

Opnamevoorwaarden rijks e-depot

Versie 1.1

Laatst bijgewerkt op 29 november 2019

Inhoudsopgave

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

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 Handreiking voorkeursformaten van het Nationaal archief.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. Aan de hand van dat formulier doen wij een impactanalyse: we onderzoeken onder meer in hoeverre de set digitale informatieobjecten en de bijbehorende metadata voldoen aan onze opnamevoorwaarden en aan de lijst van voorkeursformaten. U bent zelf verantwoordelijk voor het realiseren van een export. Mogelijk moet u de leverancier van uw documentmanagementsysteem, procesapplicatie of netwerkschijven vragen om een exportmodule te leveren of te bouwen. Deze opnamevoorwaarden beschrijven de eisen waaraan de export moet voldoen.

2. Ordening van het archief in aggregatieniveaus

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. Het e-depot dwingt echter een eigen ordening af.

Het e-depot kent vijf aggregatieniveaus, van hoog naar laag:

- archief
- serie
- dossier
- record
- bestand

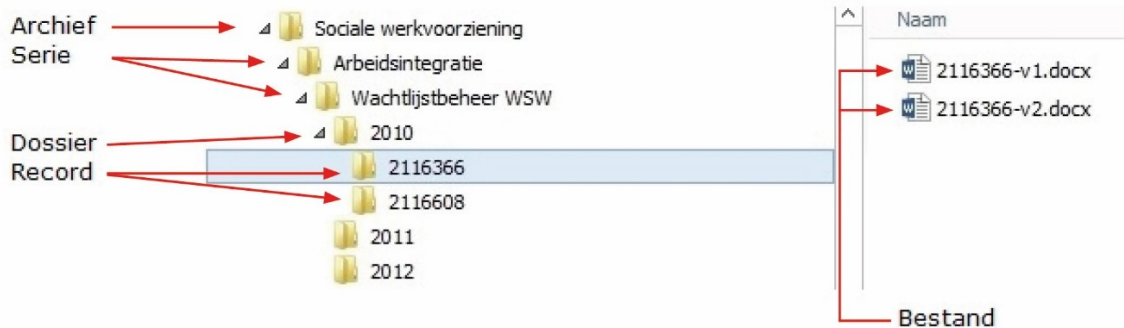
De niveaus serie, dossier en record kunnen getrapt voorkomen. Bijvoorbeeld een hoofdserie met subseries, een hoofddossier met subdossiers, of een samengesteld record dat uit meerdere records bestaat. De niveaus serie en record zijn niet verplicht. Een dossier kan onder een archief, serie of dossier hangen. Het niveau record komt overeen met een archiefstuk of document. Een record kan alleen onder een dossier of record hangen. Het niveau bestand komt overeen met een computerbestand. Het niveau bestand hangt altijd onder een record of dossier. Het aggregatieniveau bestand heeft een aparte status (zoals verderop zal blijken).

¹ Download de actuele versie op <https://www.nationaalarchief.nl/archiveren/kennisbank/handreiking-voorkeursformaten-nationaal-archief>

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

Om een set informatieobjecten te kunnen opnemen in het e-depot moet de ordening van het documentmanagementsysteem of de procesapplicatie worden vertaald naar de ordening van het e-depot. Dit is onderdeel van de metadatamapping.

Zo'n vertaling kan er bijvoorbeeld zo uitzien:

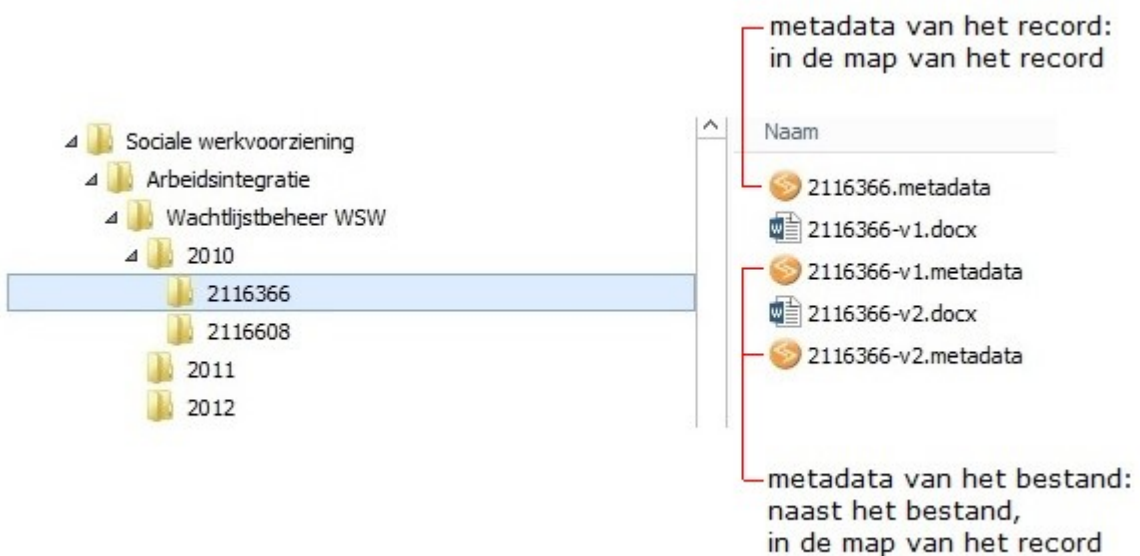


Bij de aansluiting van een nieuw systeem op ons e-depot begeleiden we u bij de eerste mapping. Bij vervolgaanleveringen voert u de mapping zelf uit en toetsen wij of deze aan de eisen voldoet.

3. Sidecar-structuur

De export moet een zogenoemde sidecar-structuur hebben. In deze structuur zitten de archiefbescheiden in een directory-structuur: de aggregatieniveaus archief, serie, dossier en record zijn mappen in een hiërarchische mappenstructuur. De bestanden bevinden zich in de mappen van de dossiers of records waar ze onder vallen. 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 toevoeging binnen dossiers niet wordt ondersteund.

5. Sidecarstructuur bij aanvullingen

Wanneer u na overdracht van een archief nieuwe series of dossiers wilt toevoegen aan een bestaand archief, moet deze aanvulling in dezelfde sidecarstructuur worden aangeboden. Ook voor de bovenliggende aggregatieniveaus archief en serie moeten mappen worden gemaakt met bijbehorende metadatabestanden. Het identificatiekenmerk in het metadatabestand moet gelijk zijn aan het identificatiekenmerk van het bestaande archief of de bestaande serie in het e-depot. Let op: eventuele wijzigingen in de overige metadata ten opzichte van de eerste aanlevering worden door het e-depot genegeerd!

Als een vervollexport een dossier bevat met hetzelfde identificatiekenmerk als een bestaand dossier, dan wordt de vervollexport door het systeem afgekeurd.

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:

- Een mapnaam moeten gelijk zijn aan het identificatiekenmerk (element 2) in de bijbehorende metadata.
- 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.
- Het e-depot maakt gebruik van Linux en dat is hoofdlettergevoelig.
- Padnamen die langer zijn dan 255 tekens kunnen bij het maken van de export en het transport ervan voor problemen zorgen.

7. Beperking op de omvang van mappen en bestanden

Om te voorkomen dat er performance problemen optreden bij het inlezen in het e-depot, mag een export niet meer dan 50.000 bestanden bevatten. Dit aantal is exclusief de metadatabestanden. Wanneer u meer dan 50.000 bestanden wilt overdragen, moet u de overdracht opsplitsen. U moet de verschillende sets dan wel tegelijk aanleveren. Houd u bij de tweede set bestanden rekening met de eisen die worden gesteld aan een aanvulling (zie paragraaf 5).

Het systeem van het e-depot accepteert geen bestanden met een omvang van 0 byte. Een bestand met bestandsgrootte 0 byte is óf een corrupt bestand óf een bestand zonder inhoud. Mocht u een dergelijk bestand aantreffen in uw set informatieobjecten, neem dan contact op.

8. 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 ToPX 2.3.1³, een XML-schema. Wanneer u archiefbescheiden aan ons overdraagt, levert u eerst een representatieve testexport aan. Of uw metadatabestanden voldoen aan het ToPX-schema kan gecontroleerd worden met behulp van een XSD⁴.

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

9. De inhoud van de metadatabestanden

Om het archief in het e-depot te kunnen beheren, conserveren en beschikbaar stellen moet het van voldoende metadata worden voorzien. Om te bepalen welke metadata *minimaal* meegeleverd moeten worden hanteren wij het Toepassingsprofiel Metadatering Lokale Overheden (TMLO) 1.1. Welke metadata verplicht zijn en op welk aggregatieniveau zij moeten worden vastgelegd is mede afhankelijk van de situatie. Bij overbrenging moeten de metadata gebruikt kunnen worden door ons collectiebeheersysteem. Bij uitplaatsing moeten de metadata gebruikt kunnen worden door uw presentatie- en zoekstelsel. In beide gevallen moeten de metadata bruikbaar zijn binnen het e-depot en uitgewisseld kunnen worden via het koppelvlak.

We stellen de volgende eisen aan de metadata:

- U levert minimaal de metadata die in TMLO verplicht zijn.
- Bij overbrenging is het archief op alle aggregatieniveaus voorzien van metadatering over openbaarheid (element 18).
- Bij overbrenging is het archief voorzien van metadatering over gebruiksrechten (element 16).
- Bij uitplaatsing is het archief op alle aggregatieniveaus voorzien van metadatering over vertrouwelijkheid (element 17).
- Bij uitplaatsing is het archief op dossierniveau voorzien van metadatering over de selectie met behulp van het eventplan 'Vernietigen' of 'Overbrengen' (element 13).
- Bij overbrenging begint het identificatiekenmerk (element 2) op het aggregatieniveau archief met het nummer dat u van ons aangeleverd krijgt.
- Het archief en elke serie binnen het archief heeft een identificatiekenmerk dat uniek is binnen deze groep.
- Elk dossier en elke record binnen het archief heeft een identificatiekenmerk dat uniek is binnen deze groep.
- Een identificatiekenmerk wordt ontleend aan het bronsysteem.
- Een identificatiekenmerk (element 2) en een naam (element 4) mogen uit niet meer dan 255 karakters bestaan.
- Bij overbrenging geldt dat alle overige metadataelementen elk uit niet meer dan 2000 tekens mogen bestaan.
- Bij overbrenging mogen er op het aggregatieniveau bestand geen inhoudelijke metadata worden vastgelegd. De metadata van een bestand in een overgebracht archief zijn namelijk niet zichtbaar en doorzoekbaar voor de bezoekers van de studiezaal en website van de Groninger Archieven.

³ Download de specificatie van ToPX 2.3 op https://www.nationaalarchief.nl/sites/default/files/field-file/specificatie_topx_2_3.pdf en de wijzigingen in versie 2.3.1 op https://www.nationaalarchief.nl/sites/default/files/field-file/versie-informatie_topx2_3_1.pdf

⁴ Download de xsd op https://www.nationaalarchief.nl/sites/default/files/field-file/topx-2.3_1.xsd

Bij uitplaatsing kunnen metadata op het niveau bestand wel worden opgevraagd via een koppelvlak. U kunt deze metadata dan wel zichtbaar en doorzoekbaar maken vanuit uw eigen systeem.

In uw documentmanagementsysteem of procesapplicatie is op alle niveaus inhoudelijke en technische metadata opgeslagen. Niet alleen de structuur van uw systeem moet vertaald worden naar ToPX, er moet ook voor elk niveau in uw systeem worden onderzocht welke metadata noodzakelijk en wenselijk zijn voor duurzame toegankelijkheid en hoe ze naar ToPX 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.

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. Bij de mapping maakt u gebruik van ons Formulier metadatamapping.⁵

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. Om de fysieke integriteit te controleren maken we gebruik van een checksum. Het Forum Standaardisatie adviseert minimaal het gebruik van een SHA2-algoritme, bijvoorbeeld SHA-256 of SHA-512, om een checksum te genereren.⁶ Als uw systeem een ander algoritme gebruikt, neem dan vooraf contact op. Het algoritme en de waarde worden vastgelegd in de metadata in element 21.7.

11. Transport en controle

Voor u een grote hoeveelheid bestanden gaat overdragen, willen we eerst een testexport opnemen. Daarvoor maakt u in overleg met ons een representatieve selectie uit de over te dragen bestanden. 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 checksum voor de complete export

Vervolgens controleren we of de export voldoet aan de opnamevoorwaarden.

Na goedkeuring levert u de complete export aan. De Groninger Archieven verzenden de definitieve export vanuit de digitale ontvangstruimte naar de Quarantaine Zone (QZ) van het e-depot. Dit is de plek van waaruit de archiefbescheiden en de metadata in de export zullen worden opgenomen in het e-depot.

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.

⁵ Download de actuele versie op <https://www.groningerarchieven.nl/over-ons/diensten/e-depot>

⁶ Zie <https://www.forumstandaardisatie.nl/standaard/sha-2>

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 onze ISILCode (NL-GnGRA), gevolgd door een nummer toegekend door de Groninger Archieven, gevolgd door het identificatiekenmerk van het archief, gevolgd door een underscore, gevolgd door datum en tijdstip van creatie (in formaat jjjjmddThhmm). Daarachter kunt u eventueel naar eigen invulling informatie toevoegen.

12. Inleiding en verantwoording bij overbrenging

Bij overbrenging van archief is het gebruikelijk om het archief voor de studiezaalbezoeker 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.