

DARE Use of Persistent Identifiers



Table of contents

- Table of contents
 - Document information
 - Document History
 - Abstract
 - Werkgroep
1. Managementsamenvatting
 2. Probleemstelling
 3. Eisen en randvoorwaarden
 - OAI-PMH 2.0 protocol
 - W3C-standaard
 - URN
 - Functionele randvoorwaarden DARE
 4. Mogelijke oplossingen
 - DOI/Handle implementatie
 - Mogelijke oplossing binnen de url/handle/doi-implementatie zou kunnen verlopen via een Third party Registration Agency, zoals CrossRef.
 - Conclusie
 - URN implementatie
 - NBN
 - Nadere uitwerking: Het harvest-model:
 - Toepassing voor repositories
 - Fasering
 - Conclusie
 5. Conclusies
 6. Literatuur
- Bijlage 1 Begrippen
- Bijlage 2. Voorstel voor gefaseerde implementatie URN:NBN in DAREverband.
- Basis Functionaliteit (realisatie uiterlijk 1.1.2008)
 - A. We moeten overeenstemming krijgen over de URN.
 - B. Er moet een centrale resolver worden ontwikkeld. Verantwoordelijk: DANS
 - C. Iedere lokale repository moet aan ieder object een persistente identifier toekennen.
 - D. Darnet resolving via centrale resolver
 - Extra functionaliteit (realisatie na 1.1.2008)
 - 1. Volledig overgaan op afgesproken persistente identifiers
 - 2. Volledig ondersteunen van 'verhuizen van objecten'
 - 3. Eigenaarschap en meerdere lokaties van een object
 - 4. Lokale resolving
 - 5. Wereldwijde resolving



Document information

Title: DARE 2 SHARE
Persistence
Subject: Cool URI's don't change
Moderator:
Version: 0.6
Date published: 2007-04-05
Excerpt: Write an excerpt here

(Optional information)

Type:
Format:
Identifier:
Language: NL
Rights:
Tags:

Document History

Date	Version	Owner	Changelog	PDF
2007-04-05	0.6		1. In het schema: "Huidige Identifiers voor "information assets"" (pg 5) zijn wijzigingen opgenomen voor de Technische Universiteit Eindhoven en de Rijks Universiteit Groningen. 2. Op pg 7 is een betere verwijzing opgenomen naar: Hans-Werner Hilse and Jochen Kothe, Implementing Persistent Identifiers [Hilse/Kothe, 2006]. 3. Op de kaft is het DARE-logo opgenomen	

Abstract

The abstract describes what the application profile is about. It should contain a problem definition, the standards described by the application profile and the goal of the application profile.

*What makes a cool URI?
A cool URI is one which does not change. What sorts of URI change?*

URIs don't change: people change them.

(c)1998 Tim Berners-Lee

<http://www.w3.org/Provider/Style/URI.html>

Werkgroep

De werkgroep persistente identifiers was als volgt samengesteld:

- Niels Rozenboom, Universiteit Leiden
- Maarten Hoogerwerf, DANS
- Peter Ruijgrok, Universiteit Utrecht Laurens Sesink, DANS
- Hubert Krekels, Wageningen Universiteit
- Maurice Vanderfeesten, SURF
- Marlon Domingus, Universiteit Leiden (vz)

1. Managementsamenvatting

Momenteel is er binnen de Nederlandse instellingen die participeren in DARE geen uniforme manier voor de implementatie van persistente identifiers van de metadata records binnen de afzonderlijke institutionele repositoria. Gerelateerd aan metadata records wordt persistentie van identifiers hier bedoeld in de betekenis van zowel 'naming' als 'resolving'.

Resolving van de digitale objecten (bitstreams), gekoppeld aan deze metadata, gebeurt vooralsnog in het geheel niet op een wijze die persistent is. Hoewel binnen het oai-pmh-protocol alleen de eis geldt dat een verwijzing naar een metadata-record persistent is, wordt in dit rapport ook de persistentie. Er zijn verschillende redenen waarom het voor de hand ligt om zeker op een nationale schaal, maar beter nog om op een internationale schaal uniformiteit na te streven op het gebied van persistentie van identifiers. Beheer(s)baarheid is een belangrijk argument bij de betrokken instellingen vanuit hun rol als (oai)data-provider, maar ook vanuit hun rol als (oai)service-provider. Vanuit DARE is er het belang om naar robuustheid te streven van de IR-gerelateerde nationale infrastructuur. Tenslotte is het voor de eindgebruiker wenselijk om te kunnen vertrouwen op de robuustheid en betrouwbaarheid van DARE-diensten met als zwakste schakel juist de naming en resolving van resources binnen de DARE-services en data.

Met persistente identificatie in DARE-verband wordt technische gezien feitelijk bedoeld: identificatie van resources (zowel metadata en bitstreams) op een persistente, (server)locatie-onafhankelijke manier (lees: urn-implementatie, geen url-implementatie).

Om dit te implementeren is naast techniek een model nodig. NBN is hiervoor een voor de hand liggende kandidaat. In dit rapport wordt een voorstel gedaan voor een identifier van het type: URN:NBN:

```
URN:NBN:landcode:identificerende_string
```

Het heeft als voordeel op het DOI-alternatief dat het gebaseerd is op standaarden en niet op een concrete toepassing van een leverancier. De werkgroep beschouwt de URN:NBN-oplossing als de meest toekomstvaste en generieke oplossing. De werkgroep verkeert tevens in de veronderstelling dat alle betrokken partijen met een minimum aan implementatie-inspanning de vruchten kunnen dragen van deze oplossing. Hierbij vervullen de Koninklijke Bibliotheek en DANS mogelijk een cruciale rol, respectievelijk als NBN-registry/agency en ontwikkelaar nationale URN:NBN-resolver. Tevens stelt de werkgroep een gefaseerde implementatie voor, zoals nader toegelicht in Bijlage 2.

Tenslotte beveelt de werkgroep aan om na een nationale oplossing samen met Duitsland en eventuele andere landen te kijken naar de mogelijkheid van een global URN:NBN-resolver. M. Domingus

2. Probleemstelling

Momenteel is er binnen de Nederlandse instellingen die participeren in DARE geen uniforme manier voor de implementatie van persistente identifiers van de metadata records binnen de afzonderlijke institutionele repositories. Gerelateerd aan metadata-records wordt persistentie van identifiers hier bedoeld in de betekenis van zowel 'naming' als 'resolving'. (Zie voor de definities van de in dit rapport gebruikte begrippen in Bijlage 1: Begrippen.)

Resolving van de digitale objecten (bitstreams), gekoppeld aan deze metadata, gebeurt vooralsnog in het geheel niet op een wijze die persistent is. Hoewel binnen het oai-pmh-protocol alleen de eis geldt dat een verwijzing naar een metadata-record persistent is, wordt in dit rapport ook de persistentie. Er zijn verschillende redenen waarom het voor de hand ligt om zeker op een nationale schaal, maar beter nog om op een internationale schaal uniformiteit na te streven op het gebied van persistentie van identifiers. Beheer(s)baarheid is een belangrijk argument bij de betrokken instellingen vanuit hun rol als (oai)data-provider, maar ook vanuit hun rol als (oai)service-provider. Vanuit DARE is er het belang om naar robuustheid te streven van de IR-gerelateerde nationale infrastructuur.

Tenslotte is het voor de eindgebruiker wenselijk om te kunnen vertrouwen op de robuustheid en betrouwbaarheid van DARE-diensten met als zwakste schakel juist de naming en resolving van resources binnen de DARE-services en data. Vanuit DARE zijn de volgende doelstellingen geformuleerd (Maurice Vanderfeesten, Identifier Impact van huidige situatie. November 2006.) die gerealiseerd zouden moeten worden met een persistent identifier nieuwe stijl:

- **Fundament voor kennisinfrastructuur aanleggen die robuust is en toekomstbestendig.** De lapjesdeken aan identifiers en de organisatorische afspraken die er nu liggen, zijn de meeste instellingen het over eens dat deze niet robuust genoeg is, laat staan toekomst vast is voor de zeer lange termijn.
- **Basis creëren voor eindgebruiker om snel bij het informatie object te komen.** Of dit nou rechtstreeks is, of via een jump-off pagina mag de instelling zelf weten. Op dit moment is er in Nederland niet eenduidige afspraak hierover. De een gebruikt een url, de ander een resolutie service elders.
- **Streven naar een pragmatische oplossing die ook te implementeren valt.** Zoals we nu zien is er bij de invoering van de repository gekozen voor een zeer pragmatische oplossing. Het streven was om zo snel mogelijk een DAREnet op te bouwen dus er was geen tijd om te beslissen wat te doen met de identifier en werd gekozen voor het meest voor de hand liggende.
- **Streven naar meer uniformiteit om informatie objecten te benoemen.** Door de diversiteit aan identifiers is er moeilijk te spreken over een uniforme regel. In de OAI header wordt een oai urn-achtige identifier gebruikt. Deze wordt zelf verzonden door de instelling zelf en heeft geen binding met een resolver. Op dit moment is er in Nederland niet eenduidige identifier syntax.
- **Streef naar een URL onafhankelijke oplossing.** Op dit moment worden er door alle instellingen URL's gebruikt om het object te identificeren en te localiseren. De URL wordt samen met de identifier aangeboden.

Voor wat dit laatste punt betreft zie ook: **What is wrong with URLs?**[§ 1.3 Hilse/Kothe, 2006]

So what is wrong with the currently used URL-approach? We already mentioned that from a user's point of view links appear to be 'broken'. This is mostly due to one of the following:

1. (a) *the document is no longer available, or the document is available, but was*
2. (b) *relocated so that it is now accessible using a different domain name,*
3. (c) *relocated on the same Internet server so that it is now accessible with an other file system path or query specification.*

Tevens is door DARE een inventarisatie gemaakt van de huidige stand van zaken van implementaties van identifiers (ibidem):

Huidige Identifiers voor "information assets"

Instelling	Systeem	Type identifier	Toelichting
Amsterdam University Press	ARNO	db volgnummers + persistentie afspraken (niet resolvable)	ARNO's kunnen elk bepalen welke URL ze gebruiken. Wel kun je bij elke ARNO aan de URL zien wat van van wat voor type het is (jop/obj)
Erasmus Universiteit Rotterdam	Dspace	Handle	
Fonryts Hogescholen	ARNO	db volgnummers + persistentie afspraken (niet resolvable)	ARNO's kunnen elk bepalen welke URL ze gebruiken. Wel kun je bij elke ARNO aan de URL zien wat van van wat voor type het is (jop/obj)
KNAW	i-Tor	db volgnummers	
NWO	i-Tor	URLs zijn "Cool URIs". Ze veranderen niet	
Open Universiteit Nederland	Dspace	Handle	
Radbout Universiteit Nijmegen	Dspace	Handle	
Rijks Universiteit Groningen	Wildfire	intern PNN en DBI	
Technische Universiteit Delft	CMS+OAI+PMHT oolkit	In Delft gebruiken ze in de OAI- repository de OAI-identificer. Bijvoorbeeld oai:tudelft.nl:001234 waarbij 001234 een uniek nummer is uit ons CMS	
Technische Unversiteit Eindhoven	Eigen ontwikkeling op basis van het Vubissysteem	Vubis recordnummer. In de geharveste records persistente URL	
Universiteit Leiden	Dspace	Handle	
Universiteit Maastricht	ARNO	URL die in identificer zit, direct naar de file danwel jump off page in de repository wijst in plaats van naar een tussenliggende resolver.	
Universiteit Twente	ARNO	db volgnummers + persistentie afspraken (niet resolvable)	ARNO's kunnen elk bepalen welke URL ze gebruiken. Wel kun je bij elke ARNO aan de URL zien wat van van wat voor type het is (jop/obj)
Universiteit Utrecht	Dspace	Handle	
Universiteit van Amsterdam	ARNO	een systeem met stabiele (persistente) URL's wat ze zelf hebben ontworpen	
Universiteit van Tilburg	ARNO	db volgnummers + persitentie afspraken (niet resolvable)	ARNO's kunnen elk bepalen welke URL ze gebruiken. Wel kun je bij elke ARNO aan de URL zien wat van wat voor type het is (jop/obj)
Vrije Universiteit Amsterdam	Dspace	Handle	
Wageningen Unversitent & ResearchCentrum	CMS+OAI+PMHT oolkit	oai:domein:repository/db volgnummer ; persitent technisch:ja organisationeel:nee; niet resolvable	

3. Eisen en randvoorwaarden

Voor de wenselijke oplossing van een te implementeren identifier zijn er een aantal typen (technische) randvoorwaarden: vanuit het OAI-PMH 2.0 protocol, de oplossing dient een W3C-standaard te zijn nu en in ieder geval in de voorzienbare toekomst en er zijn een aantal functionele randvoorwaarden waar vanuit de in DARE participerende instellingen op zal moeten worden geanticipeerd en worden hieronder door de werkgroep voorgesteld. Hieronder volgt een opsomming van de eisen en randvoorwaarden.

OAI-PMH 2.0 protocol

Vanuit het OAI-PMH 2.0 protocol gelden 2 eisen voor de identifier:

1. A unique identifier unambiguously identifies an item within a repository; the unique identifier is used in OAI-PMH requests for extracting metadata from the item.
2. The format of the unique identifier must correspond to that of the URI (Uniform Resource Identifier) syntax.

W3C-standaard

URN

All URNs have the following syntax (phrases enclosed in quotes are REQUIRED):

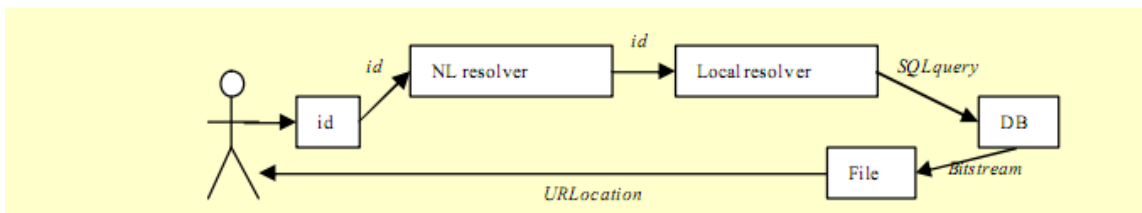
`<URN> ::= "urn:" <NID> ":" <NSS>`

where <NID> is the Namespace Identifier, and <NSS> is the Namespace Specific String. The leading "urn:" sequence is case-insensitive.

Functionele randvoorwaarden DARE

1. Er dient een unieke identifier te zijn voor elke metadata-record dan wel een jump off page. Deze identifier dient zowel te identificeren als toegang te verschaffen tot de resource (in dit geval zijnde de metadata)
2. Er dient per bitstream een unieke identifier te zijn. Deze identifier dient zowel te identificeren als toegang te verschaffen tot de resource (in dit geval zijnde de bitstream)
3. De unieke identifier heeft een syntax die verhuizingen van de resource (zowel metadata als bitstream(s)) toestaat zonder dat een verhuizing om redenen van identificatie of resolving een wijziging van de syntax van de identifier vereist.
4. de resolving van de identifier, dus het toegang verlenen tot de betreffende resource (zowel metadata als bitstream(s)) dient op generieke wijze te geschieden.
5. Om na verhuizing van een combinatie van metadata-record en bijbehorende

bitstream(s) de resolving goed te laten plaatsvinden is er een resolver buiten de specifieke in DARE participerende partijen vereist, die de rol van nationale resolver vervult voor DARE-doeleinden. Zie ook een schematische voorstelling van deze situatie hieronder.



4. Mogelijke oplossingen

Technische gezien komen eigenlijk alleen twee oplossingen in aanmerking voor het beschreven probleem en de bijbehorende randvoorwaarden: URN en DOI/Handle.

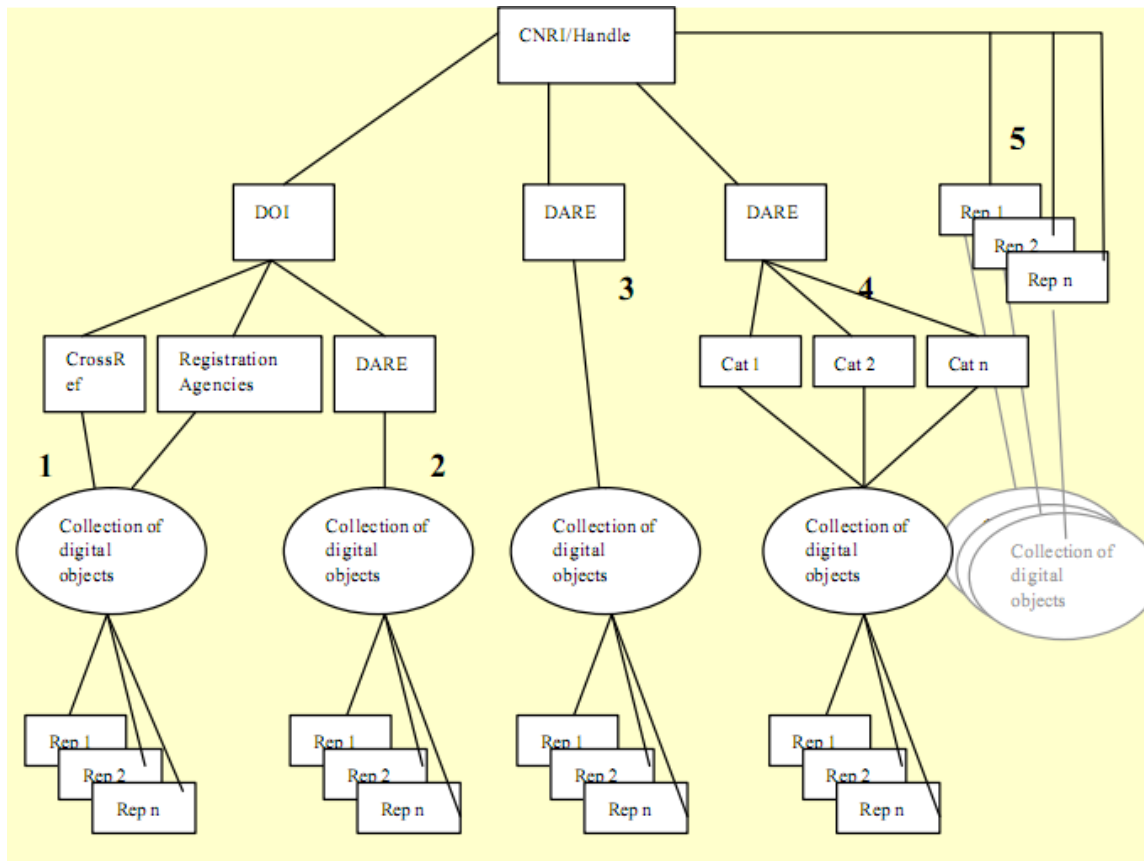
Hieronder wordt geschetst hoe deze oplossingen zouden kunnen werken en wat de sterke / zwakke punten zijn van de oplossing. (Zie voor een zeer uitgebreid en volledig overzicht van verschillende benaderingen van implementatie van persistente identifiers ook: Hans-Werner Hilse and Jochen Kothe, Implementing Persistent Identifiers [Hilse/Kothe, 2006].)

DOI/Handle implementatie

Een digital object identifier (doi), zoals veelal gebruikt door uitgeverij om een specifieke publicatie te identificeren, is in feite een implementatie van een handle. Een handle is een wereldwijde url-based implementatie van een secure name resolution. Het Handle Systeem kent een prefix toe aan een partij die handles mag toekennen aan (in de regel) metadata en zorgt voor een wereldwijde resolving van die url in handle -syntax naar de (server)locatie die door de betreffende handle-toewijzer is De syntax van een handle is: <Handle> ::= <Handle Prefix> "/" <Handle Suffix> Voorbeeld: "10.1045/april2006-paskin" is an identifier (also known as a Digital Object Identifier (DOI), an implementation of the Handle System) for an article published in D-Lib Magazine. It is defined under the prefix (naming authority) "10.1045", and its suffix (local name) is "april2006-paskin". Bron: <http://www.handle.net/overviews/handle-syntax.html>

Voor meer informatie zie ook Bijlage 1: "Begrippen".

Hieronder is een visualisatie van de beschreven opties bij gebruik van het DOI/Handle systeem.



Mogelijke oplossing binnen de url/handle/doi-implementatie zou kunnen verlopen via een Third party Registration Agency, zoals CrossRef.

Deze optie gaat uit van een Third party Registration Agency, zoals CrossRef. Het voordeel is dat met een dergelijke RA al het werk uit handen wordt genomen. CrossRef zorgt namelijk dat je op eenvoudige wijze een persistent-identifier aan je wetenschappelijke content kunt koppelen. CrossRef kan worden gezien als een betrouwbare en ervaren organisatie die DOI's registreren. Verder biedt deze organisatie services, zoals het gebruik van OpenURL, reference linking en biedt het een doorzoekbare metadata database. Het nadeel is dat we er ook naar betalen. RA's zijn vrij om hun eigen prijs te bepalen. RA's betalen op hun beurt een paar cent per DOI naam aan de IDF. CrossRef rekent voor local hosting \$20.000, plus jaarlijks \$0.06 - \$1.00 per DOI.

Wil een repository zich aansluiten bij CrossRef dan zijn een van de eisen van CrossRef dat de repository de eigenaar is van het artikel. Dit is vaak niet het geval met Open Access, bij uitgeverij vaak juist wel. Verder gaat CrossRef uit van full-text werken en niet van datasets, video's, audio bestanden etc.. Het is overigens de vraag of CrossRef zal toestaan dat voor DSpace gebruikers is het overigens relatief eenvoudig handles om te zetten naar DOI door een DOI-prefix / DSpace-prefix / DSpace-suffix. <http://hdl.handle.net/1887/4550> wordt bijvoorbeeld <http://dx.doi.org/10.1575/1887/4550>.

Conclusie

Deze optie heeft voordelen: er is een bestaande technische infrastructuur en er is een organisatie die de relevante diensten kan beheren en als toekennende instantie kan functioneren. Vraag is echter of alle gewenste resources (zowel metadata als digitale objecten) een doi toegekend zullen kunnen krijgen en wat de mate van verwarring bij de eindgebruiker is als aan hetzelfde digitale object meerdere doi's worden toegekend (bijvoorbeeld een vanwege een uitgever en een andere vanwege een repository).

URN implementatie

Een ander alternatief is een urn-implementatie. De werkelijke uitdaging zou zijn de urn, die reeds goede diensten heeft bewezen als identifier sec, ook resolvable te maken. In Duitsland is daar reeds enige tijd goede ervaring mee en is er op nationale schaal zelfs al een urn-implementatie in productie. Zie bijvoorbeeld: Een repository in Kassel: <https://kobra.bibliothek.uni-kassel.de/> Kent een identifier toe aan een publicatie, bijvoorbeeld:

urn:nbn:de:hebis:34-2007032817560

dat op nationale schaal wordt geresolved door de Deutsche Nationalbibliothek via:

<http://nbn-resolving.de/resolverdemo-eng.php/>

Het voorbeeld levert of een url:



Of het document in het repository in Kassel.

NBN

Binnen de Duitse URN-implementatie is er tevens gebruik gemaakt van het National Bibliography Number (NBN). NBN is geen internetstandaard, maar een manier om unieke en persistente identifiers te gebruiken zoals toe te kennen door nationale bibliotheken. Zie ook: <http://www.rfc-archive.org/getrfc.php?rfc=3188/> Voor de syntax van NBN is in 2001 voorgesteld:

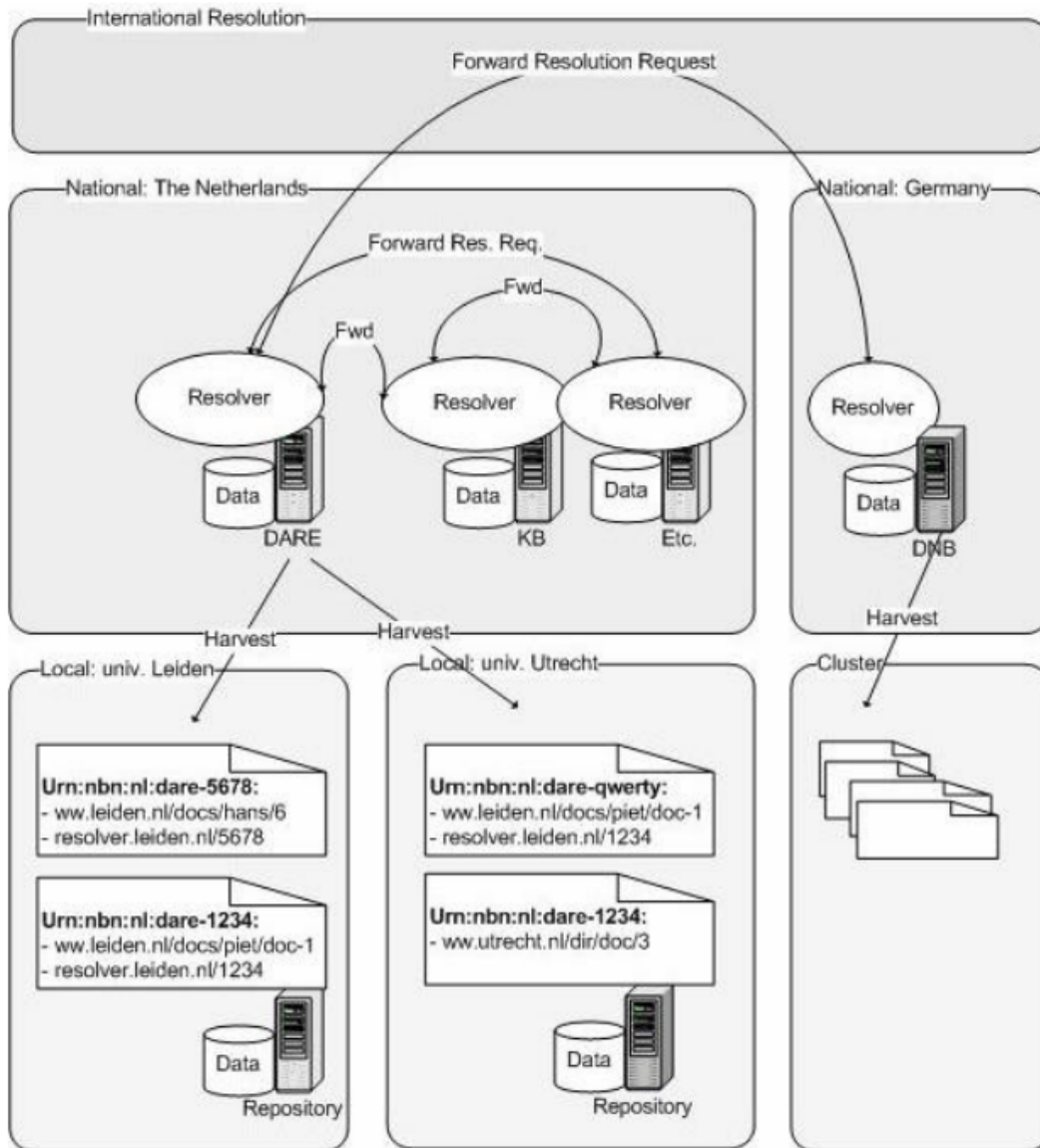
*Declaration of syntactic structure: *The namespace specific string will consist of three parts: prefix, consisting of either a two-letter ISO 3166 country code or other registered string and sub-namespace codes, delimiting characters (colon (:), or hyphen (-), and NBN string assigned by the national library.

NBN is een toepassing binnen het URN-framework. De combinatie van URN en NBN inclusief een urn-resolver biedt dus een goede oplossing voor een DARE identifier, zoals omschreven. Voorwaarden hierbij zouden zijn dat de implementatie van een urn-resolver bij een betrouwbare partner wordt ondergebracht en dat de Koninklijke Bibliotheek als NBN toekenner kan en wil functioneren.

Data Archiving and Networked Services (DANS) - <http://www.dans.knaw.nl/nl/> heeft aangegeven om onder bepaalde voor de hand liggende condities de ontwikkeling van een nationale URN-resolver voor zijn rekening te willen nemen. Ook het beheer van deze nationale resolver zou een taak van DANS kunnen worden.

De Koninklijke Bibliotheek zou binnen Nederland de rol van NBN-registrant / toekenner kunnen vervullen. De toekenning van NBN's door de KB zou voortvloeien uit de individuele instellingen [in DARE participerende instellingen] moeten worden gerealiseerd en met de te ontwikkelen nationale URN:NBN-resolver. De KB zou hiermee een belangrijke innovatie faciliteren op het gebied van identificatie en resolving van Nederlandse resources.

Voor een schematisch overzicht van een URN:NBN-oplossing op nationale schaal en zelfs wereldwijde schaal zie:



Nadere uitwerking: Het harvest-model:

Door middel van een apart MetadataPrefix kan via OAI informatie over identifiers (bijv. urn, locatie en status) worden geharvest. Voordelen van deze methode zijn:

- het OAI-protocol is bij de verschillende repositories reeds geïmplementeerd
- het is reeds toegepast door de Duitse nationale bibliotheek
- het combineert het gemak van decentraal beheer / verantwoordelijkheid met centrale informatie. Een snelle greep uit de voordelen:
 - Decentraal worden identifiers bijgehouden en geupdate
 - Centraal wordt geresolved
 - Centraal worden applicaties ontwikkeld
 - Centraal worden verhuizingen gedetecteerd

Op dit moment bestaat er geen Global URN resolutie mechanisme. Wel ziet IANA [Internet Assigned Numbers Authority - <http://www.iana.org> wereldwijd toe op het toekennen van URN namespaces 'for Public Identifiers' - zie: <http://www.iana.org/assignments/urn-namespaces/> (NBN is bij IANA geregistreerd als een van deze public identifiers).

Toepassing voor repositories

Vorstel: gebruik als generieke identifier de urn met gebruikmaking van NBN met bijvoorbeeld de syntax

URN:NBN:landcode:identificerende_string

Deze oplossing voldoet nl. aan alle randvoorwaarden zoals gedefinieerd (toelichting vereist inzake toepassing NBN) en is niet gebonden aan een specifieke implementatie van een leverancier (DOI/Handle), maar is gebaseerd op algemene standaarden. Nadere opmerkingen mbt de syntax van een urn (Met dank aan Lucas van Schaik voor nuttige ideeën en adviezen op dit punt.)

- identificerende_string: laat deze bijvoorbeeld eindigen op een controlegetal
- gebruik na de landcode geen instelling-specifieke code, dit ivm naming- en resolvingkwaliteiten na verhuizingen van de metadata/bitstreams
- gebruik als codering voor de landcode bijvoorbeeld de ISBN-landcodes of een andere ISO-
- landcode (2 dan wel 3 digits)
- wellicht is het nuttig om in de syntax een aanduiding aan te brengen van de aard van de resource die met de urn wordt geïdentificeerd.

Fasering

Het ligt voor de hand de implementatie van een URN:NBN te faseren. Zie voor een voorstel daartoe Bijlage 2. In essentie zou je eerst URN:NBN willen implementeren voor de metadata en pas daarna voor bitstreams. Vervolgens zou je ook de implementatie van URN:NBN voor metadata kunnen faseren door eerst de toekenning en presentatie van deze identifiers te realiseren en pas daarna de volledige resolving functionaliteit.

Conclusie

NBN:URN voldoet aan alle voorwaarden die aan de te ontwikkelen DARE-identifier zijn

gesteld. Er is reeds ervaring met een URN:NBN-oplossing op nationale schaal, voor zover het metadata betreft. Met de juiste partijen zouden ook de bitstreams kunnen worden geïdentificeerd en geresolved met de URN:NBN-identifier.

5. Conclusies

Met persistente identificatie in DARE-verband wordt technische gezien feitelijk bedoeld:

identificatie van resources (zowel metadata en bitstreams) op een persistente, (server)locatie-onafhankelijke manier (lees: urn-implementatie, geen url-implementatie). Om dit te implementeren is naast techniek een model nodig. NBN is hiervoor een voor de hand liggende kandidaat.

In dit rapport wordt een voorstel gedaan voor een identifier van het type: URN:NBN. Het heeft als voordeel op het DOI-alternatief dat het gebaseerd is op standaarden en niet op een concrete toepassing van een leverancier. De werkgroep beschouwt de URN:NBN-oplossing als de meest toekomstvaste en generieke oplossing. De werkgroep verkeert tevens in de veronderstelling dat alle betrokken partijen met een minimum aan implementatie-inspanning de vruchten kunnen dragen van deze oplossing. Hierbij vervullen de Koninklijke Bibliotheek en DANS mogelijk een cruciale rol, respectievelijk als NBN-registry/agency en ontwikkelaar nationale URN:NBN-resolver

Tevens stelt de werkgroep een gefaseerde implementatie voor, zoals nader toegelicht in Bijlage 2.

Tenslotte beveelt de werkgroep aan om na een nationale oplossing samen met Duistland en eventuele andere landen te kijken naar de mogelijkheid van een global URN:NBN-resolver.

6. Literatuur

[Hakala, 2003] Hakala, Juha, (NBN), *Identification of Electronic Resources*, 2003, <http://metadata-wg.mannlib.cornell.edu/forum/2003-01-29/juhahakala.ppt/>

[Hoogerwerf, 2006] Hoogerwerf, Maarten, *Persistent Identification and linking of Datasets*. Student Thesis, 2006

[Paskin, 2006] Paskin, Norman, (DOI) *Naming And Meaning: Key To The Management Of Intellectual Property In Digital Media*, 2006, 0609079IPDM

[Weibel, 2006] Weibel, Stuart, (OCLC), <http://weibel-lines.typepad.com/> And various e-mail and telephone contact.

[Hilse/Kothe, 2006] Hans-Werner Hilse and Jochen Kothe, *Implementing Persistent Identifiers*
Published by Consortium of European Research Libraries, London, <http://www.cerl.org/>
European Commission on Preservation and Access, Amsterdam, <http://www.knaw.nl/ecpa/> November 2006. ISBN 90-6984-508-3
Available as PDF at <http://www.cerl.org/> and <http://www.knaw.nl/ecpa/> Identifier urn:nbn:de:gbv:7-isbn-90-6984-508-3-8 (resolving service <http://nbn-resolving.de/>)
(<http://nbn-resolving.de/urn:nbn:de:gbv:7-isbn-90-6984-508-3-8/>)

[OAIPMH 2.0] *The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14. Document Version 2004/10/12T15:31:00Z*
<http://www.openarchives.org/oai/2.0/openarchivesprotocol.htm>

Tim Berners-Lee, Roy T. Fielding, Larry Masinter. (January 2005). "Uniform Resource Identifier (URI): Generic Syntax". Internet Society. RFC 3986; STD 66. <http://gbiv.com/protocols/uri/rfc/rfc3986.html>

Bijlage 1 Begrippen

Harvester [OAIPMH-definitie]

A harvester is a client application that issues OAI-PMH requests. A harvester is operated by a service provider as a means of collecting metadata from [repositories](#).

Zie: OAI-PMH 2.0: <http://www.openarchives.org/oai/openarchivesprotocol.html#definitionsconcepts/http://www.openarchives.org/OAI/openarchivesprotocol.html#DefinitionsConcepts>

Repository [OAIPMH-definitie]

A repository is a network accessible server that can process the 6 OAI-PMH requests in the manner described in this document. A repository is managed by a data provider to expose metadata to [harvesters](#). To allow various repository configurations, the OAI-PMH distinguishes between three distinct entities related to the metadata made accessible by the OAI-PMH.

*resource - A resource is the object or "stuff" that metadata is "about". The nature of a resource, whether it is physical or digital, or whether it is stored in the repository or is a constituent of another database, is outside the scope of the OAI-PMH.

*item - An [item](#) is a constituent of a repository from which metadata about a resource can be disseminated. That metadata may be disseminated on-the-fly from the associated resource, cross-walked from some canonical form, actually stored in the repository, etc.

*record - A [record](#) is metadata in a specific [metadata format](#). A record is returned as an XML-encoded byte stream in response to a protocol request to disseminate a specific metadata format from a constituent item.

Zie: OAI-PMH 2.0: <http://www.openarchives.org/oai/openarchivesprotocol.html#definitionsconcepts/>

<http://www.openarchives.org/oai/openarchivesprotocol.html#definitionsconcepts/Item> [OAIPMH-definitie]

An item is a constituent of a repository from which metadata about a resource can be disseminated. An item is conceptually a container that stores or dynamically generates metadata about a single resource in multiple formats, each of which can be harvested as [records](#) via the OAI-PMH. Each item has an [identifier](#) that is unique within the scope of the repository of which it is a constituent.

Zie: OAI-PMH 2.0: <http://www.openarchives.org/oai/openarchivesprotocol.html#definitionsconcepts/>

Unique Identifier [OAIPMH-definitie]

A unique identifier unambiguously identifies an item within a repository; the unique identifier is used in OAI-PMH requests for extracting metadata from the item. Items may contain metadata in [multiple formats](#). The unique identifier maps to the item, and all possible [records](#) available from a single item share the same unique identifier.

The format of the unique identifier must correspond to that of the [URI \(Uniform Resource Identifier\)](#) syntax. Individual communities may develop community-specific [URI](#) schemes for coordinated use across repositories. The scheme component of the unique identifiers must not correspond to that of a recognized URI scheme unless the identifiers conform to that scheme.

Repositories may implement the [oai-identifier](#) syntax described in the accompanying [Implementation Guidelines](#) document.

Unique identifiers play two roles in the protocol:

1. Response: Identifiers are returned by both the [ListIdentifiers](#) and [ListRecords](#) requests.

2. Request: An identifier, in combination with a [metadataPrefix](#), is used in the [GetRecord](#) request as a means of requesting a [record](#) in a specific metadata format from an item.

Note that the identifier described here is not that of a resource. The nature of a resource

identifier is outside the scope of the OAI-PMH. To facilitate access to the resource associated with harvested metadata, repositories should use an element in metadata records to establish a linkage between the record (and the identifier of its item) and the identifier (URL, URN, DOI, etc.) of the associated resource. The mandatory Dublin Core format provides the identifier element that should be used for this purpose.

Zie: OAI-PMH 2.0: <http://www.openarchives.org/oai/openarchivesprotocol.html#definitionsconcepts/>

Record [OAIPMH-definitie]

A record is metadata expressed in a single format. A record is returned in an XML-encoded byte stream in response to an OAI-PMH request for metadata from an item. A record is identified unambiguously by the combination of the [unique identifier](#) of the item from which the record is available, the [metadataPrefix](#) identifying the metadata format of the record, and the [datestamp](#) of the record. The XML-encoding of records is organized into the following parts:

- *header* - contains the unique identifier of the item and properties necessary for selective harvesting. The header consists of the following parts:
 - the [unique identifier](#) - the unique identifier of an item in a repository;
 - the [datestamp](#) - the date of creation, modification or deletion of the record for the purpose of [selective harvesting](#).
 - zero or more [setSpec](#) elements - the [set](#) membership of the item for the purpose of [selective harvesting](#).

- an optional status attribute with a value of deleted indicates the withdrawal of availability of the specified metadata format for the item, dependent on the repository support for deletions.
- *metadata* - a single manifestation of the metadata from an item. The OAI-PMH supports items with multiple manifestations (formats) of metadata. At a minimum, repositories must be able to return records with metadata expressed in the [Dublin Core](#) format, without any [qualification](#). Optionally, a repository may also disseminate other formats of metadata. The specific metadata format of the record to be disseminated is specified by means of an argument - the [metadataPrefix](#) - in the [GetRecord](#) or [ListRecords](#) request that produces the record. The [ListMetadataFormats](#) request returns the list of all metadata formats available from a repository, or for a specific item (which can be specified as an argument to the [ListMetadataFormats](#) request).
- *about* - an optional and repeatable container to hold data about the metadata part of the record. The contents of an about container must conform to an XML Schema. Individual implementation communities may create XML Schema that define specific uses for the contents of about containers. Two common uses of about containers are:
 - *rights statements*: some repositories may find it desirable to attach terms of use to the metadata they make available through the OAI-PMH. No specific set of XML tags for rights expression is defined by OAI-PMH, but the about container is provided to allow for encapsulating community-defined rights tags.
 - *provenance statements*: One suggested use of the about container is to indicate the provenance of a metadata record, e.g. whether it has been harvested itself and if so from which repository, and when. An XML Schema for such a provenance container, as well as some supporting information is available from the accompanying [Implementation Guidelines](#) document.

Identifier [W3C-definitie]

An identifier embodies the information required to distinguish what is being identified from all other things within its scope of identification. Our use of the terms "identify" and "identifying" refer to this purpose of distinguishing one resource from all other resources, regardless of how that purpose is accomplished (e.g., by name, address, or context). These terms should not be mistaken as an assumption that an identifier defines or embodies the identity of what is referenced, though that may be the case for some identifiers. Nor should it be assumed that a system using URIs will access the resource identified: in many cases, URIs are used to denote resources without any intention that they be accessed. Likewise, the "one" resource identified might not be singular in nature (e.g., a resource might be a named set or a mapping that varies over time).

Zie: Tim Berners-Lee, Roy T. Fielding, Larry Masinter. (January 2005). "Uniform Resource Identifier (URI): Generic Syntax". Internet Society. RFC 3986; STD 66. <http://gbiv.com/protocols/uri/rfc/rfc3986.html>

Resource [W3C-definitie]

This specification does not limit the scope of what might be a resource; rather, the term "resource" is used in a general sense for whatever might be identified by a URI. Familiar examples include an electronic document, an image, a source of information with a consistent purpose (e.g., "today's weather report for Los Angeles"), a service (e.g., an HTTP-to-SMS gateway), and a collection of other resources. A resource is not necessarily accessible via the Internet; e.g., human beings, corporations, and bound books in a library can also be resources. Likewise, abstract concepts can be resources, such as the operators and operands of a mathematical equation, the types of a relationship (e.g., "parent" or "employee"), or numeric values (e.g., zero, one, and infinity).

Zie: Tim Berners-Lee, Roy T. Fielding, Larry Masinter. (January 2005). "Uniform Resource Identifier (URI): Generic Syntax". Internet Society. RFC 3986; STD 66. <http://gbiv.com/protocols/uri/rfc/rfc3986.html>

<http://gbiv.com/protocols/uri/rfc/rfc3986.html> URL [W3C-definitie]

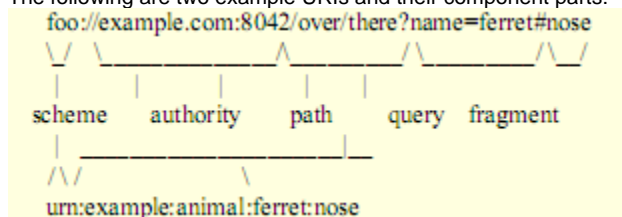
The term "Uniform Resource Locator" (URL) refers to the subset of URIs that, in addition to identifying a resource, provide a means of locating the resource by describing its primary access mechanism (e.g., its network "location").

Zie: Tim Berners-Lee, Roy T. Fielding, Larry Masinter. (January 2005). "Uniform Resource Identifier (URI): Generic Syntax". Internet Society. RFC 3986; STD 66. <http://gbiv.com/protocols/uri/rfc/rfc3986.html>

URI [W3C-definitie]

A Uniform Resource Identifier (URI) is a compact sequence of characters that identifies an abstract or physical resource. This specification defines the generic URI syntax and a process for resolving URI references that might be in relative form, along with guidelines and security considerations for the use of URIs on the Internet. The URI syntax defines a grammar that is a superset of all valid URIs, allowing an implementation to parse the common components of a URI reference without knowing the scheme-specific requirements of every possible identifier. This specification does not define a generative grammar for URIs; that task is performed by the individual specifications of each URI scheme.

The following are two example URIs and their component parts:



Zie: Tim Berners-Lee, Roy T. Fielding, Larry Masinter. (January 2005). "Uniform Resource Identifier (URI): Generic Syntax". Internet Society. RFC 3986; STD 66. <http://gbiv.com/protocols/uri/rfc/rfc3986.html>

URI, URL, and URN [W3C-definitie]

A URI can be further classified as a locator, a name, or both. The term "Uniform Resource Locator" (URL) refers to the subset of URIs that, in addition to identifying a resource, provide a means of locating the resource by describing its primary access mechanism (e.g., its network "location"). The term "Uniform Resource Name" (URN) has been used historically to refer to both URIs under the "urn" scheme [[RFC2141](#)], which are required to remain globally unique and persistent even when the resource ceases to exist or becomes unavailable, and to any other URI with the properties of a name.

An individual scheme does not have to be classified as being just one of "name" or "locator". Instances of URIs from any given scheme may have

the characteristics of names or locators or both, often depending on the persistence and care in the assignment of identifiers by the naming authority, rather than on any quality of the scheme. Future specifications and related documentation should use the general term "URI" rather than the more restrictive terms "URL" and "URN" [RFC3305]. scheme authoritypathquery fragment

URN [W3C-definitie]

Uniform Resource Names (URNs) are intended to serve as persistent, location-independent, resource identifiers. This document sets forward the canonical syntax for URNs. A discussion of both existing legacy and new namespaces and requirements for URN presentation and transmission are presented. Finally, there is a discussion of URN equivalence and how to determine it.

All URNs have the following syntax (phrases enclosed in quotes are REQUIRED): <URN> ::= "urn:" <NID> ":" <NSS> where <NID> is the Namespace Identifier, and <NSS> is the Namespace Specific String. The leading "urn:" sequence is case-insensitive. Moats, R., URN Syntax. RFC 2141, May 1997. <http://www.ietf.org/rfc/rfc2141.txt>

NBN [W3C-definitie]

National Bibliography Number (NBN) is a generic name referring to a group of identifier systems utilised by the national libraries and only by them for identification of deposited publications which lack an identifier, or to descriptive metadata (cataloguing) that describes the resources. In many countries legal (or voluntary) deposit is being extended to electronic publications. J. Hakala, Using National Bibliography Numbers as Uniform Resource Names. RFC: 3188. October 2001. <http://www.ietf.org/rfc/rfc3188.txt>

Handle / Handle System® [IDF-definitie]

The [resolution](#) component of the [DOI System](#). A general-purpose distributed information system designed to provide an efficient, extensible, and secured global name service for use on networks such as the Internet. The Handle System includes an open set of protocols, a namespace, and a reference implementation of the protocols. The DOI System is one implementation of the Handle System; hence a DOI name is a [handle](#). The Handle System is made up of [local handle services](#).

Identifier [IDF-definitie]

1. (1) An unambiguous string or "label" that references an entity (e.g. ISBN 0-19-853737-9).
2. (2) A numbering scheme: a formal standard, an industry convention, or an arbitrary internal system providing a consistent syntax for generating individual labels or identifiers (1) denoting and distinguishing separate members of a class of entities (e.g. ISBN, or DOI® Syntax NISO Z39.84).
3. (3) An infrastructure specification: a syntax by which any identifier (1) can be expressed in a form suitable for use with a specific infrastructure, without necessarily specifying a working mechanism (e.g. URI).
4. (4) A system for implementing labels (identifiers (1)) through a numbering scheme (identifiers (2)) in an infrastructure using a specification (identifiers (3)) and management policies (e.g., [DOI System](#)).

Zie: http://www.doi.org/handbook_2000/glossary.html#hs

The Handle System is only one component of the DOI System

The Handle System provides a general-purpose global name service enabling secure name resolution over the internet, designed to enable a broad set of communities to use the technology to identify digital content independent of location. The DOI System utilises the Handle System as one component in building an added value application, for the persistent, semantically interoperable, identification of intellectual property entities.

- The DOI System provides a ready-to-use system of several components: a specified numbering syntax, a resolution service (based on the Handle System), a data model system (including the indecs Data Dictionary), and policies and procedures for the implementation of DOI names through a federation of Registration Agencies.
- One component of the DOI System is the Handle System, and its implementation in the DOI System has been supplemented by expanded technical infrastructure and features specific to DOI System applications.

Zie: <http://www.doi.org/factsheets/doihandle.html>

DOI [IDF-definitie]

A DOI name - a digital identifier for any object of intellectual property. A DOI name provides a means of persistently identifying a piece of intellectual property on a digital network and associating it with related current data in a structured extensible way.

Zie: <http://www.doi.org/faq.html#1/>

DOI: URI implementation

The Uniform Resource Identifier (URI) specification is IETF RFC 2396, URI Generic Syntax, currently under revision as RFC 2396 bis. URIs formally encompass URNs as a sub set. In practice, the URI specification defines (1) an implementation more often called the Uniform Resource Locator, a location on a file server, commonly accessed using the http protocol though other protocols are allowed; (2) a syntax for referencing in XML, through which e.g. ISBNs can be specified as URIs. This provides a single framework which can accommodate any other identifier for referencing, but it is not as such persistent (since persistence is not determined by the specification but by the practical implementation). Conflating these two causes confusion. URLs, as currently understood, are demonstrably not persistent; redefining them as URIs doesn't fix that.

- URL implementation. Users may resolve DOI names using the URL syntax through the DOI System proxy server (<http://dx.doi.org>). A DOI name of the form doi:10.123/456 would be resolved from the address: "<http://dx.doi.org/10.123/456>". Any standard browser encountering a DOI name in this form will be able to resolve it. The use of the DOI proxy server does not interfere with any http requirements, so DOI names may be used with other http-based mechanisms such as OpenURL, PURL, parameter passing, etc. The proxy server is maintained by the IDF and the DOI System community for use by all. However performance and functionality can be improved beyond the level achieved through http by using the optional native resolver protocol (Handle System) not requiring the use of http. DOI® resolution in native resolver form does not require the use of the DNS (Domain Name System), though does of course when used with the proxy http resolver.
- URI syntax implementation. In the URI specification, the network path of the URI is implicitly DNS based; there are no real provisions to include systems that are not DNS based. Original URI specifications, and good design practice, assume the URI to be opaque (that is, it is not assumed that software can parse the body of the URI but that it would simply recognize the name of the scheme and hand it off to

some other software that understood the scheme). The current URI specification, however, assumes that the initial URI parser will look into every URI, no matter what the scheme, looking for certain meaningful characters such as dot and slash. This version of the URI proposed in RFC 2396 bis is so restrictive that it is difficult to see what system could make use of it.

- A specification for the DOI System as a URI exists as an Internet Draft: this document defines the 'doi' Uniform Resource Identifier (URI) scheme for DOI names, which allows a DOI name to be referenced by a URI for Internet applications. The current revision of the URI specification, plus ongoing debate within the IETF and W3C communities on several proposed URI specifications, have delayed the processing of this Draft. DOI System implementation does not depend on implementation of this specification.

Zie: <http://www.doi.org/factsheets/doiidentifierspecs.html>

DOI: URN implementation

The URN (Uniform Resource Name) specification is RFC 2141 URN Syntax. In practice, the URN specification defines (1) a formal registration process as a urn namespace, e.g., urn:doi:10.1000/1 and (2) accompanying specifications to implement a series of functional requirements for such namespaces.

- Namespace referencing. One may specify any existing identifier as a URN: e.g urn:isbn:123456789, but this has no advantage over the simpler isbn:12345678. Such identifiers may be implemented using a specially written URN plug-in and resolved to URLs: functionally this gives nothing beyond the functionality achieved by coherent management of the corresponding URLs.
- URN implementation. In order to implement the functional requirements, the URN architecture assumes an additional network service: a DNS-based Resolution Discovery Service (RDS) to allow a client to deal with a previously unknown URN type by finding the specific service appropriate to the given URN scheme. URN resolutions are then delegated to that scheme-specific resolution service. However no such deployed RDS schemes currently exist: browsers cannot action URN strings without some additional programming in the form of a "plug-in". The lack of any wide-spread infrastructural support will require any URN implementation to develop their own resolution mechanisms, such as plug-ins or proxy servers. Resolution mechanisms which require functionality beyond 1 URN to 1 URL also require the creation of data models. Several such implementations have been developed for specific uses where deployment to a closed group of users may be achieved; these carry no guarantee of ready interoperability with other deployments, which may require a different plug in for each implementation and may use conflicting data approaches.
- DOI names do not require a plug-in but offer this as an option. The CNRI Handle System web browser plug-in (native resolver, available in binary) delivers URN functional requirements in Windows-based browsers. URN plug-ins and Handle System plug-ins share the problem of any new functionality of deploying the software to users. The Handle System has significant advantages: (1) it is a global supported resolution service; (2) the plug-in is freely available, widely tested and proven across multiple applications; (3) unlike URN plug-ins, it is part of a suite of freely available and managed supporting software configured to provide coherent server-side support, including the Local Handle Service and Handle System Client Libraries. These are available across platforms and with several added security features such as trusted resolution and distributed administration. DOI System functionality can therefore be delivered through http, browser plug in, or incorporation of HANDLE.NET[®] software in a dedicated application.
- The DOI System is not registered as a formal URN, despite fulfilling all the functional requirements, since URN registration appears to offer no advantage to the DOI System. It requires an additional layer of administration for defining a DOI name as a URN namespace (the string urn:doi:10.1000/1 rather than the simpler doi:10.1000/1) and an additional step of unnecessary redirection to access the resolution service, already achieved through either http proxy or native resolution. If RDS mechanisms supporting URN specifications become widely available, the DOI System will be registered as a URN.

Zie: <http://www.doi.org/factsheets/doiidentifierspecs.html>

Bijlage 2. Voorstel voor gefaseerde implementatie URN:NBN in DAREverband.

Peter Ruijgrok

Basis Functionaliteit (realisatie uiterlijk 1.1.2008)

A. We moeten overeenstemming krijgen over de URN.

Verantwoordelijk: Werkgroep DOI in samenspraak met Willie Wortel en KANO

Alhoewel de werkgroep URN's van het formaat 'NBN' voorstelt, moet het mogelijk zijn om persistente identifiers te ondersteunen van andere formaten omdat de repositories andersoortige Persistente Identifiers kunnen bevatten, bijvoorbeeld na een verhuizing van objecten naar een repository toe.

Gevolg is dat, in theorie, een Universiteit volledig kan kiezen voor een DOI of een Handle-systeem maar toch past binnen de gekozen oplossing. Dit omdat een DOI, Handle of ander systeem gerepresenteerd kan worden in een URN en daarmee nog steeds centraal geresolved kan worden.

Theoretisch is daarmee de noodzaak weg om 'allemaal tegelijk' over te schakelen naar de URN's van het formaat NBN. Binnen grenzen kan een universiteit zijn eigen tijdschema bepalen om over te schakelen op URN/NBN .

B. Er moet een centrale resolver worden ontwikkeld. Verantwoordelijk: DANS

- Deze centrale resolver moet op de hoogte worden gebracht van alle persistente identifiers (URN/NBN, DOI's, etc) welke opgeslagen zijn in een lokale repository.
 - Deze centrale resolver zet aangeboden persistente identifiers (URN/NBN, DOI's etc) om naar de volledige URL waar het object (item of bitstream) beschikbaar is.
- Hoe de centrale resolver weet waar een persistente identifier object wordt aangeboden moet nog in detail uitgewerkt worden.

Er wordt gedacht aan het harvesten via het OAI-PMH protocol van de lokale repositories, waarbij de lokale repository aangeeft welke persistente identifiers er zijn opgeslagen en wat de volledige URL is naar het object. Bij verhuizingen van een object verschijnt dan de persistente identifier 'vanzelf' bij de centrale resolver middels een harvest. Daarmee ontstaat de mogelijkheid dat een object opgeslagen is in meerdere repositories en dat is geen enkel probleem. De resolver is daarvan op de hoogte gebracht middels de harvest en zal tijdens de resolving de repository kiezen die het object als eerste heeft aangeboden.

- Als nevenfunctie moet de resolver de history van een object bijhouden en resolvingstatistieken leveren.

C. Iedere lokale repository moet aan ieder object een persistente identifier toekennen.

Verantwoordelijkheid: Universiteiten, lokale repository eigenaren

Het technisch meest belangrijke daarbij is dat de repository in staat is om tijdens het harvesten (via OAI-PMH), de persistente identifiers (en lokale URL) mee te geven in de `dare_didl` container (exacte XML nader te specificeren) en in de speciale nog te specificeren 'persistente identifier' xml specificatie.

Het meest voor de hand liggend is om de persistente identifier als extra metadata veld op te nemen in de repository bij ieder object. In de eerste fase is het toekennen van een persistente identifier aan het object als geheel voldoende. Toekenning van persistente identifiers aan de bitstream mag maar is nog niet verplicht. Dit omdat bekend is dat enkele universiteiten moeite hebben hiermee, als gevolg van op dit moment tekortschietende repository-software hiervoor. Een lokale resolver is in principe nog niet noodzakelijk, maar is wel aan te raden. De centrale resolver kan tijdelijk direct verwijzen naar bv. de Jump-Of-Page, waarmee het lokale resoluten uitgesteld kan worden naar een later moment.

D. Darenet resolving via centrale resolver

Verantwoordelijk: DANS

De methode waarop Darenet doorlinked naar de metadata en de full-text moet geleid worden via de centrale resolver. Hiermee ontstaat in ieder geval al een 'grootverbruiker' van de centrale resolver.

Extra functionaliteit (realisatie na 1.1.2008)

1. Volledig overgaan op afgesproken persistente identifiers

Verantwoordelijkheid: Universiteiten, lokale repository eigenaren

Nadat we overeenstemming hebben over geaccepteerde persistente identifiers (waarschijnlijk URN/NBN, maar wellicht ook een DOI verpakt in een URN) moet iedere lokale repository dit volledig gaan ondersteunen. Bestaande objecten moeten een persistente identifier krijgen die we afgesproken hebben (URN/NBN, DOI, ...) en alle nieuwe objecten moeten hier uiteraard ook aan voldoen.

Er wordt dan van de lokale repositories verwacht dat ze afstappen van applicatie-specifieke oplossingen zoals het handle systeem (Dspace), OAI identifiers (CMS) of database volgnummers (ARNO, iTor) naar de overeengekomen persistente identifier.

2. Volledig ondersteunen van 'verhuizen van objecten'

Verantwoordelijkheid: DANS + Universiteiten, lokale repository eigenaren

In de eerste fase ondersteunen we de mogelijkheid dat objecten voorkomen in meerdere repositories (harvesting). Hoe een object daarna 'verdwijnt' uit een lokale repository moet nader worden afgesproken.

Middels harvesting via het OAI-PMH protocol kan wellicht een 'delete' worden doorgegeven aan de centrale resolver, waarmee kenbaar gemaakt

wordt dat het object niet meer geleverd kan worden. Echter.... dit mag uitsluitend als het object inmiddels al elders beschikbaar is. De centrale resolver moet dit registreren en actie ondernemen indien een object volledig zou verdwijnen.

3. Eigenaarschap en meerdere lokaties van een object

Verantwoordelijkheid: DANS

Indien een object in meerdere repositories voorkomt is de eerste verschijning van een object de eigenaar van dit object. Tijdens resolving zal de centrale resolver altijd doorverwijzen naar het object in de repository van de eigenaar dan dit object.

Er moeten een mechanisme bedacht worden waarmee het eigenaarschap van een object kan worden overgedragen aan een andere repository, zonder dat het object verwijderd wordt. Indien een object voorkomt in meerdere repositories moet wellicht tijdens het resolving-proces kunnen kiezen uit welke repository het object moet worden gehaald. Anderzijds kan je denken aan het aanbieden van alternatieve lokaties indien de primaire repository van de eigenaar het object (tijdelijk) niet kan aanbieden.

4. Lokale resolving

Verantwoordelijkheid: Universiteiten, lokale repository eigenaren.

Teneinde niet te veel centrale administratie te krijgen is een lokale resolver wenselijk.

Hiermee kan een universiteit lokale wijzigingen doorvoeren zonder dat dit gevolgen heeft voor de administratie in de centrale resolver.

5. Wereldwijde resolving

Verantwoordelijkheid: ?

Aansluiting verkrijgen bij de overkoepelende URN-resolving, waardoor een eenduidige plek bekend is waar iedereen wereldwijd naartoe kan zonder de exacte lokatie van 'onze' centrale resolver te kenne.

Meewerken aan het 'clickable' maken van URN's in browsers zoals IE en Firefox. De 'wereld' moet weten hoe een URN omgezet wordt naar een lokatie van een object, zonder dat daarvoor een 'begin URL" nodig is. Daarmee wordt bereikt dat het opnemen van een URN op een webpagina volstaat om door te linken naar de metadata of de fulltext. Momenteel moet de volledige URL naar de centrale resolver inclusief het betreffende URN worden opgenomen. Op termijn is dit niet wenselijk.