



Forum Standaardisatie

Expertadvies

Open Archives Initiative -  
Protocol for Metadata Harvesting (OAI-PMH)  
versie 2.0

Datum 3 februari 2010

## Colofon

Projectnaam	Expertadvies Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) versie 2.0
Versienummer	1.0 Definitief
Locatie	-
Organisatie	Forum Standaardisatie Postbus 84011 2508 AA Den Haag forumstandaardisatie@logius.nl
Auteurs	prof. dr. T.W.C. Huibers MMC, ir. L.M. Punter

## Inhoud

<b>Colofon</b> .....	<b>2</b>
<b>Inhoud</b> .....	<b>3</b>
<b>Managementsamenvatting</b> .....	<b>4</b>
<b>1 Inleiding</b> .....	<b>5</b>
1.1 <i>Doelstelling expertadvies en achtergrond</i> .....	5
1.2 <i>Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH)</i> .....	5
1.2.1 <i>Open Archives Initiative</i> .....	5
1.2.2 <i>De standaard OAI-PMH – Protocol for Metadata Harvesting</i> .....	6
1.2.3 <i>Relatie met andere standaarden</i> .....	8
1.3 <i>Opdracht en het gevolgde proces</i> .....	8
1.4 <i>Samenstelling expertgroep</i> .....	9
<b>2 Toepassings- en werkingsgebied</b> .....	<b>11</b>
2.1 <i>Versie van de standaard</i> .....	11
2.2 <i>Functioneel toepassingsgebied</i> .....	11
2.3 <i>Organisatorisch werkingsgebied</i> .....	12
<b>3 Toetsingscriteria</b> .....	<b>13</b>
3.1 <i>Openheid</i> .....	13
3.1.1 <i>Besluitvorming</i> .....	13
3.1.2 <i>Beschikbaarheid</i> .....	13
3.1.3 <i>Intellectueel eigendom</i> .....	14
3.1.4 <i>Hergebruik</i> .....	14
3.2 <i>Bruikbaarheid</i> .....	14
3.2.1 <i>Volwassenheid</i> .....	14
3.2.2 <i>Functionaliteit</i> .....	15
3.2.3 <i>Concurrerende standaarden</i> .....	16
3.3 <i>Potentieel</i> .....	17
3.4 <i>Impact</i> .....	17
<b>4 Advies aan Forum en College</b> .....	<b>19</b>
4.1 <i>Samenvatting van de toetsingscriteria</i> .....	19
4.2 <i>Advies</i> .....	19

## Managementsamenvatting

Dit rapport bevat het advies van de expertgroep OAI-PMH aan het Forum en College Standaardisatie over het opnemen van de OAI-PMH versie 2.0 op de lijst met open standaarden voor "pas toe of leg uit".

OAI-PMH is een standaard voor harvesting van metadata uit repositories. Met metadata worden kenmerken van en extra opgeslagen informatie over een document of ander object bedoeld. Te denken valt aan auteursgegevens, titel, uitgever, taal, etc. Een repository is een bibliotheek met documenten/objecten (ook wel 'content' genoemd), bijvoorbeeld een (digitaal) archief.

OAI-PMH maakt het mogelijk om deze metadata (dus niet de documenten/objecten zelf) uit verschillende repositories te verzamelen. Vanuit een centraal systeem kan dan gezocht worden naar documenten/objecten in de verschillende aangesloten repositories. Als functioneel toepassingsgebied wordt voorgesteld: *"Harvesting (het volledig verzamelen en uitwisselen) van metadata van in repositories opgenomen (digitale) documenten of objecten."*

Met betrekking tot de toetsingscriteria concludeert de expertgroep dat:

- De standaard voldoende open is. De standaard wordt beheerd door het Open Archives Initiative, die partijen de mogelijkheid biedt om inspraak te hebben.
- De standaard voldoende bruikbaar is. De standaard biedt de gewenste functionaliteit binnen het gedefinieerde functioneel toepassingsgebied en wordt als zodanig ook geadviseerd door Digitaal Erfgoed Nederland (DEN). De standaard is bovendien zeer stabiel – de laatste grote wijziging dateert van 2002. Binnen het functioneel toepassingsgebied zijn er geen direct concurrerende standaarden. Wel zijn er momenteel andere standaarden in gebruik die in de toekomst mogelijk vervangen kunnen worden door OAI-PMH.
- De standaard heeft potentieel om bij te dragen aan interoperabiliteit en leveranciersonafhankelijkheid. Dit heeft betrekking op de werking van contentketens (partijen die content met elkaar delen) en de onderliggende technologie. De afhankelijkheden verminderen en de interoperabiliteit neemt toe.
- Hoewel het delen van gegevens uiteraard impact heeft op de bedrijfsvoering en informatievoorziening, zijn er geen specifieke risico's verbonden aan toepassing van OAI-PMH.

OAI-PMH is een generieke standaard. Als organisatorisch werkingsgebied wordt daarom voorgesteld: *"Overheden en instellingen uit de (semi-) publieke sector."* Binnen sectoren kunnen daarnaast specifieke afspraken nodig zijn m.b.t. de uit te wisselen metadata. OAI-PMH doet hier geen uitspraken over. Een voorbeeld van een dergelijke afspraak is het Content Zoekprofiel (dat op dit moment ook in behandeling is voor opname op de lijst) en de Overheid Web Metadata Standaarden.

De expertgroep adviseert OAI-PMH versie 2.0 op te nemen op de lijst met open standaarden voor "pas toe of leg uit".

## 1 Inleiding

### 1.1 Doelstelling expertadvies en achtergrond

De staatssecretaris van Economische Zaken heeft op maandag 17 september 2007 het actieplan 'Nederland Open in Verbinding' (NOiV) aan de Tweede Kamer gestuurd. Het doel van het actieplan is om de informatievoorziening toegankelijker te maken, onafhankelijkheid van ICT-leveranciers te creëren en de weg vrij te maken voor innovatie. Kernpunten in het actieplan zijn het stimuleren van open source software en het adopteren van open standaarden.

Een van de acties uit het plan is het opstellen van een lijst met open standaarden, die vallen onder het principe 'comply or explain' ('pas toe of leg uit'). Overheden en publieke organisaties worden op basis van dit principe gestimuleerd standaarden te gebruiken die op de lijst staan. Hiermee wil de overheid de adoptie van open standaarden versnellen en de interoperabiliteit tussen overheidsorganisaties, bedrijven en burgers verbeteren.

Het College Standaardisatie besluit over de standaarden die op de lijst zullen worden opgenomen. Bij dit besluit wordt gebruik gemaakt van een toetsing van de standaard door experts en een publieke consultatie. Dit document beschrijft de beoordeling en het advies van de expertgroep die het Open Archives Initiative - Protocol for Metadata Harvesting v2.0 (afgekort: OAI-PMH) [1] heeft getoetst.

### 1.2 Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH)

#### 1.2.1 *Open Archives Initiative*

OAI is een samenwerkingsverband van een aantal personen, afkomstig uit de academische wereld. OAI ontwikkelt en promoot standaarden en interoperabiliteitsinitiatieven voor het vereenvoudigd verspreiden en uitwisselen van (digitale) informatie. Het samenwerkingsverband komt voort uit de wens om e-print archieven en onderzoeksinformatie beter te kunnen delen.

De toepassing van de standaarden wordt in toenemende mate verbreed naar niet-academische toepassingen. De fundamentele concepten en technologieën hebben een bredere relevantie en kunnen ook toegepast worden op ander digitaal materiaal dat in online bibliotheken is opgenomen.

---

1 Zie: <http://www.openarchives.org/OAI/openarchivesprotocol.html>

Op dit moment beheert het samenwerkingsverband twee standaarden:

- OAI-PMH, onderwerp van deze toetsingsprocedure. Deze standaard richt zich op het uitwisselen van metadata tussen repositories (digitale bibliotheken/databases met documenten/objecten).
- OAI-ORE, Object Reuse and Exchange. Deze standaard richt zich op het uitwisselen van de digitale objecten tussen repositories.

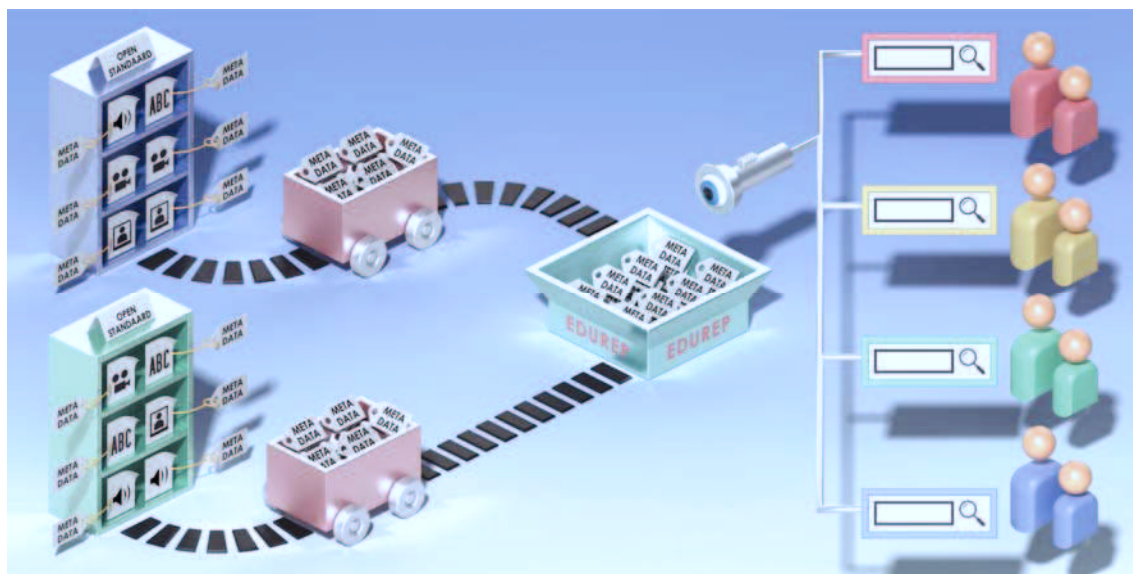
De kern van het samenwerkingsverband wordt gevormd door een executive en technical committee. Het executive committee bestaat uit twee personen van respectievelijk Cornell University en Los Alamos National Laboratory. Het technical committee bestaat uit diverse personen uit de Amerikaanse en Europese academische wereld (o.a. British Library, NASA, Library of Congress, MIT, CERN).

Financiering van het samenwerkingsverband komt hoofdzakelijk van de Andrew W. Mellon Foundation ([www.mellon.org](http://www.mellon.org)), de Coalition for Networked Information, de Digital Library Federation en de National Science Foundation. Het gaat hier in alle gevallen om (Amerikaanse) publieke instellingen. Daarnaast wordt er bijgedragen door Microsoft.

### 1.2.2

#### *De standaard OAI-PMH – Protocol for Metadata Harvesting*

In onderstaande figuur wordt grafisch weergegeven welke mechanismen ingevuld worden met OAI-PMH.



**Figuur 1 Werking OAI-PMH (bron: Afspraak Metadata Harvesting, Kennisnet/Edustandaard)**

In de figuur staan links een tweetal bronbibliotheken (repositories) met hierin digitale objecten. Dit kunnen documenten zijn, maar ook multimediaobjecten of andere digitale bestanden. Van ieder object is in iedere bibliotheek metadata beschikbaar – gegevens die de karakteristieken van het object beschrijven, vergelijkbaar met een indexkaart in de bibliotheek. Metadata kan bijvoorbeeld de titel en auteur van een object omvatten, maar bijvoorbeeld ook de gebruiksrechten of publicatiedatum.

Voor derde partijen (rechts in het figuur) kan het relevant zijn om te zoeken in meerdere digitale bibliotheken. Bijvoorbeeld om objecten te kunnen zoeken die aanwezig kunnen zijn bij meerdere organisaties / uitgevers. In plaats van te zoeken in de verschillende bibliotheken wil men de zoekvraag op één plaats kunnen uitzetten. In het voorbeeld fungeert Edurep als centrale bibliotheek (repository) met verwijzingen naar educatief materiaal. Edurep wordt beheerd door Kennisnet.

Opgemerkt moet worden dat de digitale documenten zelf niet zijn ondergebracht in deze centrale bibliotheek (repository). Slechts de metadata uit de achterliggende bronbibliotheken is opgenomen. Gebruikers kunnen dus – bij wijze van spreken – zoeken in de indexkaarten van meerdere bibliotheken. Als zij vervolgens het juiste document hebben gevonden verwijst de centrale repository (bibliotheek) hen door naar de bronbibliotheek, waar het document kan worden opgehaald.

**PMH is het protocol dat wordt gebruikt voor het verzamelen (harvesting) van metadata uit bibliotheken/repositories.** *In de figuur wordt het weergegeven als de rails en de karretjes tussen deze bibliotheken.*

De volgende bevestigingen zijn mogelijk:

- **Identify:** wie bent u? welke informatie kan ik bij u ophalen?
- **ListMetaDataFormats:** welke (metadata-) taal spreekt u?
- **ListSets:** welke rubrieken heeft u?
- **GetRecord:** geef één specifiek metadata item
- **ListRecords:** geef meerdere metadata items (die voldaan aan een bepaald criterium)
- **ListIdentifiers:** geef meerdere item identifiers (die voldoen aan een bepaald criterium)

Er zijn ook een aantal zaken die OAI-PMH nadrukkelijk *niet* regelt:

- Het soort metadata dat verzonden wordt. Hoewel minimaal het (generieke) Dublin Core moet worden ondersteund, kan ieder (XML-gebaseerde) vorm van metadata worden uitgewisseld. Denk aan een formaat dat specifiek is voor een bepaalde sector. In het primair en voortgezet onderwijs wordt bijvoorbeeld het Content Zoekprofiel toegepast;
- De uitwisseling van documenten zelf;
- De interne verwerking van de metadata.

Dit alles maakt OAI-PMH tot een bijzonder generieke standaard, die vele toepassingsmogelijkheden kent. Hoewel initieel de standaard is opgezet vanuit de wens om gegevens tussen academische verzamelingen te kunnen delen, wordt de standaard in Nederland ook binnen veel andere domeinen toegepast. Voorbeelden zijn de educatieve sector en de erfgoed sector. Zo heeft DEN (Digitaal Erfgoed Nederland) OAI-PMH opgenomen als standaard op 'De Basis', een lijst met standaarden voor de erfgoedsector [2].

---

2 <http://matrix.den.nl/row.aspx?matrixid=register&rowid=40>

### 1.2.3 *Relatie met andere standaarden*

OAI-PMH refereert niet aan andere standaarden op de lijst voor pas toe-of-leg uit.

Een aantal standaarden uit de lijst met veelgebruikte standaarden wordt toegepast binnen OAI-PMH. Het gaat dan om: XML, URI, URL, ISO 8601, HTTP, UTF-8.

Andere gerefereerde standaarden zijn: XML namespace, XML schema, Dublin core, MARC en RFC1807.

## 1.3 **Opdracht en het gevolgde proces**

De expertgroep heeft niet als opdracht gehad het ontwikkel- en standaardisatieproces over te doen. Het doel van de expertgroep was te toetsen of de standaard voldoet aan een vastgestelde set van toetsingscriteria.

De toetsingscriteria zijn afkomstig uit het rapport "Open standaarden: het proces om te komen tot een lijst met open standaarden", dat op 14 mei 2008 is vastgesteld door het College Standaardisatie.

De opdracht aan de expertgroep was dan ook om te toetsen in hoeverre de standaard voldoet aan deze criteria en op basis hiervan te adviseren of de standaard – al dan niet onder bepaalde voorwaarden – op de lijst met open standaarden voor "pas toe of leg uit" opgenomen moet worden.

De standaard is aangemeld door Stichting Kennisnet, als 'Afspraak Metadata Harvesting'. Dit is een standaard die t.b.v. het primair en voortgezet onderwijs wordt ingezet in de educatieve sector. De 'Afspraak Metadata Harvesting' definieert twee zaken:

- Het gebruik van OAI-PMH voor de harvesting van metadata;
- Het gebruik van de Content Zoekprofiel-standaard voor het vastleggen van metadata.

-  
In overleg met het Bureau Forum Standaardisatie zijn een voorzitter en mogelijke leden voor de expertgroep aangezocht. Een uitnodiging naar een eerste groep mensen begin 2009 leidde helaas niet tot voldoende aanmeldingen voor deelname aan de expertgroep.

Medio 2009 is daarom een sessie belegd met de indiener, Kennisnet, de voorzitter en Surf Foundation. Surf Foundation is een belangrijke partij in de educatieve sector, die ook OAI-PMH inzet als standaard. Tijdens deze sessie is besloten in te zetten op OAI-PMH als te behandelen standaard en het functioneel toepassingsgebied duidelijker te definiëren. Hiervoor is gekozen omdat Content Zoekprofiel al in een eerder stadium is behandeld in een expertgroep (en thans nog in procedure is) als standaard voor de lijst. Als toegevoegde standaard in de 'Afspraak Metadata Harvesting' blijft dan alleen OAI-PMH over. Deze standaard kent potentieel een generieker functioneel toepassingsgebied (niet alleen maar harvesting van educatieve content) en een breder organisatorisch werkingsgebied (niet alleen maar primair/voortgezet onderwijs) dan de 'Afspraak Metadata Harvesting'.



Naar aanleiding van de sessie is een nieuwe groep mensen uitgenodigd. Naast partijen uit de educatieve sector zijn ook partijen uit de erfgoed- en archiefsector benaderd om deel te nemen aan de expertsessie. Uiteindelijk heeft de expertsessie plaatsgevonden op dinsdag 8 december 2009. De resultaten van deze expertsessie zijn vastgelegd in dit document.

#### **1.4 Samenstelling expertgroep**

Voor de expertgroep zijn personen uitgenodigd die vanuit hun persoonlijke expertise of werkzaamheden bij een bepaalde organisatie direct of indirect betrokken zijn bij de standaard. Daarnaast is een onafhankelijke voorzitter aangezocht om de expertgroep te leiden en als verantwoordelijke voor het uiteindelijke expertadvies.

Als voorzitter is gekozen prof. dr. Theo Huibers MMC. Theo Huibers is als deeltijd hoogleraar Information Retrieval verbonden aan de Human Media Interaction Groep van de Universiteit Twente. Zijn onderzoek richt zich op de werking van information retrieval systemen, de mogelijke toepassingen en de impact voor organisaties. Daarnaast werkt hij voor Thaesis, een adviesbureau op het gebied van organisatie-inrichtingsvraagstukken en optimalisering van waardeketens. Als adviseur is hij gesprekspartner voor snel veranderende organisaties zoals uitgeverijen en IT-dienstverleners.

Bij de bijeenkomst waren, naast de voorzitter, de volgende experts aanwezig:

- Jeroen Hamers, expert onderwijsstandaarden, namens Kennisnet (indiener)
- Aard Goedewagen, adviseur digitale duurzaamheid, namens NEN en Justitie;
- Hub Kockelkorn, projectleider ICT bij Museon;
- Martin Gilissen, hoofd I&A bij Ons Middelbaar Onderwijs, een vereniging van in totaal 33 scholen voor voortgezet onderwijs in Noord-Brabant;
- Pierre van Rooden, namens Teleac;
- Jelle van der Meulen, technisch specialist bij Ultraware (softwarebedrijf), mede namens uitgeverij Iddink en Bloqs (educatieve uitgeverijen);
- Edwin Verwoerd, business consultant bij ThiemeMeulenhoff, mede namens de Groep Educatieve Uitgeverijen (GEU);
- Frans Berkhof, namens SLO.

Naast de experts waren de volgende personen aanwezig:

- Bart Knubben, namens het Bureau Forum Standaardisatie
- Matthijs Punter, procesbegeleider (TNO)

Verschillende personen, die zijn uitgenodigd uitgenodigd, konden om diverse redenen niet aanwezig konden zijn bij de bijeenkomst van de expertgroep. Zij zijn gevraagd voorafgaand aan de sessie eventuele discussiepunten aan te leveren. Voor zover deze zijn doorgegeven, zijn deze meegenomen in de discussie tijdens de bijeenkomst van de expertgroep.

Tenslotte zijn rondom een aantal toetsingscriteria vragen gesteld aan Herbert van de Sompel, een Vlaming die werkzaam is bij het Los Alamos National Laboratory en lid is van het executive committee van OAI-PMH. De antwoorden op deze vragen zijn betrokken bij dit expertadvies.

## 2 Toepassings- en werkingsgebied

### 2.1 Versie van de standaard

De exacte versie van de standaard waar de expertgroep zich op gericht heeft is 'The Open Archives Initiative Protocol for Metadata Harvesting' versie 2.0 (14 juni 2002; beschikbaar via <http://www.openarchives.org/OAI/openarchivesprotocol.html>), in dit document verder afgekort tot OAI-PMH.

De expertgroep heeft bepaald voor welke toepassingen OAI-PMH in principe gebruikt moet worden (functioneel toepassingsgebied) en welke organisaties moeten voldoen aan OAI-PMH binnen het functioneel toepassingsgebied (organisatorisch werkingsgebied).

### 2.2 Functioneel toepassingsgebied

Het functioneel toepassingsgebied van OAI-PMH is door de expertgroep als volgt bepaald:

**Harvesting (het volledig verzamelen en uitwisselen) van  
metadata van in repositories opgenomen (digitale)  
documenten of objecten**

Dit is een breed functioneel toepassingsgebied: er is geen afbakening m.b.t. de soort metadata die uitgewisseld wordt of de aard van de digitale objecten (anders dan dat ze in bibliotheken/repositories zijn opgenomen). Het kan hierbij dus gaan om e-books of e-papers die zijn opgenomen in een bestand van een academische bibliotheek, maar ook om vergunningen die opgenomen zijn in een database van een gemeente. Per sector zijn dan aanvullende afspraken nodig, bijvoorbeeld m.b.t. het toegepaste metadata formaat. OAI-PMH wordt in de praktijk vooral ingezet voor publieke informatie, hoewel ook uitwisseling van niet-publieke informatie mogelijk is. Dan zijn doorgaans aanvullende afspraken nodig over toegang, autorisatie, beveiliging, etc.

Belangrijke overwegingen bij het functioneel toepassingsgebied zijn:

- Er is vastgesteld dat het gaat om *harvesting*. Dit gaat verder dan het opvragen van een enkel metadata element. Het doel van *harvesting* is het volledig verzamelen en uitwisselen van metadata.
- Het gaat om *harvesting van metadata* – en daarmee dus niet om de documenten of objecten zelf.
- Deze metadata moet opgenomen zijn in een bibliotheek of *repository*. Er moet dus sprake zijn van een bepaalde collectie of verzameling van objecten.
- De nadruk ligt op publiek toegankelijke digitale documenten. Tegelijkertijd kan er ook metadata zijn van analoge documenten of objecten, zoals schilderijen of boeken. Hoewel dit niet uitgesloten wordt door het functioneel toepassingsgebied ligt de focus wel op digitale documenten of objecten. Daarom is dit tussen haakjes opgenomen.

## 2.3 Organisatorisch werkingsgebied

In relatie tot het hiervoor gekozen functioneel toepassingsgebied heeft de expertgroep het volgende organisatorische werkingsgebied gekozen:

<b>Overheden en instellingen uit de (semi-) publieke sector</b>
---

In relatie tot het gekozen functioneel toepassingsgebied is er volgens de expertgroep geen nadere afbakening van het organisatorisch werkingsgebied noodzakelijk. Het organisatorisch werkingsgebied is hiermee gelijk aan alle organisaties waarop het pas toe of leg uit principe van toepassing is.

Belangrijke (deels overlappende) sectoren waar gebruik van de standaard nu al plaats vindt zijn:

- Archiefsector (Nationaal archief, gemeentelijke archieven, etc.)
- Bibliotheekwereld (Koninklijke Bibliotheek, academische bibliotheken, etc.)
- Erfgoedsector (musea, collecties)
- Educatieve sector (primair, voortgezet en hoger onderwijs)

Dit is geen uitputtende lijst. Ook in andere situaties waar harvesting van metadata plaats vindt zou naar mening van de expertgroep in de toekomst OAI-PMH ingezet kunnen worden.

Dit laat onverlet dat – voor gebruik binnen specifieke sectoren – aanvullende sectorspecifieke afspraken nodig zijn over het uitgewisselde metadataformaat. OAI-PMH doet hier geen uitspraken over, anders dan dat minimaal een beperkte vorm van Dublin Core moet worden ondersteund.

## 3 Toetsingscriteria

Door de expertgroep is de standaard getoetst aan de vier door het College Standaardisatie vastgestelde toetsingscriteria. In dit hoofdstuk worden de criteria en de bijbehorende toetsingsaspecten weergegeven (cursief), evenals de mening van de expertgroep op deze punten.

### 3.1 Openheid

#### 3.1.1 *Besluitvorming*

*De standaard is goedgekeurd en zal worden gehandhaafd door een non-profit organisatie. De lopende ontwikkeling gebeurt op basis van een open besluitvormingsprocedure die toegankelijk is voor alle belanghebbende partijen (consensus of meerderheidsbeschikking enz.);*

Besluitvorming over de standaard vindt plaats bij het Open Archives Initiative. Dit is een non-profit samenwerkingsverband van een aantal academici. Er is een online community/maillinglist waar mensen zich bij kunnen aanmelden. Deelnemers daaraan zijn afkomstig uit diverse landen.

Uiteindelijke besluitvorming loopt via het *technical committee* en het *executive committee*. Deelnemers aan deze organen komen uit de non-profit / academische wereld.

#### 3.1.2 *Beschikbaarheid*

*De standaard is gepubliceerd en over het specificatiedocument van de standaard kan vrijelijk worden beschikt of het is te verkrijgen tegen een nominale bijdrage. Het moet voor een ieder mogelijk zijn om het te kopiëren, beschikbaar te stellen en te gebruiken om niet of tegen een nominale prijs.*

De standaard is gepubliceerd op de website van het Open Archives Initiative en kan vrijelijk, kosteloos worden gebruikt.

Daarnaast wordt de standaard beschikbaar gesteld via organisaties die een afgeleide sectorstandaard gebruiken, waarin OAI-PMH is opgenomen. De Afspraak Metadata Harvesting van Kennisnet/Edustandaard is een voorbeeld van een dergelijke standaard.

### 3.1.3 *Intellectueel eigendom*

*Het intellectuele eigendom – met betrekking tot mogelijk aanwezige patenten – van (delen) van de standaard is onherroepelijk ter beschikking gesteld op een "royalty-free" basis.*

Er zijn (voorzover bekend) geen intellectueel eigendomsrechtelijke claims (zoals octrooien) die het gebruik van OAI-PMH beperken.

### 3.1.4 *Hergebruik*

*Er zijn geen beperkingen omtrent het hergebruik van de standaard*

Door het Open Archives Initiative worden er geen beperkingen gesteld aan het hergebruik van de standaard. Eerder is het omgekeerde het geval: het OAI heeft – volgens haar website – als missie het gebruik van dergelijke standaarden te stimuleren.

## **3.2 Bruikbaarheid**

### 3.2.1 *Volwassenheid*

*Is de standaard voldoende uitgekristalliseerd?*

De laatste revisie van de standaard dateert van 2002. Sinds die tijd zijn er geen grote wijzigingen meer geweest. De wijzigingen die er waren, hadden betrekking op verbeteringen in de formuleringen of nadere toelichtingen in de diverse ondersteunende documentatie.

Door de expertgroep wordt aangegeven dat de standaard de gewenste functionaliteit goed afdekt. Het is dan ook de verwachting dat in de nabije toekomst er geen grote wijzigingen zullen zijn.

*Is verdere ontwikkeling en het onderhoud van de standaard verzekerd?*

De verdere ontwikkeling en het onderhoud worden belegd bij het Open Archives Initiative. Hier staan een aantal grote (Amerikaanse) publieke instellingen achter. Daarnaast wordt de standaard ondersteund door tal van organisaties, die samenkomen in de community. Hiermee is het onderhoud in voldoende mate verzekerd.

*Is er een methode waarmee conformiteit aan de standaard bepaald kan worden?*

Via de website van het Open Archives Initiative wordt een testtool aangeboden (<http://www.openarchives.org/pmh/tools>). Hiermee kan conformiteit aan het protocol worden vastgesteld.

De expertgroep wijst er wel op dat in de praktijk aanvullende (functionele) tests nodig zijn. Daarbij gaat het dan bijvoorbeeld om de juiste ondersteuning van sectorspecifieke metadataformaten, de volledigheid van de aangeboden bestanden en de inhoudelijke juistheid van de aangeboden metadata. Dit wordt uiteraard niet getoetst door de generieke testtool.

*Is er voldoende praktijkervaring met het gebruik van de standaard?*

Er is ruime praktijkervaring met OAI-PMH. Via de website van het Open Archives Initiative is een lange lijst beschikbaar van organisaties die het protocol implementeren in hun bibliotheek of zoekstelsel. In Nederland wordt de standaard gepromoot via het DEN en zijn er diverse implementaties, zowel in de archief-/erfgoedwereld als in de educatieve sector.

*Is er nu en in de toekomst voldoende ondersteuning door (meerdere) marktpartijen voor de standaard?*

Dit is wel de verwachting:

- Enerzijds is er nu in de praktijk al ondersteuning door verschillende marktpartijen.
- Anderzijds is de standaard eenvoudig van opzet en gemakkelijk op te pakken door marktpartijen die nog geen eerdere ervaring hebben met het protocol.

Er is tevens een open source implementatie van de standaard beschikbaar<sup>3</sup>.

*Is de verwachting van het toekomstig gebruik van de standaard positief?*

In toenemende mate zal content ontsloten moeten worden. Musea, archieven en erfgoedinstellingen spelen daarbij in op hun veranderende rol. Daarnaast zijn er ook buiten deze context toepassingen te verwachten waarvoor harvesting van metadata een belangrijke rol speelt, zoals de transparantieprojecten bij de overheid en educatieve toepassingen (zoals Edurep).

Deze ontwikkelingen vragen om standaarden voor uitwisseling. OAI-PMH is volgens de expertgroep een voor de hand liggende open standaard. Daarmee is de verwachting voor het toekomstig gebruik van OAI-PMH dan ook positief.

### 3.2.2 *Functionaliteit*

*Voldoet de standaard aan de functionele eisen die aan de werking van de standaard gesteld worden binnen het voorgestelde toepassingsgebied?*

Naar mening van de expertgroep voldoet de standaard ruimschoots aan de functionele eisen die voor het voorgestelde functioneel toepassingsgebied worden gesteld. Een belangrijk gegeven hierbij is dat de standaard sinds 2002 niet meer ingrijpend gewijzigd is en in de praktijk voor veel toepassingen goed blijkt te voldoen.

---

<sup>3</sup> zie <http://www.openarchives.org/pmh/tools/tools.php>

### 3.2.3 Concurrerende standaarden

*Verhoudt de standaard zich goed ten opzichte van eventuele concurrerende standaarden?*

In de expertgroep is gekeken naar drie soorten standaarden die een relatie hebben met OAI-PMH:

- **RSS / Atom**

Deze standaarden worden vooral gebruikt voor syndicatie van websites (en andere content verzamelingen). Potentieel heeft dit een overlapping met OAI-PMH.

In de praktijk vullen beide standaarden elkaar echter goed aan:

- RSS/Atom wordt voornamelijk gebruikt voor *syndicatie* van content ten behoeve van eindgebruikerstoepassingen. Voorbeelden zijn het gebruik van 'feeds' van weblogs of websites.
- OAI-PMH wordt gebruikt voor (meer uitgebreide) *harvesting* van content uit grote bibliotheken. Het gaat daarbij om veel grotere archiefsystemen met documenten, multimedia of samengestelde content. De archiefsystemen zijn bovendien vaak gestructureerd opgebouwd.

Bovendien geldt dat RSS en Atom geen vrijheid bieden t.a.v. de te gebruiken metadata velden (deze zijn voorgedefinieerd).

- **Sitemaps**

Deze standaard wordt gebruikt voor het publiceren van een overzicht / structuur van website of andere contentverzameling. Deze sitemap kan vervolgens ingelezen worden door een zoekmachine. Google heeft onlangs aangegeven voor deze toepassing alleen nog Sitemaps te ondersteunen, daar waar in het verleden ook OAI-PMH toegepast kon worden.

- **Internet / Informatie Publicatie Modellen (IPM)**

Dit is een verzameling specifieke koppelvlakken, gebruikt binnen het programma Overheid heeft Antwoord voor diverse transparantieprojecten. Het gaat dan om het publiceren van informatie over o.a. vergunningen, bekendmakingen en decentrale regelgeving in landelijke zoeksystemen. De standaard heeft twee aspecten:

- een semantisch gedeelte, dat beschrijft hoe content items voorzien moeten worden van metadata (bijvoorbeeld: welke kenmerken heeft een vergunning).
- een transportmechanisme, dat beschrijft hoe metadata elementen uitgewisseld dienen te worden. Op dit moment bestaat dit uit het ophalen van een indexbestand en een metadata bestand per content item. Dit transportmechanisme is specifiek voor de toepassingen die gebruik maken van IPM.

OAI-PMH heeft een mogelijke overlap met dit transportmechanisme. Naar mening van de expertgroep is het verstandig om te kiezen voor een generiek transportmechanisme, zoals OAI-PMH. Dit maakt het mogelijk content systemen snel en eenvoudig aan te sluiten op meerdere repositories met generieke technologie.



De regelgeving rondom comply-or-explain biedt daarnaast voldoende aanknopingspunten om te komen tot een geleidelijke overgang, mocht dit nodig blijken.

#### - **Metadata**

OAI-PMH bepaalt niet welk metadata formaat gebruikt moet worden. Minimaal moet een vorm van Dublin Core worden gebruikt. In sommige sectoren worden echter meer uitgebreide metadata schema's gehanteerd, zoals binnen de educatieve sector (Content Zoekprofiel, LORELOM, etc.).

De overheid stelt OWMS (Overheids Web Metadata Standaard) als metadateringsstandaard voor overheden voor. De expertgroep concludeert dat dit geen overlap vormt met OAI-PMH, juist doordat PMH het toe te passen metadata schema open laat.

### **3.3 Potentieel**

*Draagt het opnemen van de standaard op de lijst bij aan het vergroten van de leveranciersafhankelijkheid? Draagt het opnemen van de standaard op de lijst bij aan het vergroten van de interoperabiliteit?*

Naar mening van de expertgroep draagt het opnemen van OAI-PMH op de lijst bij aan interoperabiliteit en leveranciersafhankelijkheid:

- Interoperabiliteit tussen aanbieders binnen contentketens (onderwijs, erfgoed, overheidsdienstverlening etc.) wordt vergroot indien er een standaard is voor harvesting van metadata. Een aanbieder kan zo gemakkelijk aansluiten op meerdere partijen die metadata verzamelen en verspreiden.
- Gebruik van OAI-PMH voorkomt toepassing van proprietary standaarden, die potentieel leveranciersafhankelijkheid met zich meebrengen. Leveranciersafhankelijkheid kent hierbij twee dimensies: afhankelijkheid van toeleveranciers of afnemers van metadata (bijvoorbeeld een bepaalde repository) en afhankelijkheid van leveranciers van software (die specifieke eigen mechanismen gebruiken).

Tegelijkertijd plaats de expertgroep ook een kanttekening: voor toepassing van OAI-PMH in een bepaalde sector kunnen nadere afspraken noodzakelijk zijn – met name daar waar het gaat om het toe te passen metadata formaat. Opname van een standaard zoals Content Zoekprofiel (eveneens in behandeling voor de lijst met open standaarden) is daarom complementair aan opname van OAI-PMH.

### **3.4 Impact**

*Wat is de impact van de standaard op de bedrijfsvoering, informatievoorziening, ICT en beveiliging en privacy van de gebruikers van de standaard?*

#### Ten aanzien van bedrijfsvoering en informatievoorziening

Het beschikbaar stellen van metadata aan andere repositories biedt nieuwe mogelijkheden voor contentaanbieders om hun content te verspreiden. Tegelijkertijd ontstaan er nieuwe soorten toepassingen m.b.t. het vergaren en beschikbaar stellen van metadata. De educatieve contentketen (o.a. Edurep) is hier een voorbeeld van.

Naast kansen geeft dit ook risico's: er moet goed bekeken worden *welke* content beschikbaar wordt gesteld. En wat zijn de gevolgen voor het bedrijfsmodel van de organisatie? Deze hangen echter vooral samen met de digitalisering en het beschikbaar stellen van gegevens, dan met de standaard zelf.

De technologische risico's zijn beperkt: de standaard voorziet in technische foutafhandeling, maar dit zegt niets over de functionele/semantische juistheid van de aangeboden gegevens. Hiervoor zijn aanvullende implementatietests nodig. Daarnaast kan het delen van gegevens een extra beslag betekenen op servercapaciteit; hier zal dan in voorzien moeten worden.

#### Ten aanzien van beveiliging

Er zijn geen specifieke beveiligingsrisico's te verwachten bij het gebruik van de standaard. Wel zal bij de implementatie van de standaard hier de nodige aandacht aan moeten worden besteed.

- Organisatorisch zal bepaald moeten worden of gegevens beschikbaar gesteld worden. OAI-PMH is een middel. Opname van OAI-PMH op de lijst verplicht niet tot het beschikbaar stellen van gegevens op zich.
- Onderliggend aan OAI-PMH kunnen diverse netwerken en bijbehorende standaarden worden gebruikt. OAI-PMH schrijft enkel het gebruik van HTTP en HTTPS (standaard internettechnologie) voor. Een organisatie kan er voor kiezen om – in samenhang hiermee – extra beveiligingsmaatregelen te treffen, bijvoorbeeld door het gebruik van private netwerken.
- OAI-PMH gaat over de uitwisseling van metadata en niet over het transport van documenten/objecten zelf. Afhankelijk van de context kunnen documenten extra beveiligd moeten worden, bijvoorbeeld indien er vooraf betaald moet worden voor toegang tot het document. Dit valt echter ook buiten de scope van OAI-PMH.

#### *Hoe gemakkelijk is de migratie naar de standaard?*

Het gemak waarmee gemigreerd kan worden naar de standaard is afhankelijk van een aantal factoren:

- Indien er sprake is van een goed gestructureerd bronsysteem dan is het doorgaans eenvoudig om metadata beschikbaar te stellen voor de content in dat systeem. Indien dit niet het geval is, dan zijn er mogelijk extra stappen nodig om de metadata vast te leggen of op te slaan.
- Als er op dit moment al sprake is van de uitwisseling van metadata, dan is waarschijnlijk een aanpassing in de systemen nodig. Deze aanpassing kan verschillen van systeem tot systeem. Voordeel is dat er een open source implementatie van de standaard beschikbaar is, welke als toevoeging op veel soorten content systemen gebruikt kan worden [4].

Daarnaast speelt het al eerder aangegeven aspect: per sector zijn mogelijk aanvullende afspraken nodig over de aard van de uitgewisselde metadata. In de praktijk blijken hier al goede richtlijnen voor te liggen, waaronder de eerder genoemde voorbeelden (CZP, LORELOM, OWMS, etc.). Als deze afspraken er al zijn, dan zijn er bij toepassing van OAI-PMH geen specifieke migratieproblemen te verwachten.

---

4 Zie <http://www.openarchives.org/pmh/tools/tools.php>

## 4 Advies aan Forum en College

### 4.1 **Samenvatting van de toetsingscriteria**

Samengevat is het oordeel op de toetsingscriteria als volgt:

#### *Openheid*

De standaard is voldoende open. Er is een community waarin iedereen kan participeren en inspraak heeft. De standaard wordt beheerd door een non-profit samenwerkingsverband, dat wordt ondersteund door diverse (vooral: Amerikaanse) kennisinstellingen.

#### *Bruikbaarheid*

De standaard is stabiel en volwassen. Binnen de archiefwereld, erfgoedsector en het onderwijs wordt de standaard al toegepast. De geboden functionaliteit sluit goed aan bij de behoeften op het gebied van metadata harvesting. Het functioneel toepassingsgebied maakt het onderscheid met potentieel concurrerende standaarden voldoende duidelijk. Er is voldoende ondersteuning, ondermeer vanuit open source software.

#### *Potentieel*

Het opnemen van de standaard op de lijst voor comply-or-explain draagt bij aan de interoperabiliteit en leveranciersafhankelijkheid binnen contentketens.

#### *Impact*

De technologische risico's zijn gering. Het aanbieden van gegevens kan impact hebben op de bedrijfsvoering en het bedrijfsmodel van organisaties. Dit staat echter los van de toegepaste standaard. Migratie naar OAI-PMH levert in de meeste gevallen geen grote problemen op.

Afhankelijk van de toepassing kunnen aanvullende beveiligingsmaatregelen nodig zijn, die in combinatie met OAI-PMH kunnen worden ingezet.

### 4.2 **Advies**

De expertgroep adviseert het college de standaard OAI-PMH versie 2.0 zonder voorbehouden op te nemen op de lijst met open standaarden met het in hoofdstuk 2 vastgestelde functioneel toepassingsgebied en organisatorische werkingsgebied.