

MDTO: Specificatie Submission Information Package (SIP)

Versie 0.9

Datum	15-03-2021
Status	Concept

Colofon

Projectnaam
Projectleider(s)
Contactpersoon

V.C. Hooft
T
F +31-70-331 5477
vincent.hooft@nationaalarchief.nl
Postbus 90520 | 2509 LM Den Haag

Auteurs
Versie
Bijlage(n)

V.C. Hooft
0.9

CONCEPT

Inhoud

Colofon—2

1 Inleiding—5

2 Begrippen—6

3 Submission Information Package—6

3.1 Bronsysteem—**Fout! Bladwijzer niet gedefinieerd.**

3.2 Export—**Fout! Bladwijzer niet gedefinieerd.**

3.3 Structuur Aanlevering—7

3.4 Pakbon—**Fout! Bladwijzer niet gedefinieerd.**

3.5 Metagegevens MDTO—9

3.6 Aanlevering additionele metagegevens—**Fout! Bladwijzer niet gedefinieerd.**

CONCEPT

1 Inleiding

In dit document vindt u de structuur van het Submission Information Package (SIP) dat benodigd is voor de uitwisseling van informatieobjecten en bestanden tussen een bronsysteem en doelsysteem.

Het aanleveren van digitaal archief van een archiefvormer van een DMS naar een E-depot van een archiefinstelling is een voorbeeld van deze uitwisseling.

Een ander voorbeeld is het uitwisselen van informatie tussen ketenpartners, bijvoorbeeld van DMS naar DMS

Gebruik

Dit document is in combinatie te gebruiken met de specificatie aanlevering aan een e-depot. Het biedt een standaard werkwijze voor de aanlevering van informatie van een bronsysteem naar een doelsysteem.

Doelgroep

De specificatie is vooral bedoeld voor informatie/business analisten, informatiearchitecten, ontwikkelaars

Totstandkoming

Deze SIP specificatie is het resultaat van de overleggen van de redactiegroep koppelvak MDTO.

Verantwoording en beheer

De eisen zijn door de redactiegroep koppelvak MDTO vastgesteld en zijn deel van de standaard MDTO.

Het Nationaal Archief beheert dit document.

1 Begrippen

Bronstelsysteem:

Systeem waar de informatieobjecten voor uitwisseling bewaard worden. Dit omvat alle applicaties tot aan de verzending van de SIP. Bijv. een DMS met een exporttool.

Doelsysteem:

Systeem waar de informatie naar verplaatst wordt om bewaard te worden. Dit omvat alle applicaties die gebruikt worden om de SIP te ontvangen en op te nemen. Bijv. een QZ met een E-depot.

Sidecar: xml bestand met bijbehorende metagegevens per informatieobject of bestand

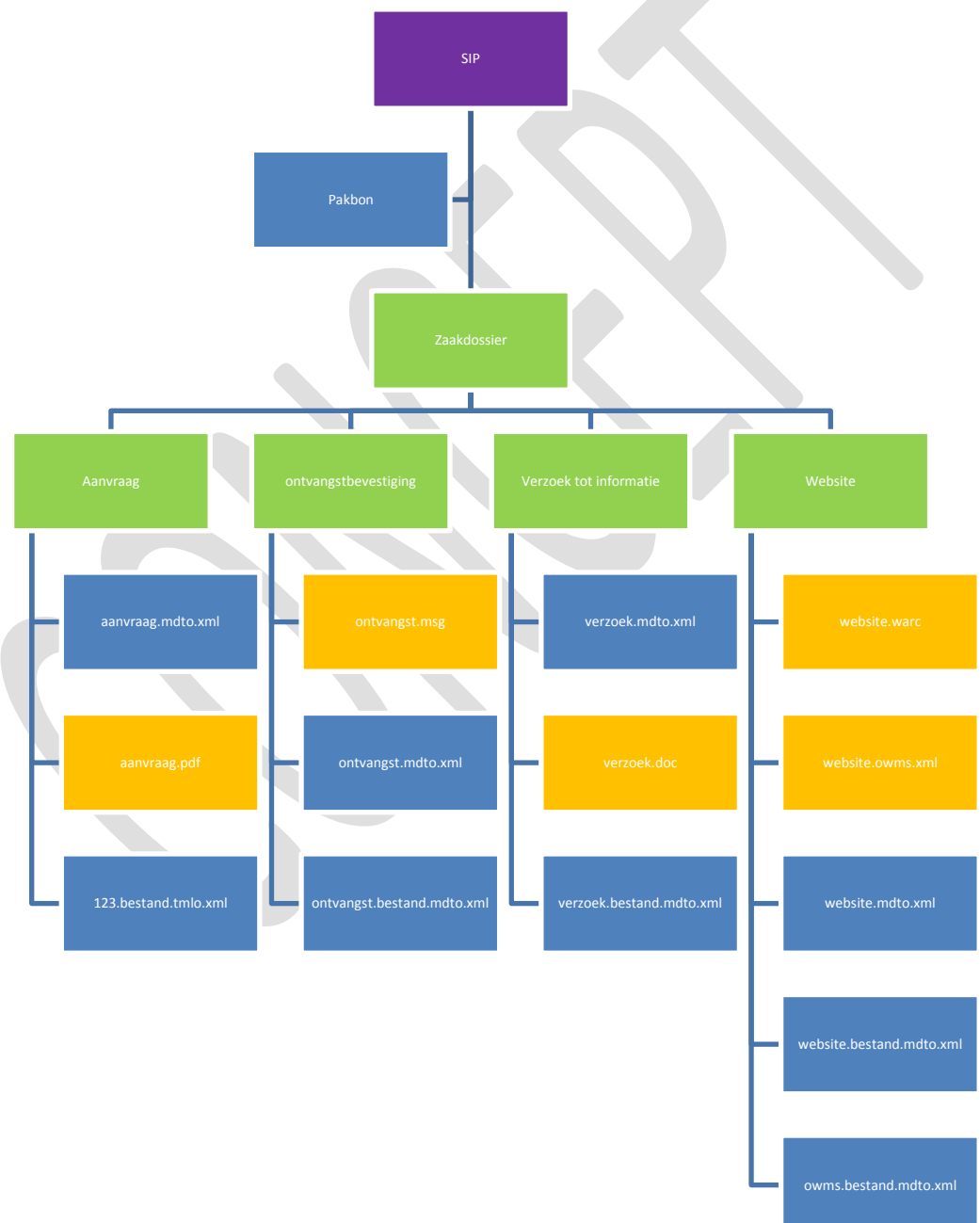
Submission Information Package (SIP):

Een set informatieobjecten en bestanden met bijbehorende inhoudelijke en technische metadata, bedoeld voor opname in een doelsysteem.

2 Submission Information Package (SIP)

2.1 Structuur SIP

In onderstaande afbeelding is een voorbeeld van de structuur weergegeven van de informatieobjecten en bestanden in een SIP. Elk informatieobject kan een aggregatie zijn.



2.2 Pakbon

De ontvangende partij moet van elke aangeleverde SIP kunnen vaststellen:

- van wie deze afkomstig is
- wat de inhoud is.

Bij de aanlevering van de export in een SIP levert de aanleverende organisatie daarom een zogenaamde pakbon mee. De inhoud en het formaat van aanlevering van de pakbon spreekt u af met uw contactpersoon bij de ontvangende partij. Het formaat varieert van een op .xml gebaseerd bericht tot een e-mail. Hierin staat in ieder geval:

- Uniek identificatienummer.
- Naam/Titel aanlevering
- Locatie doelsysteem, zoals gecommuniceerd door de ontvangende partij.
- Hashcode SIP, specificatie hashcode wordt geleverd door de ontvangende partij.
- Time stamp creatie SIP
- Organisatie/Archiefvormer.
- Contactpersoon/ e-mail
- Aantal informatieobjecten en aantal bestanden.
- Aantal bestanden in de export met extensie ongelijk aan .mdto.xml.
- Omvang van de content in de export in bytes (ofwel de optelsom van de grootte van de bestanden met extensie ongelijk aan .mdto.xml).
- Bijzonderheden materiaal: hier vermelding bijzondere formaten e.d

NB. De voorkeur gaat uit naar een xml bestand.

2.3 Sidecar-structuur

De export uit het bronsysteem moet een sidecar-structuur hebben.

In deze structuur zit de inhoud in een directory-structuur: elk aggregatieniveau van de informatieobjecten is een map. Een bestand kan op elk aggregatieniveau in de map zitten. Ook als daar ook nog een aggregatie van informatieobject beschikbaar is. Denk aan een e-mail met bijlagen of een website met de webpagina's. a Elk informatieobject en elk bestand heeft zijn eigen MDTO metagegevensbestand. In de sidecar moet het versienummer van de betreffende MDTO-versie zijn opgenomen (zie MDTO.xml).

De plaats van het metagegevensbestand dat bij een aggregatie (map in de SIP) hoort, is in deze aggregatie. De plaats van het metagegevensbestand dat bij een bestand hoort, is naast het bestand.

Een informatieobject is uniek en kan maar één keer in de SIP voorkomen. Er is altijd maar één hiërarchische relatie. De hiërarchische relatie met de bovenliggende aggregatie moet ook meegeleverd worden in de MDTO-metagegevens.

De aggregaties zijn de informatieobjecten die een hiërarchische relatie kennen. Dit is niet de 1 op 1 ordening zoals die in het bronsysteem bestaat. Als een dossier in de ordening in het bronsysteem onder meerdere informatieobjecten valt dan wordt deze maar 1 keer als record in de structuur opgenomen. De originele ordening is te vinden in de relaties.

2.4 Aggregatieniveau

Elke aanlevering kent in ieder geval één aggregatieniveau. Er is altijd sprake van een informatieobject, met meestal nog een bestand. Er kunnen meer aggregaties geleverd worden, met boven en onderliggende aggregaties.

Elke aanlevering kent een bovenliggende aggregatieniveau in het doelsysteem. Naar dit informatieobject moet in de informatieobjecten die hier direct onder geplaatst worden in de MDTO-metadata een verwijzing zijn opgenomen. Neem in de pakbon op in welke collectie, bijv. archief, deze aanlevering moet worden geplaatst.

2.5 Naamgeving

De sidecar met MDTO-metadata bestaat uit een voorgeschreven formaat (XML) en een bestandsnaam. Dit ziet er zo uit: <naam>.MDTO.xml of bij een bestand: <bestandsnaam>.bestand.MDTO.xml.

Elk informatieobject of bestand moet een naam hebben. Deze naam moet in de aanlevering uniek zijn en hoeft **niet** met de titel overeen te komen.

Voor de <informatieobjectnaam>. MDTO.xml of <bestandsnaam>. MDTO.xml geldt in de aanlevering een limiet op het aantal karakters van 255.

De namen van het informatieobject, bestand en de sidecars mogen niet de volgende karakters bevatten: < > : " / \ | ? * # & . of spatie

Schoon, indien nodig, de informatieobject- en bestandsnamen in de SIP van deze karakters voor aanlevering.

Let op: De titel van een informatieobject staat altijd in de metagegevens. Deze kan wel leestekens bevatten of langer zijn dan 255 karakters. Verwar dit niet met de naam van het informatieobject zoals meegegeven voor de aanlevering.

2.6 Metagegevens MDTO aanleveren

MDTO metagegevens moeten in een MDTO metagegevens bestand worden aangeleverd. De aanleverende organisatie zorgt ervoor dat waar mogelijk na redelijke inspanning metagegevensvelden een waarde hebben. Dit kan in voorkomende gevallen extra inspanning vragen van de aanleverende organisatie, bijvoorbeeld door het combineren van metagegevens.

Verplichte metagegevens moeten altijd worden aangeleverd.

Voor het vullen kan gebruik worden gemaakt van sjablonen die beschikbaar zijn op de website van het Nationaal Archief <link>.

Met behulp van de .xsd controleert een doelsysteem de aangeleverde MDTO.xml bestanden.

2.7

Additionele metagegevens aanleveren

Als naast MDTO ook een set met andere metagegevens meegeleverd wordt bij een informatieobject doet u dat door het toevoegen van een eigen bestand voor elke set van metadata. Zoals voor elk bestand kent dit bestand een sidecar met de MDTO-metadata voor bestand.

Voor elk metadatabestand gelden de volgende eisen:

- Elk bestand maakt gebruik van een gedefinieerde standaard en wordt conform deze standaard opgeleverd. (De standaard mag ook een eigen schema zijn.)
- Elk bestand maakt gebruik van een standaardformaat uit de lijst van voorkeursformaten. Denk hierbij aan: XML, JSON, CSV, RDF.
- Metagegevens die in meerdere sets voorkomen worden in elke metagegevens set meegeleverd. Dus als het veld naam in MDTO en in RGBZ of OWMS voorkomt wordt dit veld in beide standaarden gevuld met de metagegevens.
- Additionele metagegevens voor bestanden zijn enkel technisch van aard (bijv. PREMIS).
- De naamgeving is overeenkomstig de naamgeving van de MDTO.xml bestanden.

De documentatie, zoals een definitiebestand, van een niet-erkende (eigen) standaard of metagegevensschema, levert u als informatieobject mee .