



-Metadatabeleid e-depot
Brabants Historisch Informatie Centrum-



BRABANTS HISTORISCH INFORMATIE CENTRUM

Metadatabeleid digitaal archief

Versie 1.0

Aanleiding

Bij het BHIC werken we aan verschillende projecten om de duurzame toegankelijkheid van digitale archieven te waarborgen. Om deze projecten op een goede manier te ondersteunen, is een helder metadatabeleid van groot belang. In dit beleid bepalen we onder andere welke metadata deel uitmaakt van het Archival Information Package (AIP), hoe we verschillende soorten metadata conserveren en beheren, en hoe we de metadata toegankelijk maken voor gebruikers.

Om het beheer van metadata van digitale archiefbescheiden goed te laten verlopen, zijn er verschillende kaders opgesteld. Zo zijn er wetten en regelingen zoals de Archiefwet 1995, Archiefregeling en Archiefbesluit 1995 die hierbij een rol spelen. Daarnaast is er het Referencemodel for an Open Archival Information System (OAIS), een internationaal model dat garant staat voor duurzame toegankelijkheid van informatie. Ook zijn er richtlijnen zoals de NEN-ISO 23081 die principes voor het beheren van metagegevens beschrijven en MDTO (Metagegevens voor duurzaam toegankelijke overheidsinformatie)¹, dit is een norm voor het vastleggen en uitwisselen van eenduidige metagegevens om de duurzame toegankelijkheid van overheidsinformatie mogelijk te maken.² Het is belangrijk dat we deze kaders meenemen bij het opstellen van ons metadatabeleid om zo de duurzame toegankelijkheid van archiefbescheiden te garanderen.

Het metadatabeleid van BHIC draait om één cruciale vraag: welke metadata maken deel uit van het Archival Information Package? Door hier een weloverwogen standpunt over in te nemen, zorgen we ervoor dat we de duurzame toegankelijkheid van onze archiefbescheiden kunnen garanderen.

In het kader van ons metadatabeleid moeten we de volgende vragen beantwoorden:

- Wat is het belang van metadata?
- Welke soorten metadata onderscheiden we?
- Hoe gaan we om met wijzigingen van de metadata in overgebracht archief?
- Hoe gaan we om met aanvullingen op overgebrachte archieven en series binnen archieven?
- Hoe gaan we om met de metadata in overgebracht archief bij migratie naar een ander e-depotsysteem?
- Hoe maken we de metadata toegankelijk voor de gebruiker?

Wat is het belang van metadata?

Metadata zijn cruciaal voor het duurzaam toegankelijk houden van archiefbescheiden. Dit komt doordat metadata:

¹ <https://www.nationaalarchief.nl/archiveren/mdto> (geraadpleegd 4-5-2023)

² MDTO vervangt het Toepassingsprofiel Metagegevens Lokale Overheden (TMLO), dat specifiek gericht is op lokale overheden.

- De inhoud van archiefbescheiden beschrijven en zo de vindbaarheid van de archiefbescheiden garanderen.
- De vorm van archiefbescheiden beschrijven, inclusief de technische specificaties van bestanden, waardoor de beschikbaarheid en leesbaarheid van de archiefbescheiden gewaarborgd blijven.
- De archiefbescheiden identificeren en de relaties tussen de verschillende stukken beschrijven, wat ervoor zorgt dat de oorspronkelijke ordening bewaard blijft of gereconstrueerd kan worden.
- De openbaarheid, gebruiksrechten en privacyrechten van de archiefbescheiden beschrijven, waardoor alleen geautoriseerde gebruikers toegang hebben tot de archiefbescheiden.
- De ontstaansgeschiedenis van de archiefbescheiden en de context waarbinnen deze gefunctioneerd hebben beschrijven, waardoor de archiefbescheiden interpreteerbaar en betrouwbaar blijven.
- De geschiedenis van het beheer van de archiefbescheiden sinds de creatie tot en met het heden beschrijven, waardoor de betrouwbaarheid van de archiefbescheiden gewaarborgd blijft.
- De integriteit van de archiefbescheiden bewaken door een checksum vast te leggen in de metadata.

Bij de creatie van een digitaal archiefstuk krijgt het al een deel van deze metadata mee, gedeeltelijk handmatig en gedeeltelijk geautomatiseerd. Gedurende de levensduur van het archiefstuk worden er meer metadata toegevoegd.

Wat beschouwen wij als de AIP en in hoeverre maken de metadata daar deel van uit?

In de archiefwereld is het begrip Archival Information Package (AIP) van groot belang. Het AIP omvat niet alleen het digitale object zelf, maar ook de bijbehorende metadata. Deze metadata zijn onmisbaar voor duurzame bewaring en toegankelijkheid van het object. In het OAIS-model wordt een onderscheid gemaakt tussen het Submission Information Package (SIP), dat door de archiefvormer wordt aangeleverd, en het AIP dat door de archiefbewaarplaats wordt gecreëerd. De archiefbewaarplaats bepaalt zelf welke metadata er onderdeel uitmaken van het AIP, maar dit is minimaal de informatie die nodig is voor duurzame bewaring en toegang.

Om archiefbescheiden duurzaam te bewaren en toegankelijk te maken, legt het BHIC alle metadata op de verschillende aggregatieniveaus vast in het AIP.

Welke soorten metadata onderscheiden we?

In de Nederlandse archiefwereld wordt er onderscheid gemaakt tussen originele metadata en metadata die ontstaan tijdens het beheer van het archief. De grens kan worden gelegd bij de overgang van de semi-statische naar de statische fase, bij overbrenging. Bij digitaal archief wordt er steeds vaker archivering by design toegepast en minder in de drie fasen gedacht, maar de Archiefwet gaat nog steeds uit van een moment van 'bevriezing' van de archiefbescheiden: de overbrenging naar

een archiefbewaarplaats. Om deze reden is het zinvol om onderscheid te maken tussen metadata die worden toegekend voor en na overbrenging.

Er zijn twee soorten metadata die onderscheiden kunnen worden: metadata toegekend door de archiefvormer vanaf het moment van creatie tot het moment van overbrenging, aangevuld met metadata toegekend door het BHIC tijdens de pre-ingestfase en metadata toegekend tijdens het beheer van de archiefbescheiden na overbrenging naar de archiefbewaarplaats.

Een ander onderscheid in metadata wordt gemaakt in de Model Architectuur voor Rijks Archiefinstellingen (MARA).³ Dit onderscheid lijkt op het onderscheid tussen Representation Information en Preservation Description Information in het OAIS-model. De MARA onderscheidt manifestatie-afhankelijke metadata (MAM) en manifestatie-onafhankelijke metadata (MOM). MAM heeft betrekking op de verschijningsvorm van een informatieobject, bijvoorbeeld informatie over bestandsformaat, checksums en wie welk stuk heeft ingezien. MOM heeft daarentegen geen betrekking op de verschijningsvorm, maar bijvoorbeeld op zorgdragers, archiefvormers, openbaarheid en inhoudelijke beschrijving.

In de archiefwereld wordt het begrip manifestatie gebruikt om verschillende verschijningsvormen van een archiefstuk aan te duiden. Bijvoorbeeld, een Word-document en een inhoudelijk identieke PDF-A vormen twee verschillende manifestaties, met verschillende MAM, maar identieke MOM.

TMLO maakt ook onderscheid tussen metadata van aggregaties en metadata van bestanden. In theorie komt de metadata van aggregaties overeen met MOM en de metadata van bestanden met MAM. Echter, de MARA rekent meer metadata tot MAM dan TMLO doet. In MDTO is het mogelijk om MOM vast te leggen bij zowel aggregaties als bestanden, maar MAM kan alleen bij bestanden worden vastgelegd. Het is belangrijk om dit onderscheid te begrijpen bij het vastleggen van metadata, omdat het helpt bij het organiseren en beheren van archiefbescheiden en bij het garanderen van de toegankelijkheid op de lange termijn.

Wij hechten er aan om de begrippen MOM en MAM te handhaven. Als we het OAIS-concept AIP gebruiken en vasthouden aan de Archiefwet en aan het onderscheid tussen MOM en MAM kunnen we vier groepen metadata definiëren:

1. Primaire MOM en MAM: Manifestatie-onafhankelijke en manifestatie-afhankelijke metadata toegekend door de archiefvormer vanaf het moment van creatie tot het moment van overbrenging van de archiefbescheiden naar de archiefbewaarplaats. Dit zijn de originele, te preserveren metadata. Tevens wordt hieronder verstaan metadata die door het BHIC zijn toegekend tijdens de pre-ingestfase.
2. Secundaire MAM: Manifestatie-afhankelijke metadata toegekend tijdens de omvorming van een SIP tot een AIP bij de overbrenging naar de archiefbewaarplaats.
3. Tertiaire MAM: Manifestatie-afhankelijke metadata toegekend tijdens het beheer van de archiefbescheiden ná overbrenging naar de archiefbewaarplaats. Dit zijn technische metadata van nieuwe AIP Versies die zijn ontstaan als gevolg van preserveringsacties.

³ [https://www.noraonline.nl/wiki/MARA_\(Model_Architectuur_voor_Rijks_Archiefstellingen\)](https://www.noraonline.nl/wiki/MARA_(Model_Architectuur_voor_Rijks_Archiefstellingen)) (geraadpleegd 4-5-2023).

4. Preservation Process History: Manifestatie-onafhankelijke metadata toegekend door de beheerder van de archiefbewaarpplaats bij het beheer van de archiefbescheiden ná overbrenging. Dit zijn logbestanden van beheeracties en beschikbaarstellingen.

De SIP bevat metadata van groep 1, terwijl de originele AIP metadata bevat van de groepen 1 en 2. Als er een nieuwe versie van een AIP wordt gemaakt, ontstaat metadata van groep 3. De vierde groep metadata, Preservation Process History, wordt apart gehouden als Data Management Information en maakt geen deel uit van de AIP.

Dankzij dit onderscheid kunnen we garanderen dat de metadata die nodig is voor duurzame toegankelijkheid van archiefbescheiden altijd deel uitmaken van de AIP. Daarnaast zorgen we ervoor dat de originele versie van een archiefbestanddeel wordt bewaard met de metadata zoals die waren op het moment van overbrenging. Tot slot wordt de geschiedenis van het beheer van de archiefbescheiden vastgelegd, samen met bijbehorende metadata. Door deze aanpak kunnen we ervoor zorgen dat de archiefbescheiden duurzaam toegankelijk blijven voor de toekomst.

Welke systemen gebruiken we om de verschillende soorten metadata te preserven en beheren?

Om metadata te beheren en te behouden maken we gebruik van verschillende systemen. Het OAIS-model en de MARA zijn twee belangrijke benaderingen hiervoor.

Het OAIS-model gaat uit van twee systemen: een Archival Storage waarin de AIP's worden bewaard en een Data Management waarin de Descriptive Information en Preservation Process History worden beheerd.

De MARA benadrukt het belang van het opslaan van gegevens in één systeem en het scheiden van inhoudelijke metadata van informatie over de opslagplaats. Het e-depot wordt gezien als de digitale bewaarplaats voor overgebracht materiaal, terwijl de MOM van overgebrachte en uitgeplaatste archiefbescheiden in het collectiebeheersysteem worden beheerd.

Hoewel beide benaderingen verschillende keuzes maken voor het beheer van metadata, gebruikt het BHIC voornamelijk het OAIS-model in de praktijk. Dit houdt in dat zowel de MAM als de MOM in de Archival Storage worden bewaard, terwijl de Preservation Process History en Descriptive Information worden beheerd in het Data Management.

Dit leidt tot de volgende onderverdeling:

- Primaire MOM en MAM worden gepreserveerd in de Archival Storage.
- Secundaire en tertiaire MAM worden gegenereerd, beheerd en gepreserveerd in de Archival Storage.
- Preservation Process History en Descriptive Information worden beheerd in het Data Management.

Hoe gaan we om de oorspronkelijke ordeningsstructuur van het archief?

Wanneer overheidsorganen archiefbescheiden ontvangen of maken, moeten ze altijd de ordeningsstructuur vaststellen en de samenhang met andere archiefbescheiden kunnen aantonen. Dit kan op drie manieren:

- door de archiefbescheiden in de originele ordeningsstructuur op te nemen in de Archival Storage,
- door de originele structuur vast te leggen in de AIP in de primaire MOM,
- of door wijzigingen in de structuur te loggen.

Als BHIC hebben we ervoor gekozen om altijd ten minste de oorspronkelijke structuur vast te leggen in de metadata en indien mogelijk ook te behouden in de structuur van het archief. Zo kunnen we altijd achterhalen hoe de archiefbescheiden waren geordend en hoe ze zich verhouden tot andere bescheiden.

Hoe gaan we om met wijzigingen van de metadata in overgebracht archief?

Als we het OAIS-model strikt toepassen, moeten we altijd nagaan of wijzigingen in metadata gevolgen hebben voor de metadata in de AIP. De AIP moet minimaal de metadata bevatten die nodig zijn voor duurzame toegankelijkheid. Als we de metadata in de AIP willen wijzigen, moeten we een nieuwe AIP-editie maken. Als dit niet nodig is, kunnen we alleen de beschrijving van het pakket aanpassen in het datamanagement.

Het MARA-model verschilt van het OAIS-model. Het MARA-model maakt geen onderscheid tussen edities, versies en afgeleiden en gebruikt het begrip "manifestatie". Dit begrip beperkt zich niet tot de AIP, ook een DIP is een manifestatie. Volgens het MARA-model leidt het creëren van een nieuwe manifestatie alleen tot nieuwe MAM, terwijl de oude MOM niet opnieuw worden opgeslagen. Wij volgen het MARA-model op deze manier.

Hoe gaan we om met aanvullingen op overgebrachte archieven en series binnen archieven?

Bij een klassieke inventaris worden een archief en de series daarin gedateerd op basis van de archiefstukken die eronder vallen. Maar bij digitaal archief is dit lastiger, omdat archiefvormers periodiek afgeronde dossiers willen overbrengen zonder te wachten op het afronden van een heel archiefblok. Hierdoor zullen er regelmatig aanvullingen plaatsvinden op bestaande archieven en series, waardoor de einddatum niet altijd bekend is. Om de toegankelijkheid van de archieven en series te waarborgen, is het belangrijk dat de datering op archief- of serieniveau klopt. Het BHIC past daarom bij aanvullingen de metadata van een bestaand archief of serie aan en wijzigt indien nodig de primaire MOM in de bewaaromgeving. Zo kunnen gebruikers altijd de juiste datering vinden en wordt het archief goed toegankelijk gehouden.

Hoe gaan we om met de metadata in overgebracht archief bij migratie naar een ander e-depotsysteem?

Wanneer een archiefinstelling overstapt naar een nieuw e-depotsysteem, moeten in eerste instantie alle metadata die deel uitmaken van de Archival Information Package (AIP) gemigreerd worden. Dit omvat de primaire MOM en MAM en de secundaire en tertiaire MAM uit de Archival Storage, evenals de tertiaire MAM. Deze moeten worden geëxporteerd als een Dissemination Information Package (DIP), zodat ze kunnen worden geïmporteerd als Submission Information Package (SIP) in de Archival Storage van het nieuwe e-depotsysteem. Daarnaast moet ook de Preservation Process History uit het Data Management gemigreerd worden naar het Data Management van het nieuwe

systeem. Op deze manier kan alle metadata behouden blijven en kan de archiefinstelling blijven voldoen aan de eisen voor duurzame toegankelijkheid van het archief.

Hoe maken we de metadata toegankelijk voor de gebruiker?

Wanneer archief wordt overgebracht naar het e-depot, wordt van een deel van de MOM en MAM een afgeleide opgeslagen in Mylex Overheid (MLO). Dit wordt de Descriptive Information genoemd, die als basis dient voor toegang. Om de toegang te realiseren worden twee portals gebruikt: een ambtenaren- en medewerkersportaal voor de archiefvormers en medewerkers van het BHIC (e-depot.bhic.nl), en een publieksportaal voor derden (digitaaldepot.bhic.nl).