



**Expertadvies voor het wijzigen van de status van OAI-PMH
van 'verplicht' naar 'aanbevolen'.**

Concept ter openbare consultatie

Datum 28 juli 2017

Het Forum Standaardisatie wordt gevraagd om in te stemmen met:

Het Forum Standaardisatie adviseert het Nationaal Beraad Digitale Overheid om de huidige status 'verplicht' van de standaard OAI-PMH op de lijst open standaarden te wijzigen naar 'aanbevolen'.

Ook wordt geadviseerd om in het functioneel toepassingsgebied kleine wijzigingen door te voeren. Hiermee wordt het toepassingsgebied als volgt:

"Het vraaggestuurd aanbieden en ophalen van verzamelingen metadata uit bibliotheken met (digitale) documenten of andere objecten gericht op leermaterialen en cultureel erfgoed, met als doel het opnemen van deze metadata in een centrale bibliotheek en/of het verrijken van eigen gegevens."

Analyse en Forumadvies

Aanleiding

In 2016 heeft Verdonck, Klooster & Associates (VKA) een evaluatie uitgevoerd van het Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Het doel van deze evaluatie was om de kwaliteit van de informatie over de standaard op de lijst te waarborgen inzicht te krijgen in de huidige en toekomstige ontwikkelingen. Bij de evaluatie is onder andere gekeken naar het gebruik, het functioneel toepassingsgebied (*is deze duidelijk en zo concreet mogelijk geformuleerd?*), het belang van de standaard (*wat is de toegevoegde waarde van de standaard in de praktijk?*) en de openstaande adviezen (*zijn de bij opname geadviseerde adviezen opgevolgd?*).

Op basis van de evaluatie is geadviseerd om te onderzoeken of OAI-PMH naar de aanbevolen verplaatst moet worden, en zo ja, onder welke voorwaarden.

Op verzoek van het Forum Standaardisatie heeft VKA dit onderzoek uitgevoerd.

Totstandkoming van het advies

De inhoudelijke grondslag voor dit advies is de eerder uitgevoerde evaluatie, het bijbehorende Rapport van bevindingen - Evaluatie OAI-PMH¹ en een aantal interviews met experts die voor het huidige onderzoek zijn gehouden. Op verzoek van het Forum Standaardisatie zijn voornamelijk organisaties benaderd uit de erfgoedsector.

¹ <https://www.forumstandaardisatie.nl/sites/bfs/files/proceedings/FS%20161214.3A%20Evaluatie%20advies%20OAI-PMH.pdf>.

Betrokken experts

Leon Bok - Rijksdienst voor het Cultureel Erfgoed
Maarten Brinkerink - Beeld en Geluid
Bob Coret – Coret Genealogie
Jeroen Hamers - Kennisnet
Enno Meijers – Koninklijke Bibliotheek
Hans Schraven - Rijksdienst voor het Cultureel Erfgoed
Lex Slaghuis – Open State Foundation
Robert Thiessen - Picturae
Sjors de Valk – Netwerk Digitaal Erfgoed
Joop Vanderheiden - Rijksdienst voor het Cultureel Erfgoed
Frans van der Zande – Rijksdienst voor het Cultureel Erfgoed
Maarten Zeinstra - Kennisland

Over de standaard

OAI-PMH is een mechanisme om op een laagdrempelige en eenvoudige wijze, platform- en gegevensafhankelijk, uitwisseling tussen repositories te realiseren. Een repository is een bibliotheek met documenten of objecten (ook wel 'content' genoemd), bijvoorbeeld een (digitaal) archief. Het OAI-Protocol voor Metadata Harvesting (OAI-PMH) definieert het mechanisme om metadataseten vanuit die repositories bijeen te brengen. OAI-PMH is een technische standaard.

De standaard definieert een aantal 'vragen' die aan een digitale bibliotheek gesteld kunnen worden. Bijvoorbeeld: 'welke content zit er in deze bibliotheek?', 'geef de metadata van deze verzameling documenten', etc. De bibliotheek levert vervolgens deze metadata aan het zoekstelsel. Het zoekstelsel kan dit vervolgens weer opnemen in een totaalindex, waar gebruikers in kunnen zoeken.

OAI-PMH specificeert niet hoe de digitale content vervolgens door de gebruiker uit de onderliggende bibliotheek wordt gehaald en welk formaat de digitale content heeft. Hiervoor zijn andere standaarden nodig (zoals FTP en PDF/A).

Gebruik van OAI-PMH binnen de erfgoedsector

OAI-PMH wordt binnen de erfgoedsector beschouwd als de norm voor de samenwerking van instellingen in de erfgoedsector. De standaard kent dan ook een hoge adoptiegraad binnen de sector. Vrijwel alle middelgrote en grote instellingen (bijv. de Koninklijke Bibliotheek, het Nationaal Archief, het Rijksmuseum en Beeld en Geluid) passen OAI-PMH toe voor de uitwisseling van metadataseten.

Ook bij kleinere instellingen is het gebruik van OAI-PMH gangbaar, al is de adoptiegraad hier minder hoog. Bij kleinere instellingen kan o.a. gedacht worden aan gemeenten en organisaties die specifieke collecties beheren zoals Stichting Papua Cultureel Erfgoed en de Vereniging De Hollandsche Molen. De lagere adoptiegraad komt voornamelijk voort uit het feit dat er doorgaans minder middelen beschikbaar zijn voor OAI-implementaties. Hierdoor wordt bij deze instellingen vaker gekozen voor goedkopere en traditionelere oplossingen zoals CSV-dumps. Dit is een ad-hoc oplossing

waarbij een kopie van de gegevens wordt gemaakt in een zeer eenvoudige bestandsindeling.

De hoge adoptiegraad is geleidelijk ontstaan en ook op dit moment zijn OAI-PMH implementatietrajecten nog aan de orde. Belangrijke aanjagers van OAI-PMH zijn aggregatieplatformen als Europeana geweest die OAI-PMH vereisen. Dergelijke aggregatieplatforms zijn een soort verzamelbibliotheken waar gegevens van andere collecties bij elkaar worden gevoegd. Hier zit een gelaagdheid in waarbij grote platforms als Europeana beschouwd kunnen worden als de top van een piramide.

Alle grote leveranciers (o.a. Picturae en DE REE en DEVENTit) van collectiebeheersystemen in Nederland hebben de implementatie van het protocol opgenomen in hun software.

Ervaringen met het gebruik van OAI-PMH

Er is een hoge mate van consensus binnen de erfgoedsector over de toegevoegde waarde van OAI-PMH. Het gaat dan met name om de mogelijkheid die het organisaties biedt om op redelijk laagdrempelige wijze hun collectie informatie zo breed mogelijk te verspreiden naar de buitenwereld, in het bijzonder naar aggregatieplatforms.

Desondanks worden er ook een aantal nadelen genoemd omtrent het gebruik van OAI-PMH. Ten eerste ontstaan bij het toepassen van OAI-PMH feitelijk kopieën van metadata die heen en weer worden verplaatst tussen organisaties. De actualiteit en het eigenaarschap van deze kopieën kunnen vervolgens niet gegarandeerd worden. Daarbij wordt de gegevensverrijking die mogelijk plaatsvindt door uitwisseling, niet bij de bron opgeslagen waardoor deze verrijking niet structureel wordt. Hierdoor ontstaat een situatie waar op meerdere plaatsen kopieën van dezelfde gegevens bijgehouden worden.

Verder geeft OAI-PMH als protocol slechts beperkt houvast voor de inhoudelijke metadata die het uitwisselt. OAI-PMH wordt geïmplementeerd met verschillende doelformaten van data, bijvoorbeeld Dublin Core (de minimumstandaard) EDM of A2A. Uitwisseling van gegevens tussen platformen met dergelijke verschillende implementaties leidt tot verlies van informatie. De interpretatie van gegevens kunnen namelijk verloren gaan doordat dezelfde gegevens andere benaming kennen tussen formaten.

Om dit recht te trekken zijn handmatige aanpassingen benodigd, zogeheten 'mappings'. Dit is niet zozeer een tekortkoming van OAI-PMH, eerder een resultaat van de openheid van het protocol waar bij implementaties gebruik kan worden gemaakt van verschillende uitwisselformaten/doelformaten.

Ontwikkelingen in de erfgoedsector

In de laatste jaren is er volgens de betrokken experts in de erfgoedsector sprake van een trend waarbij de aandacht voor 'rijkere uitwisseling' van gegevens van instellingen toeneemt. Waar voorheen werd volstaan met een OAI-PMH implementatie om gegevens beschikbaar te stellen aan de buitenwereld wordt er momenteel in toenemende mate gekozen voor uitgebreidere opties. Het gaat hier bijvoorbeeld om specifieke API's

waarmee gegevens uitgebreider en specifiekere benaderbaar worden voor een breder publiek.

OAI-PMH is als protocol niet ontwikkeld voor dergelijke toepassingen, maar de standaard wordt hier desondanks soms wel voor ingezet. Bij uitwisseling met eindgebruikers loopt men al snel tegen beperkingen aan, met name doordat OAI-PMH geen resultaten geeft in formaten die gangbaar zijn voor een doorsnee eindgebruiker.

Verder gaan via OAI-PMH de gegevens vaak door meerdere aggregatieplatforms voordat een eindgebruiker ze raadpleegt. Als eigenaar van de oorspronkelijke bron van de gegevens heb je hierdoor weinig grip op hoe deze gegevens uiteindelijk worden aangeboden. Hierdoor ontstaat een risico dat de uiteindelijke gebruiker een suboptimale en onvoorspelbare gebruikservaring kent.

Mede door deze ervaringen met het toepassen van OAI-PMH en de ontwikkelingen die plaatsvinden op het gebied van de eisen die worden gesteld aan gegevensuitwisseling kijkt men in de erfgoedsector naar alternatieven voor OAI-PMH.

Als belangrijkste alternatief, ook al is er niet zozeer sprake van een 1-op-1 vervanging van de functionele toepassing, wordt Linked (Open) Data genoemd. Hiermee kunnen gegevens bij de bron blijven om van daaruit opengesteld te worden voor derden. Linked (Open) Data heeft de potentie om knelpunten die men nu ervaart bij het toepassen van OAI-PMH weg te nemen.

Wel is het algemene beeld zo dat dit nog een geruime tijd zal duren voordat men daadwerkelijk op grote schaal implementaties van Linked Data in de erfgoedsector kan verwachten. Hier worden diverse redenen voor aangewezen. Zo wordt de erfgoedsector niet beschouwd als heel revolutionair ten aanzien van het inzetten van technologie, en zijn er twijfels over de volwassenheid van Linked Data om in de praktijk al toegepast te kunnen worden. Over het algemeen is de verwachting dat er eerst diverse succesvolle praktijkimplementaties van Linked Data nodig zullen zijn om de adoptie echt aan te wakkeren.

Verwacht wordt dat OAI-PMH de komende jaren nog een de facto-standaard zal blijven binnen de erfgoedsector. Ook bij opkomst van een alternatief verwachten de meeste experts dat er nog altijd een (zij het op kleinere schaal dan op dit moment) plek zal zijn voor OAI-PMH toepassingen. Een toepassing die hiervoor wordt genoemd is bijvoorbeeld het intern bijeen verzamelen van gegevens van meerdere collecties alvorens deze via een protocol als Linked Data aan te bieden naar de buitenwereld.

Gebruik van OAI-PMH binnen het onderwijsveld

OAI-PMH is in 2010 door Kennisnet aangemeld voor de 'pas toe of leg uit'-lijst, de lijst met verplichte open standaarden. De motivatie was dat om educatieve content vindbaar te maken, het nodig is deze content te voorzien van metadata die centraal beschikbaar gesteld wordt. Er is een centrale repository gecreëerd waarin alle metadata en verwijzingen naar educatieve content vanuit diverse bronnen worden opgeslagen. Gebruikers (m.n. docenten) hoeven maar in één repository te kijken als ze

materiaal te zoeken. Plaatsing op de lijst met de verplichte status was een middel om de belangstelling, het draagvlak en het aantal gebruikers te vergroten.

Binnen het onderwijsveld heeft OAI-PMH de gewenste positie ingenomen, de standaard heeft aantoonbare meerwaarde en wordt breed gebruikt. Er is voor het onderwijsveld een afspraak gebaseerd op het OAI-protocol opgesteld, die wordt beheerd door Kennisnet (Platform Edustandaard). Kennisnet biedt handleidingen en verwijzingen voor implementatie en gebruik van OAI-PMH. Een beheer-rol voor Kennisnet buiten het onderwijsdomein past niet bij de missie van Kennisnet: 'Laat ict werken voor het onderwijs'.

Kennisnet kan zich in dit stadium goed vinden in het veranderen van de status van het protocol OAI-PMH van 'verplicht' naar 'aanbevolen'. Mede omdat Kennisnet ook zelf toepassing van gebruikte standaarden evalueert. Werk van het Forum, is voor zo ver dat het onderwijsveld betreft, daarmee dubbel werk. Stimuleren van adoptie is voor het onderwijsveld inmiddels niet meer nodig.

Conclusie

Geadviseerd wordt om de status van het OAI-PMH protocol op de lijst open standaarden te wijzigen van 'verplicht', naar 'aanbevolen'. Dit sluit beter aan bij de ervaringen met het gebruik, recente ontwikkelingen, verwachtingen voor de toekomst en de huidige adoptiegraad.

OAI-PMH is een de facto-standaard binnen de erfgoedsector waarmee een belangrijk motief voor een 'verplichting' is komen te vervallen. Verder is de toenemende vraag naar 'rijkere' gegevensuitwisseling een aanjager voor het toepassen van andere oplossingen naast OAI-PMH of ter vervanging van OAI-PMH. Desondanks blijft een specifieke functionele toepassing van OAI-PMH binnen de erfgoedsector waardevol en breed geadopteerd.

In grote lijnen gelden dezelfde argumenten voor gebruik in het onderwijsveld. Ook hier is de adoptiegraad van OAI-PMH hoog en heeft het een aantoonbare waarde. Een status 'aanbevolen' is op basis van deze argumenten in de huidige situatie het meest passend.

Aanvullende adviezen

In het evaluatie advies OAI-PMH van 2016 is voorgesteld om het functioneel toepassingsgebied als volgt aan te passen:

"Het vraaggestuurd aanbieden en ophalen van verzamelingen metadata uit bibliotheken met (digitale) documenten of andere objecten gericht op leermaterialen en kunsthistorische objecten, met als doel het opnemen van deze metadata in een centrale bibliotheek."

Tijdens het onderzoek binnen de erfgoedsector is dit voorstel ook besproken. De experts zijn van mening dat het functioneel toepassingsgebied juist is geformuleerd. Wel plaatsen zij twee kanttekeningen:

1. De formulering 'kunsthistorische objecten' is een te nauwe beschrijving. Een aanzienlijk aantal repositories (bibliotheken) bevatten gegevens die binnen het cultureel erfgoed domein vallen en worden uitgewisseld op basis van OAI-PMH. Hierbij kan bijvoorbeeld worden gedacht aan stamboomgegevens, verklarende "woordenlijsten" of tijdschrift publicaties. De definitie kunsthistorische objecten suggereert dat dergelijke repositories niet in het functioneel toepassingsgebied vallen, terwijl de praktijk anders uitwijst.
2. De formulering van het doel van de toepassing is niet volledig. Hoewel het *opnemen van metadata in een centrale bibliotheek* als een belangrijke functie wordt beschouwd, namelijk de functie voor aggregatieplatformen wordt OAI-PMH ook veel toegepast tussen onderlinge repositories die elkaars gegevens verrijken zonder dat een van deze repositories te kenmerken valt als een 'centrale bibliotheek'.

Op basis van deze kanttekeningen wordt geadviseerd om het functioneel toepassingsgebied als volgt te formuleren:

"Het vraaggestuurd aanbieden en ophalen van verzamelingen metadata uit bibliotheken met (digitale) documenten of andere objecten gericht op leermaterialen en cultureel erfgoed, met als doel het opnemen van deze metadata in een centrale bibliotheek en/of het verrijken van eigen gegevens.

Tot slot verdient het aanbeveling om bij de opname op de lijst nadere informatie te geven over het implementeren van OAI-PMH in de erfgoedsector. Hierbij kan worden gedacht aan een bepaalde "minimum implementatie" die wordt aanbevolen. Specifiek kan het hier bijvoorbeeld gaan om welke methoden beschikbaar gesteld dienen te worden. Dit kan mogelijk een bijdrage leveren aan verdere standaardisatie van OAI-PMH implementaties in de erfgoedsector. Zoals eerder aangegeven zorgt het 'open' karakter van OAI-PMH voor diverse implementaties die de standaardiserende werking verminderen.

Referenties

<https://www.forumstandaardisatie.nl/sites/bfs/files/proceedings/FS%20161214.3A%20Evaluatie%20advies%20OAI-PMH.pdf>