



## Voorstel URI strategie voor de basisregistraties

Aan Brede bijeenkomst op 30 oktober 2014  
Van Werkgroep URI-Strategie Platform Linked Data  
Datum 9 oktober 2014  
Status Derde concept

### Inleiding

Deze notitie bevat een voorstel voor een URI-Strategie voor de Basisregistraties. Het advies is tot stand gekomen in de Werkgroep URI-Strategie die in de context van het Platform Linked Data is opgericht. Dit advies sluit aan bij een onderdeel van het voorstel voor "een systematiek voor unieke objectidentificatie in het Stelsel van Basisregistraties" van de Projectgroep Utrecht.

### Aanleiding

De basisregistraties zorgen ervoor dat authentieke data van de overheid duurzaam op een centrale plaats beschikbaar is voor hergebruik. In de praktijk blijkt dat deze data toch nog moeilijk te verkrijgen is of niet gemakkelijk in samenhang kan worden gebracht met andere data. Er wordt nog veel gewerkt met kopieën, met de daarbij horende hoge kosten voor verzamelen, converteren en synchroniseren, of door het bouwen van dure landelijke (keten)voorzieningen.

Toch zijn objecten in de basisregistraties voorzien van betrouwbare identificatie. Personen hebben een BSN nummer, panden hebben een BAG identificatie enzovoort. Dit zorgt ervoor dat je naar deze objecten kunt verwijzen en ook de verwijzingen van iemand anders begrijpt. In plaats van te kopiëren of synchroniseren, zou je ook rechtstreeks kunnen linken naar een object uit een basisregistratie, als deze objecten op internet te vinden zijn. Door de data over een object op een vast adres op het internet te publiceren, kun je dit adres als identificatie van het object gebruiken en op een duurzame manier naar het object linken. Zo'n link die als identificatie wordt gebruikt, wordt ook wel een URI genoemd (Uniform Resource Identifier). Deze manier van data koppelen wordt Linked Data genoemd. Linked Data vormt zodoende een soort klittenband waarmee datasets moeiteloos aan elkaar kunnen worden geplakt.

Dit klittenband kan een simpele vorm krijgen: een URI voor elk object in de basisregistraties. Elke URI vormt één haakje of oogje van het klittenband. De URI's zorgen ervoor dat je goed kunt verwijzen naar objecten uit de basisregistraties net als de basisregistraties zelf ervoor zorgen dat een goed hergebruik van gegevens mogelijk is. Maar aan toekomstvaste, bruikbare URIs gaat heel wat denkwerk vooraf. De conclusies van dit denkwerk leggen we vast in deze URI-strategie.

Zonder URI strategie zal er steeds opnieuw worden nagedacht en gediscussieerd over het te gebruiken URI patroon en de bijbehorende beslissingen. Bovendien bestaat het gevaar dat gekozen oplossingen niet voldoende persistent en schaalbaar zijn. Tenslotte zal er een diversiteit aan URIs ontstaan, wat een negatief effect heeft op de begrijpelijkheid van en het vertrouwen in data uit het stelsel van basisregistraties.

Idealiter wordt de URI-strategie, gericht op de technische implementatie van de identificatie van authentieke data, ingebed in een bredere Linked Data Strategie, waarin ook organisatorische en beheeraspecten worden meegenomen.

T 033 460 41 00  
E [info@geonovum.nl](mailto:info@geonovum.nl)  
I [www.geonovum.nl](http://www.geonovum.nl)

bezoekadres  
Barchman Wuytierslaan 10  
3818 LH Amersfoort

postadres  
Postbus 508  
3800 AM Amersfoort



## Uitgangspunten

Een aanzet voor de URI-strategie<sup>1</sup> is gemaakt in de Pilod Linked Open Data<sup>2</sup>. Dit advies is daarop gebaseerd.

De URI strategie is gericht op (in volgorde van belangrijkheid):

**Persistentie**, door te zorgen voor oplossingen die ook stand houden als de organisatie eromheen wijzigt.

**Schaalbaarheid**, door te zorgen dat beheerskosten te overzien blijven, ook als toepassingen groeien.

**Begrijpelijkheid**, zodat de afspraken gemakkelijk worden opgepakt en overgenomen.

**Vertrouwen**, zodat organisaties te bewegen zijn om zelf strategisch te kiezen voor het linken naar andermans Linked Data en publicatie van eigen Linked Data.

**Machine-leesbaarheid**, zodat er met Linked Data werkende oplossingen kunnen worden gebouwd.

**Menselijke leesbaarheid**, om ertoe bij te dragen dat men oplossingen vertrouwt en begrijpt.

## Scope van de URI-Strategie in dit advies

Dit advies beperkt zich tot de URI-strategie voor het stelsel van Basisregistraties. Op basis daarvan kan elke basisregistratie zelf aan de slag met de implementatie en het beheermodel.

Afhankelijk van het gekozen model kan het nodig zijn om één partij aan te wijzen als beheerder van de domeinnamen, in dit geval dus van de 13 huidige basisregistraties. Mogelijk worden het er in de toekomst meer.

## Advies URI strategie voor de basisregistraties

Het is nodig om een URI patroon te kiezen dat in de basisregistraties wordt gevolgd. Het geadviseerde patroon is: `http://{domain}/{type}/{concept}/{reference}`

Dit patroon is overgenomen uit "10 Rules for persistent URIs", een veelomvattend rapport van de EU met aanbevelingen voor URIs en een waardevol overzicht van de laatste best-practices<sup>3</sup>.

Er moet een domein `{domain}` gekozen worden dat voor de URIs van de basisregistratie-objecten gebruikt zal worden. Het domein is bij voorkeur exclusief gereserveerd voor dit doel en bevat bij voorkeur geen organisatiename. Het `{path}` kan indien nodig worden gebruikt om URI's uniek te maken.

## Voorstel Domeinnaam: keuze uit twee varianten

Wij adviseren om een keuze te maken uit twee varianten die denkbaar zijn als domeinnaam voor de URI's van objecten in de basisregistraties. Hierbij is de BAG als voorbeeld genomen.

### Variant 1: Subdomein van "overheid.nl"

Elke basisregistratie krijgt een eigen subdomein op "overheid.nl". Voorbeeld: "<http://bag.overheid.nl>".

Voordelen:

Dit model levert uniforme URI's op voor alle objecten in basisregistraties. Door consequent "overheid.nl" als onderdeel van de URI's te gebruiken worden de URI's duidelijk herkenbaar en wekken ze vertrouwen bij de gebruikers van deze URI's. (Vergelijk de huisstijl van de rijksoverheid).

<sup>1</sup> <http://www.pilod.nl/wiki/Boek/URI-strategie>; en <http://www.bp4mc2.org>

<sup>2</sup> <http://www.geonovum.nl/onderwerpen/linked-en-open-data/platform-implementatie-linked-open-data>

<sup>3</sup> [https://joinup.ec.europa.eu/sites/default/files/D7.1.3%20-%20Study%20on%20persistent%20URIs\\_0.pdf](https://joinup.ec.europa.eu/sites/default/files/D7.1.3%20-%20Study%20on%20persistent%20URIs_0.pdf)



Nadelen:

- Beheerders van basisregistraties moeten de domeinnaam en de implementatie afstemmen met de beheerder van het overheid.nl-domein (i.c. Ministerie van BZK). De beheertaak vereist weinig inspanning omdat het aantal basisregistraties beperkt is. Op dit moment gaat het om 13 domeinnamen.
- Aandachtspunt is hoe om te gaan met andere registers die geen basisregistratie zijn. Sommige registers hebben al een ander adres. Verhuizen naar overheid.nl is mogelijk, maar niet vanzelfsprekend gewenst. Deze oplossing is dus minder schaalbaar, maar voor niet-basisregistraties kan eventueel ook de tweede variant gekozen worden.

**Variante 2 - Naambasisregistratie.nl**

Voorbeeld: <http://data.bag.nl>. Elke (basis-)registratie regelt haar eigen domein.

Voordelen:

Er is geen centraal register of centraal beheer nodig. Elke beheerder van een basisregistratie kan zelfstandig een domeinnaam kiezen en implementeren, dus dit is een zeer schaalbare oplossing.

Nadelen:

- Geen uniforme URI's; het stelsel straalt geen samenhang uit en is niet herkenbaar als overheidsregistratie.
- Veel voor de hand liggende domeinen zijn niet beschikbaar. Er zal in zulke gevallen een andere domeinnaam gekozen moeten worden die minder voor de hand ligt. Bijvoorbeeld "data.bagregister.nl"

## Voorstel voor het deel van de URI dat voor en na de domeinnaam komt

Voor het tweede, {type}, deel van het patroon adviseren we om één van de volgende drie waarden te gebruiken:

'id': geeft aan dat de URI een object in een register identificeert.

'doc': geeft aan dat de URI leidt naar informatie over een object in het register. De 'id' URI van een object verwijst automatisch door naar deze 'doc' URI.

'def': geeft aan dat de URI leidt naar de definitie van een term in een ontologie.

Voor het derde, {concept}, deel van het patroon adviseren we de klassenaam te gebruiken van het basisregistratieobject of een andere term die aangeeft om wat voor soort object het gaat. Dit maakt de URI begrijpelijker en kan helpen bij het uniek maken van de URI.

Voor het vierde, {reference}, deel van het patroon adviseren we de al in gebruik zijnde identifier van het object uit de basisregistratie te gebruiken. Bijvoorbeeld het 'lokaalID' uit de BGT, of de 'BAGIdentificatie' uit de BAG.

Voorbeeld van een complete URI:

Variante 1: <http://bag.overheid.nl/id/Pand/0307100000367968>

Of

Variante 2: <http://data.bagregister.nl/id/Pand/0307100000367968>



## Voorstel voor ontwikkeling tot standaard

**Het federatief overleg van de projectgroep Utrecht** zal via sponsoring vanuit het Forum het Platform Linked Data vragen om de URI-strategie verder uit te werken voor andere overheidsregisters en daarbij ook organisatorische en beheeraspecten mee te nemen. De federatie beheert de gemeenschappelijke afspraken voor de berichtenstandaard. In een eerdere reviewronde zijn veel vragen en issues voor de URI-strategie als objectidentificatie gekomen. Bij de opdracht aan het Platform zullen deze issues worden meegenomen. Als de strategie zo een bredere scope en een breed draagvlak krijgt, kan de URI-strategie zich ontwikkelen tot een standaard die voldoet aan de voorwaarden om door het Forum Standaardisatie op de 'Pas-toe-leg-uit-lijst' kan worden geplaatst.

## Gevraagde besluiten aan de PSB

Een beslissing wordt gevraagd op de volgende punten:

- Een domeinstrategie te kiezen uit de twee beschreven varianten
- Akkoord te gaan met voorstel deel 2
- Akkoord te gaan met voorstel deel 3

## Bijlage

Deze bijlage gaat in op een drietal voor de hand liggende vragen over de URI strategie: wat doen andere landen, hoe wordt de URI strategie straks beheerd, en hoe zorgen we dat de URI strategie gaat landen.

### URI strategieën in het buitenland

#### 1. UK

The UK was an early adopter of open data and of linked data in particular. Its data portal, [data.gov.uk](http://data.gov.uk), went online on 30th September 2009, 4 months after the US portal at [data.gov](http://data.gov). The UK Government adopted a policy that explicitly recognises the value of URIs as identifiers and it is clear from the policy that government URI data sets are expected to persist over the long term.

The UK example shows a URI such as <http://transport.data.gov.uk/id/road/B3178> - (the B3178 is a road in Devon)

#### 2. Estonia

Estonia, a country known for its advanced use of e-Government services, has recently published its Framework of Websites version 1.0xlii document, part of the Interoperability Framework of the State Information System version 3.0xlili. These documents do not address URIs in the sense discussed in this document, i.e. as identifiers used in data sets.

Although there is no official URI policy, Estonian's Framework of Websites document does, however, offer guidelines on Web site construction. All public administrations in Estonia are required to publish a Web site and within it there are certain sections that must be present at defined URLs.

The introduction of URIs as identifiers in the Estonian public sector should be viewed in the larger context of naming and identification practices which are recognised as being extremely important for interoperability and data quality. One current large programme, where identifier issues play major part, is Estonia's Computerised



Census 2020. This involves close linking of 24 data from a number of base registers and several projects are underway that review and improve identifier systems in these registers.

### 3. Italy

The Agenzia per l'Italia Digitale has recently published a set of guidelines for achieving semantic interoperability in the public sector through Linked Open Data. As part of their Linked Open Government Data roadmap, the agency is planning to gradually open-up data about administration, public contracts, geo-data and taxonomies, people with key positions in the public sector, public services and organisational units.

Unlike the suggestion of the UK guidelines, the Italian team decided not to use the government sector in the URI and refer directly to the concept name.

Voorbeeld <http://spcdata.digitpa.gov.it/{concept name}/{natural key}>

#### **Beheer van de URI strategie**

Als de URI strategie een 'standaard' of 'aanbeveling' wordt die voor de basisregistraties, en wellicht ook andere registraties, geldt, is het van belang dat deze URI strategie ook beheerd wordt. Bij dit beheerorgaan kan men terecht met vragen over de URI strategie en het zorgt voor doorontwikkeling / governance, enzovoort.

Dit zou geregeld kunnen worden zoals dat ook voor de gemeenschappelijke afspraken berichtenstandaard (GAB) van het stelsel wordt gedaan; beide zijn immers standaarden die bottom-up zijn ontwikkeld vanuit een community. De projectgroep Utrecht werkt voor deze federatie en kan haar rol ook voor objectidentificatie via de URI strategie vervullen. De URI strategie kan analoog aan de berichtenstandaard federatief beheerd worden in een orgaan waarin vertegenwoordigers zitten van de basisregistraties en standaardenorganisaties die nu al betrokken zijn bij de URI strategie. Anderen kunnen desgewenst later aansluiten.

#### **Adoptie van de URI strategie**

Om te zorgen dat de URI strategie gaat 'landen' is het nodig adoptie ervan actief te bevorderen. Onze suggestie is om het Platform Linked Data de opdracht te geven om hier een voortrekkersrol in te vervullen. Dit platform is een actieve community van Nederlandse gebruikers van Linked Data. Daarnaast zal de adoptie geborgd worden door opname in de gemeenschappelijke afspraken voor de berichtenstandaard. Hierdoor zal de URI strategie kunnen landen in de belangrijkste gebruikte standaarden voor de basisregistraties, nl. StUF, SUWImI en de Geostandaarden.

#### **Werkgroep URI-Strategie**

Linda van den Brink Geonovum

Hans Overbeek KOOP

Arjen Santema Kadaster

Marco Brattinga Ordina

Jeroen Baltussen DLG en Geonovum

Ine de Visser Geonovum

Martijn Odijk IenM BAG

Janette Storm Kadaster BAG

Arnold Reinders BZK

Marjan Bevelander IenM Linked Data