



Standaard Duurzame Bestandsformaten

Versie	Wanneer	Wie	Wat
0.3.1	01-07-2013	Orville Mac Donald	Commentaar verwerken van H.Vons, S.Buistraan, S. Ujzanovitch, R. v/d Belt, A. de Jonker, L. v/d Dussen.
0.4	05-07-2013	Orville Mac Donald	Commentaar verwerkt.
0.4.1	30-09-2013	Orville Mac Donald	
0.4.2	09-10-2013	Orville Mac Donald	
0.4.3	09-12-2013	Orville Mac Donald	Lijst met formaten geüpdatet.
0.4.4	05-02-2014	Orville Mac Donald	Totale update
0.5	04-04-2014	Sofie Buistraan	Tekstuele aanpassingen en aanpassingen in de structuur van de tekst.
0.6	15-04-2014	Sofie Buistraan	Verwerken opmerkingen expertiseteam
0.7	28-04-2014	Sofie Buistraan	Verwerken opmerkingen inhoudelijke ronde: inspectie, ADW, Lennart van der Dussen, Simon Algera
0.8	14-05-2014	Sofie Buistraan	Verwerken opmerkingen ronde 2: diensten, stadsdelen, IM D&I
2.0	02-10-2014	Orville Mac Donald	Verwerken commentaar CIO office, IBA, Stadsdeel Zuid, Zuidoost, Nieuw west, Centrum, West, DBI, DRO. Paragraaf toetsingscriteria toegevoegd.

Inhoud

1	Inleiding	1
2	Duurzaam opslaan van informatie	2
3	Knelpunten bij het duurzaam opslaan van digitale informatie	3
4	Oplossing voor het duurzaam opslaan van digitale informatie	4
5	Standaard duurzame archiveringsformaten	6
5.1	Documentsoorten en hun voorkeurs- en acceptabele duurzame archiveringsformaten	6
5.2	Toetsingscriteria voor voorkeurs- en acceptabele duurzame archiveringsformaten	
5.3	Pas toe of leg uit	10

1 Inleiding

Het College van B&W nam op 8 januari 2013 het besluit dat per 1 januari 2015 (streefdatum) alle organisatieonderdelen van de gemeente Amsterdam alleen nog digitaal archiveren.

Overheidsorganen zijn verplicht om hun 'archiefbescheiden', ongeacht de vorm, in goede, geordende en toegankelijke staat te brengen en te houden.¹ Digitaal archiveren is het digitaal opslaan en bewaren van originele, procesgebonden informatie in duurzame, geordende en toegankelijke staat gedurende de voorgeschreven bewaartermijn. Na verloop van de voorgeschreven bewaartermijn wordt de informatie vernietigd. Permanent te bewaren informatie wordt overgebracht naar de archiefbewaarplaats (voor digitale informatie is dat het E-depot) van het Stadsarchief Amsterdam.

Deze 'standaard duurzame bestandsformaten' betreft het duurzaam opslaan van digitale informatie. Duurzaam betekent in dit kader dat de informatie zolang de wet dat voorschrijft reproduceerbaar is.

Digitale informatie moet o.a. duurzaam worden opgeslagen zodat de informatie toegankelijk is en:

- de bedrijfsvoering snel en ongestoord plaatsvindt;
- de overheid zich publiek kan verantwoorden;
- rechten en plichten van burgers kenbaar zijn;
- historische informatie behouden blijft.²

Bij het uitvoeren van de werkzaamheden wordt door de concernonderdelen informatie gegenereerd en ontvangen die opgeslagen/gearchiveerd moet worden. Het concernonderdeel dat de informatie genereert of ontvangt is de archiefvormer. De archiefvormer is tot het moment van overbrengen naar het Stadsarchief verantwoordelijk voor het archiveren van de informatie, waaronder het duurzaam opslaan van de informatie. Na het overbrengen van de informatie ligt de verantwoordelijkheid voor het archiveren, waaronder het duurzaam opslaan van de digitale informatie, bij het Stadsarchief.

Leeswijzer

In hoofdstuk 2 wordt het belang van de duurzame opslag van informatie beschreven. In hoofdstuk 3 worden de knelpunten en risico's ten opzichte van de duurzaamheid en integriteit van de opgeslagen informatie beschreven. In hoofdstuk 4 wordt beschreven dat een standaard archiveringsformaat bijdraagt aan de oplossing voor de duurzame opslag van informatie. In hoofdstuk 5 is de standaard voor duurzame archiveringsformaten opgenomen, zoals die door de concernonderdelen dienen te worden toegepast volgens het 'pas toe, leg uit' principe. De standaard duurzame archiveringsformaten kan door nieuwe inzichten en/of technologische ontwikkelingen worden gewijzigd.

¹ Artikel 3 van de Archiefwet 1995.

² Zwagerman, G., en Lucker, P. (2013). *E-depot SAA als generieke voorziening Concern*, Beleidsnotitie. (Amsterdam) 2013. pagina 1. en <http://www.nationaalarchief.nl/informatiebeheer-archiefvorming>, geraadpleegd op 06-08-2013.

2 Duurzaam opslaan van informatie

Bij het uitvoeren van hun werkzaamheden genereren en ontvangen medewerkers van de gemeente Amsterdam informatie. De informatie wordt zo snel mogelijk opgeslagen in het zaakdossier waarin de informatie thuishoort, zodat deze toegankelijk is voor en te gebruiken is door iedereen die de informatie nodig heeft.³ Naast het gebruik van de informatie bij het uitvoeren van verschillende werkprocessen kan de informatie worden ingezet voor het afleggen van verantwoording, het voeren van bewijs en voor cultureel-maatschappelijke doeleinden. Informatie die vandaag wordt opgeslagen moet geopend, gelezen en gebruikt kunnen worden zolang de wettelijke bewaarplicht geldt. De toegankelijkheid van de informatie wordt onder andere gewaarborgd door het duurzaam opslaan van de informatie.

³ Uitgangspunt 4 bij Digitaal Archiveren: Amsterdam voert haar taken proces- en projectgericht uit, waarbij het niveau van archivering het zaakdossier is: Procesgericht werken, zaakgerichte dossiervorming; & Uitgangspunt 6: De behandelend medewerker is verantwoordelijk voor de inhoud van het dossier en het volgens de regels plaatsen van documenten in het dossier; Bustraan, S. *Uitgangspunten voor Digitaal Archiveren* (Amsterdam) 2014. pagina 10 en 11.

3 Knelpunten bij het duurzaam opslaan van digitale informatie

De archiefvormer genereert, ontvangt en slaat digitale informatie op in verschillende bestandsformaten. Bijvoorbeeld tekstdocumenten, power-point presentaties, spreadsheets, e-mails, databases, (audio)visuele informatie en computer aided design (CAD) bestanden. De gegenereerde of ontvangen informatie moet duurzaam worden opgeslagen. De duurzaamheid van de informatie komt in gevaar door:

- Technologische ontwikkelingen;
- Vendor lock-in;⁴

Technologische ontwikkelingen

De technologie ontwikkelt zich constant, onder andere om aan de steeds veranderende behoefte van de gebruiker te voldoen. Bestandsformaten blijven constant evolueren met als gevolg dat bestandsformaten in onbruik raken. Bestandsformaten raken bijvoorbeeld in onbruik doordat:

- Software veroudert of omdat er een nieuwe versie van de software op de markt komt;
- Er leesbaarheidsproblemen ontstaan;
- Er betere ondersteuning is voor nieuwe bestandsformaten of omdat een ander bestandsformaat meer functionaliteiten biedt;
- Bestandsformaten niet wijd gedistribueerd zijn;
- Bepaalde platformen en/of software het bestandsformaat niet accepteren;
- Het bestandsformaat slecht of niet gedocumenteerd is;

Informatie die vandaag wordt opgeslagen in een specifiek bestandsformaat kan, na het in onbruik raken van dat bestandsformaat, niet meer worden gereproduceerd. Het niet kunnen reproduceren van de informatie betekent dat de informatie verloren gaat. De informatie kan niet meer kan worden geopend, gelezen en/of bewerkt. De informatie kan dan niet meer worden gebruikt in het primair proces, niet meer worden ingezet ter verantwoording en bewijs en kan geen cultureel- maatschappelijke functie meer vervullen. Wanneer informatie verloren gaat voordat de wettelijke bewaartermijn is verstreken, wordt er niet voldaan aan de eisen die vanuit de Archiefwet worden gesteld, namelijk het in goede, geordende en toegankelijke staat hebben en houden van de informatie.

Vendor lock -in

Leveranciers willen hun positie op de markt behouden en zien graag dat gebruikers afhankelijk zijn van de door hen geproduceerde software. Het gevolg hiervan is dat bestandsformaten specifiek gekoppeld kunnen zijn aan één type software. Dit begrip staat beter bekend als vendor lock-in. De gebruiker is gebonden aan de mogelijkheden van de software die door de leverancier zijn vastgesteld. Dit beperkt de flexibiliteit en vrijheid van gebruikers. Wanneer van een ander softwarepakket gebruik gemaakt gaat worden of wanneer de ontvanger van de informatie geen beschikking heeft over de software waarin de informatie tot stand is gekomen is de informatie niet reproduceerbaar.

⁴ Vendor lock-in maakt een klant afhankelijk van een leverancier voor producten en diensten, omdat hij niet in staat is om van leverancier te veranderen zonder substantiële omschakelingskosten. [bron: http://nl.wikipedia.org/wiki/Vendor_lock-in]

4 Oplossing voor het duurzaam opslaan van digitale informatie

Zoals hierboven is beschreven is de duurzame toegankelijkheid van de informatie niet vanzelfsprekend en is deze zelfs in gevaar als er geen maatregelen worden getroffen. Het duurzaamheidsprobleem wordt op dit moment het beste opgelost door de informatie op te slaan in een standaard bestandsformaat dat eveneens geschikt is als archiveringsformaat.

Standaard bestandsformaat

Een bestandsformaat is een standaard formaat wanneer?

- Het bestandsformaat bijvoorbeeld een open standaard is. Open standaarden zijn niet gebonden aan een specifieke leverancier, wat vendor lock - in tegengaat. Een open standaard is uitwisselbaar tussen verschillende software/systemen;
- Het formaat onder de 'bewezen technologie' valt; het formaat heeft een grote gebruikersgroep en heeft zijn nut en noodzaak bewezen. Bewezen technologie heeft als voordeel dat er van bestaande kennis gebruik gemaakt kan worden. De standaarden die zijn geselecteerd als duurzaam archiveringsformaat, zoals opgenomen in tabel 1 in hoofdstuk 4, hebben in het verleden hun bruikbaarheid bewezen en bewijzen deze nog steeds;

Standaard archiveringsformaat

Niet alle standaard bestandsformaten zijn geschikt om informatie te archiveren.

Een standaard bestandsformaat is ook een geschikt archiveringsformaat wanneer het formaat tenminste aan de volgende criteria voldoet:

- Het archiveringsformaat is in principe een open standaard;
- Het archiveringsformaat heeft voldoende marktondersteuning en is wijdverspreid;
- Het archiveringsformaat geeft een goed gestructureerde opslag van informatie met zo min mogelijk informatieverlies;
- Een (non-profit) beheerorganisatie heeft het archiveringsformaat goed gedocumenteerd;
- Het archiveringsformaat is uitwisselbaar en niet afhankelijk van een specifieke leverancier;
- Informatie opgeslagen in het archiveringsformaat moet geen verwijzingen bevatten naar externe informatiebronnen of -objecten.
- Het archiveringsformaat biedt de mogelijkheid tot insluiten van (zelf gedefinieerde) metadatavelden;
- Het archiveringsformaat bewaart de integriteit⁵ van het document en de informatie daarin;
- Het archiveringsformaat kent een robuust foutopsporings- en verbeteringsmechanisme: fouten in bitopslag worden automatisch gedetecteerd en zijn herstelbaar.

Het opslaan van de informatie in een duurzaam archiveringsformaat dient in het werkproces te worden geïntegreerd. Elk definitief exemplaar van een document wordt door de archiefvormer direct in het archiveringsformaat opgeslagen. Bij overbrenging

⁵ Integriteit van het archiveringsformaat betekent dat een document intact moet zijn en niet zodanig veranderd of gecorrumped dat de betekenis van de informatie ervan niet meer duidelijk is.

accepteert het Stadsarchief alleen die informatie die volgens de standaard in het duurzame archiveringsformaat is opgeslagen. Op het moment dat informatie is opgeslagen in een duurzaam archiveringsformaat kan de informatie niet meer worden gewijzigd, waarmee de integriteit van de informatie wordt gewaarborgd.

Niet alle informatie wordt in een duurzaam archiveringsformaat opgeslagen, omdat:

- Niet voor alle formaten een geschikt duurzaam archiveringsformaat op de markt is;
- De functionaliteiten van het originele bestandsformaat waarin de informatie is gegenereerd of ontvangen behouden moeten blijven. Bijvoorbeeld wanneer een spreadsheet (rekenblad) met complexe rekenfunctionaliteit in MS Excel is gegenereerd en de rekenfunctionaliteiten duurzaam behouden moeten blijven. De functionaliteiten gaan verloren als de informatie wordt opgeslagen in PDF/A. PDF/A is een duurzaam archiveringsformaat, maar biedt geen ondersteuning voor (complexe) functionaliteiten.

In bovenstaande gevallen dient de informatie opgeslagen te worden in het formaat waarin de informatie is gegenereerd of ontvangen, ter voorkoming van onherstelbaar informatieverlies.

5 Standaard duurzame archiveringsformaten

Alle concernonderdelen dienen hun documenten op te slaan in de voorgeschreven duurzame archiveringsformaten, zoals opgenomen in tabel in 1 in dit hoofdstuk. Nieuwe inzichten en technologische ontwikkelingen kunnen zorgen voor een aanpassing van de voorgeschreven duurzame archiveringsformaten. Wanneer een concernonderdeel niet aan het voorgeschreven formaat kan voldoen dient daarover verantwoording te worden afgelegd volgens het pas toe, leg uit principe. Het 'pas toe, leg uit' principe wordt beschreven in § 5.3.

5.1 Documentsoorten en hun voorkeurs- en acceptabele duurzame archiveringsformaten

In tabel 1 op pagina 7 staat per documentsoort aangegeven wat het duurzame archiveringsformaat is waarin de informatie moet worden opgeslagen. Er wordt onderscheid gemaakt tussen het duurzame archiveringsformaat dat de voorkeur heeft en welke formaten acceptabel zijn als duurzaam archiveringsformaat. In de laatste kolom van de tabel wordt verwezen naar een richtlijn en/of standaard. Niet voor alle soorten documenten is reeds een richtlijn beschikbaar, als dit het geval is wordt de richtlijn nog ontwikkeld, door het Stadsarchief en het programma Digitaal Archiveren 2015, en toegevoegd zodra deze klaar is.

Vanuit het perspectief van beheer en van efficiënt werken is er gekozen voor een beperkte set aan archiveringsformaten. Dit betekent dat elke medewerker kennis kan hebben van het formaat waarin de informatie moet worden opgeslagen en dat binnen het hele concern dezelfde soort informatie in hetzelfde formaat wordt gearhiveerd. De informatie is dan beter toegankelijk voor iedereen. Ook hard- en software moet beheerd worden: zodat deze toegankelijk zijn en blijven. Hoe meer hard- en software er gebruikt wordt hoe complexer het beheer. Dienst ICT stelt de benodigde hard- en software beschikbaar en is verantwoordelijk voor het beheer van de programmatuur.

5.2 Toetsingscriteria voor voorkeurs- en acceptabele duurzame archiveringsformaten

De archiveringsformaten worden geplaatst op de lijst voor voorkeur of acceptabele formaten. De plaatsing van formaten op de voorkeur of acceptabele lijst is afhankelijk in welke mate zij aan de criteria voldoen. Voorkeursformaten zijn de formaten die in de ideale situatie geschikt zijn als duurzaam archiveringsformaat. Acceptabele formaten voldoen in mindere mate aan alle criteria, maar zijn duurzaam genoeg. Op de lijst van de acceptabele formaten kunnen ook formaten staan die bijvoorbeeld geen open standaard zijn, maar wel als een archiveringsformaat worden gezien.

Eerder in hoofdstuk 4 zijn een aantal criteria opgesomd. Deze zijn nu samengevoegd tot vier toetsingscriteria waaraan een archiveringsformaat moet voldoen. Dit is gedaan om het onderscheid te kunnen hanteren tussen de voorkeurs- en acceptabele archiveringsformaten. Deze vier toetsingscriteria zijn:

Archief	Interoperabiliteit	Techniek	Vendor-lock in
Het archiveringsformaat heeft voldoende marktondersteuning en is wijdverspreid	Het archiveringsformaat is in principe een open standaard.	Het archiveringsformaat geeft een goed gestructureerde opslag van informatie met zo min mogelijk informatieverlies.	Het archiveringsformaat is niet afhankelijk van een specifieke leverancier
Een (non-profit) beheerorganisatie heeft het archiveringsformaat goed gedocumenteerd	Het archiveringsformaat is uitwisselbaar	Informatie opgeslagen in het archiveringsformaat moet geen verwijzingen bevatten naar externe informatiebronnen- of objecten	
Het archiveringsformaat onder de 'bewezen technologie' valt; Het formaat heeft een grote gebruikersgroep en heeft zijn nut en noodzaak bewezen.		Het archiveringsformaat biedt de mogelijkheid tot insluiten van (zelf gedefinieerde) metadatavelden	

Tabel 1: Toegangs criteria

Archiveringsformaten komen op de voorkeurslijst te staan zodra deze voldoen aan de vier toetsingscriteria. Niet voldoen betekent dat de archiveringsformaten worden opgenomen onder de acceptabele formaten.

Soort document	Voorkeur	Acceptabel	Richtlijn/standaard
Afbeeldingen		- Joint Photographic Experts Group JPEG (.jpg) - Tagged Image File format (.tiff) - Portable Network Graphics (.PNG)	
Audio	Waveform Audio File Format (.wav)	- MPEG 1-Layer 3 (.mp3) - Windows Media Audio (.wma)	Richtlijn Audio en Video
Containers		- Disk image (.iso) - ZIP	
Databases	- Comma Separated Values (.csv) - Extensible Markup Language (.xml) i.c.m. met bijbehorend .xsd schema.	- Acces (.mdb) - Open Document Database (.odb)	
Datasets (statistische gegevens)			
Email		- MSG – unicode - EML	Beleid e-mail archiveren (in besluitvorming) Generiek E-mail protocol (in ontwikkeling)
Geografische data	- GML - HTML		Handreiking Archiveren Ruimtelijke Plannen – Geonovum.nl
Geografische vectordata Geografische rasterdata Geografische sensordata			
Presentatie		- MS Powerpoint (.ppt, .pptx) - PDF/A-1a of -1b - OpenDocument (.odp)	
Spreadsheets (rekenbladen)		- PDF/A-1a of -1b - MS Excel (.xls, .xlsx) - Open Document Spreadsheet (.ods)	
Tekst-documenten, platte tekst	- PDF/A-1a of -1b - Text (.txt)	Open Document Text (.odt)	Richtlijn Tekstdocumenten Handleiding Digitaal Document Beheer
Vectorafbeeldingen*		PDF/E	

Video (bewegend beeld)		▪ MPEG-2 ▪ Audio Video Interleaved (.avi) ▪ Matroska (.mkv) ▪ Quicktime (.mov)	Richtlijn Audio en Video
Websites		▪ Web Archive (.warc)	Archiveren van Websites & Sociale media (in ontwikkeling)

Tabel 2: Standaard duurzame archiveringsformaten

Toelichting vectorafbeeldingen.

De complexiteit die vectorafbeeldingen met zich meebrengen zorgen ervoor dat de oplossing niet direct voor de hand ligt. De eventuele oplossingen waar de nadruk op kan worden gelegd zijn:

- 1) **Oplossing 1 Visuele reproductie** van het bronbestand waarbij deze wordt 'platgeslagen' en eventueel functionaliteitsverlies plaatsvindt.
- 2) **Oplossing 2 Technisch leesbaar** van het bronbestand met eventueel functionaliteits- en informatieverlies en het bronbestand wordt omgezet naar een duurzaam archiveringsformaat.
- 3) **Oplossing 3 Semantisch bronbestand** waarbij het origineel niet wordt omgezet naar een ander duurzaam archiveringsformaat en de volledige functionaliteit behouden blijft.

Elke oplossing heeft zijn eigen dynamiek en de vraag is in hoeverre deze door het SAA kan worden ondersteund. De oplossingen zijn erg sterk gebonden aan de informatiebehoefte van de afnemer. Oplossing 1 zal het aantrekkelijkst voor de burger zijn. Oplossing 2 & 3 zijn juist van belang voor de diensten en bouwbedrijven.

Het komt voor dat vectorafbeeldingen een relatie met externe bronbestanden hebben die continu worden geactualiseerd. Dit valt vanuit het oogpunt van het Stadsarchief niet beheren, omdat informatie onveranderd wordt opgeslagen. Eén van de toegangscriteria is ook dat bestanden volledig autonoom fungeren.

In overleg met de diensten uit de Geo- en Gis sector uit de gemeente is er voorgesteld om oplossing 3 te realiseren voor eventueel hergebruik. In dit geval zullen de Geo en Gis diensten dit zelf moeten organiseren en zijn zelf verantwoordelijk voor het beheer. Dit betekent dat het SAA een visuele reproductie maakt op een bepaald moment, deze beheert en vervolgens beschikbaar stelt.

5.3 Pas toe of leg uit

Pas toe:

Alle concernonderdelen dienen de voorgeschreven duurzame archiveringsformaten toe te passen, zoals opgenomen in tabel 1 in § 5.1

Leg uit:

Wanneer een concernonderdeel niet kan voldoen aan het toepassen van de voorgeschreven duurzame archiveringsformaten en hier dus van afwijkt, moet dit worden uitgelegd en vastgelegd. Een goede uitleg bevat per afwijking de volgende onderdelen:⁶

Onderdeel	Toelichting
Specificatie	Welk(e) aspect(en) van de standaard betreft de 'leg uit'?
Oorzaak	Wat is de reden dat er (nog) niet aan kan worden voldaan?
Gevolg	Welke gevolgen/beperkingen heeft het niet of niet volledig voldoen aan de standaard?
Alternatieven	Zijn zowel de voorkeurs als acceptabele archiveringsformaten geen optie voor gebruik?
Maatregelen	Welke maatregelen zijn of worden genomen om alsnog aan de standaard te kunnen voldoen?
Planning	Op welke termijn zullen de maatregelen zijn geïmplementeerd?

⁶ [bron: onderdelen zijn gebaseerd op de lijst van de webrichtlijnen, <http://www.webrichtlijnen.nl/wat-en-waarom/pas-toe-of-leg-uit>].