

## Pagina 3

### <Titel> Selectiecriteria voor bestandsformaten

Op deze pagina vind je de criteria die helpen bij de selectie van bestandsformaten voor het opslaan van informatie met oog voor de [duurzame toegankelijkheid van deze informatie](#). Deze criteria gebruik je bij de aanschaf van applicaties en het inrichten van werkprocessen. Ze zijn daarmee niet bedoeld voor reeds gevormd archief.

Om te kunnen bepalen welke bestandsformaten geschikt zijn, is het belangrijk de principes die aan zo'n keuze ten grondslag liggen te verduidelijken en onderbouwen.

De module beschrijft daarom de eigenschappen van bestandsformaten die bijdragen aan duurzame toegankelijkheid. Ze overstijgen specifieke informatiesoorten en bestandsformaten. Op deze manier kun je zelf beoordelen welke bestandsformaten voldoen en welke niet. Zo kun je keuzes maken wanneer je voor je eigen organisatie aanvullingen moet doen op de lijst voorkeurs- en acceptabele formaten </link module 3>. Een praktisch stappenplan hiervoor vind je in module 4 </link module 4 Voorkeursformaten in de praktijk>.

Door de criteria hier op te nemen geven we ook verantwoording van de gemaakte keuzes voor de lijst voorkeurs- en acceptabele formaten. Deze formaten scoren namelijk hoog. Afwijken van de lijst en criteria brengt mogelijk risico's voor de duurzame toegankelijkheid van de opgeslagen informatie met zich mee.

#### <Blok> Bronnen

De criteria zijn gebaseerd op ervaringen uit de praktijk, [de 'Sustainability Factors' geformuleerd door de Library of Congress \(LoC\)](#), het [Digital Preservation Framework for Risk Assessment and Preservation Planning van The National Archives and Records Administration \(NARA\)](#), de verplichte en aanbevolen standaarden op de lijst van het [Forum Standaardisatie](#), de eisen in de [Archiefregeling 2009](#) en good practices van (inter)nationale organisaties, zie bijvoorbeeld het werk van de werkgroep [International Comparison of Recommended File Formats](#) van The Open Preservation Foundation.

De lijst zal regelmatig herzien worden en daarvoor zijn ervaringen uit het veld (zowel overheid als archief) nodig. Jouw praktijkkennis kan daarbij helpen </link feedback pagina>. Draag informatiesoorten en/of bestandsformaten aan of geef feedback op de huidige lijst.

#### <Blok> Essentiële overweging

##### *Essentiële kenmerken*

De beste keuze voor een bestandsformaat is afhankelijk van de inhoud en vorm van de informatie, de context waarin ze is gecreëerd en onderhouden moet worden, en het doel van opslaan. Niet elk bestandsformaat dat onder een informatiesoort beschreven staat als voorkeur, is automatisch geschikt voor alle typen informatie. Een voorbeeld: CSV (comma separated value) is een voorkeursformaat voor het opslaan van spreadsheets. Het voldoet namelijk aan alle selectiecriteria. Maar dynamische elementen zoals formules en diagrammen kunnen niet worden weergegeven in een CSV bestand. Wanneer je een spreadsheet met dergelijke elementen opslaat als CSV, verlies je dus veel informatie.

Daarom is het van belang dat je een bestand opslaat in een formaat dat de [essentiële kenmerken](#) van het informatieobject intact houdt. Dit zijn de "kenmerken van een informatieobject die haar uiterlijk, gedrag (functionaliteit), kwaliteit en bruikbaarheid bepalen. Ze zijn te onderscheiden in categorieën als inhoud, context (metagegevens), verschijning (lay-out, kleur, teksteigenschappen zoals bijvoorbeeld diakritische tekens), gedrag (bv. interactie, functionaliteit) en structuur (bv. paginering, segmentering)", [aldus JISC](#). The National Archives and Records Administration (NARA)

heeft uitgebreide beschrijvingen van de essentiële kenmerken van verschillende informatiesoorten, zie hiervoor de [GitHub met 'risk analyses and preservation plans'](#).

## **<titel uitklapmenu> Criterium 1: Open standaarden</titeluitklapmenu>**

Kies voor bestandsformaten die gebaseerd zijn op open standaarden. De beschikbaarheid van standaarden en specificaties voor bestandsformaten zorgt ervoor dat een grote verscheidenheid aan software kan worden ontworpen die bestanden in deze formaten kan creëren, lezen, valideren en tonen aan gebruikers.

Het [Forum Standaardisatie](#) is de voornaamste aanjager van het gebruik van open standaarden binnen de overheid. Zij beschrijven een standaard als 'een afspraak over informatie of een proces'. Ze noemen een standaard open wanneer:

1. de benodigde documentatie laagdrempelig beschikbaar is;
2. er geen hindernissen zijn op het terrein van intellectueel eigendomsrecht;
3. er voldoende inspraakmogelijkheden zijn voor stakeholders tijdens (door)ontwikkeling van de standaard;
4. en de onafhankelijkheid en duurzaamheid van de standaardisatieorganisatie verzekerd zijn'.

## **<titel uitklapmenu> Criterium 2: Adoptie en ondersteuning</titeluitklapmenu>**

Selecteer een formaat dat op grote schaal gebruikt en ingezet wordt door ontwikkelaars en gebruikers. Het volledig verdwijnen van ondersteuning voor het openen en tonen van een bestand is dan namelijk minder waarschijnlijk, omdat het voor leveranciers (financiële) voordelen heeft langdurig ondersteuning te bieden aan de gebruikersgroep, ook wanneer het bestandsformaat verouderd. Grootgebruik maakt het daarbij waarschijnlijker dat er tools voor validatie, conversie en emulatie verschijnen zonder dat de archiefsector hierin hoeft te investeren.

Een actieve en goed georganiseerde gemeenschap van gebruikers, softwareontwikkelaars, leveranciers levert hier dus voordelen. De leeftijd van een bestandsformaat speelt hier ook een rol; verouderde formaten (waarbij een actieve gemeenschap van gebruikers en hoge mate van adoptie zijn weggevallen) zijn logischerwijs ongewenst. Zeker wanneer een formaat (niet volledig) open is (criterium 1), is grootgebruik en ondersteuning van groot belang. Microsoft Office formaten zijn hier voorbeelden van.

## **<titel uitklapmenu> Criterium 3: Onafhankelijkheid van externe soft- of hardware en interoperabiliteit </titeluitklapmenu>**

Technologie is constant in verandering. Software en hardware krijgen periodieke updates of worden vervangen door nieuwe versies. Kies daarom voor bestandsformaten die onafhankelijk zijn van specifieke (versies) software en/of hardware, omdat deze anders ook gepreserveerd moeten worden. Dit is een intensieve en dure onderneming.

Niet voor alle informatiesoorten is deze onafhankelijkheid eenvoudig. Bij geo-data bestaat meestal een afhankelijkheid van bepaalde software om toegang te verkrijgen tot de informatieobjecten. Kijk in deze gevallen goed naar de middelen die nodig zijn om de informatieobjecten te blijven openen en tonen.

Het gebruiken van bestandsformaten zonder strikte afhankelijkheden van soft- en hardware bevordert ook de interoperabiliteit van informatieobjecten, ofwel de mogelijkheid om deze uit te wisselen. De interoperabiliteit heeft invloed op de duurzame toegankelijkheid van informatieobjecten, omdat ze in hun levenscyclus vaak verplaatst moeten worden van systeem naar systeem of uitgewisseld tussen organisaties. Overweeg hierbij gebruik te maken van open formaten </link met uitleg> of exportformaten die het mogelijk maken informatieobjecten in of uit gesloten of propriëtaire omgevingen te exporteren. Of let erop dat de propriëtaire omgeving de

mogelijkheid biedt te exporteren in een open formaat. Denk hierbij aan ESRI bestanden van ArcGIS, die ook kunnen worden geëxporteerd in GeoJSON.

## **<titel uitklapmenu> Criterium 4: Transparantie</titeluitklapmenu>**

Digitale informatie beschouwen we als transparant wanneer het mogelijk is deze te bekijken en begrijpen met basishulpmiddelen, zoals teksteditors. Een bestandsformaat waarin de aanwezige informatie op een simpele en directe manier kan worden gerepresenteerd zal namelijk makkelijker te migreren zijn. Ook het ontwikkelen van software om de informatie op de juiste wijze te kunnen weergeven (renderen) wordt hier simpeler door. Kies daarom voor bestandsformaten die dit toelaten. Een voorbeeld is een e-mail in .eml waarvan je de inhoud ook buiten je e-mail applicatie om kan lezen, bijvoorbeeld in Kladblok.

Transparantie wordt negatief beïnvloed door encryptie (zie criterium 7) </link criterium 7> en compressie </link Bijlage Compressie>.

## **<titel uitklapmenu> Criterium 5: Zelf-documenterend</titeluitklapmenu>**

Selecteer een bestandsformaat dat het mogelijk maakt om essentiële technische metagegevens in het informatieobject op te slaan. Hierdoor zijn deze metagegevens gemakkelijker te beheren en te controleren op integriteit en bruikbaarheid. Ook bij migratie van systeem naar systeem is het eenvoudiger om deze metagegevens betrouwbaar mee te migreren. Bij conversie is het van belang dat deze metagegevens niet verloren gaan. Beschrijvende en administratieve metadata worden meestal apart opgeslagen.

Denk bijvoorbeeld aan het TIFF-formaat voor rasterafbeeldingen (Tagged Image File Format). In de structuur van dat bestandsformaat zijn gegevens (tags) opgenomen die het weergeven van de afbeelding mogelijk maken.

## **<titel uitklapmenu> Criterium 6: Patenten en licenties</titeluitklapmenu>**

Wanneer je kiest voor een bestandsformaat met een open standaard volgens de definitie van het Forum Standaardisatie, is deze in principe vrij van patenten en licenties. Wanneer er geen bestandsformaat met een open standaard beschikbaar en/of gangbaar is, is het belangrijk te bekijken welke patenten en licenties er nog rusten op het bestandsformaat. Deze kunnen namelijk onvoorspelbare (hoge) kosten met zich meebrengen, wat een negatieve invloed heeft op duurzame toegankelijkheid. Kies daarom wanneer mogelijk voor een bestandsformaat dat vrij is van patenten en licenties.

Een patent of licentie hoeft niet altijd een probleem te zijn. Als een bestandsformaat op grote schaal geadopteerd is, is dat een goede indicatie dat het patent of de licentie geen grote invloed heeft op de bruikbaarheid van het formaat, ook op langere termijn. Hierbij weegt criterium 2, adoptie </link criterium 2>, dus het zwaarst.

## **<titel uitklapmenu> Criterium 7: Beschermingsmethoden</titeluitklapmenu>**

Digitale informatie wordt regelmatig beschermd met technische beschermingsmethodes zoals wachtwoorden, versleuteling en/of digitale certificaten. Dit kan een risico voor duurzame toegankelijkheid vormen, omdat in het geval van wachtwoorden en sleutels, [deze ook blijvend bewaard moeten worden](#). Voor digitale certificaten geldt dat deze mogelijk afhankelijk zijn van een externe entiteit, wat er in de praktijk op neerkomt dat er een verbinding met een netwerk of het internet nodig is om het bestand te gebruiken. Informatieobjecten zullen tijdens hun levenscyclus verplaatst of geconverteerd moeten worden en dit wordt lastiger door deze beschermingsmethoden. Houd bestanden daarom vrij van wachtwoorden en digitale encryptie om de duurzame toegankelijkheid te optimaliseren.

Voor bestanden waarbij encryptie onvermijdelijk is, zoals bestanden met digitale handtekeningen, kan je [de handreiking over elektronische handtekeningen van de VNG](#) en de verplichte [AdES Baseline Profiles standaard bij Forum Standaardisatie](#) raadplegen.

## **<Blok> Eisen voor digitale toegankelijkheid**

Elke burger heeft het recht op digitale informatie op het web. Maar niet iedereen heeft evenveel toegang tot deze content. Zo zullen mensen met een visuele of auditieve beperking soms extra middelen nodig hebben om deze informatie te kunnen krijgen. Om hierin te voorzien zijn de [Web Content Accessibility Guidelines \(WCAG\) versie 2.1](#) vastgesteld, criteria waaraan een gepubliceerd document of website moet voldoen om digitoegankelijk te zijn. Voorbeelden zijn het contrast tussen tekst en achtergrond, ondertitels bij filmpjes, of afbeeldingen die een alt-tekst hebben waarin de afbeelding wordt beschreven. Overheidsinstanties zijn verplicht om op deze manier te publiceren. Dit geldt voor websites en apps, maar ook een Excel spreadsheet dat moet worden gepubliceerd kan digitoegankelijk worden gemaakt. Alhoewel digitale toegankelijkheid niet direct verbonden is aan duurzame toegankelijkheid, is het ook een wettelijke verplichting om al bij creatie van informatie rekening mee te houden.

Meer informatie over digitale toegankelijkheid en de eisen waaraan moet worden voldaan is te vinden op [Home | Digitoegankelijk](#), of zie de [Handreiking Duurzaam Toegankelijke PDF's](#) van het RDDI.