

Opnamevoorwaarden Groninger Archieven e-depot Versie 1.1
Laatst bijgewerkt op 8 december 2022

Inhoudsopgave

1. Inleiding.....	3
2. Ordening van het archief in aggregatieniveaus	4
3. Sidecar-structuur	6
4. Alleen afgesloten dossiers.....	6
5. Sidecarstructuur bij aanvullingen.....	6
6. Naamgeving van mappen en bestanden	7
7. Beperking op de omvang van mappen en bestanden.....	7
8. De inhoud van de metadatabestanden	7
9. De vorm van de metadatabestanden.....	9
10. Checksum voor controle op fysieke integriteit bestanden.....	10
11. Transport en controle.....	10
12. Inleiding en verantwoording bij overbrenging.....	10

1. Inleiding

Als u archiefbescheiden wil overbrengen of uitplaatsen in het e-depot, dan moeten deze aan een aantal voorwaarden voldoen. Alleen dan kunnen de Groninger Archieven garanderen dat we ze kunnen opnemen, beheren, preserveren en beschikbaarstellen. De voorwaarden hebben betrekking op:

- de structuur van het archief
- de vorm en inhoud van de metadata
- de bestandsformaten

In dit document worden de eerste twee typen voorwaarden behandeld.

Informatie over de bestandsformaten die door het e-depot worden ondersteund vindt u in de Voorkeursformaten Groninger Archieven.¹ Als u informatieobjecten wilt overdragen die niet in deze lijst voorkomen of die gecomprimeerd zijn (zoals zip-bestanden), neem dan contact met ons op.

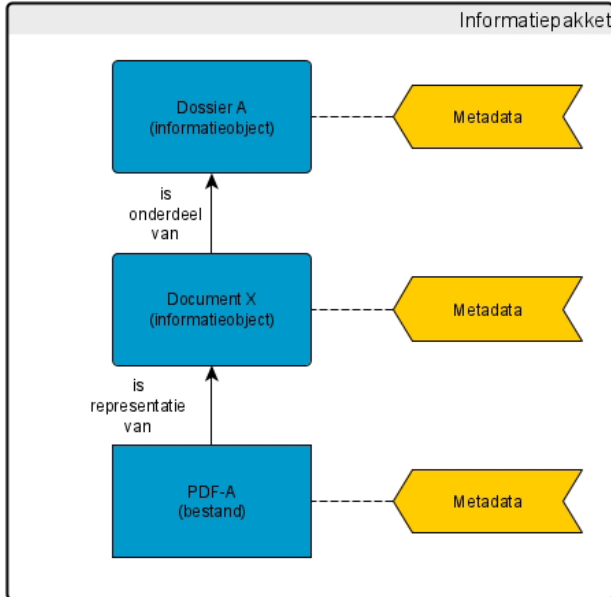
Wanneer u een archiefbestanddeel overbrengt of uitplaatst, vragen wij u om ons zelfevaluatieformulier² in te vullen en een datadump mee te sturen. Aan de hand van dat formulier en de datadump doen wij een impactanalyse: we onderzoeken onder meer in hoeverre de set digitale informatieobjecten en de bijbehorende metadata voldoen aan onze opnamevoorwaarden en voorkeursformaten. U bent zelf verantwoordelijk voor het realiseren van een export. Mogelijk moet u de leverancier van uw applicatie vragen om een exportmodule te leveren of te bouwen. Deze opnamevoorwaarden beschrijven de eisen waaraan de export moet voldoen. Kunt u aan één of meer van de eisen niet voldoen? Neem dan contact op.

¹ Download de actuele versie op <https://www.groningerarchieven.nl/over-ons/diensten/e-depot>

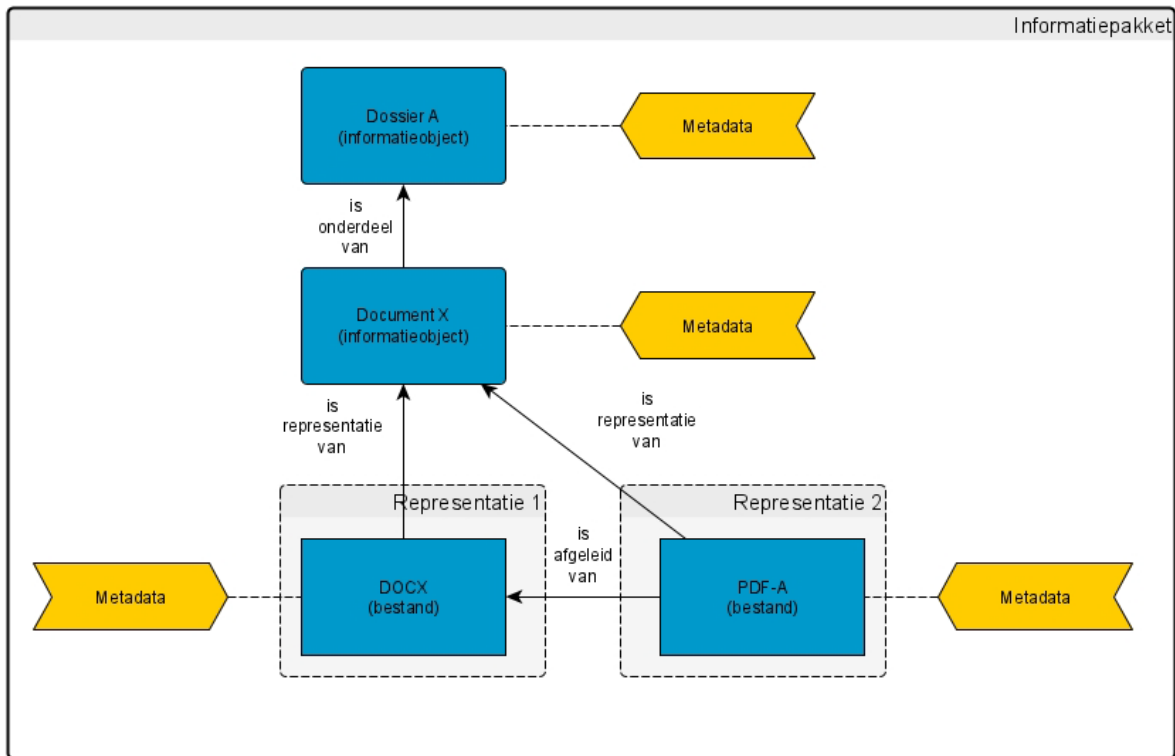
² Download de actuele versie op <https://www.groningerarchieven.nl/over-ons/diensten/e-depot>

2. Ordening van het archief in aggregatieniveaus

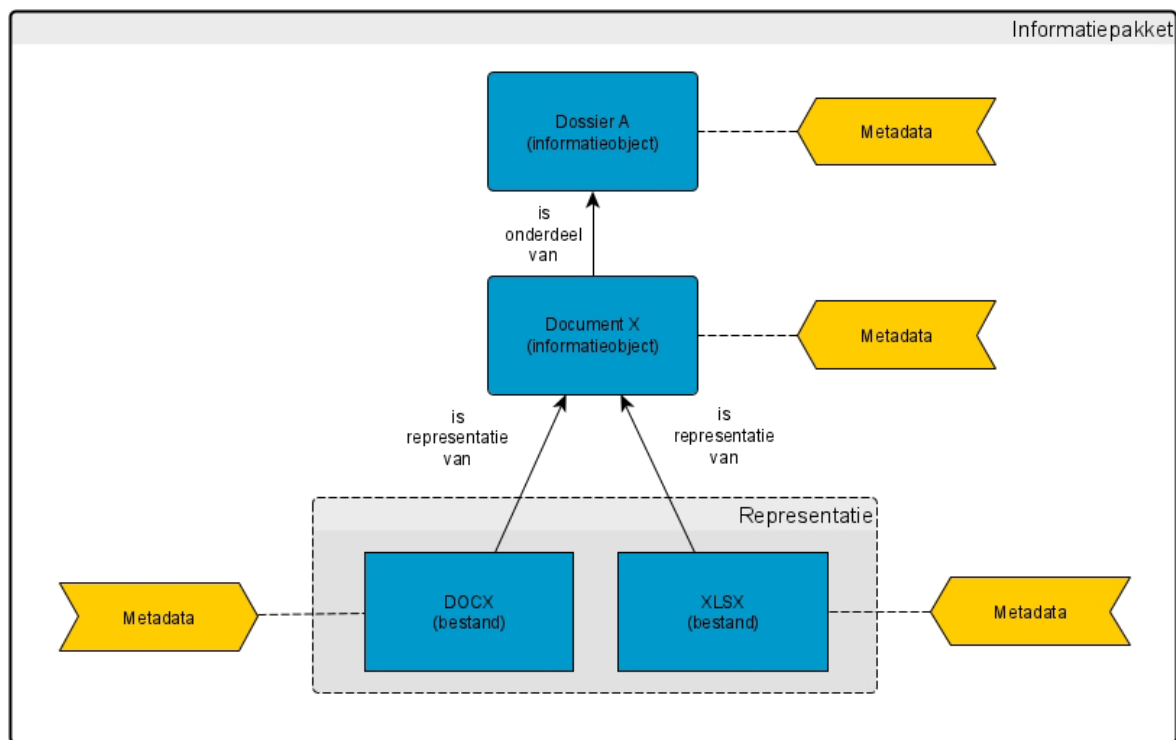
De Groninger Archieven maken onderscheid tussen informatieobjecten en de computerbestanden die de representaties ervan vormen. In de meest eenvoudige vorm heeft één informatieobject één representatie in de vorm van één computerbestand:



Maar er kunnen ook meerdere representaties naast elkaar bestaan, wanneer het ene informatieobject is opgeslagen in twee verschillende bestandsformaten:



Ook is het mogelijk dat twee computerbestanden samen één representatie van één informatieobject vormen, bijvoorbeeld een advies dat bestaat uit een rapport met een bijlage:



Informatieobjecten zijn vaak onderdeel van een hiërarchische structuur die uit verschillende aggregatieniveaus bestaat, zoals bijvoorbeeld een archief bestaande uit zaakdossiers die weer bestaan uit documenten.

Een set digitale informatieobjecten die wordt overgebracht of uitgeplaatst moet in goede geordende en toegankelijke staat verkeren. Wanneer het archief zich in een documentmanagementsysteem of procesapplicatie of op netwerkschijven bevindt, is het geordend volgens de inrichting van dat systeem. Bij de overdracht moet zoveel mogelijk de oorspronkelijke ordening ten tijde van de vorming van het archief worden gehandhaafd.

Om een set informatieobjecten te kunnen opnemen in het e-depot moet deze ordening worden vertaald naar de ordening van een metadatastandaard die door het e-depot wordt ondersteund (zie hoofdstuk 9). Omwille van het gebruiksgemak hanteren wij bij beide standaarden een maximum van zeven aggregatieniveaus.

De vertaling van de oorspronkelijke ordening naar die van de metadatastandaard is onderdeel van de metadatamapping.

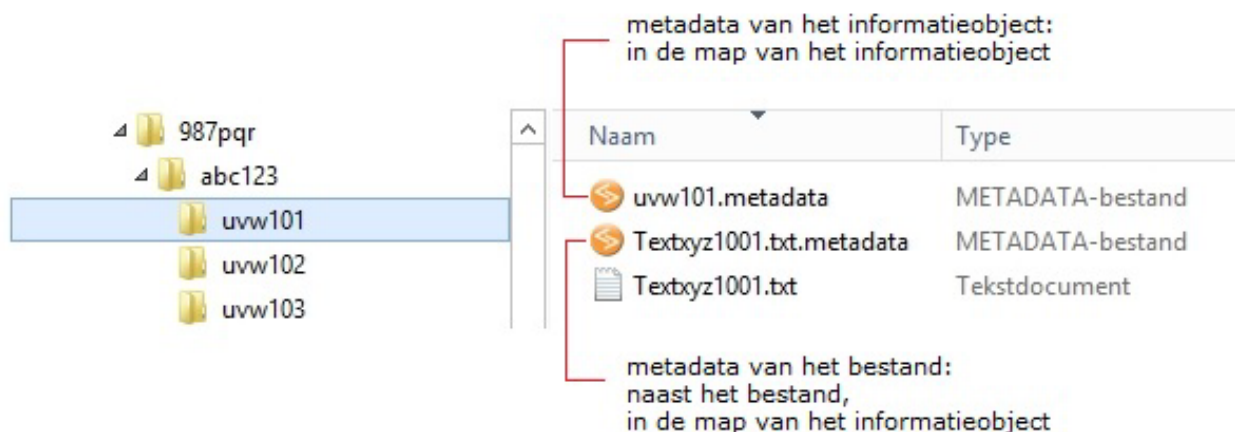
Bij de aansluiting van een nieuw systeem op ons e-depot begeleiden we u bij de eerste mapping. Bij vervolgaanleveringen voert u de mapping zelf uit en toetsen wij of deze aan de eisen voldoet. De mapping gebeurt aan de hand van ons mappingformulier.³ Het formulier geeft nadere toelichting op de structureisen van de betreffende metadatastandaard.

³ Download de actuele versie op <https://www.groningerarchieven.nl/over-ons/diensten/e-depot>

3. Sidecar-structuur

De aggregatieniveaus in uw systeem worden in de export uitgedrukt in een zogenoemde sidecar-structuur. Hierin zitten de aggregatieniveaus in een directory-structuur: de informatieobjecten zijn mappen in een hiërarchische mappenstructuur. De bestanden bevinden zich in de mappen van de informatieobjecten waarvan ze representaties zijn. Elke map en elk bestand 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 '.metadata'. Een metadatabestand dat bij een map hoort staat in deze map zelf. Een metadatabestand dat bij een bestand hoort staat naast dit bestand in dezelfde map.

Een sidecar-structuur kan er bijvoorbeeld zo uitzien:



4. Alleen afgesloten dossiers

De export mag alleen afgesloten dossiers bevatten. Nog lopende (zaak)dossiers kunnen niet in het e-depot worden opgenomen, omdat volledig geautomatiseerde toevoeging binnen dossiers op dit moment niet wordt ondersteund. Het is wel mogelijk nieuwe, afgesloten dossiers toe te voegen aan een bestaande ordening die al in het e-depot staat. Zie hiervoor 5.

5. Sidecarstructuur bij aanvullingen

Wanneer u na overdracht van een archief nieuwe informatieobjecten wilt toevoegen aan een bestaand archief, moet deze aanvulling in dezelfde sidecarstructuur worden aangeboden. Ook voor de bovenliggende aggregatieniveaus moeten mappen worden gemaakt met bijbehorende metadatabestanden. Voor elk van die aggregatieniveaus geldt dat het identificatiekenmerk in het metadatabestand gelijk moet zijn aan het identificatiekenmerk van de bestaande aggregatie in het e-depot. Let op: eventuele wijzigingen in de metadata van reeds bestaande aggregaties worden door het e-depot genegeerd! Mocht wijziging van de bestaande metadata noodzakelijk zijn, geef dit dan duidelijk aan tijdens de zelfevaluatie.

Een vervollexport mag niet een dossier bevatten met hetzelfde identificatiekenmerk als een bestaand dossier binnen dat archief.

6. Naamgeving van mappen en bestanden

De Groninger Archieven hanteren een naamgevingsconventie voor de mappenstructuur waarin digitale informatieobjecten worden opgeslagen. Daarnaast stelt de uitwisselbaarheid van informatieobjecten en metadatabestanden tussen verschillende besturingssystemen ook eisen aan de map- en bestandsnamen. Bij het maken van de sidecarstructuur moet u daarom rekening houden met het volgende:

- Namen van submappen en namen van bestanden moeten uniek zijn binnen de map waarin ze zich bevinden.
- Er kan niet vertrouwd worden op hoofdlettergevoeligheid om mapnamen of bestandsnamen van elkaar te onderscheiden. Bij de toepassing van de eis van uniciteit dient daarmee rekening te worden gehouden. Het identificatiekenmerk kan in de bijbehorende metadata wel hoofdlettergevoelig worden vastgelegd.
- Een mapnaam moet gelijk zijn aan het identificatiekenmerk in de bijbehorende metadata.
- Elk bestand moet zijn voorzien van een bestandsextensie.
- De karakters < > : " / \ | ? * # & mogen niet voorkomen in de map- en bestandsnamen.
- Map- en bestandsnamen mogen niet gelijk zijn aan **AUX, COM1, COM2, COM3, COM4, COM5, COM6, COM7, COM8, COM9, CON, LPT1, LPT2, LPT3, LPT4, LPT5, LPT6, LPT7, LPT8, LPT9, NUL** of **PRN**. Een map- of bestandsnaam mag deze strings wel bevatten.
- Er is geen tekenlimiet voor padnamen.
- Map- en bestandsnamen mogen uit maximaal 255 tekens bestaan.

7. Beperking op de omvang van mappen en bestanden

We hanteren per export een limiet van 100.000 mappen en bestanden of 500 GB data. Dit aantal is exclusief de metadatabestanden. Wanneer u meer dan 100.000 mappen en bestanden wilt overdragen, moet u de overdracht opsplitsen. U moet de verschillende sets dan wel tegelijk aanleveren. Houdt u bij de tweede set bestanden rekening met de eisen die worden gesteld aan een aanvulling (zie paragraaf 5).

Het e-depot is in staat om bestanden met een omvang van 0 bytes te verwerken. Een bestand met bestandsgrootte van 0 bytes is echter een corrupt bestand of een bestand zonder inhoud. Daarom moet in zo'n geval de integriteit worden toegelicht in de metadata.

8. De inhoud van de metadatabestanden

Om een archief duurzaam toegankelijk te houden en in het e-depot te kunnen beheren, conserveren en beschikbaar te stellen moet het zijn voorzien van voldoende en gestandaardiseerde metadata. De Groninger Archieven ondersteunen twee uitwisseldatastandaarden: het Toepassingsprofiel Metadatering Lokale Overheden (TMLO) 1.1⁴ en het metagegevensschema Metagegevens Duurzaam Toegankelijke Overheidsinformatie (MDTO) 1.0⁵. De metadata in uw bronsysteem moeten worden vertaald naar één van beide standaarden. Welke standaard het meest geschikt is, hangt af van de exportfunctionaliteit van uw bronsysteem en van de ordening van de informatieobjecten in uw bronsysteem. Bij de aansluiting van een nieuw systeem op ons e-depot begeleiden we u bij de eerste mapping. Bij vervolgaanleveringen voert u de

⁴ Download TMLO 1.1 op: <https://www.nationaalarchief.nl/archiveren/kennisbank/tmlo>

⁵ Raadpleeg MDTO 1.0 op: <https://www.nationaalarchief.nl/archiveren/mdto>

mapping zelf uit en toetsen wij of deze aan de eisen voldoet. Bij de mapping maakt u gebruik van ons Formulier metadatamapping.⁶

We stellen de volgende eisen aan de metadata:

- U levert minimaal de metadata die in TMLO of MDTO verplicht zijn.
- De metadata van de verschillende aggregatieniveaus van de informatieobjecten vormen samen een zogeheten 'gehopte' beschrijving van elk informatieobject. De metadata geven samen informatie over:
 - GEzichtspunt
 - Handeling
 - Object
 - Plaats
 - Tijd
- Als u bij creatie van het archief een classificatie hanteerde, dan is minimaal de omschrijving van de classificatie en bij voorkeur ook de code vastgelegd in de metadata.
- Beschrijvingen zijn bij voorkeur geschreven vanuit het perspectief van de archiefvormer.
- Bij overbrenging is het archief op alle aggregatieniveaus voorzien van metadatering over openbaarheid conform de MDTO-begrippenlijst beperkingGebruik opgesteld door de Architectuurcommissie van de RHC's en het NA.⁷ Dit geldt zowel bij gebruik van MDTO als bij gebruik van TMLO/ToPX.
- Bij overbrenging is het archief op een zo laag mogelijk niveau voorzien van metadatering over gebruiksrechten, conform de MDTO-begrippenlijst beperkingGebruik opgesteld door de Architectuurcommissie van de RHC's en het NA. Dit geldt zowel bij gebruik van MDTO als bij gebruik van TMLO/ToPX. Vastlegging op het laagste niveau is noodzakelijk omdat informatieobjecten waarop geen gebruiksrechten berusten niet onnodig mogen worden uitgesloten van online publicatie en downloadbaarheid.
- Bij uitplaatsing is het archief op alle aggregatieniveaus voorzien van metadatering over vertrouwelijkheid. Bij gebruik van TMLO/ToPX worden de door ToPX voorgeschreven begrippen gebruikt. Bij gebruik van MDTO worden de begrippen gehanteerd van de MDTO-begrippenlijst beperkingGebruik.
- Een informatieobject mag geen lager openbaarheidsniveau of vertrouwelijkheidsniveau hebben dan de bovenliggende aggregatie. Een beperkt openbaar dossier mag geen openbare archiefstukken bevatten. Een vertrouwelijk archief mag geen niet-vertrouwelijke dossiers bevatten.
- Bij uitplaatsing is het archief voorzien van metadata over waardering en selectie. De metadata beschrijven de te nemen actie (overbrengen of vernietigen), de bewaartermijn, de grondslag en de actiedatum of aanleiding/trigger van de actie. De metadata worden vastgelegd op het aggregatieniveau waarop de selectie plaatsvindt.
- Bij overbrenging komt het identificatiekenmerk op het aggregatieniveau archief overeen met het toegangsnummer dat u van ons aangeleverd krijgt.
- Elk ander identificatiekenmerk wordt ontleend aan het bronsysteem.
- Een identificatiekenmerk is uniek binnen de aggregatie waarvan hij deel uitmaakt. Dit is noodzakelijk omdat de mapnaam wordt ontleend aan het identificatiekenmerk en een mapnaam uniek moet zijn binnen de map waarin hij zich bevindt (zie hoofdstuk 6).

⁶ Download de actuele versie op <https://www.groningerarchieven.nl/over-ons/diensten/e-depot>

⁷ Download de actuele versie op <https://kia.pleio.nl/groups/view/184556a0-c96b-42a0-a2ee-b8f97f61b053/kennisplatform-metadata/blog/view/fb889d97-d395-41f3-96d6-bb6d61938662/voorstel-mdto-begrippenlijst-beperkinggebruik>

- Een identificatiekenmerk mag uit niet meer dan 255 karakters bestaan, omdat anders de maximale lengte van de mapnaam wordt overschreden (zie hoofdstuk 6).
- Dateringen worden vastgelegd volgens de voorkeursdateringen van de Groninger Archieven of - als dat niet mogelijk is - conform de interpretatie van de Groninger Archieven. Beide worden toegelicht in het mappingformulier.
- Bij overbrenging moeten metadata van vernietigde informatieobjecten zoveel mogelijk worden geschoond.
- Oude identificatiekenmerken en identificatiekenmerken uit andere systemen moeten in de metadata worden vastgelegd, zodat informatieobjecten ook op die nummers vindbaar blijven.
- Als u een grondslag had voor het registreren van persoonsgegevens conform de AVG, dan mogen de persoonsgegevens mee overgedragen worden. Anonimisering is niet nodig, maar bij overbrenging kan het wel nodig zijn om de openbaarheid te beperken.
- Persoonsgegevens worden indien mogelijk vastgelegd in een apart veld, zodat ze kunnen worden afgeschermd voor onbevoegde gebruikers indien de AVG dat eist.
- Metadata die bepalend zijn voor de integriteit, authenticiteit en betrouwbaarheid moeten mee worden overgebracht.

In uw documentmanagementsysteem of procesapplicatie kunnen op allerlei niveaus metadata zijn opgeslagen: inhoudelijke metadata, technische metadata en ook logfiles. Sommige metadata zijn alleen zichtbaar voor applicatiebeheerders. Niet alleen de structuur van uw systeem moet vertaald worden naar de uitwisselstandaard, er moet ook voor elk niveau in uw systeem worden onderzocht welke metadata noodzakelijk en wenselijk zijn voor duurzame toegankelijkheid en hoe ze naar de uitwisselstandaard kunnen worden omgezet. Op archiefniveau zal de meeste informatie waarschijnlijk niet met uw systeem gemapt kunnen worden, maar u kunt daarvoor vaste waarden gebruiken. Ook is het mogelijk dat u relevante informatie verspreid over meerdere systemen bewaart en dat een combinatie van bronnen nodig is.

De uitwisselstandaarden TMLO en MDTO beschrijven uitsluitend generieke metadata die noodzakelijk zijn voor duurzame toegankelijkheid. Bij uitplaatsing kan het wenselijk zijn om daarnaast ook metadata op te nemen die noodzakelijk zijn voor procesondersteuning. Bij overbrenging kan het voor de duurzame toegankelijkheid noodzakelijk zijn om ook domeinspecifieke metadata op te nemen. Opname van aanvullende metadata conform een breedgedragen standaard is mogelijk. Tijdens de (toetsing van de) metadatamapping wordt in overleg vastgesteld welke aanvullende metadata noodzakelijk of gewenst zijn en hoe deze kunnen worden overgedragen.

9. De vorm van de metadatabestanden

Om de metadata te kunnen opnemen en gebruiken in het e-depot moeten de metadatabestanden worden aangeleverd in de vorm van een XML-schema. Bij gebruik van TMLO is dat het uitwisselformaat ToPX 2.3.2⁸. Bij gebruik van MDTO is dat MDTO-XML; MDTO-RDF wordt nog niet ondersteund. Wanneer u archiefbescheiden aan ons overdraagt, levert u eerst een representatieve testexport aan. Of uw metadatabestanden voldoen aan het XML-schema kunt u controleren met behulp van een XSD⁹.

De metadatabestanden moeten worden opgeslagen met tekencodering UTF-8 zonder BOM (Byte Order Mark).

⁸ Download de specificatie van ToPX 2.3.2 op <https://www.nationaalarchief.nl/archiveren/kennisbank/metagegevens-en-het-e-depot>

⁹ Download de XSD voor ToPX 2.3.2 op <https://www.nationaalarchief.nl/archiveren/kennisbank/metagegevens-en-het-e-depot>

10. Checksum voor controle op fysieke integriteit bestanden

Wanneer we uw archief opnemen in het e-depot willen we er zeker van zijn dat er tijdens de overdracht geen bestanden verdwijnen of gemuteerd worden. Ook controleert het e-depot tijdens de opslag doorlopend op bitintegriteit. Om de fysieke integriteit te controleren maken we gebruik van een checksum. Het Forum Standaardisatie adviseert minimaal het gebruik van een SHA2-algoritme om een checksum te genereren.¹⁰ Het e-depot ondersteunt SHA-256 en SHA-512. Daarnaast ontvangen we ook graag andere checksums, zoals MD5, indien aanwezig. Als uw systeem een ander algoritme gebruikt, neem dan vooraf contact op.

Het algoritme en de waarde worden vastgelegd in de metadata op bestandsniveau.

11. Transport en controle

Voor het transport van de export maken we gebruik van een beveiligde FTP-verbinding. De export komt dan binnen in onze digitale ontvangstruimte. Hier controleren we allereerst of de export compleet is. Hiervoor hebben we een drietal gegevens nodig:

- het totaal aantal bestanden in de export exclusief metadatabestanden
- de totaalomvang van de export
- een SHA-256 of SHA-512 checksum voor de complete export

Vervolgens controleren we of de export voldoet aan de opnamevoorwaarden en aan de gemaakte afspraken, waaronder de mapping.

Wanneer uw export padnamen bevat die langer zijn dan 255 tekens, kan dat bij het transport voor problemen zorgen. Om dat te voorkomen moet u de export in een container met tar-formaat plaatsen met behulp van een tool die geen beperking aan de padnaamlengte stelt, zoals de tool 7-Zip en de tool GNU tar.

De container mag alleen de mappenstructuur bevatten, met het archiefniveau als hoogste map, en mag niet gecomprimeerd worden aangemaakt. De naam van de container begint met een aanwinstnummer toegekend door de Groninger Archieven, gevolgd door datum van creatie (in formaat jjjjmmdd), gevolgd door het batchnummer in geval van aanlevering in verschillende batches. Daarachter kunt u eventueel naar eigen invulling informatie toevoegen.

12. Inleiding en verantwoording bij overbrenging

Bij overbrenging van archief is het noodzakelijk om het archief voor het publiek van een inleiding te voorzien en de eventuele bewerking van het archief te verantwoorden. Wij spreken vantevoren met u af welke informatie wij daarvoor nodig hebben en hoe u die kunt aanleveren.

¹⁰ Zie <https://www.forumstandaardisatie.nl/standaard/sha-2>