

Aanlevervoorwaarden voor e-depot



ERFGOED LEIDEN EN OMSTREKEN

Versie: 3.0
Datum: 12-03-2025

Versiebeheer

Versie	Opmerkingen	Datum
0.1-0.3	Conceptversie afgestemd intern Erfgoed Leiden en Omstreken	06-03-2019
0.9	Conceptversie te gebruiken bij testopnames	20-05-2019
1.0	Vastgesteld door directeur Erfgoed Leiden en Omstreken	12-12-2019
2.0	Vastgesteld door directeur Erfgoed Leiden en Omstreken	26-10-2021
3.0	Vastgesteld door adjunct- directeur Erfgoed Leiden en Omstreken	12-03-2025

Afkortingen en begrippen

Afkorting of begrip	
Archiefvormer	De (overheids)organisatie of het organisatieonderdeel waar het informatieobject gecreëerd of ontvangen is.
Digitaal archief	Een samenhangende verzameling van digitale bestanden en metadata, bijvoorbeeld bouwdoSSIers of ruimtelijke plannen.
E-depot	Het geheel van organisatie, beleid, processen en procedures, financieel beheer, personeel, databeheer, databeveiliging en aanwezige hard- en software, dat duurzaam beheren en raadplegen van te bewaren digitale archiefbescheiden mogelijk maakt.
Erfgoed Leiden	Erfgoed Leiden en Omstreken (Erfgoed Leiden) voert erfgoeddiensten uit en/of beheert historische collecties voor gemeenten en gemeenschappelijke regelingen waarmee een overeenkomst inzake beheer en bewaring van archieven is afgesloten.
Ingest	Het importeren van het digitale archief in het e-depot.
Metadata	Metadata zijn gegevens die de context, inhoud, structuur en vorm van informatie en het beheer ervan door de tijd heen beschrijven.
PAX	PAX is een pakket gebruikt door Preservica om informatieobjecten bestaande uit meerdere representaties of onderdelen te bundelen in één object.
Preservica	Leverancier van de e-depot (SAAS) oplossing van Erfgoed Leiden
OPEX	Open Exchange Format (OPEX) is een uitwisselingsformaat in gebruik door Preservica. Digitale archieven die worden opgenomen in het e-depot dienen metadata aan te bieden in het OPEX formaat om een soepele ingest mogelijk te maken.
MDTO	Metadatagegevens voor duurzaam toegankelijke overheidsinformatie (MDTO, opvolger van het TMLO) is een landelijke standaard, waarin is vastgelegd welke set van kenmerken (metadata) lokale en rijksoverheden tenminste aan informatieobjecten moeten meegeven, zodat informatie in het e-depot straks makkelijk te vinden en eenduidig te interpreteren is. De MDTO-standaard voorziet informatie op eenzelfde manier van metadata, met als doel om informatieprocessen en -systemen op elkaar aan te laten sluiten. Het draagt bij aan goede vindbaarheid en duurzame toegankelijkheid.

Inhoud

Versiebeheer	2
Afkortingen en begrippen	2
Inleiding	4
1. Overbrengen of uitplaatsen.....	5
2. Voorwaarden	6
2.1 Goede, geordende en toegankelijke staat	6
2.2 MDTO en OPEX mapping.....	6
2.3 MDTO-XML schema en OPEX metadata versie 1.0	7
2.4 Sectorspecifieke metadata	7
2.5 Sidecar structuur	7
2.6 Samengestelde records: PAX.....	8
2.7 Bestandsformaten	9
2.8 Lossless compressie.....	11
2.9 Fixity en aantallen controle.....	12
2.10 Wijze van transport	12
3. Overzicht van bijlagen	12
Bijlage 1 - Guide to OPEX ingest en Monitoring.....	12
Bijlage 2 - OPEX xsd	12

Inleiding

Dit document is bedoeld om een overzicht te geven van de voorwaarden waaraan over te brengen of uit te plaatsen digitale archief naar het e-depot van Erfgoed Leiden moet voldoen. Het is vooral bedoeld voor de leverancier van de bronapplicatie of technisch specialist van de archiefvormende organisatie.

Erfgoed Leiden gebruikt de SAAS oplossing van Preservica als e-depot.

1. Overbrengen of uitplaatsen

Digitaal **overheidsarchief** kan op verschillende manieren aan Erfgoed Leiden worden aangeleverd:

Overbrenging

Te bewaren afgehandeld/afgesloten digitaal archief (geen wijzigingen na aanlevering) wordt overgebracht na 10 jaar (nu nog 20 jaar, naar verwachting in 2026 wordt dit 10 jaar). Erfgoed Leiden wordt hierna formeel de beheerder van het archief.

Vervroegde overbrenging

Te bewaren afgehandeld/afgesloten digitaal archief wordt overbracht vóór de termijn van 10 jaar (nu nog 20 jaar, naar verwachting in 2026 wordt dit 10 jaar). Erfgoed Leiden wordt bij vervroegde overbrenging formeel de beheerder van het archief.

Uitplaatsing

Te bewaren en te vernietigen digitaal archief wordt overgedragen. Een groot gedeelte van dit archief is nog niet openbaar. Na verloop van tijd zal het digitaal archief óf (gedeeltelijk) overgebracht óf (gedeeltelijk) vernietigd moeten worden. De archiefvormer blijft de beheerder van het archief, totdat het te bewaren gedeelte formeel naar Erfgoed Leiden wordt overgebracht. Dat gedeelte komt dan in het beheer van Erfgoed Leiden.

Voor alle opties hierboven, waarbij er na de initiële overbrenging vervolgoverbrengingen volgen, geldt dat in overleg wordt besloten op welke termijn deze vervolg overbrengingen zullen plaatsvinden (1 keer per jaar, 1 keer per paar jaar etc.).

Particulier archief kan ten alle tijden (in overleg) overgebracht worden. Hiervoor geldt ook dat het over te brengen digitaal archief geschoond en afgesloten moet zijn. Voor particulier archief heeft Erfgoed Leiden de *Aanlevervoorwaarden Particulier Digitaal Archief* opgesteld.

2. Voorwaarden

2.1 Goede, geordende en toegankelijke staat

Het over te brengen digitaal archief dient aan de normen van de goede, geordende en toegankelijke staat (Archiefregeling 2010) te voldoen.

De goede staat betekent dat het digitaal archief o.a. virus-vrij, zonder corrupte bestanden en zonder encryptie (of met de encryptiesleutel) aangeleverd wordt.

De geordende en toegankelijke staat betreft de ordening en toegankelijkheid van het digitaal archief.

Mappen- en bestandsnamen

Algemeen geldt:

- Beschrijf mapnamen en bestandsnamen betekenisvol en goed
- Kies een vaste volgorde voor de opbouw van de bestandsnamen
- Gebruik de originele extensie en gebruik die termen nooit in de naam
- Schrijf de datum volgens het stramien JJJJMMDD
- De tekens < > : " / | \ ? * # & ~ % { } mogen niet in de bestandsnaam voorkomen.
- De letter- en tekencombinaties CON, PRN, AUX, NUL, COM1, COM2, COM3, COM4, COM5, COM6, COM7, COM8, COM9, LTP1, LTP2, LTP3, LTP4, LTP5, LTP6, LTP7, LTP8, LTP9, .lock, _vti_ en desktop.ini mogen niet in bestandsnamen voorkomen.
- Bestands- en padnamen mogen niet langer zijn dan 255 tekens.

2.2 MDTO en OPEX mapping

Erfgoed Leiden gebruikt het MDTO (Metagegevens voor Duurzaam Toegankelijke Overheidsinformatie en OPEX (Open Exchange Format) om metadata te importeren in het e-depot.

Voordat digitaal archief kan worden overgebracht naar het e-depot, vindt er een mapping plaats tussen het MDTO en de metadatavelden van het archiefbescheiden. Een mapping verschilt per digitaal archief. Naast de MDTO velden, kan er ook metadata worden meegenomen die niet in het MDTO staan, maar specifiek zijn voor het digitaal archief. In de OPEX metadata worden gegevens als de security tag, fixity en ID opgenomen die benodigd zijn voor een juiste ingest in het e-depot.

De MDTO en OPEX mapping inclusief de sectorspecifieke metadata voor een digitaal archief vormen samen het metadataschema.

2.3 MDTO-XML schema en OPEX metadata versie 1.0

Als het metadataschema in 2.2 is samengesteld, plaatst de archiefvormer (en als nodig de leverancier) de MDTO metadata (en eventuele sectorspecifieke metadata) in het XML document tussen de OPEX metadata zodat de velden ingelezen kunnen worden in het e-depot.

De links hieronder van het Nationaal Archief geven meer informatie:

MDTO XML schema specificatie, versie informatie en xsd:

<https://www.nationaalarchief.nl/archiveren/mdto>

XML voorbeelden per aggregatieniveau:

<https://www.nationaalarchief.nl/archiveren/mdto/xml-schema>

Meer informatie over OPEX is te vinden in bijlage 1 *Guide to OPEX ingest en Monitoring*. In bijlage 2 is de OPEX xsd te vinden.

De aangeleverde XML bestanden dienen gevalideerd te zijn aan zowel de OPEX als MDTO xsd's en met UTF-8 (8-bit Unicode Transformation Format) gecodeerd te zijn.

2.4 Sectorspecifieke metadata

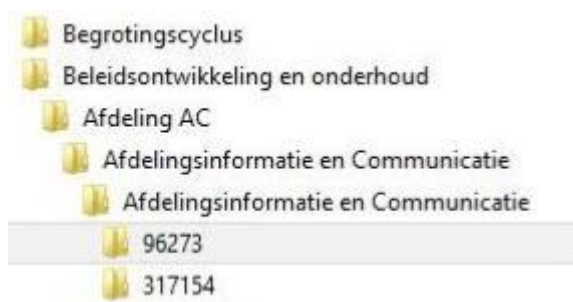
Het kan voorkomen dat niet alle benodigde metadata in het MDTO in te passen zijn. Dit kunnen sectorspecifieke metadata zijn die het generieke karakter van het MDTO overstijgen. In overleg met Erfgoed Leiden dienen deze metadata in een separaat metadataschema geplaatst te worden. Ook deze metadata moeten in hetzelfde XML bestand als de OPEX en MDTO metadata geplaatst worden (binnen het element OPEX.descriptiveMetadata).

2.5 Sidecar structuur

Digitaal archief dient door de archiefvormer aangeleverd te worden in een sidecar structuur.

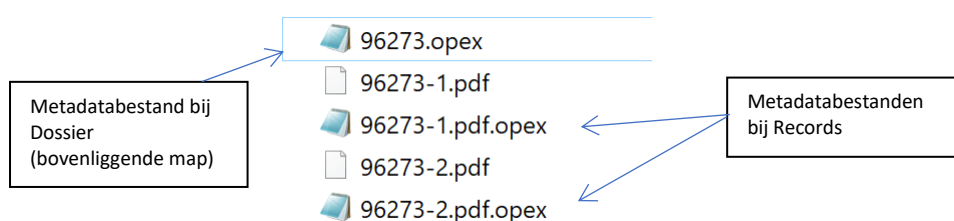
In een sidecar structuur zit de content van het digitaal archief in een directory-structuur: dat wil zeggen elk archief, serie, dossier (zaak), subdossier en document is een map en elke versie van een document is een bestand. Elke map (archief, serie, dossier, subdossier, etc.) en elk bestand (versie van een document) heeft zijn eigen metadatabestand, een zogenoemde sidecar. De naam van een metadatabestand is identiek aan die van de map of bestand waar het bij hoort, aangevuld met de extensie '.opex'. Let erop dat ook de extensie van het bestand opgenomen wordt in de bestandsnaam van het metadatabestand. Bijvoorbeeld 'bestand.docx.opex'. De plaats van een metadatabestand die bij een map hoort, is in de map (dus op een lager niveau). De plaats van een metadatabestand die bij een bestand hoort, is naast dit bestand (dus op hetzelfde niveau in een map).

In de Nederlandse archiefwereld worden veelal vijf types aggregatieniveau gebruikt. Het hoogste aggregatieniveau is *archief*. Daarna kunnen er één of meerdere aggregatieniveaus van het type *serie volgen*. Dan volgt er tenminste één aggregatieniveau van het type *dossier* (ook wel zaak). Dan volgen er één of meerdere aggregatieniveaus van het type *record*, hoewel dit aggregatieniveau niet in alle gevallen verplicht is. Hieronder een voorbeeld:



Het laagste aggregatieniveau is van het type *bestand*. We spreken hiervan indien een record uit meerdere onderdelen bestaat. Bijvoorbeeld een email met een bijlage, een 3D bestand bestaande uit een .obj, .mtl en .png bestand óf een record met meerdere versies. Voor dit aggregatieniveau hoeven geen metadatabestanden aangeleverd te worden. Het e-depot genereert de technische metadata van deze bestanden automatisch bij ingest. Tevens bundelt het systeem deze losse bestanden samen in het record, waardoor ze duurzaam als één pakket worden opgeslagen.

Hieronder als voorbeeld de inhoud van *dossier* ofwel map '96273' in het vorige plaatje met daarin bestanden en metadata:



Het dossier (mapje 96273) bevat dus:

- het metadatabestand bij het *dossier* '96273.opex'
- de *records* '96273-1.pdf' en '96273-2.pdf'
- de metadatabestanden bij deze twee *records*, te weten '96273-1.pdf.opex' respectievelijk '96273-2.pdf.opex'

Het systeem is hoofdlettergevoelig. Let er dus op dat de naam van zowel het record/de map als het bijbehorende metadatabestand op exact de wijze zijn gespeld inclusief upper en lower case.








2.6 Samengestelde records: PAX

Indien een record uit meerdere bestanden bestaat (zoals hierboven besproken) dienen deze aangeleverd te worden in een PAX. Dit kan je zien als een soort container die ervoor zorgt dat losse bestanden gebundeld worden in het e-depot.

Hieronder is een voorbeeld te zien waarin de afzonderlijke pagina's van een boek gebundeld worden. Deze worden in het e-depot opgeslagen als de preserveringskopie (Representation_Preservation). Tevens wordt er als toegangskopie (Representation_Access) een pdf versie van het boek in het record gebundeld.

```
book.pax.zip
  Representation_Preservation/
    page_001/
      page_001.tiff
    page_002/
      page_002.tiff
    page_003/
      page_003.tiff
    etc
  Representation_Access/
    complete_book/
      complete_book.pdf
```

In de aan te leveren directory structuur komt dit er als volgt uit te zien.

- ▼  book.pax.zip
- ▼  Representation_Access
 - >  complete_book
- ▼  Representation_Preservation
 - >  page_001
 - >  page_002
 - >  page_003

Indien het gebruik van PAX nodig is zal de adviseur digitale archieven dit aangeven en wordt extra documentatie en begeleiding verschaft.

2.7 Bestandsformaten

Erfgoed Leiden kan alle bestandsformaten die in het PRONOM register staan opslaan in het e-depot. Het preservatiebeleid leidt echter tot een kleinere set van voorkeursformaten en geaccepteerde formaten. Hierbij is van de volgende uitgangspunten gebruik gemaakt:

- Open bestandsformaten hebben de voorkeur. Dit zijn bestandsformaten die niet van bepaalde software afhankelijk zijn.
- Duurzame bestandsformaten hebben de voorkeur. Dit zijn bestandsformaten waarvan we verwachten dat deze in de toekomst ook nog te openen zijn.
- Indien een bestandsformaat niet open is, maar wel de de facto standaard is, wordt deze in veel gevallen geaccepteerd. Hierbij is ook in acht genomen of het in Preservica mogelijk is het bestandsformaat te normaliseren of migreren naar een open bestandsformaat.
- Erfgoed Leiden accepteert geen bestanden die in een container (bijvoorbeeld zip) worden aangeleverd. De enige uitzondering hierop is het WARC bestandsformaat en de bovengenoemde PAX.
- Bestandsformaten die niet in het PRONOM register staan worden niet geaccepteerd (m.u.v. video en audio codecs).

- Indien een bestandsformaat als 'endangered' is bestempeld in de BitList van de Digital Preservation Coalition wordt deze niet opgenomen in het e-depot, tenzij hierop een normalisatie of migratie kan worden uitgevoerd.

Toepassingsgebied	Voorkeursformaat	Geaccepteerd formaat	Toegestane Pronom ID's
Afbeelding	TIFF, JPEG2000	PNG, JPEG, GIF	fmt/353 (tiff), x-fmt/392 en fmt/151 (jpeg2000), fmt/42, fmt/43 en fmt/44 (jpeg), fmt/3 en fmt/4 (gif)
Afbeeldingen binnen tekstbestanden	PNG		fmt/11, fmt/12 en fmt/13
Vector afbeelding	SVG		fmt/91, fmt/92, fmt/413
Tekst	ODF (.odt en .ott), PDF/A-1a, PDF/A-2a, PDF/A-1b (alleen voor gescande documenten)	DOCX, PDF 1.7, EPUB	fmt/136, fmt/290 en fmt/291 (odt), fmt/95 (PDF/A-1a), fmt/476 (PDF/A-2a) fmt/354 (PDF/A-1b), fmt/412 (docx), fmt/276 (pdf 1.7), fmt/483 (epub)
Platte tekst	TXT		x-fmt/111
Audio	WAV, WAVE, RF64	MP3	fmt/6, fmt/1, fmt/2, fmt/527, fmt/703, fmt/704, fmt/705, fmt/706, fmt/707, fmt/708, fmt/709, fmt/710, fmt/711, fmt/712 (wav/wave, rf64), fmt/134 (mp3)
Audio codecs	Linear Pulse Code Modulation (LPCM), Free Lossless Audio Codec (FLAC)	MPEG-4 Advanced Audio Coding / Audio Lossless Coding (AAC/ALS)	LOC ID fdd000198 (FLAC), LOC ID fdd000011 (LPCM), fmt/1812 (AAC).
Video	MXF, MKV (alleen in combinatie met FFV1 codec)	MP4	fmt/791 (mxf), fmt/569 (mkv), fmt/199 (mp4)
Video codecs	XDCAM HD422 video codec voor MXF container, FF Video Codec 1 (FFV1) voor Matroska container	MPEG-4, Advanced Video Coding (Part 10/H.264)	LOC ID fdd000341 (FFV1), LOC ID fdd000081 (Part 10/H.264)
Email	EML	MSG	fmt/278, fmt/950 (eml), x-fmt/430

Website	WARC		fmt/1355 en fmt/1281
Spreadsheet	ODF (.ods, .ots), CSV	XLSX	fmt/137, fmt/294 en fmt/295 (ods), x-fmt/18 (csv), fmt/214 (xlsx)
Presentatie	ODF (.odp, otp)	PPTX, PUB, VSD	fmt/138, fmt/292 en fmt/293 (odp), fmt/215 en fmt/1829 (pptx), fmt/1516 (pub), x-fmt/113 (vsd)
Database	ODF (.odb), SQL, SIARD	ACCD	fmt/140, fmt/444 en fmt/424 (odb), fmt/206 (sql), fmt/1196, fmt/275 (accdb)
Data uitwisseling en code	XML, CSS, JSON, DTD, XSD, HTML		fmt/101 (xml), fmt/817 (json), x-fmt/315 (dtd), x-fmt/224 (css), x-fmt/280 (xsd), fmt/471 (html)
3D	OBJ gecombineerd met een GLB access kopie	GLB	fmt/1210 (obj), fmt/1316 (glb)
CAD en Geo	DXF, GML, GeoTIFF, GPKG, GeoJSON	KML, ESRI Shapefile bestaande uit: SHP (verplicht), SHX (verplicht), DBF (verplicht), XML Esri Shapefile Geospatial Metadata File (sterk aangeraden), PRJ (sterk aangeraden), SBN, SBX.	fmt/63 (dxf), x-fmt/227 en fmt/1047 (gml), fmt/155 (geotiff), fmt/1700 (gpkg), fmt/1367 (geojson), fmt/244 (kml), x-fmt/235 (shp), fmt/277 (shx), x-fmt/272 (dbf), fmt/1729 (xml Esri Shapefile Geospatial Metadata File), fmt/320 (prj), fmt/319 (sbn en sbx)

*Bestandsformaten die rood zijn kunnen in Preservica worden opgenomen, maar niet worden getoond. Dikgedrukte bestandsformaten kunnen opgeslagen worden, maar enkel na migratie in een ander bestandsformaat worden getoond.

Van de meeste bestandsformaten bestaan verschillende versies. Niet elke versie wordt geaccepteerd. Controleer daarom in de tabel welke Pronom ID's aangeboden kunnen worden.

Indien een bestandsformaat niet in de bovenstaande tabel is opgenomen, maar zich wel in het over te dragen archief bevindt dient de archiefvormer contact op te nemen met Erfgoed Leiden om de eventuele mogelijkheden te bespreken.

Je vindt het PRONOM register en meer informatie hier: www.nationalarchives.gov.uk/PRONOM/Default.aspx.

2.8 Lossless compressie

Het is niet toegestaan gebruik te maken van een compressietechniek op het over te brengen digitaal archief.

2.9 Fixity en aantallen controle

Bij de aanlevering van archiefbescheiden aan Erfgoed Leiden dient van elk archiefstuk een hashcode (fixity) meegegeven te worden in het SHA-256 algoritme. Deze code moet zowel in het XML schema als in een aparte Excel lijst aangeleverd worden. Met deze code controleert Erfgoed Leiden of het bestand tijdens de aanlevering is beschadigd. Indien dit het geval blijkt zal het bestand opnieuw opgevraagd worden bij de archiefvormer. Tevens moet uit dit Excel bestand duidelijk worden hoeveel en welke bestanden er aan Erfgoed Leiden geleverd zijn ten behoeve van de controle op volledigheid van de aanlevering.

2.10 Wijze van transport

Het over te brengen archief dient op een server van Erfgoed Leiden geplaatst te worden. Het transport vindt plaats via een FTPS verbinding. De leverancier van de bronapplicatie of de archiefvormende instelling dient zelf zorg te dragen voor de benodigde software om deze verbinding te kunnen leggen. De afzender partij geeft aan Erfgoed Leiden door via welke IP-adres het archief wordt verzonden. Erfgoed Leiden verschaft een gebruikersnaam en wachtwoord om de verbinding tot stand te brengen. Ook is er een handleiding beschikbaar met gedetailleerde uitleg over de werking van het transport.

Bij het transport mogen geen verborgen thumbs.db of directory.ini bestanden meegeleverd worden.

3. Overzicht van bijlagen

Bijlage 1 - Guide to OPEX ingest en Monitoring

Zie bijgevoegd document

Bijlage 2 - OPEX xsd

Zie bijgevoegd document