

Overbrengingsvoorwaarden voor digitale overheidsarchieven (gebaseerd op het TMLO)



Versie: 26-05-2023, V1.0

De overbrengingsvoorwaarden zijn onderverdeeld in onderstaande groepen:

- Selectie en vernietiging
- Openbaarheid
- Rechten van derden
- Ordeningsstructuur
- Metadata
- Digitale objecten
- Sidecarstructuur
- Aanlevering
- Begrippenlijsten
- Voorkeursformaten

De concrete invulling van de export van een bepaald archief dat wordt overgebracht, is mede afhankelijk van de inhoud en (technische) verschijningsvorm van het archief. Zo verschilt een website, bijvoorbeeld, enorm van een zaak die is gecreëerd en gearhiveerd in een zaakstelsel, en is een omgevingsvergunning weer anders dan een raadsvergadering.

Ook de manier waarop een leverancier de exportfunctionaliteit van een applicatie heeft ingericht, is van invloed op het eindresultaat en de mogelijkheden voor een zorgdrager om een export te maken.

Tot slot

Feedback op deze samenvatting is van harte welkom!

Deze kan worden verstuurd naar: casper.huisman@noord-hollandsarchief.nl

Selectie en vernietiging

- Alleen statisch archief kan worden overgebracht.
- Het archief moet in goede, geordende en toegankelijke staat zijn.
- Het archief is geselecteerd op basis van de daarvoor geldende selectievoorschriften en is geschoond.
- De waardering (bewaren of vernietigen) is in de export vastgelegd in de metadata van de archiefbescheiden.

Openbaarheid

- Eventuele beperkingen op de openbaarheid van archiefbescheiden zijn aangegeven en moeten in de export zijn vastgelegd in de metadata van de archiefbescheiden.

Rechten van derden

- Eventuele rechten van derden (zoals intellectuele eigendomsrechten, waaronder het auteursrecht) die rusten op de archiefbescheiden zijn aangegeven en kunnen worden vastgelegd in de export in de metadata van de archiefbescheiden.

Orderingsstructuur

- De originele orderingsstructuur van het archief - zoals die is gecreëerd door de zorgdrager ten tijde van de uitvoering van de werkprocessen - wordt zoveel mogelijk behouden in de preservingvoorziening (het zgn. "structuurbeginsel").
- De orderingsstructuur dient in de export te worden vastgelegd middels de metadata en de sidecarstructuur.

Metadata

- Een metadatamapping van metadata in het bronsysteem naar TMLO moet zijn gemaakt voor in ieder geval de verplichte velden in het metadatamodel.
Dat wil zeggen, de metadata die verplicht zijn volgens het XML-schema van het metadatamodel.
- TMLO-metadata worden in sidecarbestanden (XML-metadatabestanden) overgebracht en moeten voldoen aan het voorschreven TopX 2.3.2 XML-schema.
- Aanvullende (domein-, organisatie- en/of processpecifieke) metadata of gegevens naast de (gemapte) TMLO-metadata worden enkel overgebracht in overleg met het NHA.
- Aanvullende metadata of gegevens worden in XML-bestanden overgebracht (tenzij anders afgesproken met het NHA) en moeten voldoen aan een met het NHA overeengekomen structuur.
- Het veld <identificatiekenmerk> voor een informatieobject van het aggregatieniveau "Archief" heeft de waarde van het toegangsnummer in Mais Flexis (het collectiebeheersysteem van het NHA).
- Voor alle Informatieobjecten van het aggregatieniveau "Archief" moet voor het element <context\actor> de naam van de archiefvormer worden ingevuld.
- Voor alle Informatieobjecten van het aggregatieniveau "Archief", "Serie" en "Dossier" moet voor de looptijd/periode het element <dekking\InTijd> in de ToPX-metadata worden ingevuld.
- Voor alle Informatieobjecten van het aggregatieniveau "Archief", "Serie", "Dossier", "Zaak" en "Record" moet voor de openbaarheid het element <openbaarheid> in de ToPX-metadata worden ingevuld.

- Als het aggregatieniveau “Record” niet wordt toegepast voor een document, maar enkel het aggregatieniveau “Bestand”, dan moet voor “Bestand” voor de openbaarheid het element <openbaarheid> in de ToPX-metadata worden ingevuld.
- Het element <context\activiteit> wordt gedurende het werkproces ingevuld.

Hieronder de metadata waarvoor een Begrippenlijst geldt:

- <aggregatieniveau>
- <context\actor>
- <openbaarheid\omschrijvingBeperkingen>
- <bestandsformaat> heeft bij voorkeur een Pronom-ID dat afkomstig is uit het Pronom-register.¹ (geen Begrippenlijst in de bijlage: zie de voetnoot voor de toegestane waarden)
- <fysiekeIntegriteit\algoritme>

Digitale objecten

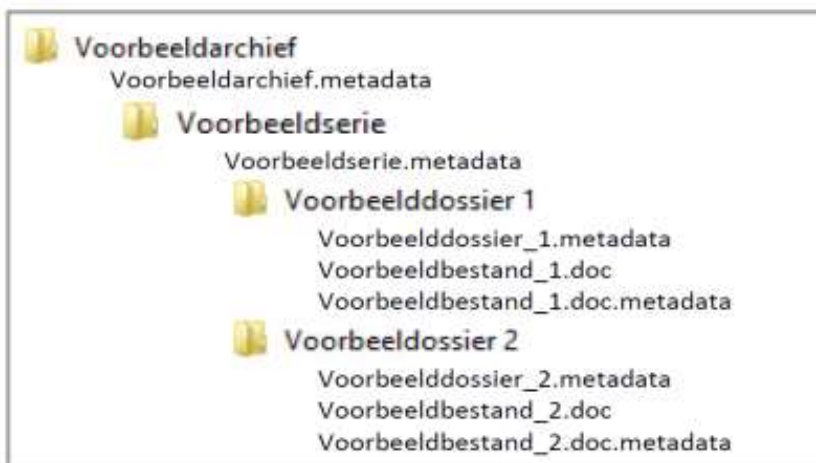
- Elk Informatieobject en Bestand moet een unieke waarde hebben voor het element <identificatiekenmerk> in de ToPX-metadata.
- Elke map en elk bestand moet een naam hebben.
- Map- en bestandsnamen mogen niet de volgende karakters bevatten: < > : “ / \ | ? * # & spatie.
- Map- en bestandsnamen mogen niet identiek zijn aan: CON, PRN, AUX, NUL, COM1, COM2, COM3, COM4, COM5, COM6, COM7, COM8, COM9, LPT1, LPT2, LPT3, LPT4, LPT5, LPT6, LPT7, LPT8 of LPT9.
- De map- en bestandsnamen mogen niet uit meer dan 255 karakters bestaan.
- De bestandsnaam voor een sidecarbestand dat bij een Informatieobject hoort, heeft voor de bestandsnaam het stramien: <informatieobjectnaam>.metadata
- Het onderdeel <informatieobjectnaam> in de bestandsnaam van het sidecarbestand is gelijknamig aan de mapnaam van de map van het Informatieobject in de sidecarstructuur waar het sidecarbestand bij hoort.
Voorbeeld: *Omgevingsvergunning_1234.metadata*
Deze bestandsnaam hoort bij de map (van het Informatieobject) met de mapnaam *Omgevingsvergunning_1234*.
- De bestandsnaam voor een sidecarbestand dat bij een Bestand hoort, heeft voor de bestandsnaam het stramien: <bestandsnaam>.metadata
- Het onderdeel <bestandsnaam> in de bestandsnaam van het sidecarbestand is gelijknamig aan de bestandsnaam van het bestand in de sidecarstructuur waar het bij hoort.
Let op: het onderdeel <bestandsnaam> is de bestandsnaam inclusief de formaatextensie van het bijbehorende computerbestand.
Voorbeeld: *constructietekening.dgn.metadata*
- Dit sidecarbestand hoort bij een bestand met de bestandsnaam *constructietekening.dgn*
- Bestanden moeten vrij zijn van kwaadaardige software zoals virussen en malware.
- Bestanden zijn niet corrupt, beschadigd of onleesbaar.
- Bestanden hebben geen omvang van 0 kB.
- De sidecarbestanden voldoen aan het XML schema van ToPX.

¹ <https://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

- De sidecarbestanden dienen in UTF-8 te zijn gecodeerd.
- Verborgene mappen en bestanden (zoals thumbs.db) worden niet aangeleverd.
- Bestanden zijn te openen.
 - Bestanden die met een wachtwoord zijn beveiligd, moeten door de zorgdrager ontsleuteld worden.
 - Bestanden die met behulp van encryptiesoftware zijn afgeschermd, moeten door de zorgdrager toegankelijk gemaakt worden.
- Informatieobjecten die uit meerdere bestanden bestaan, moeten kunnen worden weergegeven.
- Bestanden mogen niet worden aangeleverd in containerbestandsvormen als .zip- of .gz-file format, tenzij anders overeengekomen met het NHA.
- Alle bestanden die aan elkaar zijn gelinkt, worden tegelijkertijd aangeleverd.
 - Bijvoorbeeld: Object Linking and Embedding bestanden bij Excel-bestanden, of External Reference bestanden bij CAD-bestanden.
 - Let op: Voor gelinkte bestanden die te vernietigen informatie bevatten, dient met het NHA te worden overlegd over een aanpak.
- Bestanden worden, indien mogelijk, aangeleverd in bestandsformaten die voorkomen op de lijst van voorkeursformaten.²

Sidecarstructuur

Het archief wordt aangeleverd in een sidecarstructuur. Dit is een mappenstructuur waarin mappen en bestanden staan. Elke map en ieder bestand heeft een bijbehorend sidecarbestand (een metadatabestand) waarin de metadata staan die betrekking hebben op het informatieobject of het bestand. Bijvoorbeeld:



² Zie bijlage.

Aanlevering

De aanlevering van het over te brengen archief bestaat uit:

- Een TAR-bestand. Dit is een gecomprimeerd bestand dat het archief bevat.
 - De naamgeving dient als volgt te gebeuren: [Aanleveringsdatum]_[Volgnummer aanlevering]_[Omschrijving archief].tar, waarbij:
 - [Aanleveringsdatum] genoteerd dient te worden als JJJJMMDD.
 - [Volgnummer aanlevering] een heel getal moet zijn, voorzien van een voorloopnul indien kleiner dan 10.
 - Voorbeeld: 20120414_02_Omgevingsvergunningen.tar.
- E-mailbericht aan het NHA met de melding dat het archiefbestand door de zorgdrager op de aangewezen locatie in het netwerk van het NHA is geplaatst.
- Een tekstbestand behorend bij het TAR-bestand. Dit is een plat tekstbestand, bestaande uit:
 - Een korte omschrijving van de collectie.
 - De datum en het volgnummer van de aanlevering.
 - De checksum van het TAR-bestand met een van de volgende algoritmen:
 - SHA-1
 - SHA-256
 - SHA-512

Begrippenlijsten

Begrippenlijst: *Aggregatieniveaus*

Deze begrippenlijst wordt gebruikt voor het element <aggregatie>.

Waarde	Definitie
Archief	Geheel van informatieobjecten, ontvangen of opgemaakt door een archiefvormer
Serie	Verzameling van dossiers, fysieke archiefbestanddelen en/of stukken, numeriek, alfabetisch, chronologische of logisch geordend, ontstaan vanuit een identieke "handeling", dan wel een identieke vorm hebbend dan wel verwante inhoud bevattend.
Dossier	Geheel van fysieke of virtueel gekoppelde informatieobjecten die op één zaak betrekking hebben
Record	Enkelvoudig informatieobject of informatie-eenheid. Enkelvoudig wil zeggen dat het stuk niet meer dan één component bevat.

Begrippenlijst: *Archiefvormers*

Deze begrippenlijst wordt gebruikt voor het element <context\actor>.

Archiefvormer	Waarde voor het element <archiefvormer>
Gemeente Aalsmeer	Gemeente Aalsmeer
Gemeente Beverwijk	Gemeente Beverwijk
Gemeente Bloemendaal	Gemeente Bloemendaal
Gemeente Haarlem	Gemeente Haarlem
Gemeente Haarlemmerliede	Gemeente Haarlemmerliede
Gemeente Haarlemmermeer	Gemeente Haarlemmermeer
Gemeente Heemskerk	Gemeente Heemskerk
Gemeente Heemstede	Gemeente Heemstede
Gemeente Spaarnwoude	Gemeente Spaarnwoude
Gemeente Uitgeest	Gemeente Uitgeest
Gemeente Uithoorn	Gemeente Uithoorn
Gemeente Velsen	Gemeente Velsen
Gemeente Zandvoort	Gemeente Zandvoort
Provincie Noord-Holland	Provincie Noord-Holland

Omgevingsdienst Noordzee Kanaalgebied	Omgevingsdienst Noordzee Kanaalgebied
Omgevingsdienst IJmond	Omgevingsdienst IJmond
Recreatie Noord-Holland	Recreatie Noord-Holland
Recreatieschap Groengebied Amstelland	Recreatieschap Groengebied Amstelland
Recreatieschap Twiske-Waterland	Recreatieschap Twiske-Waterland
Recreatieschap Spaarnwoude	Recreatieschap Spaarnwoude
Recreatieschap Alkmaarder- en Uitgeestermeer	Recreatieschap Alkmaarder- en Uitgeestermeer
Veiligheidsregio Kennemerland	Veiligheidsregio Kennemerland
Cocensus	Cocensus

Begrippenlijst: *Openbaarheid*

Deze begrippenlijst wordt gebruikt voor het element <openbaarheid\omschrijvingBeperkingen>. Waarde	Definitie
Openbaar	Openbaar in de zin van de Archiefwet 1995
Beperkt openbaar	Beperkt openbaar in de zin van de Archiefwet 1995

Begrippenlijst: *Algoritme*

Deze begrippenlijst wordt gebruikt voor het element <fysiekeIntegriteit\algoritme>.

Waarde	Definitie
SHA-1	Cryptografisch hashalgoritme ten behoeve van authenticatie en integriteitscontrole.
SHA-256	Cryptografisch hashalgoritme ten behoeve van authenticatie en integriteitscontrole.
SHA-512	Cryptografisch hashalgoritme ten behoeve van authenticatie en integriteitscontrole.

Voorkeursformaten

Dit overzicht bevat de voorkeursformaten voor digitale objecten die worden overgebracht. Het is niet nodig speciaal voor overbrenging bestaande bestandsformaten om te zetten naar voorkeursformaten. De e-Depot voorziening van het NHA is hier speciaal voor ontworpen. De voorkeursformaten zijn in de eerste plaats bedoeld als hulpmiddel bij de aanschaf of inrichting van nieuwe informatiesystemen.

Voor overleg over bestandsformaten die niet in dit overzicht staan, kunt u contact opnemen met onze adviseurs digitale archivering.

Type data	Extensie	Versie	Toelichting	PRONOM ID (PUID) met hyperlink	Categorie formaat	Direct toonbaar	Na conversie toonbaar	Alleen toonbaar in externe viewer
Audio	WAV, WAVE		Waveform Audio	fmt/141 , fmt/142 , fmt/143	Voorkeur		X	
	WAV	2 Generic	Broadcast WAVE	fmt/527	Voorkeur		X	
	FLAC		Free Lossless Audio Codec	fmt/279	Voorkeur		X	
	MP3		MPEG 1/2 Audio Layer 3	fmt/134	Acceptabel	X		
	M4A		Apple Lossless Audio Codec	fmt/596	Acceptabel		X	
Database	SQL		Structured Query Language Data	fmt/206	Voorkeur	X		
	SIARD	2.1	Software Independent Archiving of Relational Databases	fmt/1196	Voorkeur			X
	ODB	1.2	OpenDocument Database Format	fmt/424	Voorkeur			X
Document (tekst)	PDF	2a	Acrobat PDF/A - Portable Document Format	fmt/476	Voorkeur	X		
	PDF	1a	Acrobat PDF/A - Portable Document Format	fmt/95	Voorkeur	X		

	PDF	1.7	Acrobat PDF 1.7 - Portable Document Format	fmt/276	Acceptabel	X		
	ODT	1.2	OpenDocument Text	fmt/291	Voorkeur	X		
	DOCX	vanaf 2007	Microsoft Word for Windows	fmt/412	Acceptabel	X		
E-mail	EML		Mime Email	fmt/950	Voorkeur	X		
	MSG	97-2003	Microsoft Outlook Email Message	x-fmt/249	Acceptabel		X	
Afbeelding (raster)	TIFF		Tagged Image File Format	fmt/353	Voorkeur	X		
	PNG	1.2	Portable Network Graphics	fmt/13	Voorkeur	X		
	JPEG		JPEG File Interchange Format	fmt/44	Acceptabel	X		
	JPG		Raw JPEG Stream	fmt/41	Acceptabel	X		
	JP2		JPEG 2000 part 1	x-fmt/392	Acceptabel	X		
	JPX		JPEG 2000 part 2	fmt/151	Acceptabel	X		
	GIF	89a	Graphics Interchange Format	fmt/4	Voorkeur	X		
Vector afbeelding	SVG	1.1	Scalable Vector Graphics	fmt/92	Voorkeur	X		
Presentatie	ODP	1.2	OpenDocument Presentation	fmt/293	Voorkeur	X		
	PDF	2a	Acrobat PDF/A - Portable Document Format	fmt/476	Voorkeur	X		
	PDF	1a	Acrobat PDF/A - Portable Document Format	fmt/95	Voorkeur	X		
	PDF	1.7	Acrobat PDF 1.7 - Portable Document Format	fmt/276	Acceptabel	X		
	PPTX	vanaf 2007	Microsoft Powerpoint for Windows	fmt/215	Acceptabel	X		

Spreadsheet	ODS	1.2	OpenDocument Spreadsheet	fmt/295	Voorkeur	X		
	CSV		Comma Separated Values	x-fmt/18	Voorkeur	X		
	XLSX	vanaf 2007	Microsoft Excel for Windows	fmt/214	Acceptabel	X		
Video	MXF		Material Exchange Format	fmt/200	Voorkeur		X	
	WEBM		WebM	fmt/573	Voorkeur	X		
	MP4		MPEG-4 Media File	fmt/199	Acceptabel	X		
	MKV	1-4	Matroska	fmt/569	Acceptabel		X	
Website	WARC	1.1	Web ARChive	fmt/1281	Voorkeur	X		
E-books	EPUB		ePub format	fmt/483	Voorkeur	X		