



Nationaal Archief  
*Ministerie van Onderwijs, Cultuur en  
Wetenschap*

Duurzaam vindbaar op het web:  
URI-strategie Nationaal Archief  
het gebruik van Handles en URI's

Versie 1.1

Datum 17 november 2020  
Status NA



## Colofon

Versienummer	1.1
Contactpersoon	Ed de Heer ed.deheer@nationaalarchief.nl
Versiegeschiedenis	<p>Versie 1.1 voor intern gebruik NA</p> <ul style="list-style-type: none"><li>- status aangepast voor het delen met derden als voorbeeld</li><li>- Deze versie is specifiek voor het NA</li><li>- Geen ISIL-code meer in URI's</li><li>- Overal https</li><li>- Voorkeurstaal is Nederlands</li></ul> <p>Versie 1.0 na vaststelling Stuurgroep</p> <p>Versie 0.91 aanpassingen n.a.v. opmerkingen Gelders Archief</p> <p>Versie 0.9: aangepast n.a.v. eerste bespreking in Stuurgroep Toegang</p> <ul style="list-style-type: none"><li>- managementsamenvatting aangepast</li><li>- scope aangevuld</li><li>- term "rijksarchieven" vervangen</li></ul> <p>Versie 0.8: kleinere wijzigingen n.a.v. bespreking met Gerard Kuys</p> <p>Versie 0.7 bijgewerkt n.a.v. bespreking in Architectuurcommissie en interne input. Wijzigingen:</p> <ul style="list-style-type: none"><li>- domeinnaam niet "overheid.nl" maar archief.nl</li><li>- scope (par. 1.3) toegevoegd</li><li>- bijlage bijgewerkt</li></ul> <p>Versie 0.6 bijgewerkt n.a.v. gesprek met KOOP en interne reviews.</p> <p>Versie 0.5 bijgewerkt n.a.v. gesprekken met Tresoar en interne overleggen in maart 2018. Belangrijkste wijzigingen:</p> <ul style="list-style-type: none"><li>- geen onderscheid in collectietypes vastleggen in URI, dus niet in de URI vastleggen of het NT of een beeldrecord of iets anders</li><li>- {path} gebruiken voor het tenant ID (de ISIL-code)</li><li>- eigen termen definiëren in het Engels</li></ul>



## Inhoud

Colofon 3  
Managementsamenvatting 7

<b>1</b>	<b>Inleiding 8</b>
1.1	Doel van dit stuk 8
1.2	Waarom dit beleid? 8
1.3	Scope 8
1.4	Wat zijn PID's en URI's? 8
1.5	Wat zijn API's? 8
1.6	Relatie met ander beleid 9
1.7	Verantwoordelijkheid 9
1.8	Versiebeheer 9
<b>2</b>	<b>Kaders &amp; richtlijnen 10</b>
2.1	Overheidsrichtlijnen 10
2.2	Digitale Erfgoed Referentiearchitectuur (DERA) 10
2.3	Model Architectuur Rijks Archiefinstellingen (MARA) 10
2.4	OpenAPI Specification 10
2.5	PILOD URI-Strategie 10
<b>3</b>	<b>Overwegingen 12</b>
3.1	De archiefinstellingen als collectiebeheerders 12
3.2	Nationaal Archief als dienstverlener 14
<b>4</b>	<b>Beleidskeuzes 16</b>
4.1	Het Nationaal Archief als bronhouder 16
4.2	Nationaal Archief als dienstverlener 18
<b>5</b>	<b>URI-strategie 20</b>
5.1	Het Nationaal Archief als collectiebeheerder 20
5.2	Nationaal Archief als dienstverlener 22
<b>6</b>	<b>Consequenties voor RHC's 24</b>
6.1	Handles 24
6.2	URI's 24
6.3	Services 24
<b>7</b>	<b>Literatuur en bronnen 25</b>
<b>8</b>	<b>BIJLAGE: Type URI's 26</b>



## Managementsamenvatting

Een Persistent Identifier (PID) is een unieke identificatiecode van een digitaal object die op een afgesproken plaats wordt geregistreerd<sup>1</sup>. De bedoeling van de PID is dat deze altijd blijft werken ook al verandert bijvoorbeeld de naam van de organisatie. Er zijn van oudsher verschillende soorten PID's, onder meer Handles en DOI's. Door de opkomst van Linked Open Data is er een andere vorm van persistent identifiers in gebruik gekomen, namelijk de Uniform Resource Identifier (URI). Dit document specificeert hoe het Nationaal Archief URI's wil gaan inzetten in combinatie met Handles en daarmee wil zorgen dat de digitale bronnen en diensten op internet vindbaar zijn en blijven.

Voor het gebruik van URI's zijn nog geen vastgestelde en uitgekristalliseerde overheidskaders op dit gebied. Er zijn wel enkele pioniers zoals het Kadaster, het Digitaal Stelsel Omgevingswet en het Platform Linked Open Data. Deze strategie sluit aan bij de best practices van deze organisaties. Er is bij het opstellen van deze strategie ook overleg geweest met deze organisaties.

In dit document wordt onderscheid gemaakt tussen het Nationaal Archief als bronhouder en het Nationaal Archief als leverancier van diensten op de gezamenlijke digitale infrastructuur.

<sup>1</sup> Bron: [www.pidwijzer.nl](http://www.pidwijzer.nl)

# 1 Inleiding

## 1.1 Doel van dit stuk

In dit stuk wordt het beleid geformuleerd van het Nationaal Archief ten aanzien van duurzame vindbaarheid. Daarmee maakt het Nationaal Archief duidelijk hoe het ervoor wil zorgen dat haar digitale informatie en digitale diensten vindbaar zijn en vindbaar blijven.

## 1.2 Waarom dit beleid?

Het Nationaal Archief en de Regionaal Historische Centra (RHC's) stellen steeds meer informatie en diensten beschikbaar via het internet. Dit is enerzijds collectiemateriaal zoals gedigitaliseerd en Born Digital archief, maar ook foto's, scans en kaarten en kennisproducten zoals bedoeld in de Handreiking vervanging. Vanaf 2018 zal het Nationaal Archief de gegevens van foto's en indexen via Linked Open Data gaan ontsluiten.

Het Nationaal Archief gebruikt al sinds de introductie van gahetna.nl in 2011 Handles als Persistent Identifiers op de website. Het is daarmee een van de voorlopers in de erfgoedsector in Nederland. Linked Open Data werkt met URI's en niet speciaal met Handles. Dit roept de vraag op wanneer we Handles gebruiken en wanneer URI's. Dit beleidsstuk geeft een antwoord op die vraag.

Anderzijds biedt het Nationaal Archief steeds meer services die door onszelf, door de RHC's en door derden worden gebruikt. Derden zijn bijvoorbeeld hergebruikers van archiefdata, maar ook portalen zoals Archives Portal Europe (APE), Brabant Cloud en archiefvormers. Ook voor de services die het Nationaal Archief aanbiedt is het belangrijk dat deze zoveel mogelijk altijd op dezelfde webadressen te vinden blijven. Het gebruik van gestandaardiseerde URI's en API's maakt dit mogelijk.

## 1.3 Scope

Dit beleid geldt voor de informatie en diensten die vanuit de digitale infrastructuur van het Nationaal Archief als Linked Open Data wordt aangeboden. Maar de uitvoering van deze strategie heeft ook consequenties voor de RHC's die gebruik willen gaan maken van de digitale infrastructuur die het Nationaal Archief levert. Deze 1.1 versie van de URI-strategie gaat er vanuit de Regionaal Historische Centra geen URI's munten in het archief.nl domein. Zie paragraaf 6.2 voor meer uitleg.

## 1.4 Wat zijn PID's en URI's?

Een Persistent Identifier (PID) is een unieke identificatiecode van een digitaal object die op een afgesproken plaats wordt geregistreerd. Er zijn verschillende vormen van PID's. De traditioneel meest voorkomend zijn Handle, DOI en URN-NBN. Het Nationaal Archief gebruikt sinds 2011 Handles.

Een alternatieve vorm van Persistent Identifiers zijn Uniform Resource Identifiers (URI's). Een URI is een reeks karakters die gebruikt wordt om een "resource" uniek te identificeren. De "resource" kan een fysiek object zijn, zoals een museumstuk of een boek, maar meestal gaat het om een digitale bron van informatie op het Internet: een webpagina, een tekst, een afbeelding, enz met een webadres. URI's worden vooral veel gebruikt in Linked Open Data.

Voor meer informatie over PID's zie: <http://pidwijzer.nl/>.

## 1.5 Wat zijn API's?

Een **Application Programming Interface (API)** is een verzameling definities op basis waarvan een computerprogramma kan communiceren met een ander programma of onderdeel.<sup>2</sup> Een API definieert functionaliteit zonder dat degene die

<sup>2</sup> Bron: [https://www.noraonline.nl/wiki/API\\_geraadpleegd\\_12-12-2017](https://www.noraonline.nl/wiki/API_geraadpleegd_12-12-2017)



de API aanroept precies hoeft te weten hoe de software daarachter werkt. Het e-Depot heeft bijvoorbeeld een API waarmee bestanden opgehaald kunnen worden (de zogenaamde CMIS-interface). Degene die deze API gebruikt hoeft niet te weten hoe het e-Depot precies werkt maar kan wel de records opvragen. Open data worden doorgaans via een API ter beschikking gesteld.

**1.6 Relatie met ander beleid**

Dit beleid hangt samen met

- Informatiebeleid en architectuur: deze bepalen immers de informatie strategie.
- Applicatieportfoliomanagement
- Applicatiebeheer

**1.7 Verantwoordelijkheid**

Verantwoordelijk voor de URI-strategie, inclusief het toezicht op de uitvoering ervan, is hoofd I&S.

**1.8 Versiebeheer**

Dit beleid wordt tweejaarlijks geëvalueerd en als nodig herzien.

## 2 Kaders & richtlijnen

Er zijn op het gebied van Persistent Identifiers voor de overheid nog geen vastgestelde kaders en richtlijnen. Er zijn wel een aantal standaarden en best practices van enkele pioniers.

### 2.1 Overheidsrichtlijnen

De NORA heeft geen standaarden of richtlijnen voor PID's of URI's, maar geeft wel in het stuk over gegevensbeschrijving een handreiking met een lijst van referentienamen<sup>3</sup>. Daarin wordt als richtlijn voor URI's de RFC 3986<sup>4</sup> genoemd en de URI-strategie Linked Open Data (zie par. 2.4).

Ook voor de vormgeving van API's zijn nog geen overheidsrichtlijnen vastgesteld. Er is bij het Forum Standaardisatie wel een discussiestuk opgesteld over RESTful API's, daarin wordt aangeraden voor de verwijzing naar objecten URI's te gebruiken. In het kader van het Digitaal Stelsel Omgevingswet is een URI-strategie uitgewerkt waarin ook de API's worden meegenomen.

### 2.2 Digitale Erfgoed Referentiearchitectuur (DERA)

De DERA 1.0 heeft als ontwerpvisie: Gebruik Persistent Identifiers, maar specificeert niet welke identifiers moeten worden gebruikt.

### 2.3 Model Architectuur Rijks Archiefinstellingen (MARA)

De MARA heeft als uitgangspunt dat data en metadata uniek en persistent geïdentificeerd zijn, maar specificeert niet het type identifier. In de *Visie aggregatie andere collectietypes*, die beschouwd kan worden als een aanvulling op de MARA, wordt gekozen voor de inzet van Linked Open Data voor collectieoverstijgende ontsluiting, maar bevat geen specificatie van het gebruik van URI's.

### 2.4 OpenAPI Specification

De OpenAPI Specification<sup>5</sup> (voorheen bekend als de Swagger Specification) is een gezaghebbend Open Source initiatief dat een standaard beschrijving biedt voor REST API's. Door deze standaard te gebruiken kunnen computers en mensen de service gebruiken zonder specifiek de functionaliteit te hoeven kennen. De OpenAPI Specification vindt veel navolging ook in Nederland, en bijvoorbeeld het Digitaal Stelsel Omgevingswet baseert zijn API-strategie op deze standaard.

### 2.5 PILOD URI-Strategie

De belangrijkste richtlijn voor URI's binnen het Nederlandse taalgebied is de URI-strategie van het Platform Linked Data Nederland. Deze is opgesteld door organisaties zoals KOOP en Geonovum n.a.v. een Pilot met Linked Open Data in 2012 en 2013. Deze URI-strategie is een aanzet tot een nationale strategie zoals die wordt aanbevolen voor overheidspublicaties en in het kader van de Inspire richtlijn van de Europese Commissie. Naar deze URI-strategie wordt ook verwezen door de NORA. De URI-strategie wordt inmiddels gehanteerd door meerdere overheidsorganisaties die actief met Linked Open Data bezig zijn.

De URI-strategie is speciaal bedoeld voor Linked Open Data en de objecten en concepten daarbinnen waar andere toepassingen naar kunnen verwijzen.

<sup>3</sup> Bron: [https://www.noraonline.nl/images/noraonline/fff/Referentienamen\\_definities\\_en\\_standaarden.pdf](https://www.noraonline.nl/images/noraonline/fff/Referentienamen_definities_en_standaarden.pdf)  
Geraadpleegd 26 september 2017.

<sup>4</sup> <https://www.ietf.org/rfc/rfc3986.txt>

<sup>5</sup> Bron: <https://www.openapis.org/> geraadpleegd 12 december 2017

De PILOD URI-strategie heeft een patroon voor URI's gedefinieerd dat algemeen navolging heeft gekregen bij Nederlandse overheidsorganisaties die met Linked Open Data bezig zijn. De basisopbouw van een URI volgens PILOD is:

`http://{domain}/{type}/{concept}/{reference}`

Om deze opbouw toe te lichten een voorbeeld uit de Basisregistratie voor Adressen en Gebouwen (de BAG) en de registratie daarin van het pand van het Nationaal Archief.

<https://bag.basisregistraties.overheid.nl/bag/doc/verblijfsobject/0518010000368721>

Het `{domain}` duidt het onderwerpsdomein aan, waarbinnen de URI's gelden. In dit voorbeeld is het domein de overheid en daarbinnen de basisregistraties en daarbinnen de bag<sup>6</sup>.

<https://bag.basisregistraties.overheid.nl/bag/doc/verblijfsobject/0518010000368721>

`{type}` wordt gebruikt om aan te geven wat voor soort URI het is, of het gaat om een identifier (id) of document (doc) of een definitie (def). Zie par. 5.1.1. In dit voorbeeld is het "doc", de documentatie van het verblijfsobject (het gebouw van het Nationaal Archief)

<https://bag.basisregistraties.overheid.nl/bag/doc/verblijfsobject/0518010000368721>

`{concept}` geeft een indicatie van wat voor soort concept de URI identificeert. In dit geval is dat een verblijfsobject.

<https://bag.basisregistraties.overheid.nl/bag/doc/verblijfsobject/0518010000368721>

De `{reference}` is de identificerende naam of code van het individuele object. In dit geval is dat de identifier 0518010000368721

<https://bag.basisregistraties.overheid.nl/bag/doc/verblijfsobject/0518010000368721>

<sup>6</sup> De bag URI is niet helemaal netjes – volgens de betrokkenen van de BAG zelf.

<https://bag.basisregistraties.overheid.nl/bag/> is dubbelop. Beter zou zijn:

<https://bag.basisregistraties.overheid.nl/> of <https://basisregistraties.overheid.nl/bag/>

## 3 Overwegingen

Nationaal Archief en Regionaal Historische Centra beheren samen de rijkscollectie. Daarnaast leveren ze ook diensten aan derden, met name de archiefvormers, bijvoorbeeld voor uitgeplaatst digitaal archief. Daartoe realiseren ze een gezamenlijke digitale infrastructuur waar het e-Depot een belangrijk onderdeel van is. Het Nationaal Archief fungeert als leverancier en beheerder van de digitale infrastructuur.

Hieronder is onderscheid gemaakt tussen de overwegingen in de URI-strategie met betrekking tot de collectie waarbij het Nationaal Archief "collectiebeheerder" is en de URI-strategie voor de services die het Nationaal Archief verleent als "dienstverlener".

### 3.1 De archiefinstellingen als collectiebeheerders

#### 3.1.1 Huidige situatie

Op dit moment (december 2019) is het zo dat alleen het Nationaal Archief en het Noord-Hollands Archief gebruik maken van een vorm van Persistent Identifiers. Beide gebruiken Handles. Handles zijn een vorm van Persistent Identifiers. Het Noord-Hollands Archief heeft sinds 2017 Handles op de fotocollectie in de Beeldbank. Bij het Nationaal Archief worden sinds 2011 Handles toegekend aan de volgende objecten:

- **Archiefinventarissen en archiefbestanddelen**  
 Een archiefinventaris is een "Systematisch ingedeelde toegang op een archief of in de tijd begrensd deel ervan, uitgaande van de ordening toegepast door de archiefvormer, voorzien van een inleiding ..."7  
 Een archiefbestanddeel is een "Geheel van archiefbescheiden binnen een archief, bijeengebracht met een bepaald doel en in onderlinge samenhang te raadplegen"8  
 Beide zijn digitale beschrijvingen van objecten, niet de digitale of fysieke objecten zelf.  
 Bijvoorbeeld de Handle op de inventaris van het veel geraadpleegde archief van de VOC: <http://hdl.handle.net/10648/9b54fab1-6340-474a-93d8-04bf78294de8>
- **Indexen en indexrecords**  
 Een index is een "Register van gegevens in een archiefstuk van grotere omvang of van archiefbescheiden in een archiefbestanddeel in een andere systematiek dan de volgorde waarin deze in het archiefstuk of archiefbestanddeel voorkomen onder opgave van de vindplaatsen van de gegevens en archiefbescheiden."9  
 Voor indexen en indexrecords geldt ook dat de PID's verwijzen naar beschrijvingen en niet naar objecten zelf.  
 Bijvoorbeeld de Handle op een record uit de overgekomen brieven en papieren van de VOC waarin Jan Pieterszoon Coen wordt genoemd: <http://hdl.handle.net/10648/7478936c-ab66-102e-8314-0050569c51dd>
- **Afbeeldingen (foto's,kaarten en scans)**  
 Het Nationaal Archief heeft de ambitie om 10% van de collectie te digitaliseren. Gedigitaliseerde objecten kunnen zowel scans zijn van

7 Bron: <http://archiefwiki.org/wiki/Archiefinventaris> zoals geraadpleegd op 14 maart 2017

8 Bron: <http://archiefwiki.org/wiki/Archiefbestanddeel> zoals geraadpleegd op 14 maart 2017

9 Bron: <http://archiefwiki.org/wiki/Index> zoals geraadpleegd op 14 maart 2017

documenten als van foto's of kaarten. Al deze digitale objecten krijgen een Persistent Identifier.

In geval van afbeeldingen verwijst de PID naar het digitale object, inclusief de beschrijving.

Bijvoorbeeld een portret van Multatuli (Eduard Douwes Dekker):

<http://hdl.handle.net/10648/6a116aec-1ad9-102f-a76c-003048944028>

Het Noord-Hollands Archief hanteert sinds 2017 Handles op de fotocollectie in de Beeldbank.

- **Actoren**

Een actor is een persoon of organisatie die archiefstukken maakt of beheert<sup>10</sup>.

Het Nationaal Archief houdt een registratie bij van actoren voor de centrale overheid (Rijk en zbo's) vanaf 1975. Het is de bedoeling dit register uit te breiden met actoren van de gehele overheid en eerdere periodes.

Elke beschrijving van een actor krijgt zijn eigen PID.

Bijvoorbeeld de Belastingdienst:

<http://hdl.handle.net/10648/9e8e4f2a-ffaf-451c-b0d9-cf00185bd175>

- **Bijzondere webcontent**

Het Nationaal Archief heeft bepaalde informatie op de website staan, die ze permanent beschikbaar wil blijven stellen; denk aan kennisproducten zoals de Handreiking Vervanging en TMLO, maar ook Archiefgidsen.

Door het Nationaal Archief is in 2010 gekozen om Handles toe te kennen aan deze objecten omdat:

- Het digitale objecten zijn waarvan we als Nationaal Archief kunnen en willen garanderen dat die tot in lengte van dagen via internet beschikbaar blijven – Bijvoorbeeld de teksten en plaatjes voor een webexpositie voldoen niet aan het criterium dat we ze altijd beschikbaar willen houden.
- Deze objecten stabiel zijn – de beschrijvingen kunnen misschien in detail wijzigen maar als geheel blijven ze bestaan
- Het objecten zijn waarnaar door anderen duurzaam verwezen moet kunnen worden. Bijvoorbeeld in historische publicaties is het gebruikelijk om bij bronnen te verwijzen naar het inventarisnummer en het bestanddeelnummer. Maar ook aggregatoren zoals APE en Europeana en ook hergebruikers die onze data harvesten, moeten erop kunnen vertrouwen dat de verwijzingen blijven werken.

### 3.1.2

#### *Huidig gebruik en beheer bij het Nationaal Archief*

Het gebruik en het beheer van Handles bij het NA gaat nog niet vlekkeloos.

- Er wordt in de studiezaal en op de website niet duidelijk gecommuniceerd over de aanwezigheid en het gebruik van Handles
- Uit statistieken van SURF Sara (die de Handles host) blijkt dat de Handles heel weinig worden gebruikt. Van de tientallen miljoenen handles die het NA heeft worden er slechts een twintig à dertigtal handles per dag geresolved.
- Bij steekproeven blijken behoorlijk wat Handles niet te werken
- Er is nog geen oplossing geïmplementeerd voor handles die verdwijnen, bijvoorbeeld van de archiefgidsen die straks als pdf op de nieuwe publiekswaarsite Spyridon komen
- Een deel van de handles is voor het publiek niet zichtbaar (bijvoorbeeld de Handles op het niveau van inventarisnummers)

<sup>10</sup> Bron: <http://archiefwiki.org/wiki/Actor> zoals geraadpleegd op 14 maart 2017

### 3.1.3

#### *Toekomstige situatie*

Met het realiseren van een gemeenschappelijke digitale infrastructuur willen Nationaal Archief en Regionaal Historische Centra gebruik gaan maken van Linked Open Data. Voor meer informatie over het hoe en waarom, zie het architectuurdocument: *Visie aggregatie andere collectietypes*<sup>11</sup>.

Linked Open Data breekt alle informatie op in elementaire triples: subject, predicaat, object. Daardoor ontstaan veel meer entiteiten. Niet alleen het record is een entiteit, maar ook de persoon die in een record wordt beschreven is een entiteit, en de plaats, en de gebeurtenis (bijvoorbeeld een benoeming). Linked Open Data vergt dat je naar al die entiteiten afzonderlijk kan verwijzen, en dat die verwijzing tot in lengte van dagen blijft werken. Daardoor ontstaan veel meer entiteiten waarnaar duurzaam verwezen moet kunnen worden.

Bijvoorbeeld bij onderstaand indexrecord is het niet voldoende dat men kan verwijzen naar het record met de Handle.



In bovenstaand voorbeeld betekent het dat je minimaal ook moet kunnen verwijzen naar de persoon, de functie en de aanstellingsdatum: het wie, wat en wanneer. Theoretisch gezien kun je aan al deze entiteiten ook Handles toekennen. Beter lijkt te zijn om aan te sluiten bij bestaande best practices voor Linked Open Data en naast Handles ook gebruik te gaan maken van URI's.

## 3.2 Nationaal Archief als dienstverlener

### 3.2.1

#### *Huidige situatie*

Op dit moment biedt het Nationaal Archief voor de eigen collectie verschillende services aan: men kan de EAD's harvesten op de website. Men kan de foto's harvesten van de beeldbank, en het e-Depot heeft bepaalde services zoals de zogenaamde CMIS-interface.

Er is nog geen standaardisering of uniformering in de aanroep van deze services. De naamgeving en het gebruikte adres hangt af van de software en de softwareleverancier die de services realiseert.

De huidige services worden door verschillende partijen gebruikt. Ze worden intern in de digitale infrastructuur van het Nationaal Archief gebruikt, zoals de CMIS-interface op het e-Depot, maar er worden ook interfaces door andere partijen gebruikt zoals Archives Portal Europe (APE) en er worden interfaces beschikbaar gesteld voor het publiek, zoals de OAI-PMH interface op de beeldbank.

### 3.2.2

#### *Toekomstige situatie*

In de toekomstige situatie zal het Nationaal Archief als leverancier fungeren van het generieke gedeelte van de digitale infrastructuur voor NA en RHC's.

Bij het NA zal er een ingrijpende herziening van de infrastructuur plaats vinden. Applicaties zoals de beeldbank worden uitgefaseerd, webservices moeten opnieuw worden gebouwd en er worden nieuwe webservices gerealiseerd waar niet alleen de bestaande gebruikers maar ook andere partijen gebruik van moeten kunnen maken.

<sup>11</sup> Zie Proza: <https://proza.ocw.local/otcs/cs.exe/app/nodes/17882011>

Niet alleen RHC's en Nationaal Archief zelf, maar ook andere partijen zoals zogenaamde hergebruikers van archiefdata of archiefvormers die Born Digital Archief willen uitplaatsen.

De diensten dienen zoveel mogelijk persistent te zijn en op een zelfde manier opgebouwd te zijn zodat ze duurzaam zijn en makkelijk te gebruiken.

Dit kan door ook bij webservices gebruik te maken van URI's. Dit wordt onder meer in de praktijk gebracht bij het Kadaster en het Digitaal Stelsel Omgevingswet (DSO).

Het DSO heeft een URI-strategie opgesteld waarbij de URI's voor de objecten en diensten zoveel mogelijk een zelfde gestandaardiseerde opbouw hebben.

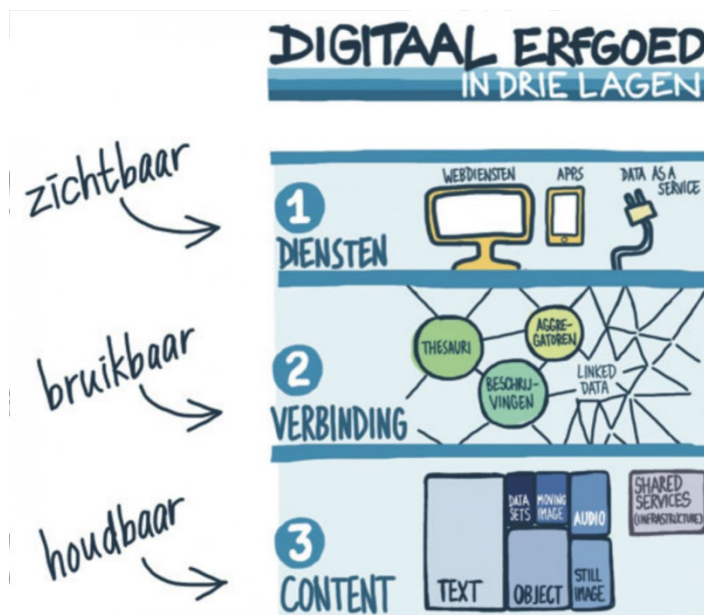
## 4 Beleidskeuzes

### 4.1 Het Nationaal Archief als bronhouder

De beleidskeuzes in dit stuk zijn de keuzes van het Nationaal Archief en (nog) niet van de RHC's. Elke RHC heeft als collectiebeheerder zijn eigen verantwoordelijkheid en eigen keuzes hierin – vergelijkbaar met de verantwoordelijkheden voor preservation en de preservation policy. Onderstaande uitgangspunten zijn zo generiek mogelijk geformuleerd zodat ze ook voor de RHC's te gebruiken zijn.

#### 4.1.1 URI-strategie

Het Nationaal Archief hanteert in navolging van de DERA en de Nationale Strategie Digitaal Erfgoed een drielagenmodel, met een zichtbaarheidslaag, een bruikbaarheidslaag en een houdbaarheidslaag. De collectie bevindt zich in de houdbaarheidslaag, daar bevinden zich de digitale objecten en de content. Vanuit de houdbaarheidslaag worden de collectieobjecten beschikbaar gesteld voor de diensten in de zichtbaarheidslaag, maar ook voor de verbindingen in de bruikbaarheidslaag.



De basis van het beheer van de collectie ligt in de houdbaarheidslaag bij de beheersystemen, met name het Collectiebeheersysteem en het e-Depot. En eventueel ook andere beheersystemen zoals een beeldbank en een bibliotheekstelsel.

Via de verbindingen in de bruikbaarheidslaag komen collectieobjecten beschikbaar voor de diensten in de zichtbaarheidslaag. De bekende diensten zijn websites, maar denk ook aan Apps op een telefoon of bijvoorbeeld een OAI-PMH endpoint. De collectieobjecten die beschikbaar worden gesteld in de zichtbaarheidslaag moeten vindbaar zijn en blijven. Daarvoor gebruiken we URI's en Handles.



Voor de digitale objecten in de rijkscollectie hanteren Nationaal Archief en Regionaal Historische Centra het volgende uitgangspunt: (bron MARA).

**PID-01 Content die blijvend op internet wordt gepubliceerd krijgt een Persistent Identifier**

Blijvend gepubliceerd houdt in dat het object altijd gevonden moet kunnen worden. Content kan een digitaal archiefstuk zijn, maar ook een beschrijvend record (inventarissen en indexrecords), een foto, een kennisproduct of bepaalde belangrijke webcontent.

Het Nationaal Archief heeft in 2011 gekozen voor Handles omdat dit de binnen de Nederlandse Erfgoedwereld het meest werd toegepast en naast de benodigde stabiliteit en robuustheid de meeste flexibiliteit bood. Voor collectiecontent zoals toegangen (inventarissen), inventarisnummers, foto's en indexrecords is het belangrijk dat een eenmaal gemaakte verwijzing altijd blijft werken. Handles worden centraal geregistreerd, inclusief de locatie van het object. Als er onverhoopt toch een wijziging plaats vindt in naam of locatie wordt dit op centraal niveau bijgewerkt en wordt er een verwijzing naar de nieuwe locatie opgenomen. De Handle verwijst zo altijd naar de huidige locatie van het object.

**PID-02 Het Nationaal Archief gebruikt voor collectie content Handles**

Het Nationaal Archief kiest ervoor Handles te gebruiken voor collectiecontent, met name: inventarissen, inventarisrecords, foto's, actoren, indexen en indexrecords.

Naast collectie content zijn er ook nog andere inhoudelijke entiteiten, zoals personen, plaatsen en gebeurtenissen. Deze worden vooral gebruikt om inhoudelijke verbindingen te leggen met andere bronnen. Om permanente verbindingen te kunnen leggen gebruiken we URI's. Hiervoor gebruiken we geen Handles om twee redenen:

- binnen Linked Open Data is het gebruikelijk om met URI's te werken, daarmee conformeren we ons aan de best practices in Nederland en daarbuiten
- de beheerlast om Handles te onderhouden zou niet meer opwegen tegen de voordelen. Handles gebruiken in plaats van URI's zou naar schatting neerkomen op honderden miljoenen Handles extra. Het registreren en periodiek checken daarvan zou veel tijd, moeite en geld kosten.

**PID-03 Entiteiten die gebruikt worden om verbindingen te kunnen leggen, krijgen een URI.**

Verbindingen mogelijk maken gebeurt vooral met Linked Open Data. Entiteiten zijn met name Persoon, Plaats en Onderwerp.

Voor de opbouw van een URI volgt het Nationaal Archief de internetstandaard RFC3986 en meer in het bijzonder de PILOD richtlijn<sup>12</sup>.

Inhoudelijke verbindingen met andere bronnen kunnen op twee manieren tot stand komen:

- Gecontroleerd door collectiebeheerders met inhoudelijke expertise
- Ongecontroleerd door externen: personen of geautomatiseerde systemen

Het toekennen van URI's maakt beide manieren van verbinding leggen mogelijk, maar de verbindingen die gecontroleerde verbindingen worden bij voorkeur

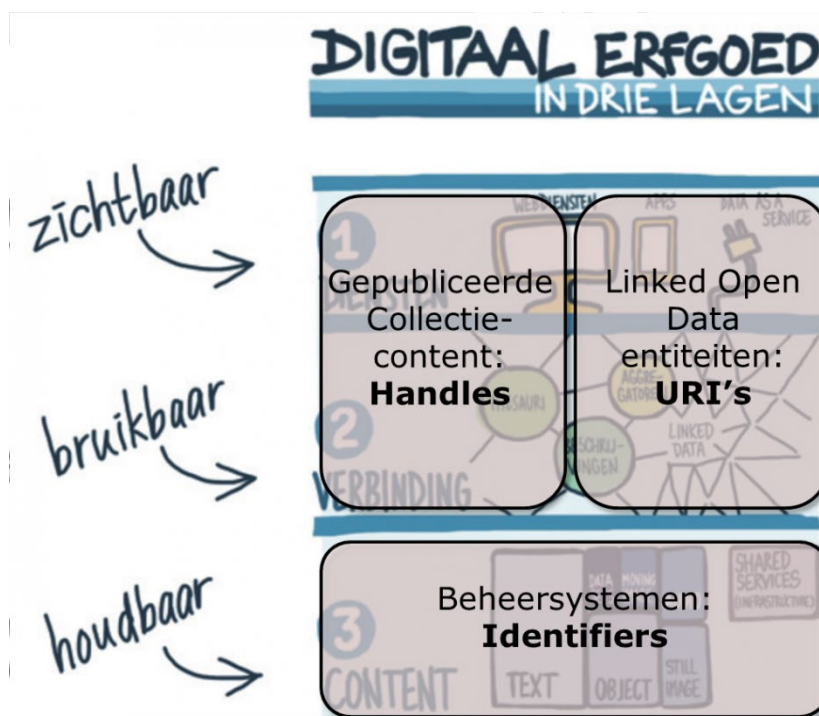
<sup>12</sup> De PILOD URI Strategie maakt ook een indeling in categorieën informatiebronnen. Deze indeling is doelbewust niet overgenomen, omdat die veel verwarring wekt. Voor meer uitleg over de PILOD-richtlijn zie: <http://www.pilod.nl/wiki/Boek/URI-strategie>

vastgelegd in de bronbeschrijving van de collectiesystemen. Denk aan het verwijzen naar thesauri zoals de GTAA en internetbronnen als GeoNames.

**PID-04 Gecontroleerde verbindingen worden vastgelegd in de beheersystemen.**

Verbindingen die door Collectiebeheerders worden gelegd, worden zoveel mogelijk vastgelegd in de bronbeschrijvingen zelf.

Dit resulteert in het volgende plaatje:



In de bruikbaarheidslaag en de zichtbaarheidslaag kunnen Handles en URI's gebruikt worden die altijd blijven bestaan los van het specifieke beheersysteem. De basis voor zowel de Handles als de URI's vormen de identifiers die afkomstig zijn uit de beheersystemen, dit zijn bij voorkeur Universally Unique Identifiers (UUID). Zie: [https://nl.wikipedia.org/wiki/Universally\\_unique\\_identifier](https://nl.wikipedia.org/wiki/Universally_unique_identifier)

**PID-05 De beheersystemen kennen aan de objecten Universally Unique Identifiers of Globally Unique Identifier toe (UUID of GUID)**

UUID en GUID's zijn gegarandeerd uniek en blijven bruikbaar ook als een beheersysteem wordt vervangen.

**4.2 Nationaal Archief als dienstverlener**

Het Nationaal Archief hanteert een URI-strategie voor de webservices die het NA zelf aanbiedt. Bij gebrek aan overheidsrichtlijnen volgt het Nationaal Archief hierbij de URI-strategie van het Digitaal Stelsel Omgevingswet.

**PID-06 Diensten die het Nationaal Archief aanbiedt als webservice krijgen een webadres volgens een vast en voorspelbaar patroon op basis van de OpenAPI Specification  
Basispatroon: service.archief.nl**

De basis URI is “service.archief.nl” omdat het services zijn die het Nationaal Archief als dienstverlener biedt. In tegenstelling tot de objecten van de rijkscollectie zijn deze services wel degelijk afhankelijk van het Nationaal Archief als organisatie.

Door ook intern zoveel mogelijk gebruik te maken van gelijkvormige API's realiseert het Nationaal Archief ontkoppeling van specifieke applicaties en maakt het de infrastructuur flexibeler en duurzamer.

**PID-07 Nationaal Archief maakt zoveel mogelijk gebruik van de API's – ook binnen de eigen infrastructuur.**

## 5 URI-strategie

### 5.1 Het Nationaal Archief als collectiebeheerder

**PID-08 De basis URI is: `https://{domain}/{type}/{concept}/{reference}`.**

Voor **{domain}** gebruiken we een domeinnaam die geschikt is voor de rijksarchiefcollectie

De URI moet zo persistent mogelijk zijn. Het is daarom niet wenselijk om in de URI te verwijzen naar de naam van de organisatie of de plaats waar het archief ligt. Beide kunnen immers veranderen.

Domeinnaam:

- <https://archieff.nl>

**{type}** wordt gebruikt om aan te geven wat voor soort URI het is, of het gaat om een non-information resource of een informationresource. Een non-information resource is iets tastbaars dat in de werkelijkheid bestaat of heeft bestaan, bijvoorbeeld een gebouw of een historische persoon, of een plaats. Een information resource is een set aan gegevens, bijvoorbeeld een record. Zie verder paragraaf 5.1.1 Soorten URI's.

Voorbeeld:

- <https://archieff.nl/id/>

Dit is een URI die verwijst naar een non-information resource: iets dat voorkomt in de buitenwereld.

**{concept}** kan gebruikt worden om aan te geven wat voor soort entiteit de URI identificeert.

We onderscheiden in ieder geval:

- Persoon (persoon, organisatie)
- Onderwerp (gebeurtenis, onderwerp)
- Plaats

Uitbreiding hierop is mogelijk.

Voorbeeld:

- <https://archieff.nl/id/persoon>

Dit is een URI die verwijst naar een non-information resource waarbij het om een persoon gaat.

De **{reference}** is de identificerende naam of code van het individuele object. Dit is het individuele geval. Een reference kan een onderverdeling bevatten.

Voorbeelden

- [https://archieff.nl/id/persoon/Hans\\_Wiegel](https://archieff.nl/id/persoon/Hans_Wiegel)

Dit is de URI van de reëel bestaande persoon Hans Wiegel. In dit geval is de collectiespecificatie {path} als parameter weggelaten.

- <https://archieff.nl/doc/nt00191>

Dit is de URI van de NA dataset van de Nadere Toegang nt00191 als geheel

- <https://archieff.nl/doc/indexrecord/2f1d4dce-af24-102d-afe5-0050569c51dd>

Dit is de URI van de beschrijving door het NA van een record in een NT met het UUID 2f1d4dce-af24-102d-afe5-0050569c51dd.

### 5.1.1

#### Soorten URI's

De PILOD-richtlijn onderscheidt drie **{type}** URI's:

- **'id'** identifier van een object (individual/instance)
- **'doc'** documentatie (metadata) over het object
- **'def'** definitie van een term in een ontologie.

"doc" en "def" worden gebruikt voor zogenaamde information resources. "id" wordt gebruikt voor een non-information resource, d.w.z. een persoon of object in de werkelijkheid.

Het Nationaal Archief hanteert de volgende uitgangspunten:

- Alleen objecten in de werkelijke wereld zijn voor archieven non-information resources en alleen daarvoor gebruiken we URI's met het type "id".
- Beschrijvingen van archiefentiteiten (bijvoorbeeld records van nadere toegangen of foto's of ook personen die in archieven genoemd worden, zijn information resources en daarvoor gebruiken we URI's van het type 'doc'.

De archiefstukken zelf (de collectieobjecten) vallen buiten deze drie categorieën omdat ze niet alleen een reëel object zijn in het depot, maar ook altijd informatie bevatten. We verwijzen naar de beschrijving met een handle.

Voor meer uitleg hierover zie Bijlage I, Type URI's.

### 5.1.2

#### Collectietype

Traditioneel onderscheiden we verschillende collectietypes zoals archiefmateriaal, bibliotheekmateriaal, beeldmateriaal, nadere toegangen en actoren. Deze collectietypes werden traditioneel veelal in eigen systemen beheerd met eigen identifiers. Deze collectietypes kwamen tot nu toe ook altijd tot uitdrukking in de webadressen. Het doel van Linked Open Data is nu juist om de beperkingen van deze gescheiden benadering te overstijgen. Het type materiaal terug laten komen in de URI zou juist het onderscheid weer bevestigen. Vandaar dat er in deze strategie voor is gekozen om niet per se het collectietype in de URI's te laten terugkomen.

### 5.1.3

#### Non-information resources

Archiefstukken en records over archiefstukken verwijzen uiteindelijk altijd naar reële in de werkelijkheid bestaande objecten, bijvoorbeeld historische personen (Multatuli), of plaatsen (Amsterdam in de 17<sup>de</sup> eeuw) of gebeurtenissen (een benoeming, een veroordeling, een aangifte, enz. enz.). In Linked Open Data noemt men dit non-information resources en die krijgen een URI zodat je ernaar kunt verwijzen en ernaar kunt linken. Non-information resources krijgen als type "id". We onderscheiden twee soorten non-information resources: geïdentificeerde en ongeïdentificeerde historische entiteiten (non-information resources).

#### **PID-09 Non-information resources (entiteiten) krijgen als URI:**

<https://archieff.nl/id/{concept}/{reference}>

De URI voor een bepaald concept (persoon of plaats enz.) met een identifier wie, wat, waar of wanneer.

We onderscheiden de volgende **{concepten}**:

- Persoon
- Onderwerp
- Plaats

De **{reference}** is het kenmerk om de historische entiteit aan te duiden. Dit is een identifier die afkomstig is uit het bronsysteem: het collectiebeheersysteem waarin de data wordt beheerd

Voorbeeld:

- <https://archieff.nl/id/persoon/ed061fdc-a7f4-102d-ad54-0050569c51dd>  
Dit is een persoon uit het NA gedeelte van de rijkscollectie met de identifier UUID ed061fdc-a7f4-102d-ad54-0050569c51dd  
We weten niet precies wie het is alleen dat het een persoon is.

#### 5.1.4

##### Termen

Essentieel voor Linked Open Data is het definiëren van de concepten die je gebruikt. Het Nationaal Archief maakt zoveel mogelijk gebruik van concepten uit bestaande vocabulaires, maar zal ook een aantal eigen specifieke (archieff)termen moeten definiëren, bijvoorbeeld "Nadere Toegang". Alle termen die we in Linked Open Data zelf munten, worden beschikbaar gesteld als URI's met de typeaanduiding "def".

##### **PID-11 Definities van eigen termen publiceren we met een eigen URI:**

<https://archieff.nl/def/...>

Eigen definities van Linked Open Data termen die we niet overnemen uit andere vocabulaires publiceren we met hun betekenis en een eigen URI.

Voorbeelden:

- <https://archieff.nl/def/Index>  
De URI van de definitie wat binnen de rijkscollectie verstaan wordt onder een index (een nadere toegang). Deze definitie konden we niet overnemen uit een bestaand vocabulaire.
- <https://archieff.nl/def/Archiefverwijzing>  
De URI van de definitie van het concept "archieffverwijzing"

##### **PID-12 Definities van eigen Linked Open Data termen zijn in het Nederlands**

Dit heeft de volgende redenen:

- Het voornaamste gebruik van Linked Open Data verwachten we door Nederlands sprekenden in het Nederlandse taalgebied.
- Het gebruik van Nederlandse termen bevordert het begrip en gebruik. Een term als "index" of "nadere toegang" is voor Nederlandstalige gebruikers begrijpelijk, een term als "Additional Finding aid" niet.
- Uniformiteit is niet mogelijk. Er blijft altijd een mix van Nederlandstalige en Engelse termen
- In het label gebruiken we altijd ook de Engelse term overnemen zodat de Engelstalige doelgroep altijd precies weet wat er wordt bedoeld

## 5.2

### **Nationaal Archief als dienstverlener**

Het Nationaal Archief is de leverancier van een infrastructuur voor de digitale rijksarchieffcollectie. In het kader daarvan bouwt het NA componenten en levert het

diensten. Voor de diensten is het belangrijk dat ze duurzaam zijn, om dat te realiseren maakt het NA zoveel mogelijk gebruik van een vaste URI-strategie.

### 5.2.1 *URI-strategie voor webservices*

Het Nationaal Archief volgt bij gebrek aan overheidsrichtlijnen de URI-strategie van het Digitaal Stelsel Omgevingswet.

**PID-13 Het URI basispatroon voor webservices is:**  
**`https://service.archief.nl/{collectie}/{service}/{versie}/{query}/...`**

De basis URI is "service.archief.nl" omdat het services zijn die het Nationaal Archief voor de rijksarchieven beschikbaar stelt.

**{collectie}** is optioneel en maakt het mogelijk om specifiek een bepaalde collectieonderdeel aan te roepen, bijvoorbeeld de beeldcollectie of de nadere toegangen. Als de collectieparameter niet wordt gebruikt biedt de service diensten over alle collecties heen, bijvoorbeeld het zoeken door al het materiaal.

**{service}** geeft aan welk type webservice het is. Het Nationaal Archief biedt in ieder geval RESTful API's aangeduid met "api" en OAI-PMH services aangeduid met "oai"

**{versie}** is optioneel en maakt het mogelijk een specifieke eerdere versie van de webservice aan te roepen. Als de versie parameter niet wordt gebruikt, wordt default de laatste versie van de webservice aangeroepen. De opbouw is altijd "v" plus een versienummer, bijvoorbeeld "v1"

**{query}** met de query parameters kan men bijvoorbeeld een specifiek filter op de resultaten aanbrengen

Voorbeelden:

- <https://service.archief.nl/beeld/oai/v2...>

Met deze aanroep zoekt men specifiek in de beeldcollectie van de rijksarchieffcollectie. Men maakt daarbij gebruik van de OAI-PMH service versie 2.

- <https://service.archief.nl/zoeken/api/v1/plaats?naam=Amsterdam/...>

Deze aanroep is over alle collecties van de rijksarchieffcollectie met het filter plaatsnaam is "Amsterdam". Men maakt daarbij gebruik van de REST service versie 1.

Op de plaats van de puntjes komt de specifieke aanroep, bijvoorbeeld een "get"

- <https://service.archief.nl/archieven/api/v1/mam/...>

Met deze aanroep bevestigt men specifiek de Manifestatie Onafhankelijke Metadata van de rijksarchieffcollectie, op dit moment gerealiseerd met de zogenaamde METS-files.

## 6 Consequenties voor RHC's

Zoals eerder in de scope van dit document gesteld is dit de URI-strategie van het Nationaal Archief en heeft iedere RHC een eigen verantwoordelijkheid in het bepalen hoe de duurzame toegankelijkheid van digitale informatie wordt veilig gesteld. Hieronder zijn de consequenties van deze strategie van het NA voor de RHC's uitgewerkt.

### 6.1 Handles

Elke RHC heeft een eigen verantwoordelijkheid om te zorgen voor duurzame toegankelijkheid van collectieobjecten. Er zijn meerdere mogelijkheden. Voor hulp bij het inzetten van Persistent Identifiers, zie [www.pidwijzer.nl](http://www.pidwijzer.nl). Als een RHC net als NA voor Handles kiest, zal het zelf een organisatieprefix moeten aanvragen bij de Handle organisatie en zal het zelf een contract moeten aangaan met SURFSara (of een andere provider) voor de hosting van de Handles.

### 6.2 URI's

De Minister heeft aangegeven op den duur uit de gemeenschappelijke regeling met de RHC's te willen treden. Hoe dit uitpakt en wanneer dit plaats vindt, moet nog worden uitgewerkt. Maar met dit onzekere toekomstperspectief zou het onlogisch zijn als het Nationaal Archief nu de verantwoordelijkheid op zich neemt voor URI's van de RHC's. URI's zijn immers in principe voor de eeuwigheid.

### 6.3 Services

Het Nationaal Archief zal services gaan aanbieden waarmee informatie van de gemeenschappelijke digitale infrastructuur wordt ontsloten. Denk aan een OAI-PMH service op de metadata van de foto's of een service om zichtexemplaren te serveren. Dit zijn technische interfaces waarbij de URI's zoveel mogelijk hetzelfde patroon hebben, nl:

`https://service.archief.nl/{collectie}/{service}/{versie}/{query}/...`

Deze patroon van deze URI is voor al het materiaal hetzelfde, ook voor materiaal van de RHC's.

Vanzelfsprekend kunnen met deze interfaces natuurlijk altijd specifiek het materiaal van één RHC wordt opgehaald.



## 7 Literatuur en bronnen

- Architectuurcommissie Nationaal Archief en RHC's, *Visie nadere toegangen en aggregatie andere collectietypes*. Versie 1.0 (definitief) 23 december 2016 (eDocs nr 1172442)
- Deelprogramma Digitaal Stelsel Omgevingswet, Kaderstellende notities URI-strategie. Versie 1.0 (Vastgesteld 20-07-2017)
- Digitaal Erfgoed Referentie Architectuur / Bram Gakeer et al. Versie 1.0 (maart 2015)
- Platform Linked Data Nederland, *Boek/URI-strategie: Aanzet tot een nationale URI-Strategie voor Linked Data van de Nederlandse overheid*.  
Online: <http://www.pilod.nl/wiki/Boek/URI-strategie>  
(geraadpleegd 26 september 2017)

## 8 BIJLAGE: Type URI's

De PILOD-richtlijn onderscheidt drie `{type}` URI's:

- `'id'` identifier van een object (individual/instance)
- `'doc'` documentatie (metadata) over het object
- `'def'` definitie van een term in een ontologie.

Het Nationaal Archief gebruikt "doc" en "def" voor zogenaamde informationresources. "id" wordt gebruikt voor non-information resource. Een information resource is in EDM en CIDOC-CRM een set aan gegevens, bijvoorbeeld een record of een webpagina.

Voor de archiefwereld is dit een lastig onderscheid om meerdere redenen:

1. Fysieke archiefstukken kun je beschouwen als een non-information resource (iets tastbaars dat daadwerkelijk wordt overgedragen) en een information resource omdat ze ook altijd informatie bevatten.
2. Bij Born Digital archief is het onduidelijk of het een non-information resource of een information resource is. Je zou kunnen zeggen dat je alleen nog maar een information resource hebt en geen "tastbaar" iets meer.
3. Nadere Toegangen (Indexen en Archiefgidsen) hebben records die alleen maar een information resource zijn en hooguit indirect verwijzen naar iets tastbaars (de fysieke archiefstukken).

Uit gesprekken met mensen van het Kadaster en de PILOD URI-strategie blijkt dat de archiefwereld hier mogelijk afwijkend is en dat de PILOD URI-strategie hier mogelijk op aangepast moet worden.

Archieven hebben in tegenstelling tot andere instellingen zoals het Kadaster niet als startpunt de reëel bestaande wereld, maar de documentaire neerslag in de archiefstukken. Vaak zijn de archiefstukken slechts globaal ontsloten. Voor sommige archieven zijn indexen gemaakt, bijvoorbeeld op persoonsnaam. Het enige dat we dan bijvoorbeeld weten is dat de naam "E. Douwes Dekker" voorkomt in de index op de Stamboeken Ambtenaren en Gouvernementsmarine (nt00286).

### Oost-Indië: Stamboeken Ambtenaren en Gouvernementsmarine, Initialen: E.

**Zoeken**

[Terug naar zoekresultaat](#)

[Meer resultaten](#)

**Initialen E.**

**Achternaam** Douwes Dekker

**Bronverwijzing** [Nummer toegang: 2.10.36.22, inventarisnummer: 910, folionummer: 510](#)  
(Register: R)

Deze E. Douwes Dekker is mogelijk de schrijver Multatuli (Eduard Douwes Dekker) die als ambtenaar in Indië diende, maar zeker weten doen we dat niet. We hebben alleen een verwijzing naar een persoon die E. Douwes Dekker heet en een verwijzing naar een archiefstuk waarvan we niet precies weten wat daar in staat. In tegenstelling tot de situatie bij andere instellingen zoals het Kadaster zijn voor archiefinstellingen vaak information resources (zoals records in Nadere Toegangen) het uitgangspunt en van daaruit kunnen we in sommige gevallen de relatie leggen naar de historische werkelijkheid, maar vaak zal dit niet mogelijk zijn.

Het Nationaal Archief volgt de PILOD-richtlijn in het onderscheid tussen URI's van het type "doc" en het type "id", maar gebruikt ook URI's van het type "doc", zonder dat er een URI van het type "id" tegenover staat.