

## **Voorkeursformaten Groninger Archieven**

Laatst bijgewerkt op 24-04-2024

**Inhoud**

Inleiding.....	4
Informatiesoorten.....	4
Standaarden .....	4
Toelichting bij het gebruik .....	5
Toelichting bij PDF.....	7
PDF Versies .....	7
PDF Profielen .....	8
Toelichting bij codecs.....	8
Voorkeursformatenlijst.....	6
Toelichting bij de tabel .....	7
Bronnen document .....	8

**Versiebeheer**

<b>Versienummer</b>	<b>Versiedatum</b>	<b>Wijzigingen</b>
<b>1.0</b>	17-06-2020	Versie voor publicatie/verspreiding
<b>2.0</b>	24-04-2024	Herziening

## Inleiding

Digitale bestanden komen tegenwoordig voor in allerlei soorten en formaten. Elke informatiesoort heeft zijn eigen bestandsformaten, en elk bestandsformaat heeft zijn voor- en nadelen. Dit document maakt onderdeel uit van ons overkoepelende conserveringsbeleid dat op dit moment in ontwikkeling is. In dit document hebben wij een lijst gemaakt met onze voorkeursformaten; een lijst die is samengesteld uit bestandsformaten waar wij zelf praktijkervaring mee hebben en formaten die afkomstig zijn uit voorkeurslijsten van partnerinstellingen. De voorkeursformaten dragen bij aan de digitale duurzaamheid door een referentiekader te bieden voor:

- de optimale overdracht van in het verleden gecreëerde bestanden en
- de optimale overdracht van in de toekomst aan te maken bestanden.

## Informatiesoorten

In dit document onderscheiden we verschillende informatiesoorten. Dit betreft de meest voorkomende informatiesoorten die wij de komende jaren verwachten te ontvangen. De lijst met informatiesoorten zullen we monitoren en bijwerken aan de hand van de informatiesoorten waar wij in de praktijk tegen aan lopen.

- o Document
- o Video
- o Afbeelding
- o Vector Afbeelding
- o Audio
- o E-mail
- o Spreadsheet
- o Presentatie
- o Database
- o Website
- o Geo-data

De voorkeursformaten zijn per informatiesoort in het schema in opgenomen.

## Standaarden

In de lijst met voorkeursformaten staan veel bestandsformaten die een open standaard<sup>1</sup> zijn. Een open bestandsformaat is vastgelegd in een specificatiedocument, waardoor het geïmplementeerd kan worden in verschillende ICT-systemen. Dit specificatiedocument is vrij te verkrijgen.

De lijst kent ook een aantal *de facto* standaarden. Een *de facto* standaard is een bestandsformaat waarvan de toepassing zo wijd verbreid is dat het niet meer is weg te denken uit de markt. Denk hier bijvoorbeeld aan de Microsoft Office bestandsformaten zoals docx en xlsx. Het commerciële bedrijf heeft dan zelf de keuze om de documentatie wel of niet te publiceren.

De overheid hanteert het principe "pas toe of leg uit", wat inhoudt dat een overheid gebruik moet maken van open standaarden. Als deze dat niet doet, dan is het verplicht om uit te leggen waarom er voor een alternatief is gekozen. Er wordt voor open standaarden gekozen vanwege de interoperabiliteit en omdat ze

---

<sup>1</sup> <https://www.forumstandaardisatie.nl/thema/open-standaarden-algemeen>

leveranciersafhankelijk zijn. Interoperabiliteit zorgt voor betere communicatie en informatie-uitwisseling tussen systemen.

Bij permanent te bewaren archiefbescheiden is een overheid verplicht om bestanden uiterlijk bij overbrenging om te zetten naar een open bestandsformaat<sup>2</sup>. Wij adviseren om zoveel mogelijk bij creatie al te kiezen voor een open bestandsformaat. Wij raden af om zelfstandig bestandsformaten om te zetten, er zijn ook applicaties waarbij er automatisch bestanden worden geconverteerd. Wij raden aan om vast te leggen welke handelingen er bij het converteren worden uitgevoerd. Elke conversie geeft namelijk risico op dataverlies<sup>3</sup>.

### **Toelichting bij het gebruik**

De lijst met voorkeursformaten is continu in ontwikkeling en wordt regelmatig bijgewerkt. Bij het gebruik kunt u alvast rekening houden met de volgende vuistregels:

- Het bestandsformaat komt niet voor op de lijst, is een legacy bestand of valt onder de categorie *In overleg*: Neem contact op om af te stemmen wanneer u met deze bestandsformaten wenst over te dragen of te gebruiken bij de vorming van nieuwe informatieobjecten.;
- Het bestandsformaat valt onder de categorie *Acceptabel*: voor in het verleden gecreëerde informatieobjecten levert dit geen problemen op bij de overdracht. Voor het gebruik van deze bestandsformaten bij de inrichting van een nieuw systeem of proces kunt u het beste contact opnemen;
- Het bestandsformaat valt onder de categorie *Voorkeur*: zowel bij de overdracht van bestaande informatieobjecten als bij de vorming van nieuwe zal de toepassing van deze bestandsformaten geen problemen opleveren. Bij twijfel kunt u altijd contact opnemen

---

<sup>2</sup> <https://wetten.overheid.nl/jci1.3:c:BWBR0027041&hoofdstuk=3&paragraaf=2&artikel=26&z=2014-01-01&g=2014-01-01>

<sup>3</sup> Archiefregeling Art. 25 lid 2 AW, verklaring conversie bij digitale B-stukken

## Voorkeursformatenlijst

Extensie	PRONOM ID	Titel	Versie	Categorie
<b>Informatiesoort: Document</b>				
PDF	fmt/276	PDF	1.7	Acceptabel
PDF	fmt/476	PDF/A	2A	Voorkeur
PDF	fmt/95	PDF/A	1A	Voorkeur
PDF	fmt/354	PDF/A	1B	Acceptabel
DOC	fmt/40	Word	97-2003	Acceptabel
DOCX	fmt/412	Word		Acceptabel
ODT	fmt/136	OpenDocument	1.0	Acceptabel
ODT	fmt/290	OpenDocument	1.1	Acceptabel
ODT	fmt/291	OpenDocument	1.2	Voorkeur
<b>Informatiesoort: Video</b>				
MOV	fmt/797	Apple ProRes		in overleg
MP4	fmt/199	MPEG-4		in overleg
MXF	fmt/791			in overleg
MKV	fmt/569	Matroska	v1-4	in overleg
<b>Informatiesoort: Video Codecs</b>				
MXF		XDCAM HD422		in overleg
MPEG-4		H.264		in overleg
MKV		FFV1		in overleg
<b>Informatiesoort: Afbeelding</b>				
PNG	fmt/13		1.2	Voorkeur
PNG	fmt/12		1.1	Voorkeur
PNG	fmt/11		1.0	Voorkeur
JPG	fmt/42			Acceptabel
JPG	fmt/43			Acceptabel
JPG	fmt/444			Acceptabel
JPX	fmt/151			Acceptabel
JP2	x-fmt/392			Acceptabel
TIFF	fmt/353		6.0	Voorkeur
<b>Informatiesoort: Audio</b>				
WAV/WAVE	fmt/6			in overleg
WAV		BWF		in overleg
MP3	fmt/134	MPEG 1/2 Audio Layer 3		in overleg
AAC	fmt/1812			in overleg
<b>Informatiesoort: E-mail</b>				
MSG	x-fmt/430 x-fmt/248, x-			Acceptabel
PST	fmt/249			Acceptabel
MBOX	fmt/720			Acceptabel
EML	fmt/950			Voorkeur
<b>Informatiesoort: Spreadsheet</b>				
CSV	x-fmt/18			Acceptabel
ODS	fmt/137	OpenDocument	1.0	Voorkeur
ODS	fmt/294	OpenDocument	1.1	Voorkeur
ODS	fmt/295	OpenDocument	1.2	Voorkeur
XLS	fmt/61	Excel	97-2003	Acceptabel
XLSX	fmt/214	Excel		Acceptabel
<b>Informatiesoort: Presentatie</b>				
PPTX	fmt/215	Powerpoint		Acceptabel
PPT	fmt/126	Powerpoint	97-2003	Acceptabel

ODP	fmt/138	OpenDocument	1.0	Voorkeur
ODP	fmt/292	OpenDocument	1.1	Voorkeur
ODP	fmt/293	OpenDocument	1.2	Voorkeur
<b>Informatiesoort: Database</b>				
SQL	fmt/206	SQL		Voorkeur
SQL	fmt/729	SQL		Voorkeur
SQL	fmt/868	SQL		Voorkeur
SIARD	fmt/1196	SIARD	2.1	Voorkeur
ODB	fmt/424	OpenDocument	1.2	Voorkeur
ODB	fmt/444	OpenDocument	1.1	Acceptabel
ODB	fmt/140	OpenDocument	1.0	Acceptabel
ACCDB	fmt/275	Access	2007	Acceptabel
MDB	fmt/241	Access	2002	Acceptabel
MDB	fmt/240	Access	2000	Acceptabel
MDB	fmt/239	Access	97	Acceptabel
MDB	fmt/238	Access	95	Acceptabel
<b>Informatiesoort: Websites</b>				
WARC	fmt/1281	WARC	1.0	Voorkeur
WARC	fmt/1355	WARC	1.1	Voorkeur
<b>Informatiesoort: Afbeelding Vector 2D/3D</b>				
SVG	fmt/91			in overleg
SVG	fmt/92			in overleg
SVG	fmt/413			in overleg
DXF	fmt/63			in overleg
DWG				in overleg
OBJ	fmt/1210	Wavefront		in overleg
PLY	fmt/831	Polygon File Format		in overleg

### Toelichting bij de tabel

Ter verduidelijking wordt hieronder uitleg gegeven bij de betekenis van de kolommen.

**Informatiesoorten:** We maken bij bestandsformaten gebruik van verschillende informatiesoorten. Aan elke informatiesoort zijn bepaalde eigenschappen gekoppeld zoals functie, eigenschappen, gedrag, opmaak en weergave.

**Extensie:** De extensie is de toevoeging aan het einde van de bestandsnaam waarmee aangegeven wordt om wat voor bestandsformaat het gaat. Binnen dat formaat kunnen er dan nog verschillende versies van het formaat bestaan.

**PRONOM ID:** PRONOM is een register van bestandsformaten en versies van dat bestandsformaat. Elke versie van een bestandsformaat krijgt een uniek PRONOM ID. Dit ID staat in deze kolom vermeld. PRONOM is ontwikkeld door The National Archives in het Verenigd Koninkrijk. **Titel:** Een korte omschrijving van het bestandsformaat.

**Categorie:** Zie *Toelichting bij gebruik*.

### Toelichting bij PDF

#### PDF Versies

Het pdf-formaat kent veel verschillende soorten versies, binnen deze versies zijn er verschillende standaarden. PDF-versies zijn achterwaarts compatibel omdat elke versie een uitbreiding is op de mogelijkheden van de eerdere versies. Dit zorgt ervoor dat het mogelijk is om oude versies van PDF te openen met recente applicaties. De versie van de PDF vertelt applicaties welke functionaliteiten er binnen de PDF verwacht kunnen worden.

### PDF Profielen

Voor de duurzame toegankelijkheid gebruiken we de PDF/A-1 en PDF/A-2 profielen. PDF/A-1 maakt gebruik van PDF versie 1.4 en PDF/A-2 maakt gebruik van PDF versie 1.7, bij PDF/A-2 is er ondersteuning voor lagen en transparantie toegevoegd en de compressie naar JPEG2000 zorgt voor een kleiner bestandsformaat. PDF/A-2 heeft dan ook de voorkeur boven PDF/A-1. PDF/A-3<sup>4</sup> heeft ook nog de mogelijkheid tot het embedden van andere bestandsformaten. Deze PDF/A profielen zijn ook vastgelegd in NEN-ISO normen ([pdf/a-1](#), [pdf/a-2](#)). Hiernaast is er ook nog een PDF/-UA ([pdf/ua-1](#)), dit is een profiel op basis van versie 1.7 met aanvullende specificaties voor digitale toegankelijkheid. Deze eisen zijn op basis van [WCAG 2.0](#) vastgesteld. Het gaat hierbij onder andere om de aanwezigheid van tags, alternatieve teksten bij afbeeldingen, ingebedde lettertypes en gebruikte kleurcombinaties. Tags zijn metagegevens over het document waarmee de documentstructuur en volgorde worden beschreven.

### **Toelichting bij codecs**

Audio en videobestanden bestaan vaak uit een containerbestand of wrapper. Binnen zo'n containerbestand zijn er meestal meerdere tracks. Deze tracks kunnen als ruwe, ongecomprimeerde data opgeslagen worden maar in de meeste gevallen worden ze gecodeerd opgeslagen. Bij het coderen is er sprake van een compressie, door deze compressie zal de bitrate kleiner worden. Met een lagere bitrate wordt het bestand kleiner maar is er ook minder informatie beschikbaar om geluid of beeld te genereren.

Er bestaan verschillende soorten coderingen, oftewel codecs. Er zijn lossless en lossy codecs, bij een lossless codec is het mogelijk om de compressie terug te vertalen naar een ongecomprimeerd formaat. Bij een lossy formaat is de compressie niet omkeerbaar. Het gebruik van deze codecs maakt de preservering van audio en video formaten complexer dan andere informatiesoorten zoals documenten en afbeeldingen. Er bestaan verschillende codecs voor verschillende toepassingen; voor een productiemaster wordt er meer gebruik gemaakt van ruwe data en dus grotere bestanden dan bij een bestand dat op een website wordt getoond. In tegenstelling tot overige informatiesoorten volgen de codecs van AV formaten elkaar veel sneller op.

Voor de preservering gaat onze voorkeur uit naar ongecomprimeerde of lossless bestanden. Er zijn een aantal codecs die veel worden gebruikt en daarom de voorkeur hebben boven onbekendere en minder ondersteunde codecs.

### **Bronnen document**

[NA Norm Voorkeursformaten](#)

Forum standaardisatie

Archiefregeling art. 26, lid 1.

---

<sup>4</sup> Op dit moment staat PDF/A-3 nog niet op onze voorkeurslijst, wij kunnen op dit moment nog niet volledig overzien wat de consequenties zijn van het embedden van bestanden binnen een PDF voor de preservering.