



Contact

W. van der Reijden
Adviseur recordkeeping

T +31 6 55 26 79 52
wout.van.der.reijden@nationaal
archief.nl

Datum

25 oktober 2018

Onze referentie

-

Auteur

W. van der Reijden

memo

algemene knelpunten gebruik toepassingsprofielen
metagegevens en mogelijke oplossingsrichtingen

Aanleiding

Een toepassingsprofiel metagegevens is een model dat metagegevens en hun onderlinge relaties specificeert die nodig worden geacht voor de duurzame toegankelijkheid van informatie. Binnen de Nederlandse overheid worden twee generieke toepassingsprofielen voor dat doel gebruikt. Het TMLO wordt gebruikt door lokale overheden, de Rijksoverheid gebruikt het Toepassingsprofiel Metagegevens Rijksoverheid (TP Rijk).

Sinds 2009 worden binnen de Nederlandse overheid toepassingsprofielen gebruikt voor het metadateren van informatie ten faveure van de duurzame toegankelijkheid. Inmiddels is er de nodige praktische ervaring opgedaan in het gebruik van de profielen (zowel TMLO als TP Rijk) en het hier van afgeleide uitwisselingsformaat ToPX.

Deze notitie beschrijft een aantal knelpunten in het gebruik van de toepassingsprofielen. Deze knelpunten kunnen onder andere ten koste gaan van de duurzame toegankelijkheid van informatie en leiden tot onevenredige inspanningen door zowel archiefvormers als archiefinstellingen. De notitie beschrijft ook mogelijke oplossingsrichtingen voor deze knelpunten. De notitie is bedoeld als een discussiestuk. De uitkomsten van de discussie kunnen worden gebruikt voor het verder uitwerken van oplossingsrichtingen.

De discussie naar aanleiding van deze notitie heeft tot doel:

- te verifiëren of deze knelpunten herkend worden;
- te weten te komen of er nog andere knelpunten zijn die ontbreken in het overzicht;
- te verifiëren of de oplossingsrichtingen herkenbaar zijn en zij inderdaad als oplossingsrichting voor de betreffende knelpunten gezien kunnen worden; en
- te weten te komen of er alternatieve oplossingsrichtingen zijn om de knelpunten aan te pakken.

Knelpunten

Hieronder volgt een korte beschrijving van knelpunten die worden ervaren door archiefvormers en archiefinstellingen. Er wordt daarbij onderscheid gemaakt tussen inhoudelijke, technische en organisatorische knelpunten.

Nationaal Archief

Onze referentie

Datum

25 oktober 2018

Inhoudelijke knelpunten

1. De profielen bevatten verplichtingen en beperkingen waar archiefvormers niet aan kunnen (of willen) voldoen. Dit leidt bij het uitwisselen van metagegevens tot aanpassingen van de originele metagegevens. Dit gaat ten koste van de kwaliteit. Zoals het invullen van dummy-waarden en op punten twijfelachtige vertalingen waardoor metagegevens een andere betekenis krijgen dan bedoeld.

De inspanningen voor deze zogenaamde mapping is met name voor archiefvormers aanzienlijk en leveren hoegenaamd geen verbetering van de metagegevens op. De eisen worden in veel gevallen niet reëel gevonden en kunnen geweld doen aan archivalistische principes als original order. Het wordt ook niet realistisch geacht dat toekomstige informatiesystemen wel volledig aan de profielen zullen voldoen. Daarvoor bieden de profielen te weinig meerwaarde voor het primaire proces (en dat bepaald waar een informatiesysteem aan moet voldoen). Dit knelpunt komt in de praktijk bij elk aansluittraject terug.

2. De toepassingsprofielen en ToPX zijn gesloten. Ze bieden geen mogelijkheid om op een systematische manier extra metagegevens toe te voegen of een andere ordening aan te houden dan een traditionele hiërarchische. Dit heeft een aantal gevolgen:
 - Bij een uitwisseling gaan waardevolle metagegevens die in het bronproces zijn vastgelegd verloren, omdat ze niet zijn gespecificeerd in de profielen en ToPX. Of ze worden op een ad hoc manier uitgewisseld waardoor de mogelijkheden om ze te gebruiken beperkt zijn. Bijvoorbeeld omdat niet is vastgelegd wat de betekenis van de extra metagegevens is (die nu in element 22 van ToPX terecht komen). Dit vraagstuk komt bij elk aansluittraject voor.
 - De traditionele hiërarchische ordening (archief-serie-dossier-archiefstuk) is afkomstig uit het papieren tijdperk. In DMS'en zaaksystemen is een dergelijke ordening nog enigszins te herkennen en is een vertaling nog te realiseren. Echter, bij andere digitale informatiesystemen is een dergelijke ordening vaak niet aanwezig en wordt samenhang bereikt door het leggen van relaties. Een goed voorbeeld hiervan zijn de digitale Handelingen van de Tweede Kamer. Of er is geen standaard ordening, zoals is gebleken uit het onderzoek naar het archiveren van e-mails. Een mapping is dan niet goed te maken, waardoor een aansluiting op inhoudelijke gronden op dit moment niet te realiseren is. Of alleen met serieuze concessies op het gebied van kwaliteit en toegankelijkheid.
3. De manier waarop een archiefinstelling metagegevens ontvangt en de manier waarop ze de metagegevens beschikbaar wil stellen (EAD, RDF of iets anders?) is verschillend. Dit betekent dat er een of meerdere transformatie plaats moeten vinden voordat metagegevens beschikbaar gesteld kunnen worden. Dit gaat ten koste van de kwaliteit van de metagegevens (omdat er altijd vertaalfouten gemaakt worden) en brengt extra kosten en inspanning met zich mee.

Technische knelpunten

4. TMLO en TP Rijk kennen nog geen standaard koppelvlakspecificatie (dit is een specificatie van de vorm en het proces voor uitwisseling van (meta)gegevens). Hierdoor is de praktische toepasbaarheid voor het uitwisselen van metagegevens beperkt. ToPX heeft wel een koppelvlak, maar die is alleen van toepassing op uitwisseling met het e-Depot van het NA. En ToPX zelf is zoals gezegd ook gesloten.
5. De manier waarop de teksten in de profielen zijn geschreven is niet gestandaardiseerd. Daardoor kan er onduidelijkheid ontstaan over de interpretatie van die teksten en zijn er geen standaard tools beschikbaar om de profielen en metagegevens te verwerken. Bijvoorbeeld voor de publicatie van de profielen en de validatie van de metagegevens.

Nationaal Archief

Onze referentie

Datum

25 oktober 2018

Organisatorische knelpunten

6. Er zijn verschillende profielen voor het Rijk (TP Rijk) en decentrale overheden (TMLO). De doorontwikkeling van TMLO kan ervoor zorgen dat de profielen meer van elkaar gaan afwijken. Dit kan de samenwerking tussen verschillende overheidslagen belemmeren (b.v. rond de Omgevingswet) en geeft een grotere beheerlast.
7. Het beheer van de profielen vergt een grote inspanning. Waarvoor landelijk slechts beperkte deskundigheid beschikbaar is. Dat betekent dat aanpassingen een lange doorlooptijd hebben.

Mogelijke oplossingsrichtingen

a) Met betrekking tot knelpunt 1 & 7:

Beperk de profielen tot het absolute minimum dat noodzakelijk is om informatie duurzaam toegankelijk te houden.

Dit betekent dat onderzocht moet worden of het aantal elementen in TMLO teruggebracht kan worden. Een uitgangspunt daarbij is dat bij de profielen alleen elementen worden opgenomen die voor alle soorten informatie en bronprocessen relevant zijn.

b) Met betrekking tot knelpunt 2:

Sta archiefvormers toe aanvullende metagegevens aan te leveren.

De archiefvormer maakt of kiest aanvullende profielen voor specifieke soorten informatie en bronprocessen (als dat nodig is). Vaak is daarvoor al een domeinstandaard beschikbaar, zoals het toepassingsprofiel voor geoinformatie. Deze moeten voldoen aan de gekozen standaard voor het beschrijven van profielen.

Deze oplossingsrichting is voorwaardelijke voor a)

c) Met betrekking op knelpunt 3:

Stem de manier waarop archiefvormers metagegevens vastleggen en beschikbaar stellen en de manier waarop archiefinstellingen op hun beurt metagegevens beschikbaar stellen op elkaar af. Zodat de noodzaak tot vertalen geminimaliseerd wordt.

Het gaat hier om de manier waarop en de vorm van beschikbaarstelling, niet de inhoud. De trend is daarbij het gebruik van Linked Data, waarbij toekomstige gebruikers meerdere datasets kunnen koppelen.¹

Nationaal Archief

Onze referentie

d) Met betrekking op knelpunt 3 & 5:

Kies een standaard voor het beschrijven van metagegevensprofielen. De standaard dient daarbij zodanig te zijn dat er al een koppelvlakspecificatie voor beschikbaar is.

Datum

25 oktober 2018

Hier zijn UML en RDF voor de hand liggende standaarden. Waarbij UML vooral bruikbaar kan zijn voor het modelleren van informatie, is een standaard als RDF mogelijk beter bruikbaar voor het op een standaard manier beschrijven van metagegevens.

Een dergelijke standaard kan worden gebruikt om zowel de standaard metagegevens als de door de zorgdragers zelf gedefinieerde metagegevens eenduidig te beschrijven.

e) Met betrekking tot knelpunt 6 & 8:

Voeg TPR en TMLO samen tot een enkel toepassingsprofiel dat gebruikt kan worden door de gehele overheid.

Dit ondersteunt de samenwerking tussen overheden en betekent daarnaast dat er één product overblijft dat beheerd moet worden.

¹ De NDE maakt voor haar gedistribueerde netwerk gebruik van Linked Data. Aggregatoren als Europeana en DBpedia hebben linked data omarmd. Het Kadaster biedt geo-datasets aan als linked data en ook de Tweede Kamer is voornemens om een pilot met linked data uit te voeren. Standaarden als Dublin Core worden als vocabulaires gebruikt in het kader van linked data.