

# Een fundament voor digitale duurzaamheid

Een mapping tussen MDTO en PREMIS ten behoeve van een Common Ground e-depot



**Kerngroep Duurzame Toegankelijkheid & Common Ground  
Mei 2024**

## **Colofon**

De Kerngroep Duurzame Toegankelijkheid & Common Ground bestaat momenteel (mei 2024) uit:

Annelot Vijn (Het Utrechts Archief);  
Bram Klapwijk (Regionaal Archief Zuid-Utrecht);  
Eva van den Hurk-van 't Klooster (RHC Eindhoven);  
Léon Masselink (Regionaal Archief Rivierenland);  
Wietse Bakker (Regionaal Archief Zuid-Utrecht).

De Kerngroep Duurzame Toegankelijkheid & Common Ground is een groep informatie- en archiefprofessionals die in netwerkverband samenwerken met als doel te formuleren hoe een voorziening voor digitale duurzaamheid (ook wel een e-depot) eruit zou kunnen zien als die conform Common Ground zou worden opgebouwd.

## Inhoud

Inleiding.....	4
1. De standaarden naast elkaar.....	6
1.1. MDTO .....	6
1.2. PREMIS .....	7
2. Praktische observaties over relatie MDTO en PREMIS .....	9
2.1. Mapping .....	9
2.2. Uitgangspunten implementatie.....	9
2.3. Technische implementatie.....	10
3. Toepassing van de mapping in de praktijk.....	12
3.1. Inleiding en uitgangspunten .....	12
3.2. Scenario 1: MDTO als uitwisselingsstandaard naast PREMIS .....	13
3.3. Scenario 2: MDTO en PREMIS in een gezamenlijke omgeving.....	16
4. Conclusie en verder onderzoek .....	18

# Inleiding

Duurzame toegankelijkheid van digitale informatieobjecten binnen Common Ground betekent meer dan het implementeren van een aantal technische modules met daartussen API's of het borgen van processen en functies. Het is ook van belang dat de juiste metagegevens worden vastgelegd. Op het gebied van metadatering is MDTO op dit moment de beschrijvingsstandaard voor digitale overheidsinformatie en PREMIS de internationale beschrijvingsstandaard voor conserveringsmetadaten.<sup>1</sup> Deze twee standaarden hebben dan ook in potentie een belangrijke rol te spelen in het borgen van duurzame toegankelijkheid binnen Common Ground.

Hierom wil de Kerngroep Duurzame Toegankelijkheid & Common Ground (hierna: de Kerngroep) met deze notitie en bijbehorende documenten antwoord geven op de volgende vraag: "Is een gezamenlijke implementatie van MDTO en PREMIS mogelijk in een informatiemodel dat geïmplementeerd is in een Common Ground-ingerichte omgeving, zodat de duurzame toegankelijkheid van digitale archiefbescheiden wordt geborgd?"

Om een dergelijk scenario mogelijk te maken, is het eerst nodig om de standaarden inhoudelijk met elkaar te vergelijken door middel van een onderlinge mapping van hun elementen. De Kerngroep heeft de afgelopen maanden gewerkt aan deze mapping. De resultaten daarvan zijn te lezen in een aantal documenten:

- Het Excel-document '**MDTO-PREMIS-Mapping-V1**' bevat zowel de mapping van MDTO naar PREMIS, als de mapping van PREMIS naar MDTO, ieder in een eigen tabblad.
- Het Word-document '**Mapping-Beschrijving-Tabellen**' bespreekt de mapping in meer detail per element.

De mapping is bedoeld om inzichtelijk te maken waar de metagegevensstandaarden elkaar overlappen, waar ze elkaar aanvullen en waar er eventuele interpretatieverschillen bestaan over beschreven entiteiten. Er is bewust gekozen voor een mapping beide kanten op, omdat wegens de diversiteit aan velden in beide standaarden hiermee het overzicht beter bewaard blijft. In het uitvoeren van de mapping zijn enkele keuzes gemaakt. Zo is de PREMIS-klasse 'Environment' voorlopig buiten de mapping gelaten. PREMIS beschrijft deze klasse als "*Technology (software or hardware) supporting a Digital Object in some way (e.g. rendering or execution).*" We liepen tegen te veel vragen en interpretatieverschillen aan om een eenduidige mappingkeuze te maken bij deze klasse.

Het document dat u nu leest bevat een inhoudelijke beschouwing van MDTO en PREMIS op basis van de uitgevoerde mapping en meerdere gevoerde gesprekken met key users van beide standaarden. In hoofdstuk 1 worden beide standaarden en hun toepassingsgebieden uitgewerkt. In het hoofdstuk 'Gezamenlijke implementatiescenario's' wordt stilgestaan bij enkele manieren om MDTO en PREMIS gezamenlijk te implementeren in de praktijk. In dat hoofdstuk zijn enkele denkrichtingen uitgewerkt, bedoeld om te leiden tot discussie en zo daadwerkelijke gezamenlijke implementatie vorm te geven. In hoofdstuk 3 zijn enkele

---

<sup>1</sup> Voor gedetailleerde informatie over beide standaarden verwijzen we naar de websites van beide standaarden: [MDTO | Nationaal Archief](#), [PREMIS Data Dictionary for Preservation Metadata, Version 3.0 \(Library of Congress\) \(loc.gov\)](#)

praktische observaties toegevoegd die voor de mapping en de implementatie relevant zijn. In hoofdstuk 4 komt naar voren welke openstaande vragen er nog zijn om de hoofdvraag van dit document te beantwoorden.

# 1. De standaarden naast elkaar

MDTO en PREMIS zijn beiden standaarden die handelen over digitale (informatie)objecten, en die zijn gericht op digitale duurzaamheid. Het toepassingsgebied is echter verschillend. In dit hoofdstuk wordt een schets gegeven van de doelen die de beide standaarden nastreven en een schets gegeven van hun huidige toepassing in de Nederlandse praktijk.

## 1.1. MDTO

### *Inleiding*

MDTO is een metagegevensstandaard voor duurzame toegankelijkheid van overheidsinformatie. Het Nationaal Archief definieert het toepassingsgebied 'op het vastleggen en uitwisselen van eenduidige metagegevens om de duurzame toegankelijkheid van overheidsinformatie mogelijk te maken.'<sup>2</sup> MDTO is bedoeld voor overheidsinformatie, maar kan ook daarbuiten worden toegepast.

Onder het uitwisselen van metagegevens wordt hier verstaan het uitwisselen tussen informatiesystemen (via een koppeling) en het beschikbaar stellen door informatiesystemen aan de gebruikers (via een weergave- of exportfunctie). Een formele overbrenging van overheidsinformatie naar een archiefdienst is een voorbeeld van uitwisseling.<sup>3</sup>

MDTO is de opvolger van de domeinstandaarden TMLO<sup>4</sup> en TMR<sup>5</sup>, die een eerste standaard boden voor de migratie van digitale informatieobjecten tussen informatiesystemen in de Nederlandse overheidscontext. In tegenstelling tot zijn voorgangers, is MDTO een overheidsbrede standaard voor digitale duurzaamheid.

### *Opbouw van de standaard*

MDTO maakt onderscheid tussen verschillende objecten waar metagegevens aan verbonden zijn. Informatieobject en Bestand worden genoemd en uitgewerkt in de standaard. Bedrijfsactiviteit, Actor, en Locatie worden genoemd, maar zijn niet verder van metagegevens voorzien in de standaard. MDTO biedt de mogelijkheid om hiervoor andere standaarden te gebruiken.

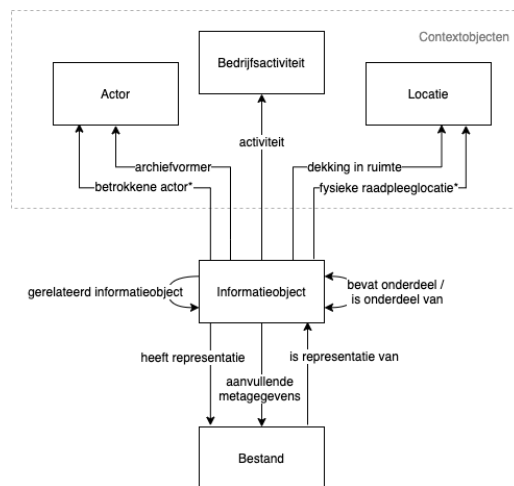
---

<sup>2</sup> Zie ook: <https://www.nationaalarchief.nl/archiveren/mdto>.

<sup>3</sup> [MDTO | Nationaal Archief](#).

<sup>4</sup> Toepassingsprofiel Metadatering Lagere Overheden.

<sup>5</sup> Toepassingsprofiel Metagegevens Rijk.



Figuur 1. De relaties tussen de objecten binnen MDTO.

### *Gebruik van de standaard*

Ondanks de nog relatief geringe actieve toepassing van de standaard, wordt er momenteel veel mee geëxperimenteerd en vindt het aansluiting bij en verankering in initiatieven als DUTO, het Gemeentelijk Gegevensmodel en de Richtlijn Metagegevens Overheidsinformatie. Daarmee lijkt het een stevig fundament te hebben binnen de Nederlandse overheidsinformatisering.

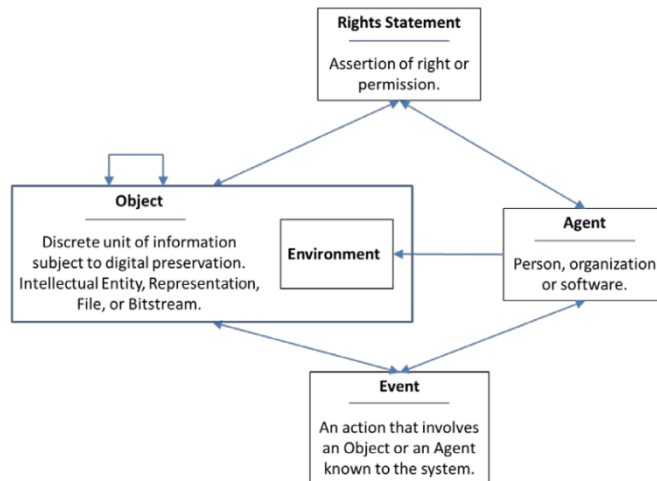
## 1.2. PREMIS

### *Inleiding*

De PREMIS Data Dictionary (PREMIS) is een metagegevensstandaard voor digitale preservatie, bedoeld om de preservatie van digitale objecten te ondersteunen en hun toekomstige raadpleegbaarheid te waarborgen. Het kan daarnaast dienen als informatiemodel in een preservatiesysteem, zoals een e-depotvoorziening of een ander soort digitale objectenbewaarplaats. PREMIS heeft daarin een nauwe relatie met het Open Archival Information System (OAIS)-model, het conceptuele model dat aan de grondslag ligt van de huidige generatie e-depotsystemen.

### *Opbouw van de standaard*

PREMIS onderscheidt vier entiteiten, namelijk Agents, Rights, Events en Objects. Deze vier entiteiten kennen een onderlinge relatie. Een Agent met bepaalde Rights mag bijvoorbeeld een Event uitvoeren op een Object. Iedere entiteit is voorzien van een eigen, uitgebreide set aan metagegevens.



Figuur 2. De vier entiteiten van PREMIS in hun onderlinge relatie.

### *Gebruik van de standaard*

Er zijn relatief weinig implementaties van PREMIS bekend in beheersystemen van Nederlandse erfgoedinstellingen. Er is enige praktijkervaring met de standaard bij Beeld en Geluid, bij het Nationaal Archief en bij de Koninklijke Bibliotheek, maar een brede implementatie lijkt nog te ontbreken. De huidige generatie e-depots is weliswaar gebaseerd op het OAIS-model, maar daarbinnen is digitale preservatie vaak nog niet of nauwelijks ingericht. De opname, de opslag en de beschikbaarstelling van informatieobjecten hebben voorrang gekregen in de productontwikkeling.



## 2. Praktische observaties over relatie MDTO en PREMIS

### 2.1. Mapping

Tijdens het mappen van de twee standaarden zijn een aantal kenmerkende overeenkomsten en verschillen naar voren gekomen. Beide standaarden kennen bijvoorbeeld een 'object' dat centraal staat en waaraan bepaalde attributen worden gekoppeld. Veel van deze attributen zijn gelijk in soort, zoals het toekennen van een naam en identificatiegegevens aan een object. Toch zijn er voorafgaand ook al enkele verschillen te noemen die voor de lezer handig zijn om in het achterhoofd te houden wanneer de mapping wordt toegepast:

- *Gelaagdheid*: Waar het MDTO een generiek 'informatieobject' en bijbehorend(e) 'bestand(en)' kent, maakt PREMIS gebruik van specifiek gedefinieerde 'Objects' in 4 categorieën ('Intellectual Entities', 'Representations', 'Files' en 'Bitstreams') en kent het expliciet 'Environments' toe waarin een object zich bevindt.
- *Reikwijdte*: Waar het MDTO zich bewust tot een minimale set aan gegevens beperkt om een doelgerichte uitwisseling van gegevens te waarborgen, heeft PREMIS als doel langdurige preservering te waarborgen. Met andere woorden: waar sommige MDTO-attributen breed inzetbaar zijn, geeft PREMIS expliciete invulling aan attributen, zoals bijvoorbeeld het beschrijven van gegevens over een digitale handtekening.
- *Opbouw van attributen*: Beide standaarden kennen een gelaagde opbouw van attributen, waarbij een deel van een attribuut opgebouwd kan zijn uit meerdere deelattributen. In een aantal gevallen wordt er wel gesproken van hetzelfde 'hoofdattribuut', maar is de opbouw van deze attributen afwijkend.
- *Positionering actor*: Het MDTO definieert een actor als een contextobject dat direct wordt gelinkt aan het informatieobject, als verantwoordelijkeActor (bij events) of als betrokkene met een relatie tot het informatieobject. Door middel van een typering bij de actor wordt duidelijk gemaakt wat de rol van de actor is (zoals 'opsteller' of 'indiener'). PREMIS, daarentegen, linkt actoren aan specifieke preservingsacties (Event), aan rechten (Rights Statement) of een 'Environment'. In vergelijking met het type 'opsteller' in MDTO zal in PREMIS een bepaalde actor genoemd worden bij het Event 'Creation'. Dit Event zal vervolgens als eerste Event aan een informatieobject worden gekoppeld.

### 2.2. Uitgangspunten implementatie

Het doel van de mapping is geweest om de gecombineerde implementatie van MDTO en PREMIS in een Common Ground-omgeving dichterbij te brengen. Hoe dit geïmplementeerd kan worden is nog niet uitgekristalliseerd. De kerngroep wil hierbij zowel onderzoeken hoe de mapping geïmplementeerd kan worden in een Linked Data-georiënteerde omgeving (bijvoorbeeld in een Triple Store) als ook in een relationeel-georiënteerde database. Dit omdat in de archiefsector de laatstgenoemde momenteel dominant is in de markt en de eerstgenoemde juist in opkomst is in zowel de archief- als de cultuursector. Aangezien de Common Ground visie van standaarden uitgaat, houden we rekening met implementatie in beide soorten systemen.

Bij de implementatie worden een aantal uitgangspunten meegenomen:

- *Geen dubbele registratie van dezelfde gegevens.* Hiermee wordt bedoeld dat een attribuut (zoals 'naam') niet twee keer wordt geregistreerd met dezelfde waarde. In een Triple Store zou dit bij voorkeur betekenen dat er slechts eenmaal de standaarden op elkaar worden gemapt en dat er geen dubbele Triples worden geregistreerd.
- *Standaarden zijn afzonderlijk van elkaar te herleiden.* Samen met het bovengenoemde punt zien wij een mogelijke uitdaging. Als het nodig is, willen we in het e-depot één van de standaarden kunnen weergeven of laten exporteren. Het moet in het systeem duidelijk zijn aan één attribuut (met waarde) meerdere standaarden gekoppeld kunnen worden.
- *Interne logica standaarden moeten aansluiten.* Beide standaarden kennen een eigen logica toe. Het MDTO kent bijvoorbeeld 'Informatieobjecten' en 'Bestanden', terwijl PREMIS spreekt over 'Intellectual Entities', 'Representations', 'Files' en 'Bitstreams'. Er moet een implementatieoplossing gevonden worden die zo min mogelijk af doet aan de logica van de standaarden.

Alle drie de hierboven genoemde punten brengen hun eigen uitdagingen met zich mee. Op technisch vlak zouden we één attribuut, met waarde, aan meerdere standaarden willen koppelen en in meerdere standaarden willen kunnen exporteren. Een link tussen matchende velden in een linked data-omgeving kan via SKOS gelegd worden, maar de technische uitwerking in de componenten moet nog uitgedacht worden. SKOS kan ook een oplossing bieden voor niet-exacte matches in de mapping in een linked data-omgeving. Daar waar een aantal velden voor de Kerngroep een exacte match zijn tussen MDTO en PREMIS, bestaat het merendeel van de mapping namelijk uit redelijke matches of near matches. Hoe het kan dat de matches zo moeilijk te maken zijn, heeft te maken met het derde punt: het verschil in de interne logica van de twee standaarden. Het MDTO kent bijvoorbeeld 'Informatieobjecten' en 'Bestanden', terwijl PREMIS spreekt over 'Intellectual Entities', 'Representations', 'Files' en 'Bitstreams'. Dit verschil in gelaagdheid maakt dat de metadatavelden die op het eerste gezicht identiek lijken te zijn, binnen de twee standaarden een andere invulling krijgen en op een andere wijze gerelateerd worden aan de beschreven entiteit.

### 2.3. Technische implementatie

Een van de doelen van de mappingexercitie is om MDTO en PREMIS gezamenlijk te implementeren in een component- en linked data-gebaseerd systeem voor duurzame toegankelijk, maar we wilden ook graag weten of een dergelijke mapping te implementeren zou zijn in een relationele database. Omdat binnen Common Ground oud naast nieuw kan bestaan en werken met standaarden gegevensuitwisseling vergemakkelijkt, ligt het voor de hand om de mapping van de landelijke en internationale standaarden MDTO en PREMIS op zowel de huidige generatie e-depotsystemen als op de in ontwikkeling zijn component-gebaseerde systemen voor duurzame toegankelijkheid toe te passen, in de aanloop naar de toekomstige uitwisseling en samenwerking tussen de twee verschillende typen systemen. De mapping implementeren in een component- en linked data-gebaseerd systeem lijkt op dit moment praktisch uitvoerbaar, dankzij de inzet van SKOS als relatieverklarende koppeling

tussen de twee standaarden.<sup>6</sup> De implementatie van de mapping in een relationele database wordt ingewikkelder. Relationele databases zijn vaak voor een speciaal onderwerp en een specifieke doelgroep ontwikkeld, waardoor deze gericht is op gebruikersgemak maar waardoor uitwisseling of migratie bemoeilijkt wordt. Hier is meer onderzoek of een gerichte doorontwikkeling nodig.

---

<sup>6</sup> Voor meer informatie over SKOS, zie de [SKOS Simple Knowledge Organization System Reference](#).

## 3. Toepassing van de mapping in de praktijk

### 3.1. Inleiding en uitgangspunten

#### *Inleiding*

Zoals hiervoor is te lezen, hebben MDTO en PREMIS als metagegevensschema's een eigen functie in het ecosysteem van digitale duurzaamheid. Hoewel noch de implementatie van MDTO, noch die van PREMIS in de Nederlandse erfgoedsector is waar ze zou kunnen zijn, zien wij voldoende reden om te verkennen hoe een gezamenlijke toepassing van MDTO en PREMIS er in de praktijk uit zou kunnen zien.

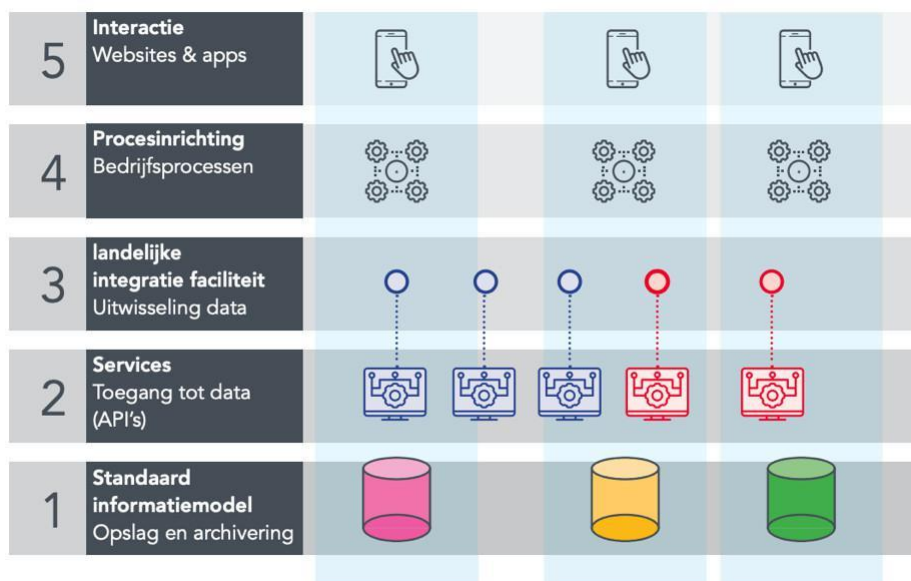
Een belangrijke voorwaarde is uiteraard dat beide standaarden een plek krijgen in de implementatieomgeving. De implementatie van met name PREMIS vereist aardige inspanning, omdat die standaard een behoorlijk precieze blauwdruk biedt voor de logging van gegevens over digitale conserveringsprocessen. Het is daardoor bijvoorbeeld minder geschikt als uitwisselstandaard, behalve van en naar systemen die al conform PREMIS zijn opgebouwd. MDTO is flexibeler op dat gebied, maar kent weer als uitdaging dat sommige MDTO-klassen en elementen wel erg breed zijn gedefinieerd. Want hoe leg je binnen MDTO expliciet vast wat je verstaat onder een informatieobject? Dit kan immers bijvoorbeeld een serie, een groep dossiers, een database, een fysiek of een digitaal informatieobject zijn. En hoewel je veel MDTO-attributen kan vastleggen in begrippenlijsten, kan iedere organisatie zijn eigen begrippenlijsten opstellen. Dat zijn factoren die de implementatie in een informatiesysteem compliceren, tenzij je een expliciete inkadering ervan gaat hanteren.

Hieronder hebben we twee scenario's uitgewerkt die we voor ons zien qua een gezamenlijke implementatie. Deze scenario's richten zich primair op de inhoudelijke en functionele (zo je wil informatiekundige) overwegingen, en minder op de specifieke technische vereisten die aan de implementatie ten grondslag liggen. De twee scenario's zijn:

- De geïsoleerde implementatie van MDTO als uitwisselingsstandaard voor overheidsinformatie naast de implementatie van PREMIS voor digitale preservatie, en
- Een gezamenlijke implementatie van MDTO en PREMIS in een geïntegreerde informatiearchitectuur.

#### *Uitgangspunten*

In het uitwerken van de implementatiescenario's gaan we uit van een leeg canvas, waarin sprake is van een systeeminrichting conform de informatiekundige visie Common Ground. Dat wil zeggen dat er een uit losse componenten samengesteld archiefbeheersysteem bestaat, waarin PREMIS en MDTO als informatiemodellen op de onderste laag van het vijflagenmodel zijn geïmplementeerd. De schets richt zich op de wijze waarop dit conceptuele archiefbeheersysteem gebruik kan maken van beide informatiemodellen, en wat dit betekent voor de inhoudelijke overlap tussen de twee modellen.



Figuur 3. Het vijflagenmodel van Common Ground. MDTO en PREMIS zouden op laag 1 worden gemodelleerd.

## 3.2. Scenario 1: MDTO als uitwisselingsstandaard naast PREMIS

### *Inleiding*

De Archiefwet knipt het informatiebeheer bij de overheid feitelijk op in twee beheerregimes: die van de archiefvormer en die van de archiefbewaarplaats, en stelt voor enkele soorten overheidsinformatie een overbrengingstermijn van maximaal twintig jaar verplicht.<sup>7</sup> De meest gebruikelijke implementatie van MDTO is als uitwisselstandaard bij migratie van de archiefvormer naar de archiefbewaarplaats. Technisch bekeken is dit een migratie van een informatieapplicatie naar een e-depotvoorziening. Zo'n migratie kan ook binnen de muren van de eigen organisatie plaatsvinden, als een overheid zijn archieftaken niet heeft ondergebracht bij een archiefdienst.

Migratietrajecten worden momenteel vooral getriggerd door de uitfasering van informatiesystemen, of door de wens tot vervroegde overbrenging, zodat de beheerslast en bijbehorende beschikbaarstelling vroegtijdig naar de 'eindbestemming' kan worden overgeheveld.

### *Een schets van de technische opname van deze bestanden*

Wanneer MDTO als uitwisselingsstandaard wordt ingezet, dan betekent dat een conversie van vooraf gemapte metadata uit het bronsysteem in een voorgeschreven XML-structuur (conform de MDTO XSD<sup>8</sup>). De metadata en de bijbehorende bestanden worden dan gezamenlijk in een

<sup>7</sup> Vanaf 2026 zal deze maximale termijn waarschijnlijk worden verkort naar 10 jaar. Voor meer informatie over hoe deze 'knip' tot stand is gekomen, lees de scriptie 'Tussen droom en daad' van H. Uulders: <https://kia.pleio.nl/attachment/entity/fdbed0bc-75d9-488f-8fb2-0ee19f928036>.

<sup>8</sup> Momenteel MDTO 1.0.1.

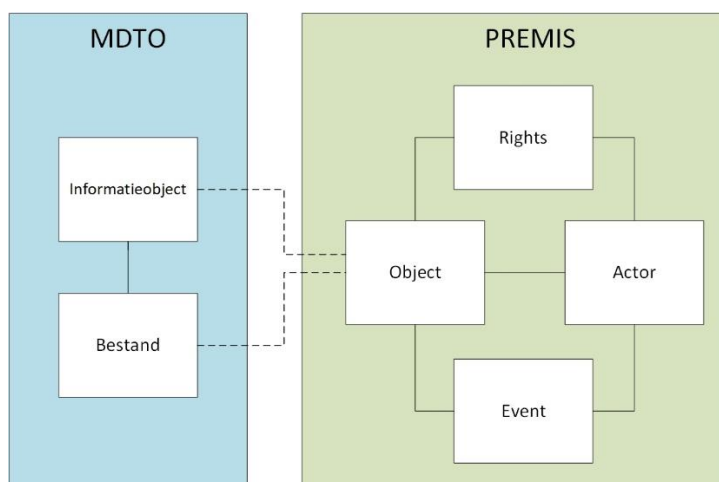
vooraf besproken indeling (side car en RIP zijn de populairste) aangeboden aan de e-depotvoorziening.

Hiermee start het opnameproces in het archiefbeheersysteem, wat ook wel de ingest wordt genoemd. Een ingestcomponent leest de XML-bestanden uit en registreert ze als datapunten in de database. Na deze fase worden de voor PREMIS relevante datapunten gekopieerd uit de 'MDTO-database' om als basisinformatie te dienen voor conserveringsprocessen. Eventuele extra benodigde metagegevens worden door vergelijkbare processen gecreëerd of geëxtraheerd en opgeslagen in de 'PREMIS-database'.

#### *Verkenning van het scenario*

Als MDTO alleen als uitwisselingsstandaard wordt gehanteerd, dan betekent dit dat er weinig conceptuele overlap is tussen de twee standaarden. MDTO beschrijft de elementen uit het *bronsysteem* die nodig zijn voor digitale duurzaamheid en PREMIS beschrijft de entiteiten die een rol spelen in digitale preservatie in een *e-depotomgeving*. De PREMIS entiteiten Actors, Events en Rights zijn immers enkel bedoeld voor actoren, events en rechten voor handelingen binnen het archiefbeheersysteem. Hoewel MDTO attributen bevat die actor, event en rechtengegevens bevatten, is het waarschijnlijk niet nuttig om deze aan elkaar te linken, gezien in dit scenario de MDTO-attributen statische gegevens bevatten uit andere informatiesystemen. De kans op inhoudelijke overlap is beperkt.

De enige conceptuele overlap tussen MDTO en PREMIS in dit scenario is tussen Informatieobject en Bestand van MDTO en Objects van PREMIS. Deze wordt hieronder verder uitgewerkt.



Figuur 4. De relaties tussen entiteiten binnen MDTO en PREMIS in scenario 1.

#### *Informatieobjecten in MDTO*

In MDTO wordt een informatieobject gedefinieerd als 'een op zichzelf staand geheel van gegevens met een eigen identiteit.' Hiermee wordt een verzameling aan elkaar gerelateerde gegevens die als eenheid wordt behandeld bedoeld.

Een informatieobject kan samengesteld zijn uit kleinere informatieobjecten, MDTO noemt dit een aggregatie. Een dossier van een vergunningaanvraag kan bijvoorbeeld zijn samengesteld uit een aanvraag, beschikking en andere relevante informatieobjecten.

### Bestanden in MDTO

In MDTO wordt een bestand gedefinieerd als een ‘geordende verzameling van gegevens in elektronische vorm, die door een elektronisch apparaat (computer, smartphone, videorecorder enzovoort) onder één naam kan worden behandeld en aangesproken.’ Bestand zal in de regel vaak overeenkomen met een computerbestand.

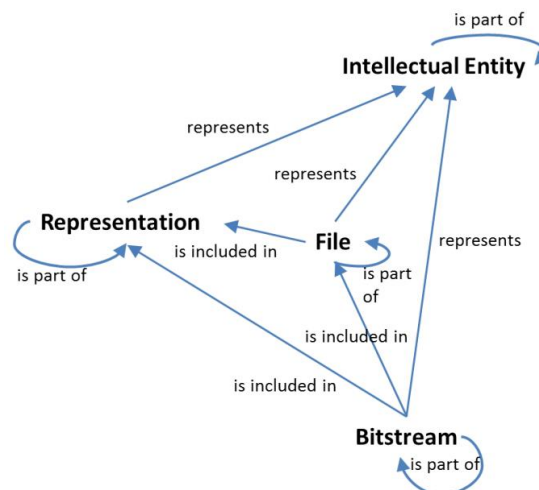
### Objecten in PREMIS

In PREMIS wordt een object gedefinieerd als een ‘discrete eenheid aan informatie die is onderworpen aan digitale preservering’.<sup>9</sup> PREMIS deelt de term Object onder in vier entiteiten: Intellectual Entity, Representation, File en Bitstream. Deze concepten helpen in het nauwer definiëren van het preservingsobject:

PREMIS Object	Definitie
Intellectual Entity	Een afzonderlijke intellectuele of artistieke creatie.
Representation	De verzameling bestanden, inclusief structurele metadata, waarin een Intellectual Entity compleet kan worden uitgedrukt.
File	Een benoemde en geordende sequentie aan bytes die bekend is bij een besturingssysteem.
Bitstream	Data binnen een bestand met eigenschappen die betekenis hebben in de context van digitale preservering.

Tabel 1. De objecten binnen PREMIS en hun definities.

Om de onderlinge verhoudingen tussen deze vier objecttypen in PREMIS zichtbaar te maken, geeft de PREMIS Data Dictionary de volgende illustratie:



Figuur 4. De vier onderdelen van Object binnen PREMIS en hun onderlinge relaties.

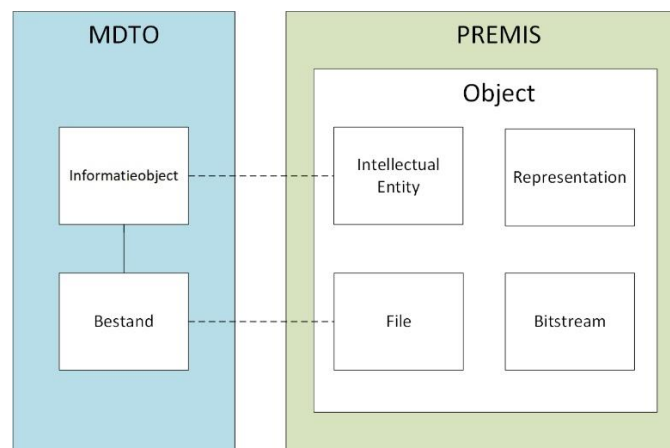
<sup>9</sup> PREMIS Data Dictionary, pagina 7.

## Informatieobject en Bestand naar Object

Informatieobject binnen MDTO komt conceptueel gezien het meest overeen met de term Intellectual Entity. Intellectual Entity kan breder worden ingezet dan hoe Informatieobject wordt bedoeld in MDTO – het kan zijn dat een digitaal informatieobject een representatie is van een fysiek boek, in dat geval is de Intellectual Entity het fysieke boek. In digitaal informatiebeheerperspectief zal een digital born informatieobject zowel een representation van een file zijn, als een intellectual entity op zichzelf. Het is namelijk de neerslag van het handelen van een actor in een bepaald proces.

Intellectual entity lijkt de meest zuivere mapping van Informatieobject. Representation komt in de buurt, maar in de preservingspraktijk wordt die term toch anders ingezet. Een Representation is een enkele instantie van een Intellectual Entity die in een bewaarplaats wordt bewaard. Een bewaarplaats kan meer dan één Representation van hetzelfde Intellectual Entity bevatten. Een voorbeeld hiervan is als er een afgeleide presentatie-exemplaar van een digitaal bestand wordt gemaakt. Beide bestanden zijn volgens PREMIS representaties van dezelfde Intellectual Entity.<sup>10</sup>

Representaties zijn in MDTO relaties tussen een intellectuele beschrijving in een Informatieobject en één of meerdere representaties in de vorm van een Bestand. Denk bijvoorbeeld aan een informatieobject dat een e-mail beschrijft, dat vervolgens zowel als een PDF-bestand en als een MSG-bestand beschikbaar is. Een representatie wordt daarmee als relatie vastgelegd tussen het generieke Informatieobject en een bestand, terwijl PREMIS dit als apart gedefinieerd object beschrijft.



Figuur 5. De verfijning van de relaties tussen entiteiten binnen MDTO en PREMIS.

### 3.3. Scenario 2: MDTO en PREMIS in een gezamenlijke omgeving

Een tweede scenario is al te voorzien, maar nog minder uitgewerkt, en dat is het scenario dat een archiefvormer en/of beheerder werkt in een volgens Common Ground ingerichte informatiebeheeromgeving waarin zowel MDTO als PREMIS zijn geïmplementeerd. Dit is het scenario van *archiving by design*, en zo je wil ook *preservation by design*. Vanaf de creatie van

<sup>10</sup> PREMIS 3.0 Data Dictionary, pagina 12.



een document worden metagegevens gelogd in zowel MDTO (over het informatieobject) als in PREMIS (over de digitale preserving). In dit scenario wordt er nog actief gewerkt aan de informatieobjecten en hun metadata in kwestie en kan er in tandem worden 'gewerkt' in beide standaarden. Er kunnen immers al preservingacties plaatsvinden in een dynamische informatiebeheeromgeving, getriggerd door hun eigen procescomponenten.

Strikt genomen kan dit scenario ook uitgaan van een beperkte functionele mapping tussen de standaarden, die alleen op Object/Informatieobject daadwerkelijk overlapt. Dit scenario beschrijft een informatieobject dat wordt geregistreerd in de beheeromgeving onder het regime van beide standaarden, en de procescomponenten die gebruik maken van de data die in de bijbehorende databases worden vastgelegd. Er zullen business rules moeten worden ingesteld die datapunten bij registratie óf automatisch 'dubbel' of verwijzingen vastleggen tussen de twee standaarden, waar ze elkaar raken. Bijvoorbeeld bij technische gegevens over Bestand/File.

De repository van PREMIS is gemodelleerd als een aparte omgeving die is gericht op langdurig beheer van informatieobjecten. Dat is in lijn met het OAIS-model, dat het e-depot als een aparte voorziening ziet ten opzichte van de creatieomgeving van informatieobjecten. In de huidige generatie e-depot voorzieningen is dit vaak geïnterpreteerd als een apart systeem waar informatie naar toe wordt geëxporteerd, waarna het in isolatie wordt beheerd. Het is de vraag hoe houdbaar een dergelijke interpretatie is in een dynamische informatiebeheeromgeving conform Common Ground, waar nog informatie kan worden toegevoegd, gewijzigd en verwijderd. Is het toevoegen van een informatieobject bijvoorbeeld hetzelfde als het aanbieden van een Submission Information Package (SIP) conform OAIS?

Dit zou een onderzoek op zichzelf waard zijn.

Dit scenario verdient verdere uitdieping, onder meer door een antwoord op de volgende vragen:

- Wat zijn de specifieke verschillen in benodigde systeeminrichting tussen implementatiescenario 1 en 2?
- Is scenario 2 te zien als een uitbreiding van scenario 1, of is een compleet andere inrichting noodzakelijk? En zo ja, welke?
- Is het OAIS-model inzetbaar binnen een Common Ground-omgeving, waarin het ook in dynamisch (informatie)beheer voorziet, of is het afhankelijk van het Ingestproces daarvoor (inclusief SIP/AIP)?
- Moeten we MDTO beschouwen als afgesloten 'pakketten metadata+bestanden' die vervolgens onder het beheerregime van PREMIS komen, of kan er sprake zijn van een dynamische MDTO-registratie die gewijzigd kan worden?

## 4. Conclusie en verder onderzoek

De hoofdvraag die aan het begin van deze notitie werd gesteld, was “Kunnen MDTO en PREMIS samen gebruikt worden in een informatiemodel dat geïmplementeerd is in een Common Ground-ingericht systeem waarin de duurzame toegankelijkheid van digitale archiefbescheiden wordt geborgd?” De gemaakte mapping, met relaties vanuit de SKOS-standaard, is een eerste stap om deze vraag met “Ja” te kunnen beantwoorden. Er is echter verder onderzoek nodig om een onderbouwd antwoord te kunnen geven op deze vraag. Zowel op conceptueel vlak als op praktisch vlak zijn er nog wat vragen te beantwoorden.

De standaarden vertegenwoordigen elk een eigen venster op de werkelijkheid. Om MDTO en PREMIS ook conceptueel beter op elkaar te laten aansluiten, is een gezamenlijk, overkoepelend begrip van entiteiten gewenst. Welk conceptueel model hangen we boven de mapping? Het uitdenken van zo'n model betekent definiëren wat onze werkelijkheid is en hoe de standaarden daarin passen. Omdat een informatieobject in principe van alles kan zijn, van een stuk papier tot een database zelf, moeten we in gesprek over hoe we de verschillende vormen van entiteit tussen de twee standaarden aan elkaar gaan relateren. Een vraag die hiermee samenhangt, is wat het betekent om onder standaarden te werken.

- Betekent dit dat een database conform een standaard ingericht moet zijn of dat je door een filter van de standaard informatie moet kunnen vinden?
- Moet je je gegevens vastleggen in een standaard maar vanuit een andere omgeving kunnen raadplegen?
- Maakt het uit of we standaardisering inrichten in de gegevensuitwisselingslaag van Common Ground of dat we die inrichten in de data laag?
- Welke eisen en randvoorwaarden moeten gesteld worden om de borging van de implementatie van een standaard duurzaam te behouden?

De beantwoording van bovenstaande vragen en het toewerken naar de implementatie van de mapping in de praktijk, brengt de volgende vragen met zich mee:

- Welke gevolgen hebben onze mappingskeuzes daadwerkelijk voor systemen? We hebben het tot nu toe alleen over theoretische implementatie, een daadwerkelijke implementatie is er nog niet.
- Als de mapping vanuit ons vakgebied op goedkeuring kan rekenen, wat kan de sector er dan mee gegeven de huidige, bestaande zaaksystemen, vakapplicaties en e-depotsystemen?

De implementatie van het uitgewerkte model kan vervolgens in verschillende mallen gegoten worden, want implementatie als uitwisselingsstandaard, implementatie als onderdeel van een relationele database of als onderdeel van een component-gebaseerd systeem vragen verschillende dingen. Hoewel het een mapping van standaarden betreft, is er dus duidelijk niet één perfecte oplossing. Als Kerngroep gingen we in den beginne uit van een volledig carte blanche Common Ground-omgeving, zoals ook elders in deze notitie genoemd als uitgangspunt, maar er moet ook rekening gehouden worden met oud naast nieuw en migratie naar een component-gebaseerde oplossing. Op basis van het uitgedachte model kunnen we onze randvoorwaarden en implementatiescenario's aanscherpen. Hierbij zullen we uitgaan

van implementatie van en uitwisseling tussen relationele databases en componentgerelateerde systemen, daar een implementatiescenario dat uitgaat van een volledig Common Ground ingerichte informatiearchitectuur nog te veel luchtfietserij is. De randvoorwaarden voor een linked data-systeem zijn relatief concreet: slaan we meerdere triples op of gebruiken we SKOS als mapping? Dergelijke randvoorwaarden zijn lastiger voor een relationele database, daarvoor zullen we bredere definities moeten gaan vastleggen.

Samengevat hopen met deze notitie een bijdrage aan de discussie te hebben geleverd, en enige ontwikkelrichtingen te hebben aangewezen. We zijn ervan overtuigd dat implementatie van MDTO en PREMIS een belangrijke bijdrage kan leveren aan het realiseren van duurzame toegankelijkheid. Het zal voor een belangrijk deel in de praktijk moeten blijken hoe een gezamenlijke implementatie van de standaarden eruit zou kunnen zien. We moedigen deze ontwikkeling aan, net als een goede, interdisciplinaire discussie hierover.