



Leidraad

Voorkeursbestandsformaten

Versie 1.1

Datum: 10-08-2022

Kenmerk: 2021-T&A-053



Versiebeheer

Versie	Datum	Auteur (s)	Toelichting
0.1	26-05-2021	Julia Emmen	Initiële versie
0.2	10-06-2021	Julia Emmen	Aangepast concept
0.3	05-06-2021	Julia Emmen	Aangepast concept nav feedback intern
1.0	20-07-2021	Julia Emmen	Definitieve versie
1.1	10-08-2022	Julia Emmen & Guido Dorssers	Jaarlijkse review 2022, zie Bijlage 1: Logboek kleine veranderingen



Inhoudsopgave

1. Inleiding	4
2. Wettelijk kader	4
3. Informatiesoorten	4
4. Standaarden	5
5. Codecs	5
6. Categorieën bestandsformaten	6
7. Toelichting bij de lijst met voorkeursbestandsformaten	6
8. Lijst met voorkeursbestandsformaten	8
9. Lijst met voorkeurscodecs	13
10. Bijlage 1: Logboek kleine veranderingen	14

1. Inleiding

Digitale bestanden komen tegenwoordig in allerlei informatiesoorten en bestandsformaten. Elk informatiesoort heeft zijn eigen bestandsformaten, en elk bestandsformaat heeft zijn eigen versies. Ieder bestandsformaat heeft voor- en nadelen ten opzichte van een ander bestandsformaat.

Omdat conserveren van digitale bestanden met sommige bestandsformaten makkelijker gaat dan met andere, heeft het HCL een lijst opgesteld met voorkeursbestandsformaten. Deze lijst met voorkeursbestandsformaten geeft aan welke bestandsformaten het beste in aanmerking komen om de informatie in het bestand langdurig te bewaren. Toepassing van deze voorkeursbestandsformaten draagt dus bij aan de duurzaamheid van de digitale bestanden.

Deze leidraad heeft een dynamisch karakter. Dit betekent dat de leidraad kan worden aangepast zodra veranderende wet- en regelgeving, digitale ontwikkeling en/of voortschrijdend inzicht daar om vraagt. Het HCL controleert jaarlijks of de leidraad voorkeursbestandsformaten nog actueel is en zal, indien nodig, de leidraad aanpassen. Deze aanpassingen worden bijgehouden in het logboek, zie Bijlage 1: **Logboek kleine veranderingen**.

2. Wettelijk kader

De Archiefregeling noemt geen specifieke bestandsformaten voor digitale informatieobjecten. In plaats daarvan stelt de Archiefregeling algemene eisen aan bestandsformaten. In artikel 26, eerste lid, van de Archiefregeling staat het volgende:

"Digitale archiefbescheiden worden, uiterlijk op het tijdstip van overbrenging, opgeslagen in een valideerbaar en volledig gedocumenteerd bestandsformaat dat voldoet aan een open standaard, tenzij dit redelijkerwijs niet van de zorgdrager kan worden verlangd. Alsdan vindt met de beheerder van de voor overbrenging aangewezen archiefbewaarplaats overleg plaats over een alternatief bestandsformaat."

De wet schrijft een uitgangspunt voor waaraan informatieobjecten moeten voldoen op moment van overbrenging, en bepaalt daarmee het uitgangspunt voor de lijst met voorkeursbestandsformaten van het HCL.

3. Informatiesoorten

Aan elk informatiesoort zijn bepaalde eigenschappen gekoppeld zoals functie, eigenschappen, gedrag, opmaak en weergave. Vandaar dat het HCL voor de lijst met voorkeursbestandsformaten onderscheid maakt tussen verschillende informatiesoorten. Het HCL hanteert de volgende onderverdeling in informatiesoorten:

1. (Raster) afbeelding



2. Audio (geluid)
3. Database
4. Document (tekstgebaseerde documenten)
5. E-mail (berichten uit het elektronisch postverkeer)
6. Presentatie
7. Spreadsheet
8. Vectorafbeelding (een vectorafbeelding is een grafische voorstelling die opgebouwd is uit eenvoudige meetkundige objecten, zoals punten, lijnen, krommen, polygonen e.d.)
9. Video (bewegend beeld)
10. Website

4. Standaarden

Het HCL heeft een voorkeur voor bestandsformaten met een open standaard omdat open standaarden ten goede komen van de digitale duurzaamheid. Een bestandsformaat is een open standaard wanneer:

- Het bestandsformaat is vastgelegd in een specificatie die bekend en vrijelijk beschikbaar is. Het gaat hierbij om een laagdrempelige beschikbaarheid van het bestandsformaat.
- Het bestandsformaat is vastgelegd in een specificatie die wordt onderhouden door een standaardenorganisatie. Verdere ontwikkeling gebeurt op basis van een open besluitvormingsprocedure: een standaardenorganisatie biedt bij de doorontwikkeling van de specificatie voldoende inspraakmogelijkheden voor stakeholders.
- Het bestandsformaat is vastgelegd in een specificatie die door iedereen gebruikt en geïmplementeerd kan worden. Er mogen bijvoorbeeld bij gebruik geen hindernissen zijn op het gebied van het intellectueel eigendomsrecht.

De Archiefregeling (art. 26 lid 1 Ar) hanteert het principe "pas toe of leg uit". Dit houdt in dat het verantwoordelijk overheidsorgaan gebruik moet maken van valideerbare en volledig gedocumenteerde bestandsformaten die voldoen aan open standaarden. Als dit niet het geval is, is het verantwoordelijk overheidsorgaan verplicht om uit te leggen waarom er voor een alternatief is gekozen. Het HCL adviseert om bij de creatie van digitale informatieobjecten te kiezen voor een bestandsformaat dat voldoet aan een open standaard.

5. Codecs

Om een mediabestand te creëren dat leesbaar is door computersoftware, worden video- en audiostromen in een container verpakt. Mediacontainers genoemd in **8. Lijst met voorkeursbestandsformaten** zijn: MPEG-4 Media File (MP4), Material Exchange Format Operational Pattern (MXF), Audio/Video Interleaved

Format (avi), en Matroska. Een mediacontainer is 'verpakkingsmateriaal' waarin de verschillende gebruikte audio- en videostromen voorkomen. Een codec is een algoritme dat wordt gebruikt om audio- of videogegevens te comprimeren en te decomprimeren. Mediagegevens kunnen gecodeerd worden met verschillende codecs. Om deze reden heeft het HCL naast de lijst met voorkeursbestandsformaten ook een lijst opgesteld waarin de voorkeur van het HCL staat voor het gebruik van codecs, zie **9. Lijst met voorkeurscodecs**. Mocht het verantwoordelijk overheidsorgaan een codec willen gebruiken die niet op deze lijst staat, dan kan het verantwoordelijk overheidsorgaan het beste met het HCL in overleg treden.

6. Categorieën bestandsformaten

Het HCL deelt bestandsformaten in de lijst met voorkeursbestandsformaten op in drie categorieën:

- *Voorkeursbestandsformaten*: bestandsformaten die de voorkeur hebben van het HCL. Hieronder vallen de bestandsformaten die valideerbare en volledig gedocumenteerd zijn en voldoen aan open standaarden.
- *Acceptabele bestandsformaten*: bestandsformaten die niet de voorkeur hebben, maar die acceptabel zijn voor inname in het e-depot.
- *In overleg*: In deze categorie vallen bestandsformaten die in overleg met het HCL kunnen worden ingenomen in het e-depot.

Wanneer een bestandsformaat in de categorie *acceptabele bestandsformaten* valt, is het niet vanzelfsprekend beter om een informatieobject om te zetten naar een voorkeursbestandsformaat. Bij conversie, het omzetten van een informatieobject van het ene medium naar het andere of van het ene formaat naar het andere, kan namelijk informatieverlies optreden. Het toepassen van conversie het verantwoordelijk overheidsorgaan het beste doen in overleg met het HCL.

Wanneer bestandsformaten niet op de lijst met voorkeursbestandsformaten staan, worden deze niet bij voorbaat volledig uitgesloten voor inname in het e-depot. Het HCL gaat dan graag met het verantwoordelijk overheidsorgaan in overleg om te kijken naar een passende oplossing. Hierbij kan bijvoorbeeld gezamenlijk een acceptabel bestandsformaat bepaald worden.

7. Toelichting bij de lijst met voorkeursbestandsformaten

Ter introductie van de lijst met voorkeursbestandsformaten wordt hier een korte uitleg gegeven bij de betekenis van de kolommen:

- **Informatiesoort**: het HCL maakt onderscheid tussen verschillende digitale informatiesoorten, zoals (raster) afbeelding, audio, database, document, e-mail, presentatie, vectorafbeelding, video of website.



- **Extensie:** de (bestands)extensie is de toevoeging na de laatste punt in de bestandsnaam die door sommige besturingssystemen wordt gebruikt om te bepalen om wat voor bestandsformaat het gaat.
- **Naam:** de naam van het bestandsformaat.
- **Versie:** de specifieke versie van het bestandsformaat.
- **Categorie:** in welke categorie de bestandsformaten vallen: voorkeur, acceptabel of in overleg.
- **PRONOM ID:** het PRONOM-register is een digitaal register waarin bestandsformaten worden bijgehouden. Elke versie van een bestandsformaat krijgt een uniek PRONOM ID. Dit ID staat in deze kolom. Het PRONOM-register monitort welke bestandsformaten onleesbaar dreigen te worden. Het register wordt onderhouden door *The National Archives of the United Kingdom*.



8. Lijst met voorkeursbestandsformaten

INFORMATIE-SOORT	EXTENSIE	NAAM	VERSIE	CATEGORIE	PRONOM ID
(raster) Afbeelding					
	.png	Portable Network Graphics	1.0	Acceptabel	fmt/11
	.png	Portable Network Graphics	1.1	Acceptabel	fmt/12
	.png	Portable Network Graphics	1.2	Voorkeur	fmt/13
	.jpeg	JPEG File Interchange Format	1.00	Acceptabel	fmt/42
	.jpeg	JPEG File Interchange Format	1.01	Acceptabel	fmt/43
	.jpeg	JPEG File Interchange Format	1.02	Acceptabel	fmt/44
	.jp2	JP2 (JPEG 2000 part 1) ¹		Acceptabel	x-fmt/392
	.jpx	JPX (JPEG 2000 part 2) ²		Acceptabel	fmt/151
	.tiff	Tagged Image File Format	6.0	In overleg	fmt/353
Audio					
	.wav .wave	Waveform Audio File Format ³		Acceptabel	fmt/141
	.wav .wave	Waveform Audio File Format ⁴		Voorkeur	fmt/142
	.wav	Broadcast Wave Format	0	Acceptabel	fmt/1
	.wav	Broadcast Wave Format	1	Acceptabel	fmt/2
	.wav	Broadcast Wave Format	2	Voorkeur	fmt/527
	.mp3	MPEG 1/2 Audio Layer 3 ⁵		In overleg	fmt/134

¹ ISO 15444-1:2019

² ISO 15444-2:2004

³ PCMWAVEFORMAT

⁴ WAVEFORMATEX

⁵ ISO 11172-3:1995



INFORMATIE-SOORT	EXTENSIE	NAAM	VERSIE	CATEGORIE	PRONOM ID
	.mka	Matroska	1 2 3 4	In overleg	fmt/569
Database					
	.sql	Structured Query Language Data	SQL 2019	Voorkeur	fmt/206
	.siard	SIARD (Software-Independent Archiving of Relational Databases)	1.0	Acceptabel	fmt/161
	.siard	SIARD (Software-Independent Archiving of Relational Databases)	2.0	Acceptabel	fmt/995
	.siard	SIARD (Software-Independent Archiving of Relational Databases)	2.1	Voorkeur	fmt/1196
	.odb	OpenDocument Database Format	1.0	Acceptabel	fmt/140
	.odb	OpenDocument Database Format	1.1	Acceptabel	fmt/444
	.odb	OpenDocument Database Format	1.2	Voorkeur	fmt/424
	.mdb	Microsoft Access Database	2000	In overleg ⁶	fmt/240
	.mdb	Microsoft Access Database	2002	In overleg	fmt/241
	.accdb	Microsoft Access Database	2007	In overleg	fmt/275
Document					
	.pdf	PDF 1.7	1.7	Acceptabel	fmt/276
	.pdf	PDF/A-1	1a	Voorkeur ⁷	fmt/95

⁶ Microsoft Access Databases zijn opgeslagen in een gesloten binair formaat, dit levert problemen op met digitale duurzaamheid. Graag in overleg treden wanneer u gebruik wil maken van dit bestandsformaat.

⁷ Voorkeur voor digital born documenten



INFORMATIE-SOORT	EXTENSIE	NAAM	VERSIE	CATEGORIE	PRONOM ID
	.pdf	PDF/A-1	1b	Voorkeur ⁸	fmt/354
	.pdf	PDF/A-2	2a	Voorkeur ⁹	fmt/476
	.pdf	PDF/A-2	2b	Acceptabel	fmt/477
	.pdf	PDF/A-2	2u	Voorkeur ¹⁰	fmt/478
	.pdf	PDF/A-3	3a	Acceptabel	fmt/479
	.pdf	PDF/A-3	3b	Acceptabel	fmt/480
	.pdf	PDF/A-3	3u	Acceptabel	fmt/481
	.pdf	PDF/A-4	4	Acceptabel	
	.pdf	PDF/A-4	4f	Acceptabel	
	.pdf	PDF/A-4	4e	Acceptabel	
	.docx	Office Open XML	ISO 29500-4:2016	Acceptabel	fmt/412
	.doc	Microsoft Office Word 97-2007 Binary File Format	97-2003	In overleg	fmt/40
	.doc	Microsoft Office Word 97-2007 Binary File Format	97-2007	In overleg	
	.odt	Open Document Format for Office Applications	1.0	Acceptabel	fmt/136
	.odt	Open Document Format for Office Applications	1.1	Acceptabel	fmt/290
	.odt	Open Document Format for Office Applications	1.2	Voorkeur	fmt/291
	.odt	Open Document Format for Office Applications	1.3	Voorkeur	
E-mail					
	.msg	Outlook Item (.msg) File Format	14.0	In overleg	x-fmt/430
	.pst	Microsoft Outlook Personal Folders File (ANSI)	1997-2002	In overleg	x-fmt/248

⁸ Voorkeur voor gedigitaliseerde documenten

⁹ Voorkeur voor digital born documenten

¹⁰ Voorkeur voor gedigitaliseerde documenten



INFORMATIE-SOORT	EXTENSIE	NAAM	VERSIE	CATEGORIE	PRONOM ID
	.pst	Microsoft Outlook Personal Folders File (Unicode)	2003-2007	In overleg	x-fmt/249
	.pst	Microsoft Outlook Personal Folders		In overleg	x-fmt/250
	.mbox	MBOX		In overleg	fmt/720
	.eml	MIME Email	1.0	Voorkeur	fmt/950
Presentatie					
	.pptx	Office Open XML		Acceptabel	fmt/215
	.ppt	Microsoft Office Powerpoint 97-2003 Binary File Format		In overleg	fmt/126
	.pdf	PDF/A ¹¹	1a 1b 2a 2b 2u 3a 3b 3u 4 4f 4e	Acceptabel	fmt/95 fmt/354 fmt/476 fmt/477 fmt/478 fmt/479 fmt/480 fmt/481
	.odp	OpenDocument Presentation Document Format	1.0	Acceptabel	fmt/138
	.odp	OpenDocument Presentation Document Format	1.1	Acceptabel	fmt/292
	.odp	OpenDocument Presentation Document Format	1.2	Voorkeur	fmt/293

¹¹ Let op: door een presentatie te converteren naar PDF/A kunnen gedragingen verloren gaan



INFORMATIE-SOORT	EXTENSIE	NAAM	VERSIE	CATEGORIE	PRONOM ID
Spreadsheet					
	.pdf	PDF/A ¹²	1a 1b 2a 2b 2u 3a 3b 3u 4 4f 4e	Acceptabel	fmt/95 fmt/354 fmt/476 fmt/477 fmt/478 fmt/479 fmt/480 fmt/481
	.csv	Comma Separated Values	RFC 4180	Acceptabel	x-fmt/18
	.ods	OpenDocument Spreadsheet	1.0	Acceptabel	fmt/137
	.ods	OpenDocument Spreadsheet	1.1	Acceptabel	fmt/294
	.ods	OpenDocument Spreadsheet	1.2	Voorkeur	fmt/295
	.xlsx	Microsoft Excel	2007 onwards	Acceptabel	fmt/214
	.xls	Microsoft Excel	97-2003	In overleg	fmt/61
Vector afbeelding					
	.svg	Scalable Vector Graphics	1.0	Acceptabel	fmt/91
	.svg	Scalable Vector Graphics	1.1	Acceptabel	fmt/92
	.svg	Scalable Vector Graphics	Tiny 1.2	Voorkeur	fmt/413
Video					
	.mp4	MPEG-4 Media File		In overleg	fmt/199
	.mxf	Material Exchange Format Operational Pattern	1a	Voorkeur	fmt/200
	.mxf	Material Exchange	1b 1c	Acceptabel	fmt/783 fmt/784

¹² Let op: door een spreadsheet te converteren naar PDF/A kunnen gedragingen verloren gaan



INFORMATIE-SOORT	EXTENSIE	NAAM	VERSIE	CATEGORIE	PRONOM ID
		Format Operational Pattern	2a 2b 2c 3a 3b 3c		fmt/785 fmt/786 fmt/787 fmt/788 fmt/789 fmt/790
	.avi	Audio Video Interleaved		In overleg	fmt/5
	.mkv	Matroska	1 2 3 4	In overleg	fmt/569
Website					
	.warc	WARC	1.0	In overleg	fmt/1355
	.warc	WARC	1.1	In overleg	fmt/1281
	.wazc	Web Archive Collection Zipped (WACZ)	1.1.1	In overleg	

9. Lijst met voorkeurscodecs

	Codec	Categorie
Video		
	V210	Voorkeur
	Motion JPEG 2000	Voorkeur
	FFV1	Voorkeur
	Digital Video Encoding (DV)	Voorkeur
	H.264	Acceptabel
	H.265	Acceptabel
Audio		
	Linear Pulse Code Modulation (LPCM)	Voorkeur
	FLAC (Free Lossless Audio Codec)	Voorkeur
	Apple lossless Audio Codec (ALAC of ALE)	Voorkeur
	WavPack	Acceptabel
	MPEG-4 Audio Lossless Coding (MPEG-4 ALS)	Acceptabel



10. Bijlage 1: Logboek kleine veranderingen

Wat is gewijzigd	Waarom	Door wie	Wanneer
<ul style="list-style-type: none"> • Bijlage 1 toegevoegd • Verandering opmaak & titel • Toevoegen informatie jaarlijkse review aan inleiding • Archiefbescheiden vervangen door informatieobjecten • Zorgdrager vervangen door verantwoordelijk overheidsorgaan • .eml veranderd naar voorkeursbestandsformaat • PDF/A-4, PDF/A-4e, PDF/A-4f toegevoegd • .wacz als bestandsformaat voor informatiesoort website toegevoegd. 	Jaarlijkse review 2022	Julia Emmen & Guido Dorssers	10-08-2022