

Afspraak Unieke Persistente Identifier voor Leermateriaal en Metadatarecord

Afspraak over het gebruik van unieke persistente identifiers in
de educatieve contentketen

Auteur(s) : Jeroen Hamers, Wim Muskee, Edwin Verwoerd,
Marjan Frijns en Jasper Roes
Versienummer : 1.0 (7 november 2011)
Totstandkoming : Dit document is tot stand gekomen in
samenwerking met vertegenwoordigers van
aanbieders en afnemers van leermateriaal en
metadatarecords en systeem leveranciers

7 november 2011

Documentgeschiedenis

1.0	7 november 2011	Eerste versie van de Afspraak Unieke Persistente Identifier voor Leermateriaal en Metadatarecord.

Inhoudsopgave

Documentgeschiedenis 2

Inhoudsopgave3

1. Inleiding 4

- 1.1 Het kader 4
- 1.2 De afspraak 5
- 1.3 Opbouw en doelgroep 5

2. Het wat en waarom van de afspraak (B) 6

- 2.1 Wat zijn unieke persistente identifiers? 6
- 2.2 Waarom unieke persistente identifiers voor leerobjecten? 6
- 2.3 Waarom unieke persistente identifiers voor metadatarecords? 6
- 2.4 Een afspraak voor de Nederlandse contentketen 6
- 2.5 Voordelen van een brede afspraak over unieke persistente identifiers 7

3. De beschrijving van de afspraak (I) 8

- 3.1 Eisen aan de unieke persistente identifier 8
- 3.2 Eisen aan de systemen 9
- 3.3 Aanbevelingen 9
- 3.4 Toegestane unieke persistente identifier types 10

4. Technische uitwerking van de afspraak (T)12

- 4.1 Eisen aan de unieke persistente identifier 12
- 4.2 Eisen aan de systemen 12
- 4.3 Aanbevelingen 12
- 4.4 Unieke persistente identifier notatie 13
- 4.5 Technische infrastructuur bij gebruik Handle 14
- 4.6 Implementatiehandleiding 15

5. Vrijwaring gebruik afspraak 16

6. Bronnen 17

7. Afkortingen / verklarende woordenlijst 18

Bijlage A: Handle structuur 19

Bijlage B: Voorbeeld metadatarecord met unieke persistente identifiers 20

Bijlage C: Uitwerkingen technische eisen unieke persistente identifiers 21

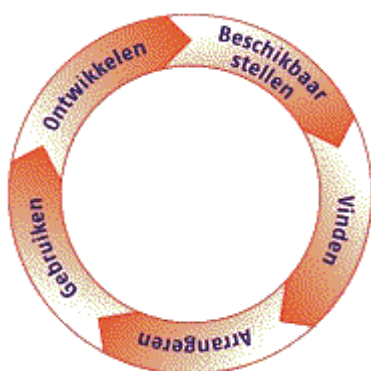
Bijlage D: Checklist voldoen aan afspraak 23

7 november 2011

1. Inleiding

1.1 Het kader

Dat internet in potentie een uitstekend medium is om educatief materiaal te ontsluiten, wordt door zo goed als iedereen in het onderwijs onderschreven. Toch worden de mogelijkheden van internet voor het onderwijs nog onvoldoende benut. Dat ligt uiteraard aan een veelheid van factoren, maar één belangrijke factor daarbij is dat het gebruik van webbased leer materiaal via internet niet altijd vlekkeloos verloopt. Er zijn tal van knelpunten die de uitwisseling ophouden of frustreren. In onderstaand model wordt de keten van educatieve content (de ECK) weergegeven. De stappen in de keten zijn het ontwikkelen, beschikbaar stellen, vinden, arrangeren tot het uiteindelijke gebruik van het leer materiaal door degene die gaat leren. Vanuit dit gebruik is er weer informatie beschikbaar die van belang is voor de ontwikkelaar en is de cirkel rond.



Figuur 1: Educatieve contentketen

De stap "vinden" is belast met het verzamelen van informatie (metadata) over een leerobject. Daartoe worden deze metadata doorzoekbaar gemaakt binnen een gemeenschappelijke voorziening (bijvoorbeeld Edurep). Dit betekent dat alle informatie die over een leerobject beschikbaar is binnen de keten voor het vinden kan worden ingezet.

In 2010 is door Kennisnet een onderzoek gedaan naar Metadatastromen¹. Vanuit dit onderzoek zijn een aantal aanbevelingen gedaan hoe binnen de Educatieve contentketen om te gaan met het groeiend aantal metadatastromen dat in de keten begint te ontstaan. De belangrijkste aanbeveling voor oplossing van de metadatastromen problematiek was 'maak leerobjecten uniek identificeerbaar en de bijbehorende metadata als geheel doorzoekbaar'. De aanbeveling 'maak leerobjecten uniek identificeerbaar' vormt de directe aanleiding voor dit project en wordt ook als uitgangspunt genomen voor de ontwikkeling van deze afspraak.

Unieke persistente identifiers zijn er echter niet alleen voor leer materiaal, maar ook voor boeken, scholen, leerlingen, vocabulaires en wetenschappelijke publicaties en kunnen er straks ook zijn voor vocabulaire termen, leerlijnen, stappen in een leerlijn, locaties, lokalen, cultureel erfgoed, etc. Dat is niet allemaal ECK problematiek, maar speelt ook in andere ketens. De afspraak in dit document probeert om deze reden zoveel mogelijk rekening te houden met bestaande unieke persistente identifiers. Daarnaast is deze afspraak ook ontwikkeld met in het achterhoofd het hergebruik van de unieke persistente identifiers uit de ECK in andere ketens.

¹ Zie: http://digitaallemateriaal.kennisnet.nl/attachments/2151909/Adviesrapportage_Metadatastromen.pdf

7 november 2011

1.2 De afspraak

Om leermaterialen en metadatarecords uniek identificeerbaar te maken is het noodzakelijk om tot goede afspraken te komen over de te gebruiken unieke persistente identifiers. Hiervoor is vanuit het Educatieve Contentketen 2 programma, samen met aanbieders en afnemers van leermaterialen en metadatarecords en systeemleveranciers deze afspraak ontwikkeld.

1.3 Opbouw en doelgroep

De opbouw van dit document is als volgt:

Hoofdstuk	Titel
Hoofdstuk 1	Inleiding
Hoofdstuk 2	Het wat en waarom van de afspraak (B)
Hoofdstuk 3	De beschrijving van de afspraak (I)
Hoofdstuk 4	Technische uitwerking van de afspraak (T)
Hoofdstuk 5	Vrijwaring gebruik afspraak
Hoofdstuk 6	Literatuurlijst
Hoofdstuk 7	Afkortingen/ verklarende woordenlijst
Bijlage A	Handle systeem
Bijlage B	Voorbeeld metadatarecord met unieke persistente identifiers
Bijlage C	Uitwerkingen technische eisen unieke persistente identifiers
Bijlage D	Checklist voldoen aan afspraak

Hierbij staat de B voor bedrijfsmatige, organisatorische kant van de afspraak en is vooral bedoeld voor bestuurders en beslissers. De I staat voor informatie waarbij de afspraak meer in detail wordt beschreven, bedoeld voor informatiemanagers. De T staat voor techniek waar de technische details van de afspraak zijn uitgewerkt, bedoeld voor technische specialisten.

2. Het wat en waarom van de afspraak (B)

2.1 Wat zijn unieke persistente identifiers?

Unieke persistente identifiers zijn codes die toegekend worden aan objecten en er daarmee voor zorgen dat het mogelijk is om objecten uniek te identificeren. Een goed voorbeeld van een unieke persistente identifier is het ISBN voor boeken of de streepjescode op bijna alle producten in de supermarkt. Op basis van het ISBN is het mogelijk om een boek te vinden en weet je zeker dat je niet een ander boek krijgt dan hetgeen je aangevraagd hebt. Door de streepjescode op producten in de supermarkt is het bij de kassa mogelijk om direct de informatie over het product op te vragen en correct op de kassabon te zetten. Zonder deze unieke persistente identifiers zou het opvragen van een boek alleen kunnen op basis van bijvoorbeeld de titel (waarbij de kans aanwezig is dat meerdere boeken dezelfde titel hebben) en het opvragen van de prijs van een artikel bij de kassa zou alleen kunnen door het product op basis van de naam op te zoeken, waarbij je hetzelfde risico loopt. Unieke persistente identifiers maken dit proces stukken eenvoudiger en de kans op fouten kleiner.

2.2 Waarom unieke persistente identifiers voor leerobjecten?

Voor leerobjecten zijn unieke persistente identifiers noodzakelijk omdat deze het mogelijk maken om metadata, zoals het eigenaarschap of de locatie van het leerobject, op verschillende plaatsen in de educatieve contentketen ondubbelzinnig naar het leerobject te kunnen laten verwijzen. Tevens kan de metadata duurzaam worden beheerd. Alle gebruikers in de educatieve contentketen kunnen op basis van de unieke persistente identifier zich altijd op de hoogte stellen van wijzigingen in de metadata.

2.3 Waarom unieke persistente identifiers voor metadatarecords?

Het toevoegen van een unieke persistente identifier aan metadatarecords is essentieel om metadatarecords van verschillende aanleverende partijen te kunnen onderscheiden en naar elkaar te kunnen laten verwijzen. Dit kan bijvoorbeeld gebruikt worden om metadata van eigenaren en reviewers uit elkaar te houden, om onderscheid te kunnen maken tussen verschillende versies van een metadatarecord van één partij, om ratings over een metadatarecord vast te leggen en om gebruikers de mogelijkheid te geven om een specifiek metadatarecord op te slaan of te delen met anderen.

2.4 Een afspraak voor de Nederlandse contentketen

De behoefte aan unieke persistente identifiers is niet nieuw en in de afgelopen decennia zijn er verschillende standaarden ontwikkeld voor unieke persistente identifiers. Vanuit het onderwijsveld in Nederland is er een behoefte geuit om specifiek voor de uitwisseling van digitale leerobjecten en metadatarecords een afspraak te maken over de unieke persistente identifiers die gebruikt kunnen worden in de educatieve contentketen. Deze afspraak is ontwikkeld voor de Nederlandse educatieve contentketen, maar maakt het mogelijk om binnen de educatieve contentketen ook unieke persistente identifiers uit andere ketens te gaan gebruiken.

7 november 2011

2.5 Voordelen van een brede afspraak over unieke persistente identifiers

Het maken van een brede afspraak over unieke persistente identifiers biedt zowel voordelen voor de keten als voordelen voor individuele ketenpartijen, zowel voor de gebruikers als voor de aanbieders. Deze afspraak kan worden gezien als één van de fundamenteën van de keten. Net als bij het bouwen van een huis, waar het fundament essentieel is om een mooi en duurzaam huis te bouwen, is ook in de educatieve contentketen een goed fundament noodzakelijk. Met een goed fundament kunnen nieuwe diensten worden aangeboden en is het mogelijk om de educatieve contentketen door te ontwikkelen tot een robuuste keten waar iedereen op kan bouwen.

Naast het creëren van een stevig fundament zijn ook de volgende voordelen te benoemen:

1. De unieke persistente identifier creëert de randvoorwaarden voor het op kunnen lossen van de bestaande zoekproblemen die gebruikers ervaren (dubbeling in zoekresultaten, niet kunnen zoeken in alle beschikbare metadata over een leerobject). Met behulp van de unieke persistente identifiers kan de centrale zoekvoorziening samen met de portalen indien gewenst zoekresultaten ontdubbelen en het mogelijk maken om te zoeken in de rijkste metadata. Het ontdubbelen zorgt er ook voor dat er binnen de keten niet de schijn wordt gewekt dat partijen leer materiaal bewust meerdere malen in de zoekresultaten proberen te krijgen om zo een hoger gebruik te creëren.
2. De unieke persistente identifier creëert de randvoorwaarden voor het up-to-date kunnen houden van het eigenaarschap van leerobjecten. Door het eigenaarschap up-to-date te houden kunnen gebruikers eenvoudig achterhalen van wie het materiaal is en op basis hiervan bijvoorbeeld besluiten het materiaal wel of niet te gebruiken.
3. De unieke persistente identifier creëert de randvoorwaarden voor het up-to-date houden van wijzigingen in de locatie van het leerobject en het metadata record. Hierdoor komen gebruikers altijd uit bij het gevonden materiaal of bij een verwijzing naar een nieuwe versie van het materiaal en wordt hun gebruikservaring verbeterd en wordt er meer leer materiaal gevonden.
4. Op het moment van schrijven van deze standaard (medio 2011) lijkt de tijd rijp te zijn om het probleem van unieke persistente identifiers op te pakken. Meerdere sectoren zijn hier nu actief mee bezig (waaronder de erfgoedsector en de wetenschappelijke sector). Hiermee kan worden samengewerkt, kan worden aangesloten en kan van worden geleerd.
5. Het voordeel van een brede afspraak is dat ketenpartijen niet worden gedwongen om meerdere oplossingen te implementeren indien ze actief zijn in meerdere ketens. De afspraak staat het hergebruik van unieke persistente identifiers uit andere ketens toe indien deze voldoen aan de gestelde voorwaarden.
6. De vijf unieke persistente identifier types waaruit binnen deze afspraak gekozen mag worden zijn alle vijf globaal uniek waardoor ze hergebruikt kunnen worden als unieke persistente identifier in andere ketens en dus niet specifiek voor de educatieve contentketen alleen geschikt zijn.
7. In het semantisch web is het leggen van relaties tussen objecten essentieel om betekenis te kunnen geven aan objecten. Om deze relaties op een goede manier te leggen is het erg belangrijk dat de verschillende objecten uniek te identificeren zijn. Deze afspraak vormt daarmee ook een goede basis om 'in te toekomst' verder te kunnen met het semantische web.

Voor het behalen van deze voordelen is het van belang dat iedereen in de keten meedoet om de robuustheid te creëren en samen de verantwoordelijkheid te dragen voor de doorontwikkeling van de keten. Daarbij is het van belang te beseffen dat bij de implementatie van de afspraak een investering vereist van alle ketenpartijen maar er in de toekomst voor zorgt dat nieuwe ontwikkelingen tegen lagere kosten kunnen worden ingevoerd.

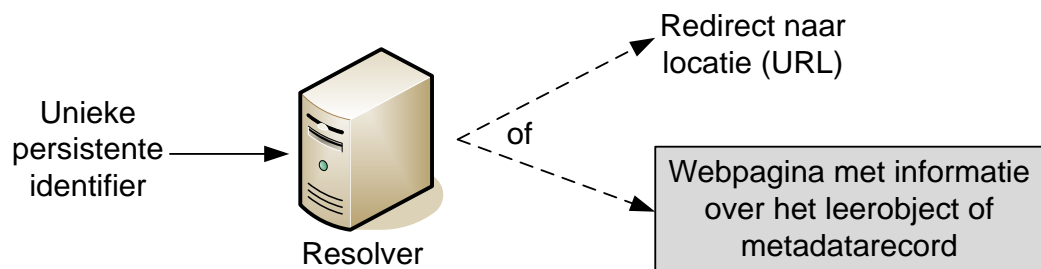
7 november 2011

3. De beschrijving van de afspraak (I)

Bij alle eisen en aanbevelingen aan de unieke persistente identifier is het van belang om in het achterhoofd te houden dat de unieke persistente identifier bij 'het werk' hoort, en niet bij een specifieke manifestatie van een werk. Een document dat zowel in een Word-versie als een PDF-versie beschikbaar is, heeft één unieke persistente identifier die is gekoppeld aan het document ('het werk'). De Word-versie en de PDF-versie zijn slechts manifestaties van het werk en hebben geen aparte unieke persistente identifier.

3.1 Eisen aan de unieke persistente identifier

1. Uniceit: De unieke persistente identifiers die toegekend worden aan leerobjecten en metadatarecords **MOETEN** globaal uniek zijn. Globaal uniek zijn betekent in deze context dat het niet mogelijk moet zijn dat een andere partij dezelfde unieke persistente identifier ook uitgeeft voor een ander object in de wereld. Dit kan bijvoorbeeld geregeld worden door gebruik te maken van een unieke prefix voor iedere partij die unieke persistente identifiers aan wil maken. De combinatie van de unieke prefix en de unieke code voor het leerobject of metadatarecord maakt de combinatie globaal uniek.
2. Unieke persistente identifiers **MOETEN** oneindig lang uniek zijn, ze mogen dus niet worden hergebruikt.
3. Persistentie: Met persistentie bedoelen we dat de unieke persistente identifier in principe zo lang mogelijk, maar ten minste 5 jaar nadat het leerobject of het metadatarecord niet meer bestaat, beschikbaar **MOET** zijn in een resolver. De resolver moet gedurende deze periode up-to-date informatie bevatten. Hierdoor is het voor de gebruiker mogelijk om de meest actuele informatie te vinden (bijvoorbeeld de melding dat het leerobject is verwijderd of een verwijzing naar een volgende versie van het leerobject).
4. Resolvabiliteit: Unieke persistente identifiers **MOETEN** in elk geval leiden naar de meest actuele:
 - 1) Locatie van het object of
 - 2) Pagina met informatie over het leerobject of metadatarecord



5. Als het leerobject of metadatarecord niet meer beschikbaar is (zie eis 3), **MOET** de resolver volgens optie 2 van eis 4 deze informatie verstrekken; bij een leerobject eventueel aangevuld met een verwijzing naar de nieuwste versie van het leerobject.
6. Actionable: Unieke persistente identifiers **MOETEN** (eventueel met behulp van een resolver) via het http-protocol te benaderen zijn.
7. Ieder leerobject dat in de keten wordt gebracht² **MOET** minimaal één unieke persistente identifier toegekend krijgen. Extra identifiers mogen worden meegegeven (zie ook aanbeveling 3 in paragraaf 3.3). Een overzicht van de toegestane unieke persistente identifiers is te vinden in paragraaf 3.4.

² Met in de keten brengen wordt het moment bedoeld waarop het leerobject of metadatarecord in de centrale zoekvoorziening wordt opgenomen.

7 november 2011

8. Aanbieders **MOETEN** voordat ze een leerobject in de keten brengen controleren³ bij de centrale zoekvoorziening of het leerobject niet al eerder in de keten is gebracht door een andere partij. Indien het leerobject al in de keten aanwezig is wordt aanbevolen om geen nieuwe unieke persistente identifier toe te kennen (zie aanbeveling 2 in paragraaf 3.3).
9. Indien bij een leerobject meerdere identifiers bekend zijn **MOETEN** alle bekende identifiers mee worden gegeven in het metadatarecord dat aan de keten wordt aangeboden
10. De unieke persistente identifier **MAG GEEN** semantische betekenis bevatten die noodzakelijk is om met het leerobject of het metadatarecord te kunnen werken. Alle semantische betekenis over het leerobject moet worden opgenomen in het metadatarecord (zie ook aanbeveling 6 in paragraaf 3.3).
11. Indien een nieuw leerobject wordt aangemaakt **MOET** er een nieuwe metadatarecord worden aangemaakt met een nieuwe unieke persistente identifier.
12. Indien een partij versiebeheer ingericht heeft voor leerobjecten **MOET** zij in het metadatarecord de relatie tussen de verschillende versies van leerobjecten vastleggen. Vooral het vastleggen van een verwijzing naar de oude versie van een leerobject bij de nieuwe versie van het leerobject is belangrijk voor de keten. In een metadatarecord dat is opgesteld volgens de NL LOM specificatie is dit mogelijk door het gebruik van veld 7. Specifiek kan de relatie 'is version of' worden gebruikt om bij versie te verwijzen naar de oude versie van het leerobject (zie ook aanbeveling 3 in paragraaf 3.3).
13. Voor unieke persistente identifiers van leerobjecten geldt dat het beheer van een unieke persistente identifier die is uitgegeven door een niet-eigenaar overgenomen **MOET** kunnen worden door de eigenaar van het leerobject (zie ook eis 3 in paragraaf 3.2).

3.2 Eisen aan de systemen

1. Systemen **MOETEN** unieke persistente identifiers toekennen aan leerobjecten. Daarnaast **MOETEN** partijen die unieke persistente identifiers toekennen aan leerobjecten en/of metadatarecords de unieke persistente identifiers in hun systeem vastleggen bij het leerobject of metadatarecord.
2. Systemen die prefixes uitgeven voor één identifier type **MOGEN NIET** een zelfde prefix uitgeven aan verschillende partijen.
3. Systemen van niet-eigenaren **MOETEN** het beheer van een unieke persistente identifier van een leerobject die uitgegeven is door een niet-eigenaar overdragen zodra de eigenaar het beheer van de unieke persistente identifier wenst over te nemen (zie ook eis 13 in paragraaf 3.1).

3.3 Aanbevelingen

1. Er wordt aanbevolen om ieder metadatarecord dat in de keten wordt gebracht⁴ te voorzien van minimaal één unieke persistente identifier. Extra identifiers mogen worden meegegeven (zie ook aanbeveling 3 in paragraaf 3.3). Een overzicht van de toegestane unieke persistente identifiers is te vinden in paragraaf 3.4.
2. Handle wordt aanbevolen als unieke persistente identifier binnen de educatieve contentketen. Partijen die nog geen unieke persistente identifier hebben geïmplementeerd wordt aangeraden om gebruik te gaan maken van Handle om unieke persistente identifiers toe te kennen aan leerobjecten en metadatarecords.
3. Er wordt aanbevolen om te streven naar maar één unieke persistente identifier per leerobject en dus geen nieuwe unieke persistente identifiers toe te voegen aan een leerobject zodra er één is toegekend (zie ook eis 8 in paragraaf 3.1).

³ Door op basis van de URL van het leerobject bij de centrale zoekvoorziening te bepalen of het leerobject al opgenomen is in de keten.

⁴ Met in de keten brengen wordt het moment bedoeld waarop het leerobject of metadatarecord in de centrale zoekvoorziening wordt opgenomen.

7 november 2011

4. Voor het versiebeheer van unieke persistente identifiers van leerobjecten wordt aanbevolen om in ieder geval als partij na te denken over hoe versiebeheer voor de leerobjecten ingericht wordt en om dit beleid vast te leggen zodat het beleid voor andere partijen inzichtelijk is (zie ook eis 12 in paragraaf 3.1).
5. Voor het versiebeheer van unieke persistente identifiers van metadatarecords wordt aanbevolen om in ieder geval als partij na te denken over hoe versiebeheer voor de metadatarecords ingericht wordt en om dit beleid vast te leggen zodat het beleid voor andere partijen inzichtelijk is.
6. Om het voor gebruikers eenvoudiger te maken om de unieke persistente identifier van een leerobject te achterhalen wordt er aanbevolen om de unieke persistente identifier van het leerobject ook in het leerobject op te nemen. Er wordt aanbevolen de unieke persistente identifier op te nemen op de plek die bij het leerobject gereserveerd is om metadata op te nemen. Door de unieke persistente identifier mee te geven in het leerobject is de kans groter dat er opgemerkt wordt dat een object al een unieke persistente identifier heeft en dit kan helpen bij het voorkomen van dubbelingen in de keten.
7. Er wordt aanbevolen om in de unieke persistente identifier geen semantische betekenis aan te brengen. Voor de keten heeft het geen toegevoegde waarde omdat alle semantische betekenis over een leerobject in het metadatarecord moet zijn opgenomen (zie eis 10 in paragraaf 3.1), binnen de organisatie maakt dit het lastiger om te garanderen dat identifiers uniek zijn en blijven.

3.4 Toegestane unieke persistente identifier types

Deze afspraak legt vast dat er binnen de educatieve contentketen meerdere unieke persistente identifier types zijn toegestaan. De volgende types zijn toegestaan: Handle, DOI, ISBN, URN en PURL (met extra eisen). Een nadere beschrijving van de toegestane types, en een voorbeeld per type is te vinden in deze sectie. Voor elk van deze types geldt dat indien hiervoor gekozen wordt er voldaan **MOET** worden aan de eisen zoals gesteld in paragraaf 3.1, niet alle types voldoen zonder extra inspanningen aan deze eisen.

De reden om meerdere unieke persistente identifier types toe te staan binnen de keten is dat partijen hierdoor in de gelegenheid worden gesteld om bestaande unieke persistente identifiers te blijven gebruiken. Het is belangrijk dat het aansluiten op de keten voor partijen zo eenvoudig mogelijk is en daarnaast is het ook van belang om geen 'eiland' te creëren voor de educatieve contentketen maar aan te sluiten bij bestaande infrastructuren en bestaande oplossingen. Er wordt namelijk binnen de educatieve contentketen veel materiaal gebruikt dat uit een andere keten komt, bijvoorbeeld de erfgoedsector en de wetenschappelijke sector. In deze ketens worden al unieke persistente identifiers gebruikt. Door meerdere unieke persistente identifier types toe te staan kan materiaal uit andere ketens eenvoudig worden gebruikt en opgenomen in de educatieve contentketen, mits er maar een unieke persistente identifier wordt gebruikt die toegestaan is in de educatieve contentketen.

Identifier	Beschrijving
Handle	Ooit opgestart als tegenhanger van het DNS systeem, nu in gebruik als stabiele structuur voor het uitdelen van globaal unieke persistente identifiers. <i>Voorbeeld: hdl:4263537/4000</i>
DOI (Digital Object Identifier)	Implementatie van Handle. DOI geeft unieke persistente identifiers uit door partijen een prefix te geven. <i>Voorbeeld: doi:10.1000/182</i>
ISBN/ISSN	Bedoeld voor boeken en publicatiereeksen, zowel papier als elektronisch. <i>Voorbeeld: urn:isbn:9789034553966</i>
URN (Uniform Resource Name)	In tegenstelling tot de URL geen verwijzing naar een locatie maar een uniforme naam voor een resource. Voorbeeld is een URN:NBN:NL (wordt gebruikt door de Koninklijke Bibliotheek en Universiteiten voor wetenschappelijke publicaties). <i>Voorbeeld: urn:nbn:nl:ui:13-367b24w783</i>

7 november 2011

PURL (Persistent Uniform Resource Locator)	Standaard waarbij met behulp van extra afspraken bovenop de URL-afspraken persistente URL's kunnen worden gebruikt als unieke persistente identifier. <i>Voorbeeld: http://purl.org/dc/elements/1.1/identifier</i>
--------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Bij het gebruik van PURL **MOET** de te gebruiken PURL worden geregistreerd bij www.edustandaard.nl en wordt er gecontroleerd of deze aan de volgende eisen voldoet:

- Geen bedrijfsnaam in PURL
- De PURL-server **MOET** voldoen aan de eisen die gesteld worden door PURL⁵
- De PURL-server dient voldoende zekerheid te bieden dat de gecreëerde unieke persistente identifiers kunnen voldoen aan de eisen die in deze afspraak zijn vastgelegd.

De volgende PURL's zijn toegestaan (een up-to-date lijst is te vinden op www.edustandaard.nl).

- purl.org
- purl.edustandaard.nl

⁵ Zie: <http://www.purl.org>

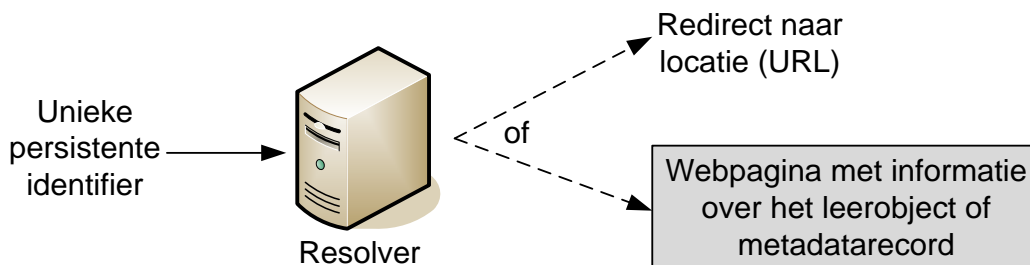
4. Technische uitwerking van de afspraak (T)

4.1 Eisen aan de unieke persistente identifier

1. De unieke persistente identifier **MOET** gebruik maken van de karakterset die voor een URI is toegestaan. Een URI mag alleen alfanumerieke karakters (de cijfers 0-9 en de letters a-z en A-Z) en de speciale karakters \$ - _ . + ! * ' () bevatten. Daarnaast zijn er een aantal gereserveerde karakters die alleen toegestaan worden met de bijzondere betekenis die ze binnen een URI hebben. Het gaat hierbij om de volgende karakters: ; / ? : @ = & #.
2. Bij het opnemen van de unieke persistente identifiers in het metadatarecord **MOET** de officiële notatiewijze van ieder unieke persistente identifier type worden gevolgd. Een uitwerking van deze schrijfwijze per unieke persistente identifier is te vinden in paragraaf 4.4
3. De specifieke resolver die gebruikt kan worden om de unieke persistente identifier te resolven **MAG NIET** worden opgenomen in het metadatarecord voor de unieke persistente identifier types Handle, DOI, ISBN/ISSN en URN. Bij het gebruik van PURL is het **VERPLICHT** om de volledige PURL op te nemen.
4. De unieke persistente identifier **MAG GEEN** semantische verwijzingen bevatten. Er mogen dus bijvoorbeeld geen bedrijfsnamen of namen van spotter-systemen in opgenomen worden.

4.2 Eisen aan de systemen

1. De resolver **MOET** publiek beschikbaar⁶ zijn.
2. Wijzigingen in locatie of eigenaar van een leerobject of metadatarecord **MOETEN** doorgegeven worden aan de resolver.
3. De resolver **MOET** in elk geval één van de volgende zaken teruggeven:
 - 1) De locatie van het object of
 - 2) Een pagina met informatie over het leerobject of metadatarecord



4.3 Aanbevelingen

1. Er wordt aanbevolen om unieke persistente identifiers niet langer te maken dan 250 karakters.

⁶Met publiek beschikbaar wordt bedoeld dat eenieder de resolver moet kunnen benaderen om informatie op te vragen, zonder enige restricties.

7 november 2011

4.4 Unieke persistente identifier notatie

Bij het opnemen van de unieke persistente identifiers in het metadatarecord moet de officiële notatiewijze van ieder unieke persistente identifier type worden gevolgd. De specifieke resolver die gebruikt kan worden om de unieke persistente identifier te resolvable **MAG NIET** opgenomen worden.

De specifieke eisen per unieke persistente identifier worden in onderstaande tabel toegelicht. Daarnaast is in de paragrafen na de tabel voor ieder van de unieke persistente identifiers een voorbeeld te vinden van het opnemen van de unieke persistente identifier in NL LOM (volgens de IEEE-LOMv1.0 binding). Andere voorbeelden zijn te vinden in de implementatiehandleiding.

Identifier	Eisen per identifier en voorbeeld
Handle	Bij het opnemen van een Handle in het metadatarecord MOET de unieke persistente identifier voorafgegaan worden door 'hdl:'. <i>Voorbeeld: hdl:4263537/4000</i>
DOI	Bij het opnemen van een DOI in het metadatarecord MOET de unieke persistente identifier voorafgegaan worden door 'doi:'. <i>Voorbeeld: doi:10.1000/182</i>
ISBN/ISSN	Bij het opnemen van een ISBN of ISSN in het metadatarecord MOET de unieke persistente identifier in de officiële ISBN-URN notatie worden opgenomen. <i>Voorbeeld: urn:isbn:9789034553966</i>
URN	Bij het opnemen van een URN in het metadatarecord MOET de unieke persistente identifier in de officiële URN notatie worden opgenomen. <i>Voorbeeld: urn:nbn:nl:ui:13-367b24w783</i>
PURL	Bij het gebruik van PURL moet de te gebruiken PURL worden geregistreerd bij www.edustandaard.nl en wordt er gecontroleerd of deze aan de volgende eisen voldoet: <ul style="list-style-type: none"> - Geen bedrijfsnaam in PURL - De PURL-server moet voldoen aan de eisen die gesteld worden door PURL⁷ - De PURL-server dient voldoende zekerheid te bieden dat de gecreëerde unieke persistente identifiers kunnen voldoen aan de eisen die in deze afspraak zijn vastgelegd. De volgende PURL's zijn toegestaan (een up-to-date lijst is te vinden op www.edustandaard.nl). <ul style="list-style-type: none"> - purl.org - purl.edustandaard.nl Bij het opnemen van een PURL in het metadatarecord MOET de volledige URL die verwijst naar de PURL-server ook meegenomen worden (dit in tegenstelling tot de andere unieke persistente identifier types waar de resolver niet opgenomen mag worden). <i>Voorbeeld: http://purl.org/dc/elements/1.1/identifier</i>

Het opnemen van de verschillende unieke persistente identifiers in de metadatarecords is hieronder weergegeven. In Bijlage B is een voorbeeld opgenomen van een metadatarecord waarin de unieke persistente identifiers voor het leerobject en het metadatarecord zijn opgenomen en daarnaast de technische locatie van het leerobject.

⁷ Zie: <http://www.purl.org>

7 november 2011

Handle:

```
<identifier>
  <catalog>uri</catalog>
  <entry>hdl:4263537/4000</entry>
</identifier>
```

DOI:

```
<identifier>
  <catalog>uri</catalog>
  <entry>doi:10.1000/182</entry>
</identifier>
```

ISBN/ISSN:

```
<identifier>
  <catalog>uri</catalog>
  <entry>urn:isbn:9789034553966</entry>
</identifier>
```

URN:

```
<identifier>
  <catalog>uri</catalog>
  <entry>urn:nbn:nl:ui:13-367b24w783</entry>
</identifier>
```

PURL:

```
<identifier>
  <catalog>uri</catalog>
  <entry> http://purl.org/dc/elements/1.1/identifier</entry>
</identifier>
```

4.5 Technische infrastructuur bij gebruik Handle

Deze paragraaf beschrijft de technische infrastructuur die ingericht dient te worden bij het gebruik van Handle. Voor de andere unieke persistente identifier types is de beschrijving te vinden in Bijlage C.

Iedere partij die gebruik wil maken van Handle **MOET** een eigen unieke prefix aanvragen bij Handle. Op basis van de eigen prefix kunnen partijen vervolgens unieke persistente identifiers geven aan hun leerobjecten en metadatarecords (hiervoor mogen uiteraard ook al bestaande unieke interne codes worden gecombineerd met de prefix).

Bij de keuze voor Handle is het noodzakelijk dat er een resolver wordt ingericht die de mogelijkheid biedt voor gebruikers om op basis van de unieke persistente identifier de locatie van het leerobject of metadatarecord te vinden. Een eis van Handle is dat er voor iedere prefix een resolver beschikbaar moet zijn. De resolver die bij een prefix hoort is ervoor verantwoordelijk dat alle unieke persistente identifiers die onder de prefix zijn uitgegeven geresolved kunnen worden. Het is mogelijk dat één resolver meerdere prefixes afhandelt.

Een partij die gebruik maakt van Handle moet dus zorg dragen voor een resolver en moet iedere unieke persistente identifier die zij aanmaakt ook aanmelden bij die Handle resolver. Daarnaast moet zij bij wijzigingen in de locatie van het leerobject of metadatarecord deze wijzigingen doorgeven aan de Handle resolver (het betreft hier het beheer van de unieke persistente identifiers).

7 november 2011

4.6 Implementatiehandleiding

Om de implementatie van de unieke persistente identifiers voor leerobjecten en metadatarecords te vergemakkelijken, en om best practices uit te wisselen, is er een implementatiehandleiding opgesteld. Deze handleiding is te vinden op www.edustandaard.nl.

7 november 2011

5. Vrijwaring gebruik afspraak

Hoewel deze afspraak met de grootst mogelijke zorg is opgesteld, kan (kunnen de rechtsopvolgers van) Edustandaard geen aansprakelijkheid aanvaarden voor de juistheid, volledigheid of bruikbaarheid van de inhoud van dit document.

De afspraak zal naar aanleiding van voortschrijdende inzichten en aanbevelingen van gebruikers aangepast kunnen worden. Eventuele kosten voortvloeiend uit deze aanpassingen zijn niet te verhalen op de (rechtsopvolgers van) Edustandaard of haar initiatiefnemers (de stichtingen SURFFoundation en Stichting Kennisnet).

De afspraak kan conform de beschreven doelstellingen worden gebruikt. Gebruik van de afspraak gebeurt voor risico van de gebruiker. Het auteursrecht van de afspraak ligt bij (de rechtsopvolgers van) EduStandaard.

De afspraak is vrij te verspreiden, te publiceren of te hergebruiken, mits de bron duidelijk vermeld wordt. Dit bestand valt onder de Nederlandse versie van de Creative Commons licentie "Naamsvermelding 3.0 Nederland". Zie <http://creativecommons.org/licenses/by/3.0/nl/>

7 november 2011

6. Bronnen

- Metadatastromen adviesrapportage,
http://digitaallemateriaal.kennisnet.nl/attachments/session=cloud_mmbase+2151909/Adviesrapportage_Metadatastromen.pdf
- Afspraak NL LOM, <http://wiki.surffoundation.nl/display/nllom/Home>
- PURL website, <http://www.purl.org>
- Handle website, <http://www.handle.nl>
- DOI website, <http://www.doi.org>
- URN:NBN:NL website, <http://wiki.surffoundation.nl/display/standards/URN-NBN>
- Educatieve contentketen website, <http://contentketen.kennisnet.nl>

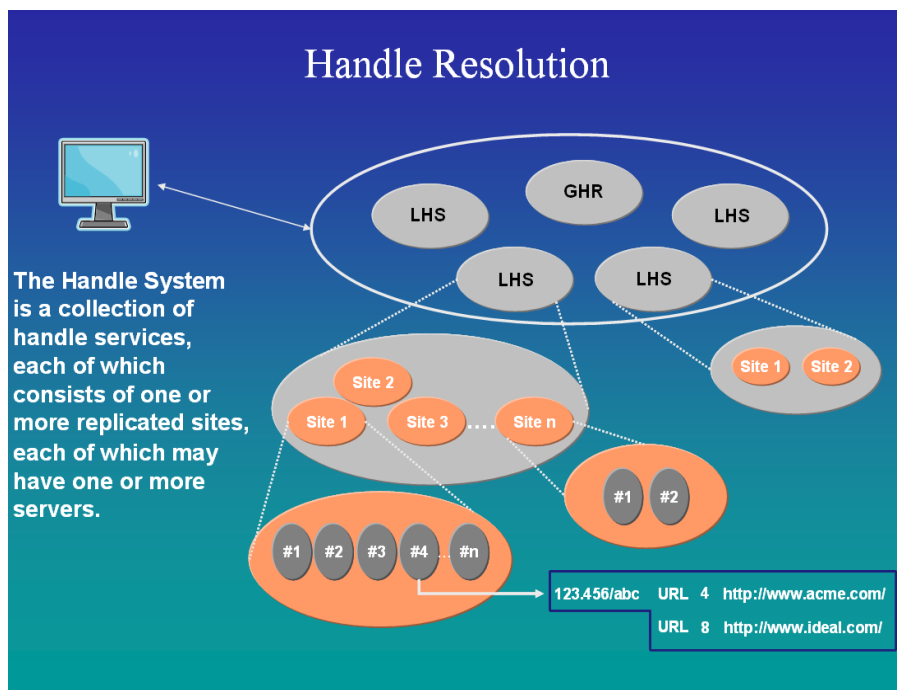
7 november 2011

7. Afkortingen / verklarende woordenlijst

Term/afkorting	Definitie
URN	Uniform Resource Name
URL	Uniform Resource Location
ISBN	Internationaal Standaard Boeknummer
ISSN	International Standard Serial Number
DOI	Digital Object Identifier
PURL	Persistent Uniform Resource Locator
prefix	Unieke code die wordt uitgegeven door een centrale organisatie. Deze code wordt voor een unieke code van een leerobject of metadatarecord wordt geplaatst om de totale identifier globaal uniek te maken.
eigenaar	De partij die als rechthebbende kan worden aangemerkt van een leerobject of metadatarecord
niet-eigenaar	Ieder partij die niet kan worden aangemerkt als rechthebbende van een leerobject of metadatarecord
leerobject	Een stuk materiaal dat door een lerende kan worden gebruikt om zijn kennis op een specifiek onderwerp te vergroten
metadatarecord	Bestand waarin relevante informatie over een leerobject op gestructureerde manier wordt beschreven. Voorbeelden van relevante informatie zijn de auteur van het leerobject en trefwoorden.
lerende	Een persoon die door middel van leermateriaal zijn kennis wil vergroten.
aanbieders	Partijen die leerobjecten beschikbaar stellen aan gebruikers
gebruikers	Een persoon of systeem dat gebruik maakt van aangeboden leerobjecten

Bijlage A: Handle structuur

Het Handle systeem is ontwikkeld door de Corporation for National Research Initiatives (CNRI) en is een infrastructuur welke kan worden gebruikt voor verschillende diensten. Handle wordt op dit moment al gebruikt om artikelen, boeken, rapporten, overheidsdocumenten, metadata, gedistribueerde leerobjecten en data sets te identificeren. Het Handle systeem wordt vaak vergeleken met het DNS systeem omdat het een soortgelijke opzet kent. Een verschil met het DSN systeem is dat het Handle systeem een twee laags hiërarchie kent in tegenstelling tot de meerlaags hiërarchie van DNS: de globale registry (GHR) en daaronder locale registries (LHS) die de prefixes afhandelen. De globale registry wordt door het CNRI onderhouden (en is uiteraard gedistribueerd over meerdere fysieke servers en locaties). De locale registries worden onderhouden door de partij die een prefix heeft aangevraagd. Uiteraard staat het deze partijen vrij om de fysieke servers en het beheer hiervan aan een derde partij uit te besteden, de partij die de prefix beheert is echter verantwoordelijk voor de prefix. In een plaatje ziet de structuur er als volgt uit:



Figuur 2: Handle systeem (uit een presentatie van Laurence Lannom, CNRI)

Het Handle systeem kent zogenaamde prefixes en derived prefixes. Een prefix is een globaal unieke code die wordt toegekend door het CNRI waaraan de Handle service (en de identifiers) geïdentificeerd kunnen worden. De prefix wordt gebruikt als het eerste deel van de identifier. Een Handle bestaat uit de prefix gecombineerd met een unieke code, met als separator een slash.

Een voorbeeld: "4263537/4000" is de Handle voor de Handle System website homepage. De Handle is gedefinieerd onder de prefix '4263537' en wordt een unieke persistente identifier omdat deze gecombineerd wordt met de unieke code voor de website homepage '4000'. Een prefix mag gebruikt worden om zogenaamde derived prefixes uit te geven. Zodra de prefix 4263537 is vastgelegd bij Handle is het mogelijk om ook de derived prefix 4263537.1 vast te leggen.

7 november 2011

Bijlage B: Voorbeeld metadata record met unieke persistente identifiers

```

<?xml version="1.0" encoding="UTF-8"?>
<lom xmlns="http://ltsc.ieee.org/xsd/LOM"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://ltsc.ieee.org/xsd/LOM
http://standards.ieee.org/reading/ieee/downloads/LOM/lomv1.0/xsd/lomLoose.xsd">
  <general>
    <identifier>
      <catalog>uri</catalog>
      <entry>hdl:4263537/4000</entry>
    </identifier>
    <title>(.)</title>
    <language>(.)</language>
    <description>(.)</description>
    <keyword>(.)</keyword>
    <coverage>(.)</coverage>
    <structure>(.)</structure>
    <aggregationlevel>(.)</aggregationlevel>
  </general>
  <lifeCycle>(.)</lifeCycle>
  <metaMetadata>
    <identifier>
      <catalog>uri</catalog>
      <entry>hdl:4263537/4012</entry>
    </identifier>
    <contribute>(.)</contribute>
    <metadataSchema>LOMv1.0</metadataSchema>
    <metadataSchema>nl_lom_v1p0</metadataSchema>
    <language>(.)</language>
  </metaMetadata>
  <technical>
    <format>(.)</format>
    <size>(.)</size>
    <location>http://example.org/file.abc</location>
    <requirement>(.)</requirement>
    <installationRemarks>(.)</installationRemarks>
    <otherPlatformRequirements>(.)</otherPlatformRequirements>
    <duration>(.)</duration>
  </technical>
  <educational>(.)</educational>
  <rights>(.)</rights>
  <relation>(.)</relation>
  <annotation>(.)</annotation>
  <classification>(.)</classification>
</lom>

```

Unieke
persistente
identifier
leerobject

Unieke
persistente
identifier
metadatareco
,

Technische locatie
leerobject

7 november 2011

Bijlage C: Uitwerkingen technische eisen unieke persistente identifiers

Deze bijlage beschrijft de technische infrastructuur die ingericht dient te worden bij het gebruik van de unieke persistente identifiers. De beschrijving voor Handle is te vinden in paragraaf 4.5.

Technische infrastructuur bij gebruik DOI

Iedere partij die gebruik wil maken van DOI **MOET** een eigen unieke prefix aan vragen bij een van de registry agencies van DOI (te vinden op www.doi.org). Op basis van de eigen prefix kunnen partijen vervolgens unieke persistente identifiers toekennen aan hun leerobjecten en metadatarecords (hiervoor mogen uiteraard ook al bestaande unieke interne codes worden gecombineerd met de prefix).

DOI is gebaseerd op het Handle systeem, bij een keuze voor DOI moeten daarom dezelfde zaken worden geregeld als bij een keuze voor Handle. Bij de keuze voor DOI is het daarom ook noodzakelijk dat er een resolver wordt ingericht die de mogelijkheid biedt voor gebruikers om op basis van de unieke persistente identifier de locatie van het leerobject of metadatarecord te vinden. Een eis van DOI is dat er voor iedere prefix een resolver beschikbaar moet zijn. De resolver die bij een prefix hoort is ervoor verantwoordelijk dat alle unieke persistente identifiers die onder de prefix zijn uitgegeven geresolved kunnen worden. Het is mogelijk dat één resolver meerdere prefixes afhandelt.

Een partij die gebruik maakt van DOI moet dus zorg dragen voor een resolver en moet iedere unieke persistente identifier die zij aanmaakt ook aanmelden bij die DOI resolver. Daarnaast moet zij bij wijzigingen in de locatie van het leerobject of metadatarecord deze wijzigingen doorgeven aan de DOI resolver (het betreft hier het beheer van de identifiers)

Technische infrastructuur bij gebruik ISBN/ISSN

Iedere partij die gebruik wil maken van ISBN **MOET** ervoor zorgen dat zij ISBN's registreert bij de daarvoor aangewezen uitgevende partij. Voor het Nederlandse taalgebied (Nederland en Vlaanderen) is dat het Bureau ISBN. Daarnaast moeten partijen bij het Bureau ISBN een uitgevers-prefix aanvragen (dit kan ook een eenmalige prefix zijn). Na een prefix te hebben aangevraagd is het mogelijk om ISBN's aan te vragen voor de leerobjecten. Bij het aanvragen moet een partij er ook voor zorgen dat ze de informatie over het leerobject opgeven bij het Bureau ISBN zodat de resolver de correcte informatie terug kan geven.

Om te voldoen aan de eis van resolvabiliteit kan bijvoorbeeld gebruik gemaakt worden van de Worldcat ISBN resolver. De partij die een ISBN registreert zal moeten controleren of de ISBN na uitgifte resolvable is. Indien dit het geval is via één van de beschikbare ISBN resolvers wordt er voldaan aan de eis van resolvabiliteit.

Technische infrastructuur bij gebruik URN

Iedere partij die gebruik wil maken van een URN **MOET** ervoor zorgen dat hij een URN kiest die resolvable is. Een voorbeeld van een resolvable URN is de URN:NBN:NL. Hiervoor is door de DANS een resolver⁸ opgezet. Bij een keuze voor een andere URN kan het zijn dat deze niet resolvable is, het staat partijen vrij om in dat geval zelf te zorgen voor een resolver.

Partijen die gebruik maken van een URN moeten ervoor zorgen dat de gebruikte unieke persistente identifier geregistreerd wordt bij de resolver van de URN en dat minimaal de informatie over de locatie van het leerobject of metadatarecord opgenomen is in de resolver.

⁸ Zie: <http://www.persistent-identifier.nl>

7 november 2011

Technische infrastructuur bij gebruik PURL

Iedere partij die gebruik wil maken van PURL als unieke persistente identifier **MOET** gebruik maken van een PURL-server die toegestaan is volgens de afspraak. Dit kan een bestaande PURL-server zijn die al is opgenomen in het overzicht van toegestane PURL-servers, echter is het ook mogelijk om een eigen PURL-server op te zetten. Deze moet dan voldoen aan de eisen die zijn opgenomen in deze afspraak. Of er wordt voldaan aan de eisen wordt gecontroleerd door de beheerders van (dan wel rechtsopvolgers van) EduStandaard.

Indien partijen een eigen PURL-server inrichten is het van belang dat zij de informatie in deze PURL-server up-to-date houden zodat de PURL blijft verwijzen naar de locatie van het leerobject of metadatarecord. Wijzigingen in de locatie van het leerobject of metadatarecord moeten dan ook door worden geven. De PURL moet dus beheerd worden.

7 november 2011

Bijlage D: Checklist voldoen aan afspraak

Om te voldoen aan de afspraak moeten de unieke persistente identifiers en bijbehorende systemen aan alle onderstaande eisen voldoen (deze eisen zijn overgenomen uit hoofdstuk 3 en 4).

Informatiekundige eisen aan de identifier

- Unicité: De unieke persistente identifiers die toegekend worden aan leerobjecten en metadatarecords **MOETEN** globaal uniek zijn.
- Unieke persistente identifiers **MOETEN** oneindig lang uniek zijn, ze mogen dus niet worden hergebruikt.
- Persistentie: Met persistentie bedoelen we dat de unieke persistente identifier in principe zo lang mogelijk, maar ten minste 5 jaar nadat het leerobject of het metadatarecord niet meer bestaat, beschikbaar **MOET** zijn in een resolver.
- Resolvabiliteit: Unieke persistente identifiers **MOETEN** in elk geval leiden naar de meest actuele:
 - 1) Locatie van het object
 - 2) Pagina met informatie over het leerobject of metadatarecord
- Als het leerobject of metadatarecord niet meer beschikbaar is (zie eis 3), **MOET** de resolver volgens optie 2 van eis 4 deze informatie verstrekken.
- Actionable: Unieke persistente identifiers **MOETEN** (eventueel met behulp van een resolver) via het http-protocol te benaderen zijn.
- Ieder leerobject dat in de keten wordt gebracht **MOET** minimaal één unieke persistente identifier toegekend krijgen.
- Aanbieders **MOETEN** voordat ze een leerobject in de keten brengen controleren of het leerobject niet al eerder in de keten is gebracht door een andere partij.
- Indien bij een leerobject meerdere identifiers bekend zijn **MOETEN** alle bekende identifiers mee worden gegeven in het metadatarecord dat aan de keten wordt aangeboden
- De unieke persistente identifier **MAG GEEN** semantische betekenis bevatten die noodzakelijk is om met het leerobject of het metadatarecord te kunnen werken.
- Indien een nieuw leerobject wordt aangemaakt **MOET** er een nieuwe metadatarecord worden aangemaakt met een nieuwe unieke persistente identifier.
- Indien een partij versiebeheer ingericht heeft voor leerobjecten **MOET** zij in het metadatarecord de relatie tussen de verschillende versies van leerobjecten vastleggen.
- Voor unieke persistente identifiers van leerobjecten geldt dat het beheer van een unieke persistente identifier die is uitgegeven door een niet-eigenaar overgenomen **MOET** kunnen worden door de eigenaar van het leerobject.

Informatiekundige eisen aan de systemen

- Systemen **MOETEN** unieke persistente identifiers toekennen aan leerobjecten en **MOETEN** deze unieke persistente identifiers ook vastleggen bij het leerobject.
- Systemen die prefixes uitgeven **MOGEN NIET** een zelfde prefix uitgeven aan verschillende identifieruitgifte systemen van één type identifier.
- Systemen van niet-eigenaren **MOETEN** het beheer van een identifier die uitgegeven is door een niet-eigenaar overdragen zodra de eigenaar het beheer van de identifier wenst over te nemen.

7 november 2011

Technische eisen aan de identifier

- De unieke persistente identifier **MOET** gebruik maken van de karakterset die voor een URI is toegestaan.
- Bij het opnemen van de unieke persistente identifiers in het metadatarecord **MOET** de officiële notatiewijze van ieder identifier type worden gevolgd.
- De specifieke resolver die gebruikt kan worden om de unieke persistente identifier te resolveren **MAG** voor de unieke persistente identifier types Handle, DOI, ISBN/ISSN en URN **NIET** opgenomen worden in het metadatarecord. Bij het gebruik van PURL is het **VERPLICHT** om de volledige PURL op te nemen.
- De unieke persistente identifier **MAG GEEN** semantische verwijzingen bevatten. Er mogen dus bijvoorbeeld geen bedrijfsnamen of namen van spotter-systemen in opgenomen worden.

Technische eisen aan de systemen

- De resolver **MOET** publiek beschikbaar zijn.
- Wijzigingen in locatie of eigenaar van een leerobject of metadatarecord **MOETEN** doorgegeven worden aan de resolver.
- De resolver **MOET** in elk geval één van de volgende zaken teruggeven:
 - 1) Locatie van het object
 - 2) Een pagina met informatie over het leerobject of metadatarecord