bei denen die Videobildqualität sekundär ist, oder die bereits als MPEG-4 vorliegen, kann das Format MPEG-4 verwendet bzw. beibehalten werden. MPEG-4 stellt jedoch immer einen Kompromiss bezüglich Qualität und Erhaltbarkeit dar und wird vom BAR explizit nur unter diesen beiden Bedingungen zur Verwendung empfohlen.

2.5 Tabellen und Datenbanken

Bei Tabellen und Datenbanken kann einerseits die Komplexität des Datenmodells den Ausschlag zur Wahl des Formats geben, andererseits können auch technische Gründe für oder gegen ein Format sprechen.

Für wenige einfache Tabellen (ohne Relationen) soll das Format CSV verwendet werden. Hierbei ist das Encoding zu beachten.

Sobald eine grosse Anzahl zusammenhängender Tabellen (mit Relationen) archiviert werden sollen, wird das Format SIARD empfohlen.

Da umfangreiche Tabellen und Datenbanken eigentlich nie selbstdokumentierend sind, ist in der Regel zusätzlich eine Dokumentation erforderlich, welche ebenfalls archiviert werden muss.

2.6 Geodaten

Bei Geodaten wird wie bei Bildern zwischen Rasterdaten und Vektordaten unterschieden. Für die Archivierung von Geodaten, die als Rasterdaten vorliegen, wird das Format TIFF+EWF.XML verwendet. Dieses besteht aus einer TIFF-Bilddatei und einer zugehörigen EWF.XML-Datei (Extended World File). Das TIFF richtet sich dabei nach dem Bildstandard TIFF, die EWF.XML-Datei ist eine XML-Datei gemäss der EWF Schema Definition. Für die Archivierung von Geodaten, welche als Vektordaten vorliegen, wird das Format INTERLIS verwendet. Zusätzlich finden bei der Archivierung von Geodaten u.U. auch weitere archivtaugliche Formate Anwendung (XML, SIARD, CSV, PDF)

Für die Archivierung von Geodaten steht zusätzliche Dokumentation zur Verfügung, die beim Bundesarchiv angefragt werden kann.

2.7 Sonstiges

2.7.1 XML/XSD

Für strukturierte (Text-)Daten kann das Format XML eingesetzt werden. Der häufigste Anwendungsfall für die Archivierung von Daten in Form von XML sind Fachapplikationen, die gewisse Daten bereits als XML gespeichert halten, oder eine Möglichkeit bieten, diese als XML zu exportieren.

2.7.2 Webarchivierung

Bei der Webarchivierung ist zuerst zu bestimmen, welche Inhalte des Webauftrittes archivwürdig sind, damit diese spezifischen Inhalte erhalten werden können. Bei Webseiten der Bundesverwaltung wird dies meist durch den Inhalt (Text) und nicht durch die Form (Layout/Design) definiert. Daher kann oft das Format PDF/A für die Archivierung verwendet werden. Webseiten mit viel Logik sollten wie Fachapplikationen behandelt werden, sodass sich hier beispielsweise SIARD oder XML als Formate anbieten können.

2.7.3 Weitere Anwendungsbereiche

Hat die Abliefernde Stelle weitere Anwendungsbereich, welche nicht in diesem Dokument behandelt werden, wird darum gebeten, mit dem Bundesarchiv Kontakt aufzunehmen, um eine optimale Lösung für die Archivierung dieser Unterlagen zu finden.

3 ARCHIVTAUGLICHE DATEIFORMATE

TEXT – unstrukturierte "Nur-Text"-Daten

FORMAT	VERBINDLICHER BEZEICHNER	TEXT – Text Plain
	PRONOM PUIDs	x-fmt/16, x-fmt/21, x-fmt/22, x-fmt/62, x-fmt/111, x-fmt/282, x-fmt/283
	DATEI-ENDUNG	Empfohlene Datei-Endung: .txt ²
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	<p>Erlaubte Zeichensatzkodierungen für "Nur-Text"-Dateien:</p> <ul style="list-style-type: none"> • ISO Latin-1 (ISO 8859-1) und ISO Latin-9 (ISO 8859-15) • Unicode 5.0 Universal Character Set (UCS) (ISO 10646:2003) • US-ASCII (ANSI X3.4-1986) bzw. US-ASCII "safe characters" 	
	<p>ISO Latin-1 (ISO 8859-1) und ISO Latin-9 (ISO 8859-15)</p> <p>Der ISO 8859 Standard ist eine Gruppe von 15 Zeichensatzkodierungen für verschiedene Alphabete</p> <p><u>Bezugsquelle:</u> International Organization for Standardization, ISO/IEC 8859-1 "Information technology -- 8-bit single-byte coded graphic character sets -- Part 1: Latin alphabet No. 1"</p>	
	<p>Unicode</p> <p>Unicode ist ein internationaler Standard, in dem langfristig für jedes sinntragende Zeichen bzw. Textelement aller bekannten Schriftkulturen und Zeichensysteme ein digitaler Code festgelegt wird.</p> <p>In Unicode finden Zeichen der wichtigsten ISO-Zeichensätze (z.B. diejenigen Zeichensätze der ISO-Normen-Serie 8859) eine 1:1-Entsprechung (das bedeutet, dass bei einer Konvertierung von ISO zu Unicode und zurück das gleiche Ergebnis herauskommt). Heute erledigen die meisten Webbrowser die Darstellung dieser Zeichensätze mit einer Unicode-kodierten Schrift in der Regel perfekt und vom Benutzer unbemerkt.</p> <p>Als alternative Bezeichnung für den Unicode-Zeichensatz wird auch der Begriff „Universal Character Set (UCS)“ verwendet, dieser Begriff stammt aus der ISO-Norm 10646.</p> <p>Die folgenden Kodierungen von Unicode sind erlaubt:</p> <p>UTF-8 (eine 8-bit Kodierung mit variabler Länge, bietet eine maximale Kompatibilität mit US-ASCII)</p> <p>UTF-16 (eine 16-bit Kodierung mit variabler Länge)</p> <p>UCS-2, UTF-7 gelten als veraltet und sollten nicht mehr verwendet werden.</p> <p><u>Bezugsquellen:</u> ISO 10646:2003 unter Publicly Available Standards</p>	
	<p>US-ASCII (ANSI X3.4-1986) bzw. "safe characters"</p> <p>Grundsätzlich ist der US-ASCII Zeichensatz gemäss ANSI X3.4-1986 Standard bzw. ISO/IEC 646-US oder ISO/IEC 646:1991-IRV (international reference version) erlaubt. Alle anderen Bezeichnungen "ASCII" sind nicht archivtauglich.</p> <p>Aufgrund "nationaler Varianten" (z.B. in den ISO/IEC 646 Standards) sind einige Zeichen des US-ASCII Zeichensatzes (z.B. @ [\] { }) unsicher und können im internationalen Austausch von Textdaten falsch übertragen bzw. falsch interpretiert werden.</p> <p>Es wird empfohlen, nur Zeichencodes zu verwenden, die als sicher gelten, die sogenannten "safe characters". Nebst den alphabetischen Zeichen ("A" bis "Z" und "a" bis "z") den Ziffern ("0" bis "9"), dem Leerschlag (" "), sind nur die folgenden Zeichen als sicher zu betrachten:</p> <p><code>! " % & ' () * + , - . / : ; < = > ?</code></p> <p><u>Bezugsquellen:</u></p> <p>Wikipedia, the free encyclopedia: ASCII</p> <p>Information technology -- ISO 7-bit coded character set for information interchange, IRV international reference version: ISO/IEC 646:1991</p>	
	EIGENTÜMER	

² Textdateien weisen manchmal andere Endungen auf (z.B., .log, .dat, .lst). Insbesondere ist dies bei strukturierten Textdateien der Fall. Es besteht kein Zwang, unstrukturierte Textdateien ggf. in .txt umzubenennen, insbesondere dann nicht, wenn in der Endung Bedeutung steckt und einen Hinweis auf deren ursprünglichen Verwendungszweck ergibt.

	ISO/IEC Standards: International Organization for Standardization ANSI Standards: American National Standards Institute
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE <p>Textdaten (einfache Texte, unstrukturiert)</p> <p>Eine unstrukturierte Textdatei ist zur Darstellung reinen Textinhaltes geeignet, der mit minimalen Strukturierungsmöglichkeiten auskommt (Zeilen) und ansonsten keine weitere Struktur- bzw. Darstellungsinformation benötigt, d.h., er enthält keine eingebetteten oder sichtbare Anweisungen zur Darstellung (Fettdruck, eingerückt, farbig, etc.) oder Informationen zur Struktur (Titel, Abschnitt, Unterabschnitt, Inhaltsverzeichnis, etc.).</p> <p>Beispiele sind einfache E-Mail-Meldungen ("Nur-Text", ohne Anhänge), Logfiles oder Kurzbeschreibungen (README.TXT).</p> <p>Sind die Textdaten strukturiert, sind andere Standards zu verwenden, beispielsweise CSV, PDF/A oder XML.</p>
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG
	<p>Unstrukturierte "Nur-Text"-Daten dürfen keine Steuerzeichen enthalten ausser der Zeilen- (LF, CR) und Seitenumbrüche (FF) sowie dem Tabulatorzeichen (TAB).</p>
	<p>Das Null-Zeichen (NUL) darf nicht vorkommen.</p>
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS <p>Dateien in anderen Zeichensätzen als die oben aufgeführten sollen in Unicode, vorzugsweise UTF-8, umgewandelt werden. Besteht Unklarheit bezüglich des Zeichensatzes des Originaltextdokuments, darf die Umwandlung nicht vorgenommen werden.</p>
	BEMERKUNGEN UND HINWEISE
	<p>Liegen grössere Mengen zu archivierender Text-Dateien mit unbekannter Zeichenkodierung, ist vorgängig das Bundesarchiv zu konsultieren.</p>

PDF/A – Portable Document Format / Archive

FORMAT	VERBINDLICHER BEZEICHNER	PDF/A – PDF/Archive
	PRONOM PUIDs	fmt/95 (PDF/A-1a), fmt/354 (PDF/A-1b), fmt/476 (PDF/A-2a) , fmt/477 (PDF/A-2b), fmt/478 (PDF/A-2u)
	DATEI-ENDUNG	Empfohlene Datei-Endungen: .pdfa, .pdf
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	ISO PDF/A-1 ISO 19005-1:2005 Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1) Dieser ISO-Standard ist eine Einschränkung von PDF 1.4.	
	ISO PDF/A-1 Korrigendum ISO 19005-1:2005/Cor.2:2011 Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1); TECHNICAL CORRIGENDUM 2 <u>Bezugsquelle:</u> International Organization for Standardization	
	ISO PDF/A-2 ISO 19005-2:2011 Document management – Electronic document file format for long-term preservation – Part 2: Use of ISO 32000-1 (PDF/A-2) Dieser ISO-Standard ist eine Einschränkung von PDF 1.7. <u>Bezugsquelle:</u> International Organization for Standardization	
	PDF 1.4 PDF Reference third edition, Adobe Portable Document Format, Version 1.4, Addison Wesley, 2001, ISBN 0-201-75839-3 PDF 1.4 Spezifikation der Version 1.4 der Seitenbeschreibungssprache PDF (Portable Document Format) von Adobe Systems Inc. unter http://www.adobe.com/devnet/pdf/pdf_reference.html .	
	PDF 1.7 ISO 32000-1:2008 Document management – Portable document format – Part 1 : PDF 1.7	
VORGABEN UND EINSCHRÄNKUNGEN	XMP XMP Specification: XMP Adding Intelligence to Media, Adobe Systems Inc., 2004 XMP Spezifikation der „Extensible Metadata Platform“. Siehe: http://www.adobe.com/devnet/xmp.html ISO-Standard: https://www.iso.org/standard/57421.html	
	EIGENTÜMER	
	ISO/IEC Standards: International Organization for Standardization Adobe Systems Inc. Standards: Adobe Systems Inc.	
	ANWENDUNGSGEBIETE	
	Druckbare „Office“ Dokumente. Eine Datei wird dann mit Vorteil im PDF/A-Format archiviert, wenn ihre Druckversion ihren Inhalt adäquat wiedergibt. Der Informationsgehalt der PDF-Version entspricht dem der Druckversion. Mit PDF/A-2 sind auch eingebettete Dateien erlaubt, sofern sie PDF/A-1 oder PDF/A-2 konform sind. Die eindeutige visuelle Reproduzierbarkeit ist daher in diesen Fällen nur noch am Computer möglich, da diese Eigenschaften nicht „ausdruckbar“ sind. Da Zeichen und Wörter im PDF-Format als Zeichen und nicht als Bildpunkte gespeichert sind, ist das PDF/A-Format dem TIFF-Format als Seitendarstellung vorzuziehen, wo immer es textuelle Zeicheninformation enthält und nicht ein einfacher Seiten-Wrapper für eine Bilddatei ist.	
VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG		

	<p>PDF/A-1 Daten müssen mindestens als PDF/A-1b validiert sein. PDF/A-2 Daten müssen mindestens als PDF/A-2b validiert sein.</p>
EMPFEHLUNGEN UND HINWEISE	<p>EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS</p>
	<p>Es gibt eine Vielzahl von PDF/A-Konvertern. Diese unterscheiden sich teilweise stark in Ihrer Ausrichtung (z.B. Fokus auf automatisierte Konvertierung oder manuelle Konvertierung, unterstützte Quellformate, Systemanforderungen usw.), weshalb hier keine pauschale Empfehlung abgegeben werden kann. Die KOST erstellt periodisch eine Studie über die verfügbaren PDF/A-Konverter.</p>
	<p>BEMERKUNGEN UND HINWEISE</p>
	<p>Dokumente im PDF/A-Format sind schwer zu verändern (auch in diesem Punkt gleichen sie einer ausgedruckten Version). Die Konversion ins PDF/A-Format ist dann vorzunehmen, wenn das Dokument nicht mehr geändert werden muss. Die Konversion ist möglichst früh vom Autor des Dokuments vorzunehmen, weil einzig der Autor beurteilen kann, ob die Druckversion das Dokument adäquat wiedergibt.</p> <p>Im Folgenden sind die wichtigsten Veränderungen von PDF/A-2 gegenüber PDF/A-1 aufgeführt:</p> <ul style="list-style-type: none"> • JPEG2000 Kompression: in PDF/A-2 ist JPEG2000 Kompression erlaubt, womit im Vergleich zur Dateigrösse eine bessere Bildqualität erzielt werden kann. • Transparenz: In PDF/A-2 ist Transparenz erlaubt. In PDF/A-1 ist Transparenz verboten, was bei der Konvertierung zu Problemen führen kann. PDF/A-2 erlaubt demnach auch die Verwendung von Wasserzeichen. • Eingebettete Dateien: PDF/A-2 erlaubt das Einbetten von Dateien, sofern sie dem PDF/A-1 oder PDF/A-2 Standard entsprechen. • Dateigrösse: PDF/A-2 Dateien können über 10 GB gross sein und damit grösser als PDF/A-1. • „Optional content“: In PDF/A-2 besteht die Möglichkeit „Layer“ ein- oder auszublenden, wodurch sich die Darstellung des Dokuments verändern kann. Dies wird oft im Bereich von Computer Aided Design (CAD) verwendet.

CSV –Text-Datei mit Trennzeichen

FORMAT	VERBINDLICHER BEZEICHNER	CSV – Text-Datei mit Trennzeichen
	PRONOM PUIDs	x-fmt/18
	DATEI-ENDUNG	Verbindliche Datei-Endung: .csv
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	<p>RFC 4180: Common Format and MIME Type for Comma-Separated Values (CSV) Files</p> <p><u>Bezugsquellen:</u> RFC4180 - Shafranovich, Y., "Common Format and MIME Type for Comma-Separated Values (CSV) Files", RFC 4180, October 2005. http://www.ietf.org/rfc/rfc4180.txt</p>	
	<p>Erlaubte Zeichensatzkodierungen für CSV-Dateien:</p> <ul style="list-style-type: none"> • ISO Latin-1 (ISO 8859-1) und ISO Latin-9 (ISO 8859-15) • Unicode 5.0 Universal Character Set (UCS) (ISO 10646:2003) • US-ASCII "safe characters" (ANSI X3.4-1986) <p>Genaue Hinweise zu den genannten Zeichensatzkodierungen sind im Standard für "Nur-Text"-Daten aufgeführt.</p>	
	<p>Obwohl verschiedene Spezifikationen und Implementierungen für CSV existieren, gibt es keinen formalen Standard, der alle existierenden Varianten von CSV-Dateien handhaben kann. Das Memo RFC 4180 beschreibt das Format, welches von den meisten Implementierungen verstanden wird. Dieses ist die Basis des BAR-Standards.</p> <p>Die Zeilen der Tabelle werden im folgenden "Records" und die Kolonnen "Felder" genannt. In einer CSV-Datei sind die Felder durch ein Trennzeichen (meist ein Komma) getrennt.</p>	
	<ol style="list-style-type: none"> 1. Jeder Record ist grundsätzlich auf einer Zeile (Ausnahme: s. Punkt 6), die durch einen Zeilenumbruch terminiert wird (ASCII LF oder ASCII CRLF oder ASCII CR), z.B.: <pre>aaa,bbb,ccc CRLF xxx,yyy,zzz CRLF</pre> 2. Der letzte Record im File kann aber muss nicht durch einen Zeilenumbruch abgeschlossen sein: <pre>aaa,bbb,ccc CRLF xxx,yyy,zzz</pre> 3. Empfehlenswert ist, die erste Zeile als Header (Kopfzeile) darzustellen, mit gleichem Format wie die übrigen Zeilen. Im Header stehen die Namen der Felder der Tabelle. Das CSV-Format beinhaltet keine Information über die Präsenz eines Headers. Diese muss extern bereitgestellt werden. (z.B. mit dem optionalen "header" Parameter des MIME Typs). Beispiel: <pre>Feldname1,Feldname2,Feldname3 CRLF aaa,bbb,ccc CRLF xxx,yyy,zzz CRLF</pre> 4. Innerhalb des Headers und jedes Records stehen mehrere Felder, die durch Kommas voneinander getrennt sind. Jede Zeile muss die gleiche Anzahl Felder beinhalten. Leerzeichen sind von Bedeutung und dürfen nicht ignoriert werden. Das letzte Feld darf nicht von einem Komma gefolgt sein: <pre>aaa,bbb,cc cc,ddd</pre> 5. Jedes Feld kann, aber muss nicht, durch Anführungszeichen (double quotes) eingefasst sein. Falls Felder nicht mit Anführungszeichen eingefasst sind, dürfen innerhalb der Felder keine Anführungszeichen auftreten: <pre>"aaa",bbb,"ccc" CRLF xxx,yyy,zzz CRLF</pre> 6. Felder, die Zeilenumbrüche (CRLF), Anführungszeichen oder Kommas enthalten, müssen in Anführungszeichen eingefasst sein: <pre>"aaa","b CRLF bb","ccc" CRLF xxx,yyy,"z,zz" CRLF</pre> 	

	<p>Nach Möglichkeit ist dieser Spezialfall zu vermeiden, weil er von vielen Programmen (z.B. Excel, Access) falsch interpretiert wird.</p> <p>7. Falls ein Anführungszeichen innerhalb eines Feldes auftritt, muss es durch ein vorangehendes markiert sein. Das Feld muss zudem mit Anführungszeichen eingefasst sein: "aaa", "b""bb", "ccc" CRLF</p>										
	<p>Zulässige BAR-Erweiterung zum RFC 4180</p> <p>8. Ein anderes Trennzeichen (Delimiter) als Komma kann gewählt werden. Es gelten analog die obigen Regeln. Folgende Trennzeichen sind gestattet:</p> <table border="1"> <thead> <tr> <th><u>Zeichen</u></th> <th><u>ASCII-Code</u></th> </tr> </thead> <tbody> <tr> <td>,</td> <td>0x2C</td> </tr> <tr> <td>;</td> <td>0x3B</td> </tr> <tr> <td> </td> <td>0x7C</td> </tr> <tr> <td>#</td> <td>0x23</td> </tr> </tbody> </table> <p>Grundsätzlich muss bei Abweichungen, die über Regeln 1.-8. hinausgehen, das Bundesarchiv vorgängig kontaktiert werden.</p>	<u>Zeichen</u>	<u>ASCII-Code</u>	,	0x2C	;	0x3B		0x7C	#	0x23
	<u>Zeichen</u>	<u>ASCII-Code</u>									
	,	0x2C									
;	0x3B										
	0x7C										
#	0x23										
<p>EIGENTÜMER</p> <p>RFC Standards: The Internet Engineering Task Force http://www.ietf.org</p>											
<p>ANWENDUNGSGEBIETE</p> <p>Tabellenartig organisierte Daten. Excel-Tabellen, wenn der Inhalt der Tabelle von Bedeutung ist (im Gegensatz zur Darstellung). Einzelne Tabellen aus kleinen Datenbanken (MS Access, MySQL, etc.). Falls die Datenbanken umfangreich sind oder mehrere Tabellen enthalten, die miteinander in Beziehung stehen, ist SIARD zu verwenden (s. Standard für relationale Datenbanken).</p>											
<p>VORGABEN UND EINSCHRÄNKUNGEN</p>	<p>VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG</p> <p>Zeilen einer CSV-Datei müssen immer die gleiche Anzahl Felder enthalten. CSV-Dateien, die von dieser Regel abweichen, sind nicht erlaubt.</p>										
	<p>EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS</p> <p>-</p>										
<p>EMPFEHLUNGEN UND HINWEISE</p>	<p>BEMERKUNGEN UND HINWEISE</p> <p>Besteht bei Excel- bzw. kleineren Datenbanken Unklarheit, ob CSV oder SIARD geeigneter ist, erteilt das Bundesarchiv gerne Auskunft.</p>										

SIARD – Software Independent Archiving of Relational Databases

FORMAT	VERBINDLICHER BEZEICHNER	SIARD Format
	PRONOM PUIDs	fmt/161 (SIARD 1.0), fmt/1196 (SIARD 2.1)
	DATEI-ENDUNG	Verbindliche Datei-Endung: .siard
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	eCH-0165: SIARD-Formatspezifikation: Version 2.1 <u>Bezugsquelle:</u> https://www.bar.admin.ch/bar/de/home/archivierung/tools---hilfsmittel/siard-suite.html oder https://kost-ceco.ch/cms/index.php?siard_de	
	eCH-0165: SIARD-Formatspezifikation: Version 1.0 <u>Bezugsquelle:</u> https://www.ech.ch/vechweb/page?p=dossier&documentNumber=eCH-0165&documentVersion=1.0	
VORGABEN UND EINSCHRÄNKUNGEN	EIGENTÜMER	
	©2005-2018 Schweizerisches Bundesarchiv	
	ANWENDUNGSGEBIETE	
	Relationale Datenbanken	
VORGABEN UND EINSCHRÄNKUNGEN	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Enthalten die Tabellen Kolonnen vom Typ BLOB (binary large object), müssen darin abgelegte Dateien ebenfalls den vom Bundesarchiv vorgegebenen archivtauglichen Formaten entsprechen.	
	Version 2.0 wird durch das Bundesarchiv nicht akzeptiert.	
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	Für die Archivierung von relationalen Datenbanken wird der Einsatz der SIARD Suite empfohlen. <u>Bezugsquelle:</u> Schweizerisches Bundesarchiv. https://www.bar.admin.ch/bar/de/home/archivierung/tools---hilfsmittel/siard-suite.html	
	Das Tool csv2siard erlaubt die Konvertierung von CSV-Dateien in eine SIARD-Datei. <u>Bezugsquelle:</u> Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST). http://kost-ceco.ch/cms/index.php?csv2siard_de	
	BEMERKUNGEN UND HINWEISE	
<p>Eine SIARD-Datei ist in aller Regel nicht ausreichend selbstdokumentierend um die Verständlichkeit der archivierten Daten gewährleisten zu können. Zusätzlich zur SIARD-Datei ist daher eine entsprechende Dokumentation zum Quellsystem und zum Datenmodell zu archivieren. Es ist für jeden Fall zu entscheiden, was alles dokumentiert werden muss. Teile der Dokumentation können sein (Beispiele):</p> <ul style="list-style-type: none"> • Grafisches Datenmodell (Entity-Relationship-Diagramm, ERD) • Datenbeschreibungen / Codelisten • Systemspezifikation / Systembeschreibung • Benutzerhandbuch / Benutzungsreglemente / Schulungsunterlagen / Screenshots 		

TIFF – Tagged Image File Format

FORMAT	VERBINDLICHER BEZEICHNER	TIFF – Tagged Image File Format
	PRONOM PUIDs	fnt/353
	DATEI-ENDUNG	Empfohlene Datei-Endungen: .tif, .tiff
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	Revision 6.0 vom 1992-06-03 und Revision 6.0.1 vom 1995-10-15. Bezugsquelle: http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf	
	EIGENTÜMER	
©1986-1988, 1992 Adobe Systems Inc., U.S.A.		
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE	
	Rasterbilder (Schwarz-Weiss, Graustufen, Farbe)	
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	TIFF-Dateien müssen als TIFF 6 validiert sein.	
	<p>PROPRIETÄRE ERWEITERUNGEN: Format-Erweiterungen durch Software-Hersteller sind nur zulässig, wenn sie sich strikt an die diesbezüglichen Anforderungen des Standards halten. Im Zweifelsfall sind durch die abliefernde Stelle schriftliche Bestätigungen des Software-Herstellers einzuholen. Insbesondere müssen proprietäre Erweiterungen („private fields and values“) oberhalb der Tag-Nummer resp. Konstanten-Nummern 32767 implementiert sein. Achtung: Auch wenn sie konform zum Standard implementiert sind, werden diese Erweiterungen vom Bundesarchiv prinzipiell ignoriert.</p> <p>Insbesondere werden auch die „TIFF Enhancements for Adobe Photoshop®“ ignoriert, was vor allem bedeutet, dass der von Adobe™ Photoshop® geschriebene Image-Source Tag 37724 ignoriert wird. Die Verwendung der zwei weiteren „Advanced TIFF“-Optionen von Adobe™ Photoshop® – ZIP/zlib- und JPEG-Kompressionen – ist ausgeschlossen. Ebenfalls ausdrücklich ignoriert werden die „TIFF Enhancements for Adobe™ PageMaker® 6.0“ und die „Kodak™ TIFF Extensions“.</p> <p>Adobe™ Photoshop „TIFF Enhancements für Adobe™ Photoshop®“: http://partners.adobe.com/public/developer/en/tiff/TIFFphotoshop.pdf</p> <p>TIFF Enhancements for Adobe™ PageMaker® 6.0: http://partners.adobe.com/public/developer/en/tiff/TIFFPM6.pdf</p> <p>Kodak™ TIFF Extensions: wenden Sie sich direkt an die Eastman Kodak Company.</p> <p>Für andere Software: siehe Dokumentation des Herstellers</p>	
	MULTIPAGE-TIFF: Die Einbindung von mehreren Seiten in eine einzelne TIFF-Datei (Multipage-TIFF, mehrere IFDs) ist nicht erlaubt.	
	Wenn dasselbe Bild in unterschiedlicher Auflösung oder Bittiefe archiviert werden soll, sind verschiedene Dateien in verschiedenen Ordnern herzustellen, damit die verschiedenen Versionen leicht separat bestellt werden können.	
	Wenn ein Dokument mit mehreren Seiten archiviert werden soll, ist es im PDF/A-Format abzuspeichern.	
	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	-	
EMPFEHLUNGEN UND HINWEISE	BEMERKUNGEN UND HINWEISE	
	<p>TIFF benutzt 4-Byte File-Offsets, weshalb die Bild-Datenmenge einer TIFF-Datei max. 4 GB (komprimiert) betragen kann.</p> <p>Liegen TIFF-Dateien im Format „GeoTIFF“ (http://trac.osgeo.org/geotiff/) vor, ist eine Konversion in das archivtaugliche Format TIFF+EWF.XML vorzunehmen.</p>	

JPEG2000 - Joint Photographic Experts Group 2000

FORMAT	VERBINDLICHER BEZEICHNER	JPEG2000
	PRONOM PUIDs	x-fmt/392
	DATEI-ENDUNG	Empfohlene Datei-Endungen: .jp2
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	ISO/IEC 15444 ISO/IEC 15444-1:2016 Information technology – JPEG 2000 image coding system – Part 1: Core coding system <u>Bezugsquelle:</u> International Organization for Standardization	
	Einige Elemente des Standards unterliegen Patentrechten gemäss der Common Patent Policy for ITU-T/ITU-R/ISO/IEC : Die Patentinhaber haben sich darauf verpflichtet, den Standard gebührenfrei zu halten.	
	EIGENTÜMER	
ISO/IEC Standards: International Organization for Standardization Joint Photographic Experts Group		
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE	
	JPEG 2000 ist ein Standard der Joint Photographic Experts Group zur verlustlosen und verlustbehafteten Codierung von Rasterbildern in hoher Qualität. Der Standard eignet sich für eine Vielfalt von Anwendungen, von digitalen (Überwachungs-) Kameras über Scanner bis zu hochauflösenden bildgebenden medizinischen Systemen. JPEG 2000 unterstützt sehr grosse Bilder.	
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Zugelassen ist nur das Dateiformat JP2 (Part 1, Annex I, ISO/IEC 15444), das die Speicherung eines einzelnen Bildes ermöglicht. Der in der JP2-Datei enthaltene JPEG 2000 Codestream muss den Restriktionen des Profile-0 ("J2P0") genügen. (Der aktuelle Standard ISO/IEC 15444-1:2016 umfasst mehrere Profile, die im bisherigen Standard ISO/IEC 15444-1:2004 nicht enthalten waren; die neuen Profile betreffen das digitale Kino, sie sind für das Archiv nicht relevant.) Das Dateiformat JPX (Part 2, Annex L und Annex M) ist nicht zugelassen. Es ermöglicht die Speicherung eines Bildes mit erweiterten Eigenschaften, deren Wiedergabe mit herkömmlichen Programmen nicht gewährleistet ist. Das Compound Image File Format JPM (Part 6) ist ebenfalls nicht zugelassen. Es ermöglicht die Speicherung von Bildern, die aus mehreren einzelnen Komponenten zusammengesetzt sind. Die Wiedergabe der zusammengesetzten Bilder kann mit herkömmlichen Programmen nicht gewährleistet werden.	
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	Nicht alle Programme, mit denen Bilder im Format JPEG 2000 gespeichert werden können, produzieren standardkonforme .jp2-Dateien. Zur Validierung kann das Programm Jpylyzer verwendet werden (http://jpylyzer.openpreservation.org).	
	BEMERKUNGEN UND HINWEISE	
	Für einzelne Bilder wird die verlustfreie (lossless) Codierung empfohlen. Für eine grössere Menge an Bildern wird eine minimal verlustbehaftete (minimal lossy) Kompression empfohlen. JPEG 2000 ist ein komplexes Format. Entsprechend zahlreich sind die Konfigurationsmöglichkeiten. Sollen grosse Mengen an Bildern mit JPEG 2000 codiert werden, empfiehlt es sich, die geeignetste Konfiguration mit Tests zu ermitteln. Dem Bundesarchiv ist bekannt, dass die offizielle Referenzsoftware Open JPEG 2000 mangelhafte Resultate liefern kann. Es kann daher empfehlenswert sein, professionelle Alternativen zu prüfen.	

WAVE – Audioformat von Microsoft

FORMAT	VERBINDLICHER BEZEICHNER	WAVE ALIAS BEZEICHNER: WAVEFORMAT, PCMWAVEFORMAT								
	PRONOM PUIDs	fmt/1, fmt/2, fmt/6, fmt/141								
	DATEI-ENDUNG	Verbindliche Datei-Endung: .wav								
	TYP	Dateiformat und Datenformat								
	VERBINDLICHE SPEZIFIKATIONEN									
	<p>Es existiert kein publizierter Standard für WAVE Dateien. Das WAVE-Format ist eine Implementierung des Resource Interchange File Formats (RIFF) der Microsoft Corporation. Dieses ist als Publikation freigegeben.</p> <p>Bezugsquellen: Multimedia Programming Interface and Data Specifications 1.0, herausgegeben durch IBM Corporation und Microsoft Corporation, August 1991. WAVEFORMAT (structure) from Microsoft.</p>									
	EIGENTÜMER									
©1991 Microsoft Corporation										
VORGABEN UND EINSCHRÄNKUNG	ANWENDUNGSGEBIETE									
	Audiodatei									
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG									
Erlaubt ist nur der default Microsoft Linear Pulse Code Modulation (LPCM) Codec. Byte-Offset 20 (0x14) muss den Wert 1 als 2-Byte Wert (short) enthalten (in der little-endian byte order).										
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS									
	Herstellung aus <u>analogen</u> Quellen									
	Folgende Abtastrate und Quantifizierungen sind bei Digitalisierungen aus analogen Quellen empfohlen:									
	<table border="1"> <thead> <tr> <th>Abtastrate</th> <th>Quantifizierung</th> </tr> </thead> <tbody> <tr> <td>48 KHz</td> <td>16 oder 24 bits</td> </tr> </tbody> </table>		Abtastrate	Quantifizierung	48 KHz	16 oder 24 bits				
	Abtastrate	Quantifizierung								
48 KHz	16 oder 24 bits									
Herstellung aus <u>digitalen</u> Quellen										
Liegen die Audiodaten bereits in digitaler Form vor, soll die bestehende Abtastrate und Quantifizierung beibehalten werden. Bei CD Audio und Audio DAT sind dies:										
<table border="1"> <thead> <tr> <th>Original</th> <th>Abtastrate</th> <th>Quantifizierung</th> </tr> </thead> <tbody> <tr> <td>CD Audio</td> <td>44.1 KHz</td> <td>16 bits</td> </tr> <tr> <td>DAT</td> <td>44.1 KHz / 48 KHz</td> <td>16 bits</td> </tr> </tbody> </table>		Original	Abtastrate	Quantifizierung	CD Audio	44.1 KHz	16 bits	DAT	44.1 KHz / 48 KHz	16 bits
Original	Abtastrate	Quantifizierung								
CD Audio	44.1 KHz	16 bits								
DAT	44.1 KHz / 48 KHz	16 bits								
BEMERKUNGEN UND HINWEISE										
Das WAVE-Format ist sehr nahe verwandt mit dem CD-Audio (CDA-)-Format. Dessen Dokumentation ist allerdings äusserst schwer zugänglich. Gängige Programme (Nero etc.) verwandeln CDA <-> WAVE.										

FFV1 Video Codec

FORMAT	VERBINDLICHER BEZEICHNER	FFV1 – FF Video Codec 1
	PRONOM PUIDs	fmt/569 (Matroska-Container ³)
	DATEI-ENDUNG	Verbindliche Datei-Endungen: .mkv
	TYP	Kompressionsverfahren (Codec) für Video
	VERBINDLICHE SPEZIFIKATIONEN	
	FFV1 Video Codec Specification Bezugsquelle: http://www.ffmpeg.org/~michael/ffv1.html	
	Draft IETF (Internet Engineering Task Force) Specification Bezugsquelle: https://tools.ietf.org/pdf/draft-niedermayer-cellar-ffv1-01.pdf	
	Matroska Specifications Bezugsquelle: https://www.matroska.org/technical/specs/index.html	
	EIGENTÜMER	
	Michael Niedermayer (Projekt FFmpeg) Das Format ist Public Domain und explizit lizenzfrei.	
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE	
	Videodaten	
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Das BAR akzeptiert nur FFV1-Dateien in der Version 3 (FFV1.3) von 2013. Als Container muss zwingend MKV (Matroska) verwendet werden. Für Audio muss zwingend WAVE als Codec verwendet werden (siehe WAVE).	
	Die „GOP size“ (ffmpeg Argument „-g“) muss zwingend auf 1 gesetzt sein.	
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	Für beste Ergebnisse sollten FFV1-Videos nach Möglichkeit direkt aus dem Originalen Videomaterial (uncompressed Video) erzeugt werden.	
	Das Slicing (ffmpeg Argument „-slices“) ist im Zweifelsfall lieber kleiner als grösser zu halten, weil ein grosser Slicing-Wert die Dateigrösse unnötig erhöhen kann, ohne dass im Gegenzug zusätzliche Encoding-Performance gewonnen wird. Um die Dateigrösse weiter zu reduzieren, kann ein Multi-pass Encoding verwendet werden (ffmpeg Argument „-pass“). Idealerweise wird das ffmpeg Argument „-slicecrc 1“ gesetzt, um zu jedem Slice Fehlererkennungs-Informationen zu speichern.	
	BEMERKUNGEN UND HINWEISE	
FFV1 ist ein verlustfreier intra-frame Videocodec (jedes Einzelbild wird komprimiert), der speziell für die Archivierung entwickelt wurde. FFV1 ist ein reiner Bildcodec, Ton wird separat als WAVE im Matroska-Container untergebracht. FFV1 unterstützt unterschiedliche Bildformate und Farbtiefen ohne Farbrunterabtastung, zudem mehrere Farbräume (intern YCbCr für YUV Daten und JPEG 2000 RCT für RGB Daten).		

³ Der FFV1-Datenstrom wird in einem Matroska-Container verpackt

MPEG-4 – Videoformat

FORMAT	VERBINDLICHER BEZEICHNER	MPEG-4
	PRONOM PUIDs	fmt/199
	DATEI-ENDUNG	Verbindliche Datei-Endungen: .mp4 .mp4v
	TYP	Dateiformat (Container) und Kompressionsverfahren (Codec) für Video und Audio
	VERBINDLICHE SPEZIFIKATIONEN	
	ISO/IEC 14496-10 Coding of audio-visual objects -- Part 10: Advanced Video Coding ISO/IEC 14496-3 Coding of audio-visual objects -- Part 3: Audio ISO/IEC 14496-14 Coding of audio-visual objects -- Part 14: MP4 file format ISO/IEC 14496-17 Coding of audio-visual objects -- Part 17: Timed Text subtitle format Bezugsquellen: International Organization for Standardization	
	EIGENTÜMER	
	ISO/IEC Standards: International Organization for Standardization	
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE	
	Videodaten mit begrenzten Qualitätsanforderungen	
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Folgende Codecs sind erlaubt: <ul style="list-style-type: none"> • Video: MPEG-4 part 10 (auch als MPEG-4 AVC oder ITU H.264 bezeichnet) • Audio: MPEG-4 AAC (Advanced Audio Coding), beschrieben in MPEG-4 part 3 Folgender Container für die Video- und Audio-Ströme ist erlaubt: <ul style="list-style-type: none"> • MP4, beschrieben in MPEG-4 part 14 Im Weiteren sind, falls vorhanden, auch Untertitel (timed text subtitles) zugelassen.	
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	MPEG-4 stellt immer einen Kompromiss bezüglich Qualität und Erhaltbarkeit dar und wird vom BAR explizit nur empfohlen, wenn Ausgangsmaterial nur als MPEG-4 vorliegt oder erhebliche Mengen von Videodaten mit geringer Qualitätsanforderung archiviert werden sollen. Im Zweifelsfall bitte vorgängig das Bundesarchiv kontaktieren.	
	Herstellung aus <u>analogen</u> Quellen Der Sampler soll direkt aus den analogen Quellen das native MPEG-4 Format herstellen. Keinesfalls darf als Zwischenschritt ein anderes, stark komprimierendes Videoformat zum Einsatz gelangen (wie z.B. WMC oder VC1).	
	Herstellung aus <u>digitalen</u> Quellen Liegen die Videodaten bereits in digitaler Form vor, soll die bestehende Kodierung möglichst beibehalten werden. <p>a) der Video Codec ist MPEG-4</p> Falls das Dateiformat bereits MPEG-4 part 14 ist, muss nichts weiter unternommen werden. Andere Dateiformate müssen in MPEG-4 part 14 umgepackt werden. Dabei ist darauf zu achten, dass keine erneute Kodierung stattfindet. <p>b) der Video Codec ist nicht MPEG-4</p> Grundsätzlich müssen die Video- und Audioströme in MPEG-4 AVC bzw. MPEG-4 AAC transkodiert werden. Da dies jedoch zu erheblichen Verlusten führen kann, muss das genaue Vorgehen vorgängig mit dem Bundesarchiv abgesprochen werden. Bei einer Transkodierung muss die Auflösung des Videoinhalts beibehalten werden, es darf kein Resizing vorgenommen werden.	

XML/XSD - eXtensible Markup Language

FORMAT	VERBINDLICHER BEZEICHNER	XML - eXtensible Markup Language
	PRONOM PUIDs	fmt/101 (XML), x-fmt/280 (XSD)
	DATEI-ENDUNG	Verbindliche Datei-Endungen: .xml, .xsd
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	Extensible Markup Language (XML) 1.0 (Fifth Edition), 2008 <u>Bezugsquelle: W3 Consortium</u>	
	Erlaubte Zeichensatzkodierungen: <ul style="list-style-type: none"> Unicode 5.0 Universal Character Set (UCS) (ISO 10646:2003) 	
VORGABEN UND EINSCHRÄNKUNGEN	EIGENTÜMER	
	W3 Consortium	
	XML ist ein frei zugänglicher Standard	
	ANWENDUNGSGEBIETE	
	XML ist ein einfaches, flexibles Textformat. XML wird zur Beschreibung strukturierter Daten und Spezifikation der Struktur bzw. der Bedeutung von Daten eingesetzt und ist ein beliebtes Format für den Austausch strukturierter Daten.	
EMPFEHLUNGEN UND HINWEISE	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Diese Vorgaben beziehen sich nur auf XML-Texte mit nicht-standardisierten Vokabularen. Für XML-Texte mit standardisierten Vokabularen sind die entsprechenden Spezifikationen massgebend (z.B. SVG, RDF/XML). <u>Wichtig:</u> Das BAR akzeptiert derzeit nur in Ausnahmefällen XML-Dateien mit standardisierten Vokabularen (z.B. SVG, RDF/XML), da dies derzeit keine akzeptierten archivtauglichen Dateiformate sind.	
	Der XML-Text muss wohlgeformt sein, d.h. den Vorgaben der Grammatik, dem Vokabular und der Syntax von XML entsprechen.	
	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
EMPFEHLUNGEN UND HINWEISE	XML-Schemadefinitionen (XSD) sollen wenn immer möglich mit den XML-Dateien mitgeliefert werden. Als Ausnahme von dieser Regel können weit verbreitete und frei verfügbare XSD betrachtet werden, z.B. XHTML. Im Zweifelsfall ist das XSD mitzuliefern.	
	BEMERKUNGEN UND HINWEISE	
	Um die spätere (Nach-) Nutzung der Inhalte der XML-Datei zu ermöglichen, müssen das verwendete, nicht-standardisierte Vokabular und die korrekte Datenstruktur unmissverständlich dokumentiert sein. Diese Dokumentation kann umgangssprachlich formuliert und mit Grafiken ergänzt sein.	
Die Wohlgeformtheit eines XML-Textes kann mit einem XML-Validator (z.B. https://validator.w3.org/) geprüft werden.		

TIFF+EWF.XML – Tagged Image File Format und Extended World File

FORMAT	VERBINDLICHER BEZEICHNER	TIFF+EWF.XML – Tagged Image File Format und Extended World File
	PRONOM PUIDs	fmt/353 (TIFF), fmt/101 (XML), x-fmt/280 (XSD)
	DATEI-ENDUNG	Empfohlene Datei-Endungen: .tif, .tiff Verbindliche Datei-Endungen: .xml, .xsd
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	„Merkblatt Spezifikation Archivtaugliches Geformat für Bild- und Grafikrasterdaten“ <u>Bezugsquelle</u> : BAR	
	EIGENTÜMER	
	TIFF: ©1986-1988, 1992 Adobe Systems Inc., U.S.A. XML ist ein frei zugänglicher Standard TIFF+EWF.XML: BAR	
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE	
	TIFF+EWF.XML ist ein durch das BAR und swisstopo gemeinsam erarbeitetes Format zur Archivierung von Bild- und Grafikrasterdaten, Thematischen Rasterdaten und Höhenrasterdaten mit einem geographischen Bezug.	
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Spezifikation und Vorgaben sind im Dokument „Merkblatt Spezifikation Archivtaugliches Geformat für Bild- und Grafikrasterdaten“ beschrieben. Dieses kann beim BAR bezogen werden. Es gelten die Vorgaben der archivtauglichen Formate TIFF und XML.	
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	Es gelten die Empfehlungen der archivtauglichen Formate TIFF und XML.	
	BEMERKUNGEN UND HINWEISE	
Es gelten die Bemerkungen und Hinweise der archivtauglichen Formate TIFF und XML.		

INTERLIS

FORMAT	VERBINDLICHER BEZEICHNER	INTERLIS – The GeoLanguage
	PRONOM PUIDs	fmt/1014, fmt/1012, fmt/654 (INTERLIS 1, 2.2, 2.3 Model Files), fmt/1013, fmt/1011, fmt/653 (INTERLIS 1, 2.2, 2.3 Transfer Files)
	DATEI-ENDUNG	Empfohlene Datei-Endungen: .xtf .xml (INTERLIS 2.x Transfer Files) Verbindliche Datei-Endungen: .ili (Model Files), .itf (INTERLIS 1 Transfer File)
	TYP	Dateiformat und Datenformat
	VERBINDLICHE SPEZIFIKATIONEN	
	INTERLIS 1 Referenzhandbuch (SN 612 030) INTERLIS 2.2 Referenzhandbuch (ehem. SN 612 031) INTERLIS 2.3 Referenzhandbuch (eCH-0031; SN 612 031)	
	EIGENTÜMER	
	Die INTERLIS-Spezifikation wird von der KOGIS (Koordination, Geo-Information und Services) gepflegt und weiterentwickelt.	
VORGABEN UND EINSCHRÄNKUNGEN	ANWENDUNGSGEBIETE	INTERLIS wird verwendet für thematischen Vektordaten und Höhenvektordaten mit einem geographischen Bezug sowie für Geodatenmodelle.
	VERBINDLICHE EINSCHRÄNKUNGEN DER ANWENDUNG	
	Für die XML-basierten Formate gelten die Vorgaben des archivtauglichen Formates XML.	
EMPFEHLUNGEN UND HINWEISE	EMPFEHLUNGEN DES BUNDESARCHIVS ZUR HERSTELLUNG DIESES ARCHIVFORMATS	
	Für die XML-basierten Formate gelten die Empfehlungen des archivtauglichen Formates XML.	
	BEMERKUNGEN UND HINWEISE	
		Für die XML-basierten Formate gelten die Bemerkungen und Hinweise des archivtauglichen Formates XML.