

Werkgroep Metadata Subwerkgroep WiSH

Bijeenkomst 26 maart 2015

1	Welkom	5 min	Rob
2	Notulen vorige keer (besluitenlijst / actielijst)	15 min	Rob
3	Kennisinfrastructuur binnen Nederland Overzicht van de verschillende partijen en de stromen van metadata, publicaties en afspraken.	15 min	Rob
4	Één loket XSD / RSS / ... / Informatie / email	15 min	Chris
	Pauze	15 min	-
5	COAR Update	15 min	Wilko
6	Issuelijst top 10 - Validatie	45 min	Rene
7	Openstaande aanbevelingen ORIS Welke issues pakken we op en hoe brengen we dat in de WiSH werkgroep (procedure).	30 min	Rob
8	Wvttk & sluiting	15 min	Rob

Werkgroep Metadata Subwerkgroep WiSH

2. Actielijst / Besluitenlijst

Actielijst

#	Datum inbreng	Actiepunt	Verantwoordelijk	Afgesloten op
1	28-10-2014	Zijn de xsd schemas ook gepubliceerd? Zo ja, waar?	Jeroen Hamers, Peter	Gedaan Later meer
2	28-10-2014	Rol Edustandaard en Wilco als inhoudelijke vertegenwoordiger aan COAR toelichten.	Jeroen Hamers	Niet gedaan
3	28-10-2014	1 Loket voor alle aangesloten partijen.	Chris	Gedaan Later meer
4	28-10-2014	URL om metadata weer op te halen wordt rondgestuurd.	Rene	Gedaan
5	28-10-2014	Gebruiksvriendelijk foutenlijst / inzicht in fouten	-	Nog niet
6	28-10-2014	Top 10 foutenlijst met uitleg voor volgende bijeenkomst.	Rene	Gedaan Later meer
7	28-10-2014	Behoeftte aan WRM. Contact opnemen met Elly van Dijk van DANS om afstemming te zoeken.	Jeroen	Niet gedaan
8	28-10-2014	Presentaties werkgroep Vocabulaires.	Jeroen	Niet gedaan
9	28-10-2014	Issuelijst op basis van aanbevelingen ORIS.	Iedereen	Gedaan Later meer

Besluitenlijst

#	Datum besluit	Besluit
1	28-10-2014	MODS-DIDL vervangen door NL-Cerif voorlopig niet, eerste huidige afspraken op order krijgen en iedereen aansluiten.
2	28-10-2014	Wilko vertegenwoordigd deze werkgroep in het COAR overleg.
3	28-10-2014	Blijven bij Semantics afspraak zoals deze nu bij Edustandaard in beheer is. Wel zal altijd gekeken worden welke wijzigingen er in COAR plaatsvinden en worden de wijzigingen in de Semantics afspraken meegenomen naar COAR door Wilko.

Vragen?

WebSite : harvester.dans.knaw.nl

Email : harvester@dans.knaw.nl

Werkgroep Metadata

Subwerkgroep WiSH

3. Kennisinfrastuctuur Nederland

Kennisinfrastructuur

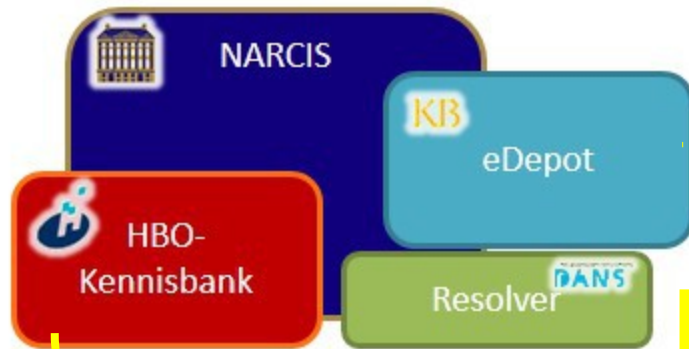
“ ... een situatie waarin alle objecten uit repositories van wetenschappelijke instellingen duurzaam kunnen worden teruggevonden in de eigen, actuele omgeving ...”

Uit: OZ.11.0408 PersID-NL Controlling Document, 29-02-2011.

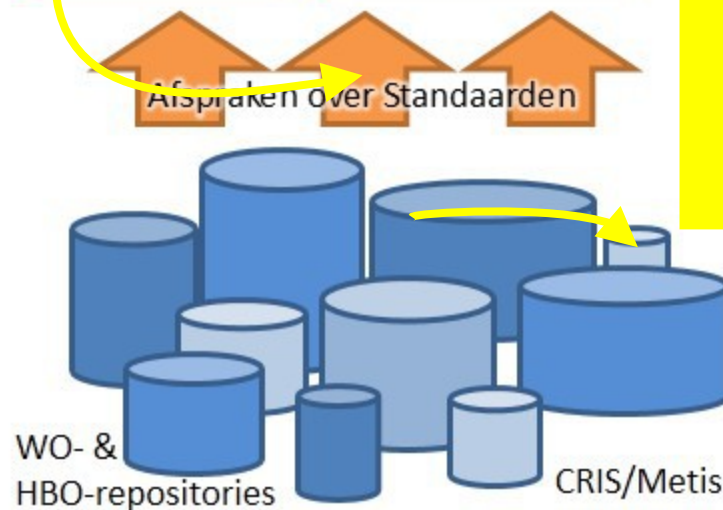
Research Life Cycle / HO-kennisketen



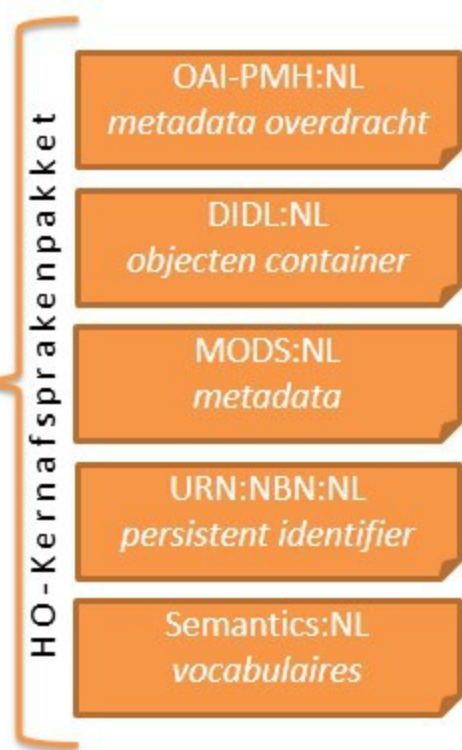
Contentafnemers/
Service providers

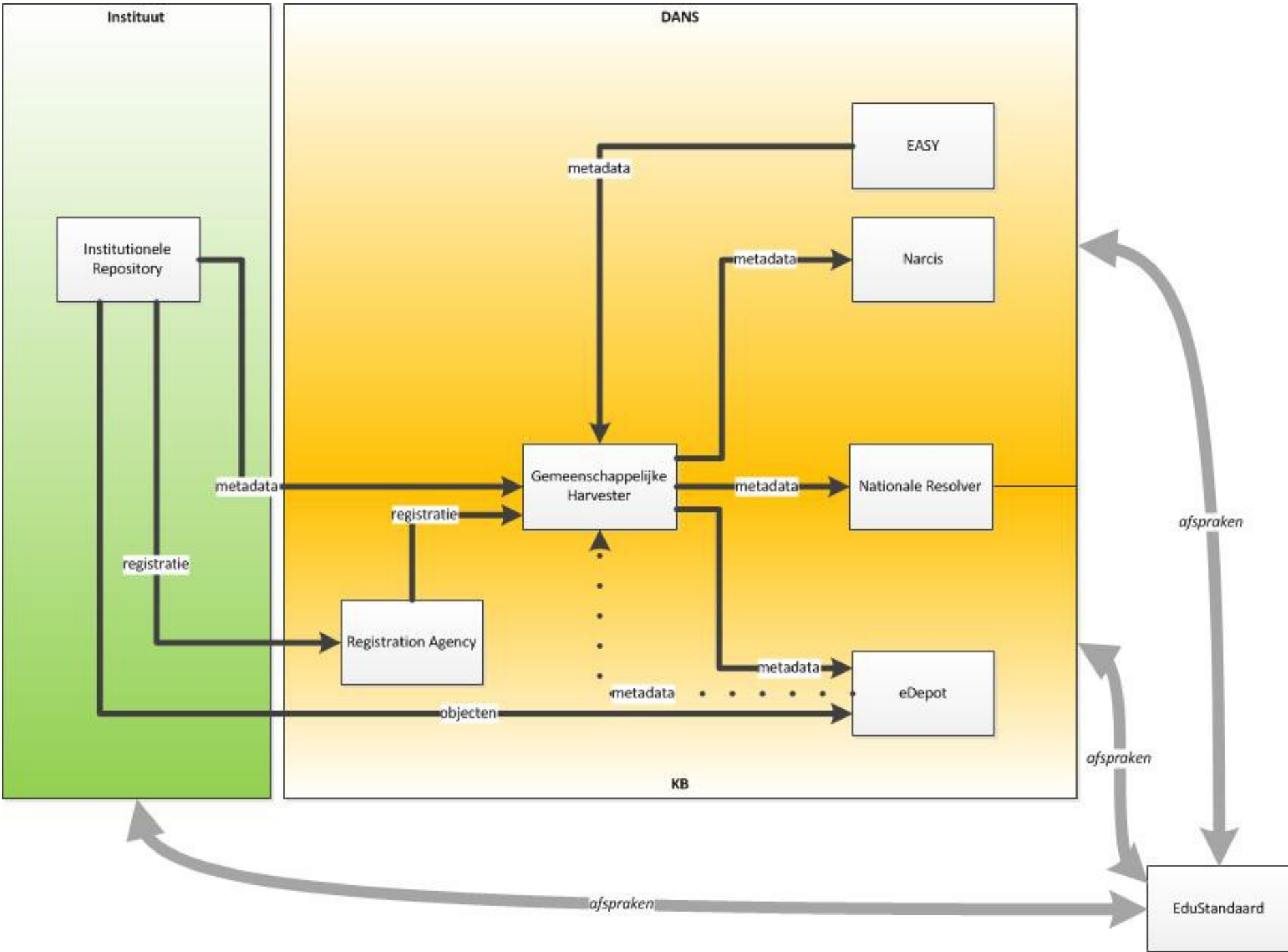


Contentleveranciers/
Data providers



Afspraken over Standaarden





Vragen?

WebSite : harvester.dans.knaw.nl

Email : harvester@dans.knaw.nl

Werkgroep Metadata

Subwerkgroep WiSH

4. Eén loket

Demo

<http://harvester.dans.knaw.nl/>

PAUZE

Bereid je issue voor

Werkgroep Metadata

Subwerkgroep WiSH

5. COAR Update

Werkgroep Metadata Subwerkgroep WiSH

6. Issuelijst top 10

GMH: top 10 validatieproblemen

Uitgangspunten GMH

- Validatie: basis Edustandaard-afspraken
- Beperkte normalisatie
- Levert zowel originele als genormaliseerde data

<http://www.edustandaard.nl/standaarden/afspraken/afpraak/semantics/1.0/>

<http://www.edustandaard.nl/standaarden/afspraken/afpraak/didl/1.0/>

<http://www.edustandaard.nl/standaarden/afspraken/afpraak/mods/1.0/>

GMH: top 10 validatieproblemen

Top 10 validatiefouten

1. MODS: Geen valide MODS genre
2. MODS: Verplichte originInfo/dateIssued ontbreekt
3. DIDL: No accessRights found for objectfile
4. DIDL: Resource location niet gevonden (urn:nbn url)
5. MODS: titleInfo/title has empty value
6. DIDL: ongeldige URI (niet toegestane tekens, bijvoorbeeld []).
7. MODS: ongeldig xs:ID (bv _123)
8. DIDL: Geen resource URL in human startpage
9. MODS: Metadata niet geldig, child ontbreekt.
10. MODS: Ongeldige taalcode

GMH: top 10 validatieproblemen

1. /mods/genre niet valide

/mods/genre wordt gebruikt voor het type van de publicatie. Verplicht element! Toegestane waarden uit lijstje ('enum').

- Edustandaard < > SurfSHARE!
- EduStandaard niet altijd eenduidig

GMH: top 10 validatieproblemen

/mods/genre: Conference [Item|Paper|Poster]

Edustandaard: (Semantics, Tabel 2 uit bijlage)

<http://purl.org/eprint/type/ConferenceItem>

<http://purl.org/eprint/type/ConferencePaper>

<http://purl.org/eprint/type/ConferencePoster>

SurfShare:

<info:eu/semantics/ConferenceItem>

<info:eu/semantics/ConferencePaper>

<info:eu/semantics/ConferencePoster>

GMH: top 10 validatieproblemen

/mods/genre: Conference [Item|Paper|Poster]

- Semantics claimt enkel info:eu namespace toe te staan
- Semantics maakt onderscheid tussen publicatietype en conference object type
- Schrijfwijze niet volgens lowerCamelCase

GMH: top 10 validatieproblemen

/mods/genre: Conference [Item|Paper|Poster]

Wijzigingsvoorstellen

- Semantics: tabellen 1&2 (publicatietypes / conference object types) samenvoegen, pas 3.3.1 aan en laat bij Toegestane implementaties alleen een verwijzing staan naar tabellen 1 & 2 (samengev.).
- In Edustandaard Mods afspraken bij “4.2.8 Publicatietype” bij encodering enkel verwijzen naar aangepast Semantics document

GMH: top 10 validatieproblemen

/mods/genre: info:eu/semantics/conferenceContribution

“contribution to a conference but not published in a conference”

- Niet in lijst goedgekeurde types.
- Wel genoemd in een voorbeeld in Semantics, MODS tabel met entitynames.

Wijzigingsvoorstel

- Alle verwijzingen naar deze waarde schrappen

GMH: top 10 validatieproblemen

/mods/genre: other

- Wat te doen als er geen passend publicatietype is of als het onbekend is?
- Hoe gaat de GMH hiermee om?

GMH: top 10 validatieproblemen

2. /mods/originInfo/dateIssued ontbreekt

Verplicht veld, W3C date-time format.

GMH: top 10 validatieproblemen

3. DIDL: no accessRights for objectfile

In het verleden was dit niet verplicht, standaard OpenAccess. Volgens Edustandaard wel verplicht.

“Bij dcterms:accessRights is één van de volgende Eprints-vocabulary-typen verplicht:

<http://purl.org/eprint/accessRights/OpenAccess>,
.../RestrictedAccess of .../ClosedAccess.

Bij een verandering in de accessRights (bijvoorbeeld bij een verlopen mandaat) is de repository verplicht het aangepaste record opnieuw aan te bieden. “

GMH: top 10 validatieproblemen

4. DIDL: Resource location niet gevonden (urn:nbn url)

Volgens DIDL afspraken hoort bij een nbn-identificer minimaal een modification date en een (component) url. Die url kan naar digitaal object of naar human start page verwijzen.

GMH: top 10 validatieproblemen

5. MODS: titleInfo/title has empty value

Titel is verplicht.

GMH: top 10 validatieproblemen

6. DIDL: ongeldige URI

URL/ URI bevat ongeldige tekens.

URLs moeten encoded opgenomen worden, bijvoorbeeld spaties als %20, [en] alleen in query-deel etc.

GMH heeft encoding en strippen spaties nu als normalisatie-stap.

GMH: top 10 validatieproblemen

7. MODS: ongeldige xs:ID

Gebruikt in /mods/name

Waarde van ID-attribuut in XML moet beginnen met [a-zA-Z] of _

Waarde van ID-attribuut moet uniek zijn.

GMH: top 10 validatieproblemen

8. DIDL: Geen resource URL in human startpage

(verplicht)

9. MODS: metadata niet geldig

(incident)

10. MODS: ongeldige taalcode

(incident)

GMH: top 10 validatieproblemen

Wijzingsvoorstellen

Wijzigingsvoorstellen

- Semantics: tabellen 1&2 (publicatietypes / conference object types) samenvoegen, pas 3.3.1 aan en laat bij Toegestane implementaties alleen een verwijzing staan naar tabellen 1 & 2 (samengev.).
- In Edustandaard Mods afspraken bij “4.2.8 Publicatietype” bij encodering enkel verwijzen naar aangepast Semantics document.
- Alle verwijzingen naar conferenceContribution schrappen
- ...

Vragen?

WebSite : harvester.dans.knaw.nl

Email : harvester@dans.knaw.nl

Werkgroep Metadata

Subwerkgroep WISH

7. Openstaande aanbevelingen ORIS

Openstaande aanbevelingen

- Totaal: 48 issues.
- DIDL: 2
- MODS: 23
- Semantics: 11
- OAI-PMH: 10
- URN:NBN:NL: 2

... en ...

- De lijst van Rene ...
- Wat we nog verder willen ...
 - DAI vervangen door ISNI (3.8?)
 - ...

Proces

- Bij Edustandaard:
- <http://www.edustandaard.nl/participeren/standaard-indienen/hoe-wijzig-ik-een-bestaande-afpraak/>

Proces EduStandaard

De volgende wijzigingen op een bestaande standaard hoeven niet aan de Standaardisatieraad worden voorgelegd:

- Verouderde verwijzingen
- Typografische wijzigingen
- Taalfouten, mits deze niet een inhoudelijke wijziging van de afspraak veroorzaken
- Typefouten

Proces Edustandaard

Procedure voor het indienen van een nieuwe afspraak

Stap 1 Vorbereiding

Informeren van Bureau Edustandaard



Aanmeldformulier invullen en memo opstellen



Documentatie, aanmeldformulier en memo opsturen naar Bureau Edustandaard

Stap 2 Beoordeling

Advies van Bureau Edustandaard ontvangen



Schriftelijk reageren op het advies

Stap 3 Besluitvorming

Presentatie geven tijdens de bijeenkomst van de Standaardisatieraad



Besluit door de Standaardisatieraad voor in beheer name van de afspraak

Stap 4 Nazorg

Vervolgstappen afspreken met Bureau Edustandaard

Vragen?

WebSite : harvester.dans.knaw.nl

Email : harvester@dans.knaw.nl

Werkgroep Metadata

Subwerkgroep WiSH

8. WVTTK & afsluiting