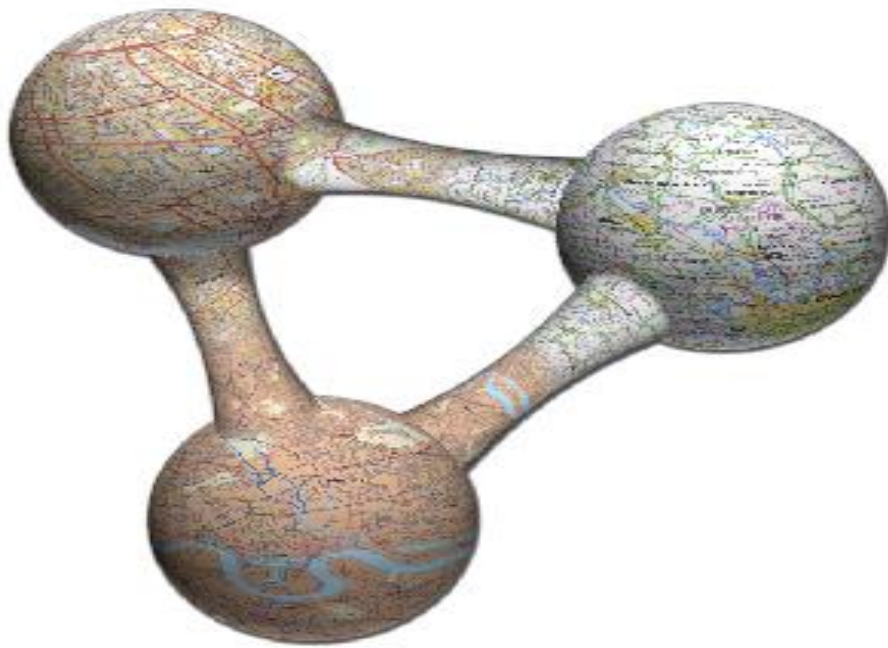




Rapport afstudeeronderzoek

Linked Open Data

De meerwaarde van Linked Open Data bij de uitwisseling van ruimtelijke waterschapsdata

Afstudeerscriptie van Elvin Martha binnen de duale opleiding Geodesie / Geoinformatica aan de Hogeschool Utrecht (2013)

COLOFON

Titel

Linked Open Data

De meerwaarde van Linked Open Data bij de uitwisseling van ruimtelijke waterschapsdata

Datum: 4 juni 2013

Opleiding

Hogeschool Utrecht

Geodesie / Geoinformatica

Cursuscode: TGEO-LAFO4-10

Student

Elvin Martha

elvinmartha@yahoo.com

Studentnummer: 1590213

Afstudeerbegeleiders

1e begeleider: Dhr. J.W. Blom MSc

2e begeleider: Dhr. Ir E. Heere

Leerbedrijf

Hoogheemraadschap Hollands Noorderkwartier (HHNK)

Bevelandseweg 1, 1703AZ Heerhugowaard

Bedrijfsbegeleider

Dhr. drs. A. Smit, Informatiemanager HHNK

a.smit2@hhnk.nl

Mede mogelijk gemaakt door:

Geodan

Netage.nl

2CoolMonkeys

Foto titelblad: <http://blog.ordnancesurvey.co.uk>



INHOUDSOPGAVE

COLOFON	2
INHOUDSOPGAVE	3
SAMENVATTING	5
VOORWOORD	6
WOORDENLIJST	7
1 INLEIDING	9
1.1 AANLEIDING	9
1.2 DOELSTELLING ONDERZOEK	10
1.3 LEESWIJZER	10
1.4 BESCHRIJVING ONDERZOEKSOMGEVING	11
1.4.1 ORGANISATIE HHNK	11
1.4.2 WATERSCHAPPEN IN NEDERLAND	12
1.4.3 WATERSCHAPSHUIS EN INFORMATIEHUIS WATER	13
2 AANPAK ONDERZOEK	14
3 OPEN DATA	15
3.1 NEDERLANDSE OPEN DATA; HET HUIDIGE AANBOD	15
3.2 OPEN GEODATA; PUBLIEKE DIENSTVERLENING OP DE KAART (PDOK)	16
3.3 OPEN WATERSCHAPSDATA; WETGEVING EN BELEID	17
3.3.1 WETGEVING	17
3.3.2 OPEN DATA BELEID	19
4 LINKED OPEN DATA	20
4.1 HET CONCEPT	21
4.2 VAN OPEN DATA NAAR LINKED OPEN DATA IN VIJF STERREN	23
4.3 WAAROM LINKED DATA?	23
4.4 DE TECHNIEK	24
4.4.1 RDF DATAFORMATEN	24
4.4.2 VOCABULAIRES	24
4.4.3 TRIPLESTORES EN SPARQL	25
4.5 NETWERK VS RELATIONEEL DATAMODEL	26
4.6 LINKED GEODATA	27
4.7 VOORBEELDTOEPASSINGEN LINKED DATA	28
4.7.1 GOOGLE KNOWLEDGE GRAPH; HET NIEUWE ZOEKEN	28
4.7.2 BBC WILD LIFE; HET WEB ALS CMS	29
4.7.3 BRANDWEER AMSTERDAM-AMSTELLAND; SNEL DE JUISTE INFORMATIE	30
4.8 NEDERLANDSE PILOT LINKED OPEN DATA	31
5 CASESTUDIE LINKED WATERSCHAPSGEODATA	32
5.1 OPEN GEODATA BIJ HHNK	32

5.2	TOEPASSINGSMOGELIJKHEDEN LINKED WATERSCHAPSDATA	35
5.3	PROOF-OF-CONCEPT LINKED WATERSCHAPSGEODATA	37
5.3.1	DE DATA	37
5.3.2	OMZETTEN NAAR LINKED DATA	38
6	CONCLUSIE EN AANBEVELINGEN	45
6.1	CONCLUSIE	45
6.2	AANBEVELINGEN VOOR HHNK	46
7	REFERENTIES	47
	BIJLAGE 1- RESULTATEN INTERVIEWS	48

SAMENVATTING

Dit rapport is het resultaat van een afstudeeronderzoek binnen de opleiding Geodesie / Geoinformatica aan de Hogeschool Utrecht. Het afstudeerbedrijf is Hoogheemraadschap Hollands Noorderkwartier (HHNK). Binnen het onderzoek is het concept Linked Open Data onderzocht met als hoofdonderzoeksvraag: "Wat is de **meerwaarde** van Linked Open Data bij de uitwisseling van ruimtelijke waterschapsinformatie?". Tevens is een stukje waterschapsdata bij wijze van proof-of-concept omgezet naar Linked Data.

Open Data is een interessante trend, waarbij overheden data die zij verzamelen voor het uitvoeren van hun taken beschikbaar stellen voor **hergebruik**. In de Verenigde Staten en Engeland heeft Open Data vanuit de overheid de afgelopen jaren een flinke vlucht genomen. Het openstellen van data wordt op het moment ook in Nederland steeds minder vrijblijvend. Vanuit Europese en Nederlandse regelgeving worden overheden gestimuleerd om hun data openbaar te maken. De **groei** van het aanbod Open Data brengt tegelijkertijd nieuwe uitdagingen met zich mee. Data worden aangeboden vanuit verschillende aanbieders en in **diverse formaten**. Om de kracht van Open Data optimaal te benutten blijkt alleen het online aanbieden van de data niet genoeg.

Het concept van **Linked Open Data** gaat dan ook een stap verder. Het neemt drempels weg voor het hergebruik van Open Data door harmonie te brengen in de wijze waarop diverse data wordt aangeboden en door **relaties** te leggen tussen de verschillende data. Data vanuit verschillende domeinen en vakgebieden worden in een **gelijksoortig datamodel** aangeboden en kunnen zo gemakkelijk met elkaar worden gecombineerd. Geodata kan zo bijvoorbeeld gemakkelijker worden gecombineerd met niet-geodata, waardoor nieuwe dwarsverbanden en inzichten kunnen ontstaan.

Er ontstaan ook onverwachte en nuttige toepassingen in de markt, zonder dat de bronhouder in kwestie in de ontwikkeling ervan hoeft te investeren. Het is nog sterker; slechts door Linked Data beschikbaar te stellen stimuleren de Nederlandse overheden een **economische groei** in de Nederlandse IT-sector. In het buitenland hebben grote organisaties zoals Google en de BBC de principes van Linked Data omarmd om de gebruikerservaring van hun diensten verder te verbeteren. Zij gebruiken het nieuwe web van data om meer **relevante content** aan te bieden.

Voor ruimtelijke data in het algemeen blijkt de technische stap naar Linked Data niet heel groot. Het effect ervan echter wel. De **grens** tussen geodata en niet-geodata wordt langzaam **vager**, doordat beiden gemakkelijker gecombineerd kunnen worden gebruikt.

Waterschappen, zo ook HHNK, blijken over het algemeen niet voorop te lopen in de Open Data discussie. Laat staan in die van Linked Open Data. Toch kan dit ze wel degelijk iets opleveren. Het aanbieden van Linked Open Data kan ervoor zorgen dat ruimtelijke waterschapsinformatie **dichter**, maar ook **gericht** bij de burger komt. Te denken valt aan informatie uit een Logger voor de waterkeringen die in een applicatie voor consumenten wordt getoond samen met andere automatisch opgehaalde informatie over een woning of bedrijfspand. Het investeren in innovatie op het vlak van Linked Open Data zal ook de **samenwerking** met (nieuwe) partners verbeteren, doordat informatie-uitwisseling veel minder hoeft te worden gezocht in het koppelen van systemen of het overdragen van data.

Linked Open Data blijkt dan ook meer te zijn dan een modewoord in de informatiewereld. Het zal als **smeermiddel** fungeren bij het gebruik en de uitwisseling van diverse soorten Open Data.

VOORWOORD

Dit rapport is de afronding van mijn duale studie Geodesie/Geoinformatica aan de Hogeschool Utrecht. Ik ben deze studie in 2010 gestart en rond hem in 2013 af.

Allereerst wil ik Jan Strijker, Sjaak van Popering en Ruben Stekelenburg bedanken, omdat zij het als managers binnen HHNK voor mij hebben mogelijk gemaakt om deze studie te volgen. Daarnaast hebben verschillende collega's binnen het Geobureau, nu afdeling Informatie & Automatisering, mij tijdens mijn opleiding geholpen met input en advies.

Verschillende personen hebben mij tijdens mijn afstudeeronderzoek geholpen in mijn zoektocht. Ik bedank hierbij Bart van Leeuwen, de Linked Data firefighter, voor zijn tips en connecties, Frans Knibbe van Geodan voor zijn hulp bij de conversie naar Linked Data, mijn afstudeerbegeleider en teamleider Arno Smit, Reind van Olst van 2CoolMonkeys, collega's Bart Bos en Kier van Gijssel bij HHNK en directeur Bedrijfsvoering bij HHNK Ton Hugens.

Mijn gezin heeft een grote bijdrage geleverd aan mijn studie, doordat ik mij veelvuldig moest vrijmaken voor opdrachten en tentamens. Hoewel ik steeds heb geprobeerd om de balans hierin te vinden is mijn succesvol afronden van deze studie zeker ook hun verdienste.

Als laatste draag ik dit afstudeeronderzoek in het bijzonder op aan mijn dochter Jaydee om haar hiermee hopelijk op latere leeftijd te inspireren om uitdagingen aan te gaan en deze te overwinnen.



Elvin Martha

WOORDENLIJST¹

API (Application Programming Interface): een set aan definities waarmee softwareprogramma's onderling kunnen communiceren.

Aquo-standaard: een semantische datastandaard, waardoor het mogelijk is om op een uniforme manier gegevens uit te wisselen tussen partijen die betrokken zijn bij het waterbeheer in Nederland.

ESRI: Internationaal leverancier van GIS-software en geodatabase management applicaties.

Geonovum: een stichting die zich inzet voor laagdrempelige toegang tot geo-informatie in Nederland en de standaarden die daarvoor nodig zijn beheerd.

GML (Geography Markup Language): een door het OGC opgestelde XML structuur voor de representatie van geografische (ruimtelijke en plaatsgebonden) informatie.

INSPIRE (Infrastructure for Spatial Information in the European Community): Europese richtlijn om de beschikbaarheid, kwaliteit, organisatie, toegang tot en uitwisseling van geo-informatie in Europa te verbeteren.

Legger: een openbaar register waarin wordt aangegeven aan welke eisen de waterstaatswerken moeten voldoen voor wat betreft ligging, vorm, afmeting en constructie.

OGC (Open Geospatial Consortium): een internationaal consortium voor geo-standaarden.

Ontologie: het doel van ontologie is het beschrijven van informatie op het Internet en de relaties tussen de gegevens.

OWL (Web Ontology Language): een standaard voor ontology vastgesteld door het W3C.

PostGIS-database: Open source database voor vastleggen van ruimtelijke data. PostGIS is een uitbreiding op de Open source relationele database PostgreSQL.

Semantiek: de wetenschap die zich bezig houdt met de betekenis van taalkundige constructies zoals woorden (woordsemantiek) en zinnen (zinssemantiek).

SKOS (Simple Knowledge Organization System): een algemeen datamodel om kennissystemen te koppelen.

SPARQL (SPARQL Protocol and RDF Query Language): een RDF zoektaal die gebruikt wordt om RDF-gebaseerde data te bevragen middels zoekopdrachten.

SQL (Structured Query Language): een gestandaardiseerde taal die gebruikt kan worden voor taken zoals het bevragen en het aanpassen van gegevens in een relationele gegevensbank.

RDF (Resource Description Framework): RDF is een netwerk datamodel bestemd voor het uitwisselen van data over het internet. RDF is één van de bouwstenen van Linked Data.

¹ Bronnen: Wikipedia, Geonovum, W3C, Waterwet

URI (Uniform Resource Identifier): een internet-protocolelement, een unieke benaming van een "bron", een stuk informatie, data of dergelijke.

W3C (World Wide Web Consortium): een internationale organisatie die zich richt op standaardisatie binnen het World Wide Web. Het W3C is ontwikkelaar en beheerder van diverse standaarden.

WCS (Web Coverage Service): een protocol voor het uitwisselen van geografische rasterdata. WCS wordt beheerd door het Open Geospatial Consortium.

WFS (Web Feature service): een protocol voor het aanbieden van services waarmee geografische informatie kan worden geraadpleegd, maar waarbij de onderliggende data ook kan worden gedownload en bewerkt (download-service). WFS wordt beheerd door het Open Geospatial Consortium.

WMS (Web Map Service): een protocol voor het uitwisselen van geogerefereerde kaartbeelden over het internet (view-service). WMS wordt beheerd door het Open Geospatial Consortium.

1 INLEIDING

1.1 Aanleiding

We gaan allereerst een stukje terug in de tijd..

Het was in het jaar 1086 toen er in opdracht van de toenmalige Engelse koning William I, ook wel bekend als William de Veroveraar, in Engeland een grootschalige gegevensinventarisatie plaats vond. Deze had als hoofddoel belasting te kunnen heffen in zijn nieuw veroverde land. Er werden circa 63.000 personen geïnterviewd over het bezit van grond en andere bezittingen in ongeveer 13.000 verschillende dorpen en nederzettingen. Deze inventarisatie ging de geschiedenisboeken in als het zogenaamde Domesday Book en staat bekend als de eerste grote inventarisatie waarbij er door de overheid data werden verzameld en geanalyseerd.

Tegenwoordig verzamelen overheden nog steeds gegevens om hun diverse taken te kunnen uitvoeren. Anders als in de tijd van William de Veroveraar gelukkig niet alleen om belasting te heffen en ook worden deze data steeds meer open gesteld aan het publiek. Open Data is een maatschappelijke trend, waarbij overheden data die zij in het kader van hun openbare taken hebben verzameld ter beschikking stellen voor hergebruik. Open Data wordt vanuit Europese en landelijke wetgeving steeds meer gestimuleerd.

Onderstaand citaat komt uit een recent nieuwsartikel ² over een door de Europese Commissie aangenomen voorstel ³. Hierin staat dat Europese overheden zo veel mogelijk data gratis beschikbaar moeten gaan stellen aan alle Europese burgers.

“We sturen vandaag een krachtig signaal aan overheden. Je data is meer waard als je het weggeeft”, aldus Eurocommissaris Neelie Kroes (Digitale Agenda). “Belastingbetalers hebben al voor deze informatie betaald. Het minste dat we kunnen doen is het teruggeven aan diegenen die het op nieuwe manieren willen gebruiken, op manieren die mensen helpen en banen en groei creëren.”

De Europese Commissie voorspelt dat het openbaar maken van overheidsdata de Europese economie 40 miljard euro zou kunnen opleveren. De nu nog niet beschikbare data noemt de Commissie “een goudmijn van ongerealiseerd economisch potentieel.”



Neelie Kroes (foto <http://www.nexstdigital.com>)

Het openstellen van data wordt dus steeds minder vrijblijvend en biedt ook economische kansen. Het aanbod van Open Data zal de komende jaren hoogstwaarschijnlijk dan ook groeien. Bij een groeiend aanbod ontstaat echter ook een nieuwe uitdaging. De grote hoeveelheden data zijn van verschillende domeinen, typen en formaten, waardoor het combineren van de data in applicaties vaak een lastige opgave is. Om optimaal gebruik te kunnen maken van Open Data is het van groot belang dat de data uitwisselbaar is. Alleen het vrijgeven van data is niet genoeg. De principes van Linked Open Data bieden hiertoe mogelijk de sleutel. Zou Linked Open Data meer zijn dan een nieuw modewoord in de informatiewereld?

² Nieuwsbericht “EU-wil-meer-overheidsdata-openbaar-maken” (www.nu.nl 10 april 2013)

³ Revisie van de “2003 Public Sector Information Directive” (Brussel, 10 april 2013)

1.2 Doelstelling onderzoek

Het doel van dit onderzoek is drieledig; als eerste zal het inzicht geven in het concept Linked Open Data. Daarnaast zal het de meerwaarde van Linked Open Data voor ruimtelijke waterschapsdata beschrijven en als laatste zal het een werkwijze beschrijven waarop Linked Open geodata vanuit het waterschapsdomein tot stand kan komen.

1.3 Leeswijzer

Hoewel de algemene delen van dit rapport goed leesbaar zullen zijn voor de minder ingewijde lezer, vereisen de meer inhoudelijke delen enige voorkennis. Het rapport is gericht op de lezer met een geodetische en/of informatica achtergrond.

Natuurlijk is het aan te raden om het hele rapport te lezen, maar het kan zomaar zijn dat slechts enkele onderdelen voor nu relevant zijn. De indeling van dit document is als volgt:

- Dit rapport geeft in hoofdstuk 1 en 2 een beschrijving van de afstudeeromgeving en de aanpak van het onderzoek.
- In hoofdstuk 3 wordt het onderwerp Open Data geïntroduceerd om vervolgens in hoofdstuk 4 het concept Linked Open Data uit te leggen.
- In hoofdstuk 5 wordt ingegaan op een praktijkvoorbeeld binnen de context van waterschap Hoogheemraadschap Hollands Noorderkwartier.
- In hoofdstuk 6 worden de conclusies en aanbevelingen beschreven.
- De gebruikte referenties zijn in het laatste hoofdstuk 7 op een rij gezet.
- In de bijlagen zijn vragenlijsten opgenomen van de afgenomen interviews

1.4 Beschrijving onderzoeksomgeving

1.4.1 Organisatie HHNK

Mijn afstudeerbedrijf is Hoogheemraadschap Hollands Noorderkwartier (HHNK). HHNK heeft een viertal kerntaken:

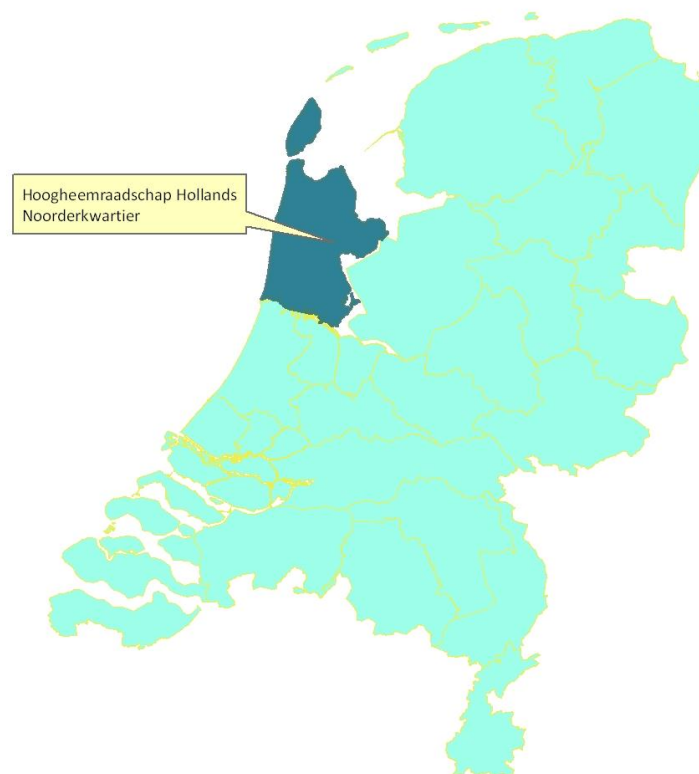
- waterkeringenbeheer: beschermen van het land tegen overstromingen;
- waterkwantiteit: regelen van de juiste waterstand;
- waterkwaliteit: schoonhouden van het oppervlaktewater en zuiveren afvalwater;
- wegenbeheer: beheer en onderhoud van polderwegen.

HHNK in cijfers (bron: website HHNK 2013)

- Een beheergebied van 196.400 hectare
- 20.116 km oppervlaktewaterlichamen
- 341 km primaire waterkeringen
- 1.486 regionale waterkeringen
- Ca 1.500 km wegen
- 550 km persleidingen
- Er wonen 1,1 miljoen personen (boven 18 jaar) in het beheergebied
- Formatie van HHNK bedraagt 495 FTE

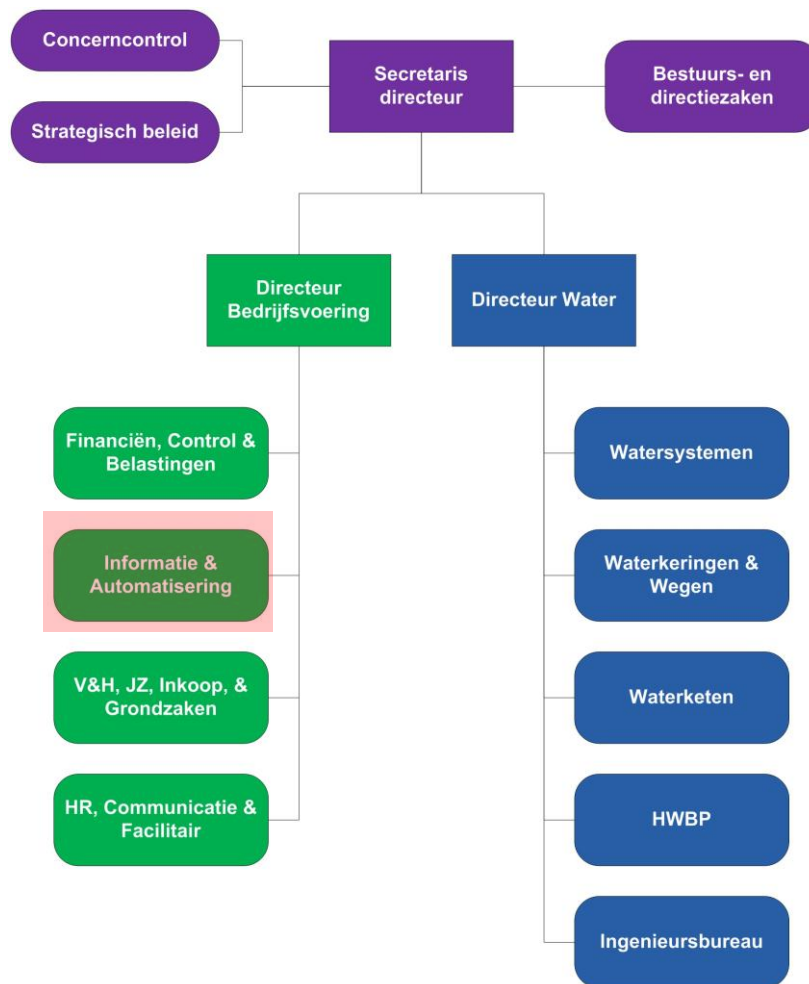


Het beheergebied van HHNK beslaat het gebied Noord-Holland boven het Noordzeekanaal en het waddeneiland Texel.



Figuur 1.1 Beheergebied HHNK

Ik ben bij HHNK werkzaam als projectleider binnen de afdeling Informatie & Automatisering (I&A), één van de ondersteunende afdelingen binnen HHNK. In onderstaand organogram staat de organisatiestructuur van HHNK weergegeven. De afdeling I&A is met een rode arcering aangegeven.



Figuur 1.2 Organogram HHNK

1.4.2 Waterschappen in Nederland

Op dit moment zijn er 24 waterschappen in Nederland. De waterschappen verschillen van elkaar in grootte maar ook in taken. Zo beheert niet ieder waterschap wegen en zijn er ook waterschappen die drinkwater leveren. Binnen alle Nederlandse waterschappen zijn totaal circa 11.000 medewerkers werkzaam.

Er vinden nog regelmatig fusies plaats tussen de waterschappen. Zo zullen per 1 januari 2014 waterschap Regge en Dinkel en waterschap Velt en Vecht samen verder gaan als waterschap Vechtstromen. De teller zakt dan naar 23 waterschappen.

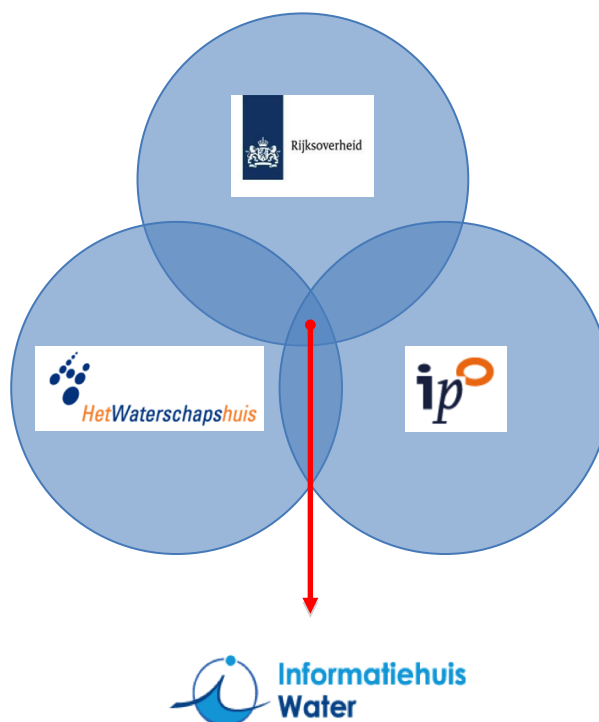
1.4.3 Waterschapshuis en Informatiehuis Water

Het Waterschapshuis

In een onderzoek rondom de informatievoorziening van een waterschap moet ook het Waterschapshuis worden genoemd. Het Waterschapshuis is een regie- en uitvoeringsorganisatie voor de waterschappen op het gebied van Informatie en Communicatie Technologie. Het Waterschapshuis heeft als doel het bevorderen van samenwerking op het gebied van ICT tussen de waterschappen en de andere overheden die actief zijn in de watersector.⁴ Zo bemiddelt het Waterschapshuis bijvoorbeeld namens de waterschappen bij de aanbesteding en realisatie van de AHN (Actuele Hoogtekaart Nederland) en ontzorgt het de waterschappen door een portaal voor waterschapsinformatie op internet te beheren (de Geovoorziening).

Informatiehuis Water

Het Informatiehuis Water (IHW) is een samenwerkingsverband tussen het Rijk, het Interprovinciaal Overleg (IPO) en de waterschappen. Het Informatiehuis Water biedt diverse producten en diensten. Deze zijn gericht op het verbeteren van de uitwisseling van informatie tussen waterbeheerders ter voorbereiding, uitvoering en evaluatie van het waterbeleid. Het Informatiehuis Water is daarbij dus een uitvoeringsorganisatie voor en door de deelnemende waterpartners.⁵ Het belangrijkste product van het IHW is de Aquo-standaard.⁶ Deze standaard heeft als doel op een uniforme manier gegevens uit te wisselen tussen partijen die betrokken zijn bij het waterbeheer in Nederland.



Figuur 1.3 Het Informatiehuis Water opereert als samenwerkingsverband tussen de 3 Nederlandse waterpartners.

⁴ Bron: www.hetwaterschapshuis.nl

⁵ Bron: www.informatiehuwater.nl

⁶ De Aquo-standaard is een semantische standaard voor het vastleggen van gegevens in de Nederlandse watersector.

2 AANPAK ONDERZOEK

De hoofdonderzoeksvraag luidt: "Wat is de meerwaarde van Linked Open Data bij de uitwisseling van ruimtelijke waterschapsinformatie?".

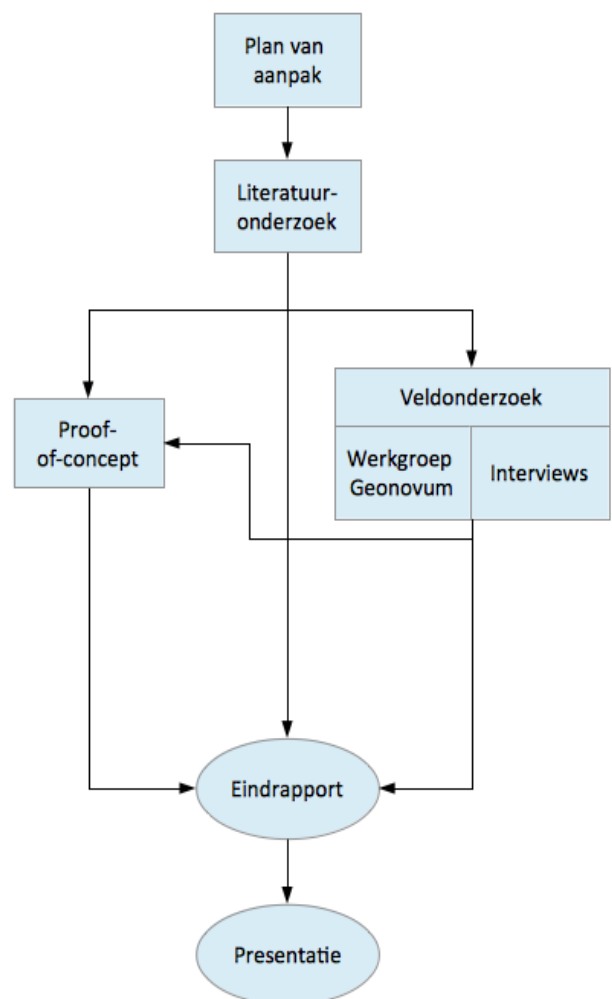
Hierbij de volgende hulpvragen:

- Wat is Open Data?
- Wat is Linked Open Data?
- Wat zijn de voordelen van Linked data?
- Wat is het verschil met de huidige uitwisselformaten voor ruimtelijke gegevens?
- Hoe komt Linked (geo)data technisch gezien tot stand?
- Wat zijn de ontwikkelingen op dit gebied binnen en buiten Nederland?
- Wat is de meerwaarde voor de Nederlandse waterschappen om hun geodata als Linked data beschikbaar te stellen?
- Wat is de meerwaarde voor andere stakeholders binnen het waterdomein van Linked waterschapsdata?
- Welke stappen moeten binnen HHNK genomen worden om Linked data te kunnen aanbieden?

Het onderzoek is gestart met een literatuuronderzoek. De geraadpleegde bronnen zijn te vinden in [hoofdstuk 7 "Referenties"](#). Om meer kennis en ervaring over het onderwerp op te doen is aangesloten bij het landelijke project "Pilot Linked Open Data⁷" van Geonovum, specifiek de actielijn Techniek. Daarnaast zijn verschillende personen binnen het waterdomein geïnterviewd. De verslagen hiervan zijn in [Bijlage 1](#) opgenomen.

Het realiseren van een proof-of-concept van ruimtelijke linked waterschapsdata maakt deel uit van de scope van dit onderzoek. Zowel het literatuuronderzoek als de kennis en ervaring van verschillende specialisten hebben geholpen dit te realiseren.

In figuur 2.1 staat de aanpak van het onderzoek geschematiseerd.



Figuur 2.1 Schematisatie onderzoek

⁷ Het project "Pilot Linked Open Data", geïnitieerd door Geonovum, heeft als doelstelling om zichtbaar te maken hoe Linked Data kan werken, door middel van het organiseren van Nederlandse pilots.

3 OPEN DATA

Open data is een concept waarbij niet-persoonsgebonden (overheids)data beschikbaar wordt gesteld voor hergebruik. In de Verenigde Staten en Engeland stelt de overheid sinds een aantal jaren op grote schaal data beschikbaar via de websites data.gov en data.gov.uk. Ook in Nederland wordt gestart met het open stellen van overheidsdata. Veel organisaties zijn momenteel bezig met het opstellen van een Open Data beleid. Een aantal ministeries en de provincies hebben voorgenomen om per 2015 overheidsdata vrij te geven aan derden.

De data zijn openbaar als wordt voldaan aan de onderstaande voorwaarden⁸;

- Er berust geen auteursrecht of andere rechten van derden op;
- De data zijn bekostigd uit publieke middelen, beschikbaar gesteld voor de uitvoering van die taak;
- De data voldoen bij voorkeur aan 'open standaarden' (geen barrières voor het gebruik door ICT-gebruikers of door ICT-aanbieders);
- Open Data is bij voorkeur computer-leesbaar, zodat zoekmachines informatie in documenten kunnen vinden.

3.1 Nederlandse Open Data; het huidige aanbod

In Nederland is het aanbod van Open Data raadpleegbaar op de website www.data.overheid.nl. Deze website is een register met datasets en verwijzingen naar de weblocaties waar de bronhouders deze data beschikbaar stellen. Hoewel (nog) niet alle Nederlandse Open Data op de website is geregistreerd geeft deze een aardig beeld van de huidige stand van zaken. Op het moment telt het register totaal 5711 verwijzingen naar Open Datasets van 50 aanbieders. Dit zijn organisaties zoals het Kadaster, TNO, gemeentes en het Rijk. In figuur 3.1 zijn een aantal van deze aanbieders met de omvang van hun aanbod op www.data.overheid.nl gevisualiseerd.



Figuur 3.1
Impressie van Open Data aanbieders op
www.data.overheid.nl

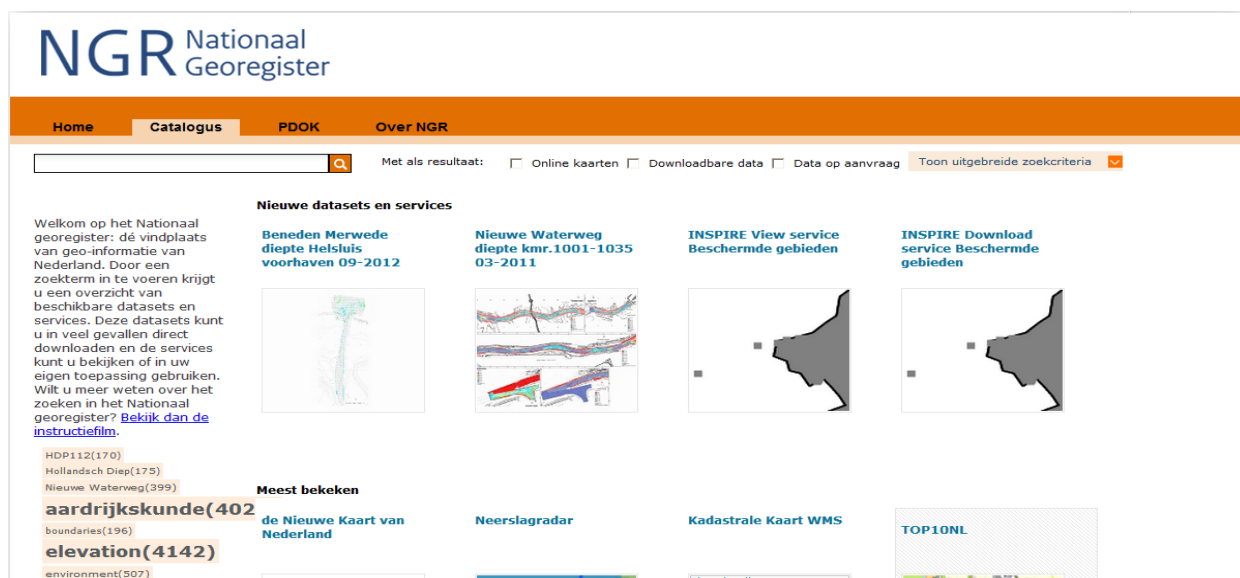
⁸ Bron: www.data.overheid.nl

3.2 Open geodata; Publieke Dienstverlening Op de Kaart (PDOK)

Publieke Dienstverlening Op de Kaart (PDOK) is een voorziening voor het ontsluiten van Open geodata van nationaal belang. PDOK is tot stand gekomen door een samenwerking tussen het Kadaster, de ministeries van Infrastructuur en Milieu en Economische Zaken, Rijkswaterstaat en Geonovum.

PDOK bestaat uit 3 onderdelen;

- allereerst biedt PDOK op het moment 42 webservices (WMS / WFS) aan via het **PDOK loket**. Voor de INSPIRE⁹-plichtige datasets is ook een downloadservice beschikbaar, waarbij de bronbestanden kunnen worden gedownload. Downloads blijken meestal in GML of ESRI Shape-formaat te zijn.
- via **PDOK kaart** kunnen gebruikers zelf een webkaart opmaken en deze in een eigen website gebruiken. Naast de PDOK services kunnen eigen services worden toegevoegd en als achtergrondkaart wordt de BRT¹⁰ topografische achtergrond of een PDOK luchtfoto aangeboden. PDOK Kaart kan hiermee een alternatief zijn voor commerciële kaarten zoals bijvoorbeeld Google Maps.
- Ten slotte beheert en exploiteert PDOK het **Nationaal Georegister (NGR)**. Het NGR is een catalogus van geodatasets in Nederland. De gegevens worden aangeleverd door de aanbieders van geo-informatie in Nederland. Anders dan bij www.data.overheid.nl zijn dit naast de Nederlandse overheden ook verschillende onderzoeksinstituten en bedrijven. Op het moment van schrijven worden er in het Nationaal Georegister 6275 geodatasets aangeboden in de vorm van een verwijzing naar een downloadlocatie of een webservice.



Figuur 3.2 Screenshot Nationaal Georegister (www.nationaalgeoregister.nl)

⁹ De Europese richtlijn INSPIRE helpt om de beschikbaarheid, kwaliteit, organisatie, toegang tot en uitwisseling van geo-informatie in Europa te verbeteren. (bron: Geonovum)

¹⁰ Basisregistratie Topografie

3.3 Open waterschapsdata; wetgeving en beleid

3.3.1 Wetgeving

Er is verschillende wetgeving met betrekking tot het openstellen van data vanuit de waterschappen: de Europese richtlijn INSPIRE, de Waterwet en het Forum standaardisatie.

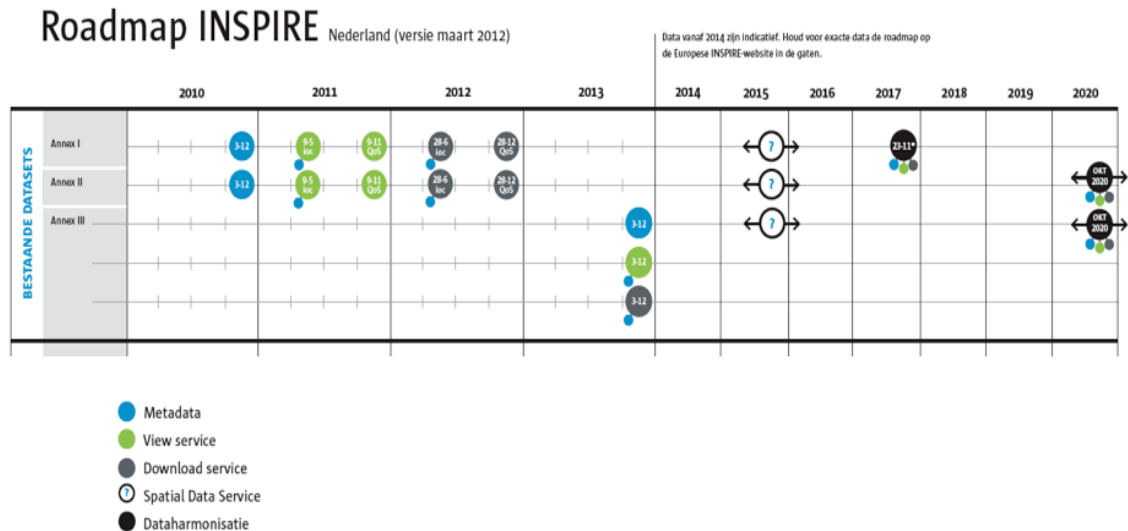
- **INSPIRE**

INSPIRE staat voor: “Infrastructure for Spatial Information in the European Community” en is een Europese Kaderrichtlijn. Hierin wordt gesteld dat gegevens over ruimtelijke ordening en milieu binnen de Europese Unie digitaal uitwisselbaar moeten worden gemaakt. Net als provincies en gemeenten moeten ook de waterschappen voldoen aan deze richtlijn. De verplichting bestaat uit de publicatie van datasets, inclusief metadata, via een (landelijke) service. De invoering van INSPIRE gebeurt in drie fasen: annex I, annex II en annex III. Binnen iedere fase worden er gegevens van verschillende thema's gepubliceerd. Bronhouders moeten de gegevens waarvoor zij zijn aangemerkt conform de gestelde specificaties beschikbaar stellen. In figuur 3.3 zijn de thema's per fase weergegeven en is aangeduid welke thema's van toepassing zijn op de waterschappen.

Thema's INSPIRE annex I, II en III	
<u>annex I</u>	<u>annex III</u>
1. Systemen voor verwijzing door middel van coördinaten	1. Statistische eenheden
2. Geografisch rastersysteem	2. Gebouwen
3. Geografische namen	3. Bodem
4. Administratieve eenheden	4. Landgebruik
5. Adressen	 5. Menselijke gezondheid en veiligheid
6. Kadastrale percelen	 6. Nutsdiensten en overheidsdiensten
7. Vervoersnetwerken	 7. Milieubewakingsvoorzieningen
 8. Hydrografie	8. Faciliteiten voor productie en industrie
9. Beschermde gebieden	9. Faciliteiten voor landbouw en aquacultuur
	10. Spreiding van de bevolking – demografie
	 11. Gebiedsbeheer, gebieden waar beperkingen gelden, gereguleerde gebieden en rapportage-eenheden
<u>annex II</u>	 12. Gebieden met natuurrisico's
 1. Hoogte	13. Atmosferische omstandigheden
2. Bodemgebruik	14. Meteorologische geografische kenmerken
3. Orthobeeldvorming	15. Oceanografische geografische kenmerken
4. Geologie	16. Zeegebieden
	17. Biogeografische gebieden
	18. Habitats en biotopen
	 19. Verspreiding van soorten
	20. Energiebronnen
	21. Minerale bronnen
 Bevat verplichting voor de waterschappen	

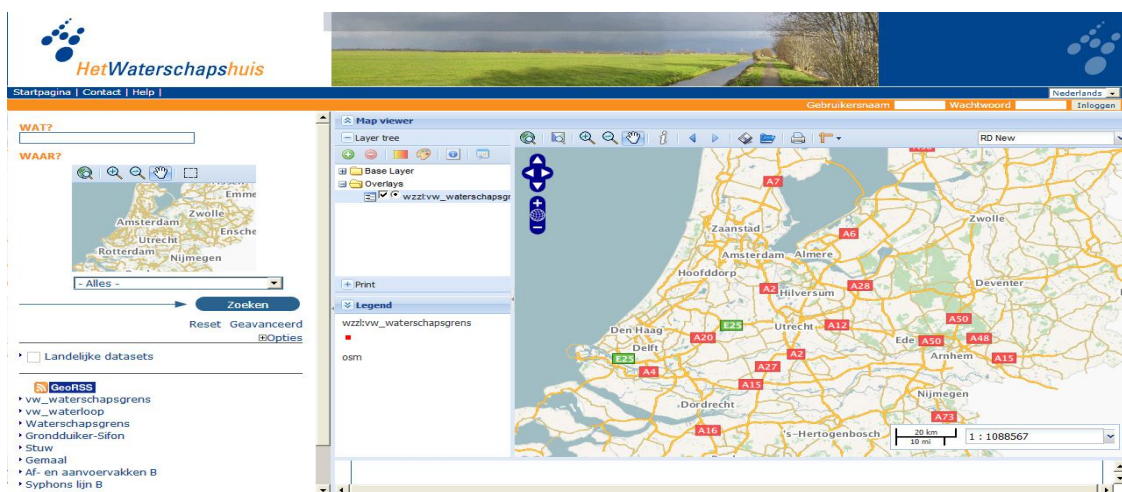
Figuur 3.3 INSPIRE thema's per fase (Afgeleid van overzicht Geonovum)

De planning (zie figuur 3.4) is dat de thema's binnen annex I en annex II eind 2012 als download service (WFS) beschikbaar moesten zijn. Voor de thema's binnen annex III geldt dat dit eind 2013 het geval moet zijn. Voor de waterschappen is de exacte invulling van hun verplichting onder annex III op dit moment echter nog niet volledig uitgekristalliseerd.



Figuur 3.4 Roadmap realisatie INSPIRE thema's (bron: Geonovum)

Om de waterschappen te ontzorgen in hun INSPIRE verplichting heeft het Waterschapshuis een portaal ontwikkeld, genaamd de Geovoorziening. Binnen de Geovoorziening worden de data per waterschap aangeboden, maar ook als een landsdekkend beeld. Dit landsdekkend beeld is echter op het moment nog niet voor alle thema's binnen annex I en annex II beschikbaar en ook zijn nog niet alle thema's als download service (WFS) beschikbaar, maar meestal als view service (WMS).



Figuur 3.5 Screenshot Geovoorziening Waterschapshuis

(<http://catalog.waterschapsservices.nl/srv/nl/main.home>)

- **Waterwet**

In de Waterwet zijn regels vastgelegd met betrekking tot het beheer en gebruik van watersystemen. De waterschappen zijn op grond van artikel 5.1 van de Waterwet verplicht om zogenaamde leggers voor alle waterstaatswerken op en vast te stellen. Een legger is een openbaar register waarin staat aangegeven waaraan de waterstaatswerken moeten voldoen wat betreft ligging, vorm, afmeting en constructie. Onder deze waterstaatswerken vallen onder andere de wateren, kunstwerken en waterkeringen. Een legger bevat onder meer een overzichtskaart, waarop de ligging van waterstaatswerken en daaraan grenzende beschermingszones staat aangegeven. De leggergegevens moeten openbaar worden gemaakt, maar de Waterwet geeft geen omschrijving over de vorm of mate van openheid van deze gegevens.

- **Forum standaardisatie; “pas toe of leg uit” lijst**

Net als andere overheden wordt van de waterschappen verwacht dat zij Open Source en Open Standaarden als alternatief overwegen. Het Forum Standaardisatie is een overheidsorgaan dat adviseert over hoe open standaarden de digitale uitwisseling van informatie tussen overheden kunnen bevorderen. Het Forum heeft een zogenaamde “pas toe of leg uit” lijst opgesteld waarop onder meer standaarden worden vermeld die gebruikt dienen te worden bij het publiceren van data. Voor geodata betreffen dit NEN 3610, GML en WMS / WFS. Geonovum¹¹ beheert de Nederlandse profielen voor deze standaarden.

3.3.2 Open Data beleid

In het rapport “Waterschappen en Open Data beleid: een afwegingskader” (Informatiehuis Water, 22 januari 2013) worden een aantal scenario’s geschetst op basis waarvan de waterschappen een Open Data beleid zouden kunnen insteken. Deze scenario’s staan in dit rapport als volgt omschreven:

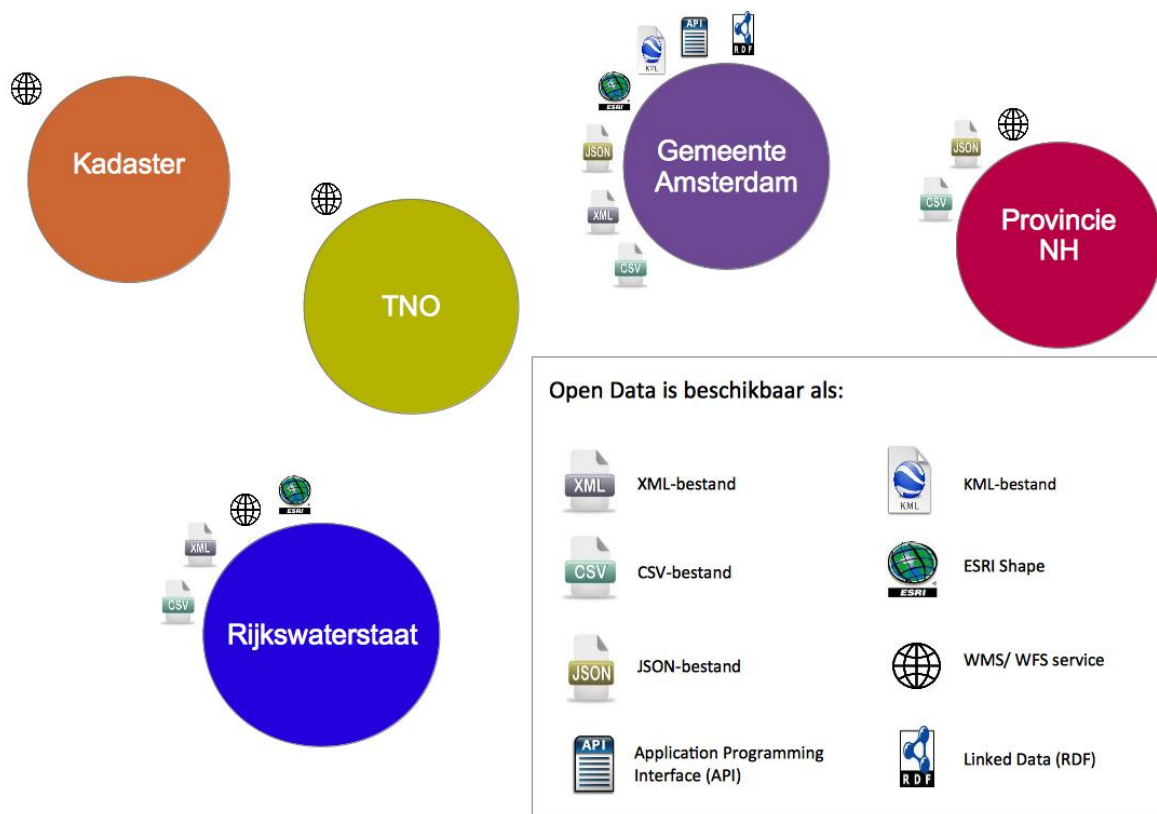
- 0: ‘As is’, Beleid zoals het nu is. Veel data gesloten, minimale kosten en baten. Met dit scenario wordt in veel gevallen niet voldaan aan de huidige wettelijke verplichtingen.
- 1: Statisch gefaseerde invoer. Oplopende, beperkte inspanning en kosten. Gegevens worden als bestanden vrijgegeven zodat gebruikers handmatige handelingen moeten verrichten. Historische gegevens worden alleen op verzoek vrijgegeven. In dit scenario kunnen gebruikers wel gegevens hergebruiken maar niet in een app.
- 2: Dynamisch gefaseerde invoer. Oplopende inspanning. Gegevens worden zo vrijgegeven dat software systemen deze automatisch kunnen gebruiken. Historische data komt ook beschikbaar als open data. In dit scenario kunnen eenvoudig toepassingen op basis van de data worden gebouwd. Kosten over de tijd verspreid, relatief hoge baten.
- 3: Volledig Open. Ook dit is een gefaseerd scenario maar met een hoog ambitieniveau. Op korte termijn worden data sets zo ontsloten dat ze herbruikbaar zijn in toepassingen, maar ook zo dat ze verwijzen naar externe datasets. Relatief hoge kosten in beperkte periode. Extra baten ten opzichte van scenario 2 op de langere termijn.

Anders dan de uitgesproken ambitie van een aantal ministeries en de provincies om vanaf 2015 hun data vrij te geven blijft het voor de waterschappen op dit moment bij een afwegingskader. Hoewel het Informatiehuis Water het tweede scenario adviseert blijft het vrijblijvend en kunnen de waterschappen hier zelfstandig invulling aan geven.

¹¹ Geonovum is een stichting die zich inzet voor laagdrempelige toegang tot geo-informatie en ontwikkelt en beheert de standaarden die daarvoor nodig zijn. (Bron: Geonovum.nl)

4 LINKED OPEN DATA

Het aanbod van Open Data groeit snel. De data zijn vanuit verschillende aanbieders beschikbaar, maar ook in verschillende (open) dataformaten. Zelfs vanuit één aanbieder komt het regelmatig voor dat de data per thema in verschillende bestandsformaten wordt aangeboden. In figuur 4.1 worden de aanbieders uit het voorbeeld in paragraaf 3.1 met elkaar vergeleken op het vlak van diversiteit in aangeboden dataformaten zoals gevonden op www.data.overheid.nl.



Figuur 4.1 Diversiteit aan gegevensformaten bij aanbieders Open data

De data zijn weliswaar open en beschikbaar, maar kan vaak door ontwikkelaars niet direct onderling worden gecombineerd (soms zelfs niet vanuit één aanbieder). Informatie uit een ESRI shape van de gemeente Amsterdam kan bijvoorbeeld niet zonder conversie naar een gemeenschappelijk formaat worden gecombineerd met informatie uit een XML-bestand van Rijkswaterstaat. Hetzelfde geldt voor informatie uit een WFS-service en een CSV-bestand die beiden door de Provincie Noord-Holland worden aangeboden. Er zullen bewerkingen en schaduwdata bases nodig zijn om binnen een toepassing gecombineerd hergebruik mogelijk te maken.

Om optimaal gebruik te kunnen maken van Open Data zijn uitwisselbaarheid, standaardisatie en context de sleutels. Hierin biedt het concept van Linked Open Data mogelijkheden.

4.1 Het concept

Tim Berners Lee, bekend als één van de grondleggers van het World Wide Web (WWW) en directeur van het World Wide Web Consortium, bedacht in 2007 het concept van Linked Data. Linked Data wordt vaak in één adem genoemd met het Semantische Web, web 3.0 of web van data. Het internet bestaat uit een web van documenten, muziek, foto's, video's, etc verbonden door middel van hyperlinks. Het concept Linked Data verbindt alle ruwe (Open) Data via het web met elkaar gebruikmakend van een gemeenschappelijk datamodel en van semantiek. We starten door in te zoomen op dit laatste woord: semantiek.

Semantiek is een ander woord voor betekenisleer. Het begrijpen van de betekenis van concepten en objecten is voor mensen vanzelfsprekend, maar voor computers is dit lastig. Bedenk dat we een computer zouden vragen om meer informatie over Mars, waarbij we de chocoladereep bedoelen. Voor een computer is dit niets anders dan een combinatie van vier letters. De computer zou in een systeem of verzameling documenten op zoek gaan naar deze lettercombinatie en de resultaten ervan tonen.

Maar wat als een computer de lettercombinatie zou kunnen herkennen als concept? In het geval van Mars zou je de onderstaande "intelligente" vraag kunnen terug krijgen.

Bedoelde u:



Mars, de planeet?



Mars, de romeinse god?



Mars, de snack?



Of Bruno Mars, de artiest?

Een tweede stap zou zijn het benoemen en herkennen van relaties tussen concepten, zodat één concept kan worden geassocieerd met een ander concept. In het geval van Mars, de snack, zouden we kunnen stellen dat Mars gemaakt wordt door het Amerikaanse bedrijf Mars Inc. Mars Inc maakt ook andere snacks zoals M&M's en Snickers, maar ook andersoortige producten zoals rijst van het merk Uncle Ben's. Rijst wordt onder meer verbouwd in Azië en China ligt in Azië en ga zo maar door. Voor mensen is het maken van dit soort associaties natuurlijk. Computers kunnen deze associaties niet maken op basis van platte tekst die in data aanwezig is. Door de data echter te verrijken met betekenis en relaties wordt dit gemakkelijker. Linked Data zorgt hierbij voor een gemeenschappelijke taal.

In een notendop betekent het concept van Linked data dat alle Open Data volgens een zelfde datamodel worden aangeboden. Concepten/locaties/personen/etc die in de data voorkomen verwijzen naar een online resource die een beschrijving geeft van het concept/de locatie/de persoon/etc. Zo is het duidelijk welke Mars uit het voorbeeld bedoeld wordt. Er is een gemeenschappelijke taal voor het beschrijven van relaties, zodat relaties zoals "bevindt zich in" of "wordt gemaakt door" altijd dezelfde betekenis hebben.

Het web van data dat is ontstaan maakt het mogelijk om ruwe data vanuit diverse datasets, en verschillende bronhouders en domeinen, gecombineerd te gebruiken. Data wordt op deze manier over het web met elkaar gelinkt, vanwaar de naam Linked Data.

Linked Data kent vier basisprincipes:

1. Relaties tussen verschillende data worden gemaakt op basis van verwijzingen naar dingen, personen of concepten. Een relatie kan bijvoorbeeld zijn: "Nederland heeft als hoofdstad Amsterdam". Alle onderwerpen die in de data voorkomen, maar ook de beschrijving van de relatie krijgen een **eigen unieke URI**. URI staat voor Uniform Resource Identifier en is een unieke identificatiecode voor een stukje informatie.
2. De URI is **online benaderbaar** via het HTTP protocol en bevat naast de naam een korte beschrijving. Het HTTP protocol is het universele toegangsmechanisme voor het internet. In het eerdere voorbeeld krijgen Nederland, Amsterdam, maar ook "heeft als hoofdstad" allen een eigen URI die kan worden bevraagd. Een fictief voorbeeld:

<http://Linkeddatanl/Nederland>

Omschrijving: Land in West-Europa dat grotendeels beneden zeeniveau ligt

<http://Linkeddatanl/Amsterdam>

Omschrijving: Nederlandse stad

<http://Linkeddatanl/heeftals hoofdstad>

Omschrijving: Geeft de hoofdstad van een land weer

3. Het datamodel is voor alle data hetzelfde. Net als dat HTML dit is voor webpagina's op internet, is het **Resource Description Framework (RDF)** dit voor data. RDF is een simplistisch netwerk datamodel dat is gemaakt voor de uitwisseling van data op het web.
4. Data wordt, net als met hyperlinks op websites, onderling verbonden. Hierbij wordt de **relatie** tussen data beschreven. In RDF is de beschrijving van een relatie vastgelegd in een zogenaamde **triple**:

Onderwerp	Bewering	Object
-----------	----------	--------

In het Engels is dit:

Subject	Predicate	Object
---------	-----------	--------

Een triple maakt feitelijk een simpele zin:

Subject	Predicate	Object
Nederland	Heeft als hoofdstad / has capital	Amsterdam

Een voorbeeld van deze statement in RDF ziet er als volgt uit:

```
<rdf:Description rdf:about ="http://Linkeddatanl/Nederland">
  <hasCapital
<rdf:Description rdf:about ="http://Linkeddatanl/Amsterdam"/>
  <hasCapital>
</rdf:Description>
```



Figuur 4.2 Visualisatie triple

4.2 Van Open Data naar Linked Open Data in vijf sterren

Berners Lee bedacht voor Open Data een puntensysteem ook wel bekend als “De vijf sterren van Open Data”. Dit puntensysteem beschrijft de mate van openheid van data op basis van herbruikbaarheid.

In onderstaande tabel¹² is deze 5-sterren rating weergegeven.

5 sterren rating	Omschrijving
★	raadpleegbaar op het Web (in welk formaat dan ook) onder een open licentie. Dit kan een afbeelding zijn, of een pdf-bestand.
★★	raadpleegbaar als gestructureerde data (zoals Excel in plaats van een plaatje van een tabel)
★★★	bruikbaar in vorm van niet-merkgebonden formats (zoals CSV in plaats van Excel)
★★★★	bruikbaar via URI's om dingen te identificeren, zodat mensen naar de data kunnen verwijzen
★★★★★	gelinkt aan andere data om context te verstrekken

De eerste 3 scores beschrijven “reguliere” Open Data, waarbij de data steeds universeler leesbaar wordt voor toepassingen. De 4^e en de 5^e ster beschrijven het toepassen van de Linked Data principes met het gebruik van URI's en het beschrijven van relaties. Vijf sterren representeert hierbij dus volwaardige Linked Open Data.

4.3 Waarom Linked Data?

In de jaren negentig is gestart met het vullen van het internet met documenten, foto's en video's. De wijze waarop wij nu informatie met elkaar kunnen delen heeft de maatschappij voorgoed veranderd. Door ruwe data niet alleen open te stellen, maar deze ook onderling met elkaar te verbinden ontstaat wederom een enorm potentieel. Voor medische of andere onderzoeken zal dit een grote meerwaarde betekenen, onderzoekers kunnen immers zonder drempels data vanuit andere vakgebieden gebruiken. Ook toepassingen voor consumenten en organisaties zullen steeds behulpzamer en nuttiger worden, doordat data van verschillende aanbieders snel en relevant kan worden gecombineerd.



Figuur 4.3 Toepassingen bloeien op met Linked Data als katalysator (bron: www.w3.org)

¹² Bron: <http://www.w3.org/DesignIssues/LinkedData.html>

4.4 De techniek

4.4.1 RDF dataformaten

Zoals eerder uitgelegd wordt alle Linked Data volgens hetzelfde RDF-datamodel beschreven. Een aantal data-formaten die volgens het RDF-model zijn opgebouwd zijn Turtle (.ttl), RDF/XML, Notation3 (N3) en JSON-LD. De formaten verschillen van elkaar in oorsprong en complexiteit, maar kunnen door elkaar worden gebruikt, doordat de regels van hetzelfde (RDF) model wordt gehanteerd.

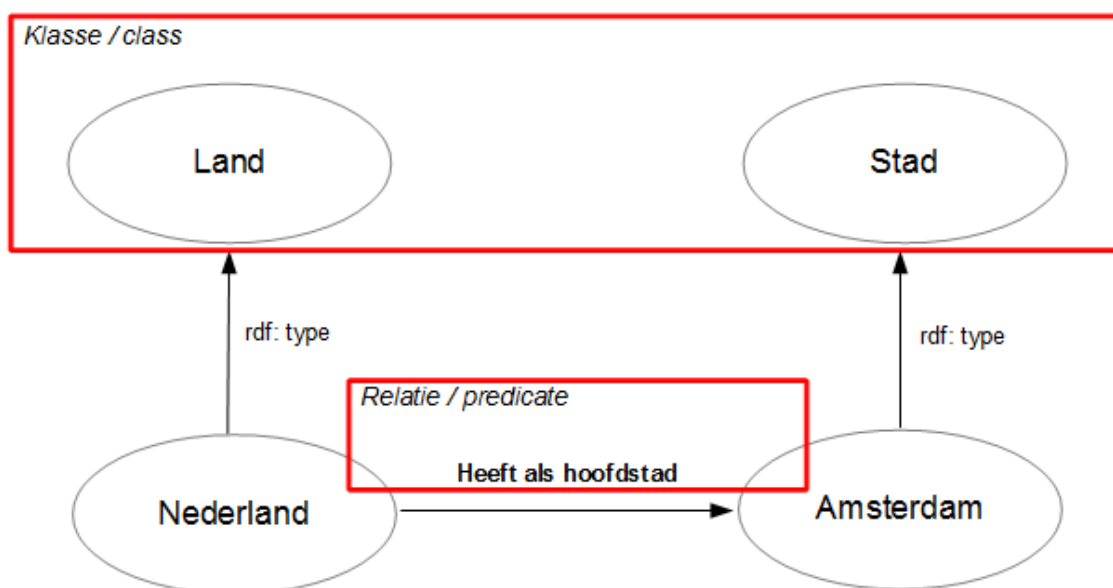
Ruwe data vanuit verschillende formaten kan worden omgezet in RDF. Hiervoor zijn verschillende tools beschikbaar. Het aanbieden van de data in RDF is hiermee een aanvulling op, en niet persé een vervanging van de huidige formaten. Binnen de verschillende vakgebieden, zoals geoinformatie, kunnen voor uitwisseling de bestaande formaten worden blijven gebruikt.

4.4.2 Vocabulaires

De aard van een concept (ook wel de klasse genoemd, bijvoorbeeld een land of stad) en de beschreven relatie (predicate, bijvoorbeeld "heeft als hoofdstad") worden vastgelegd in online toegankelijke vocabulaires, waar binnen RDF naar verwezen wordt. Dit zorgt ervoor dat bij het beschrijven van de data door iedereen dezelfde taal wordt gehanteerd. In figuur 4.4 is het verschil tussen beiden schematisch weergegeven.

Vocabulaires worden soms ook wel ontologieën genoemd. Een ontologie betekent in deze context het beschrijven van informatie op het Internet en de relaties tussen de gegevens. De scheidingslijn tussen een vocabulaire en een ontologie is binnen de Linked Data gemeenschap in de praktijk wat vaag. In het algemeen wordt gesteld dat een ontologie vanuit een gerenomeerde organisatie afkomstig is (zoals bijvoorbeeld het W3C) en vocabulaires in het veld ontstaan. Voor de duidelijkheid zullen we in dit rapport vanaf nu slechts spreken over vocabulaires.

Kort gezegd bepaalt RDF de manier **waarop** iets beschreven wordt en de vocabulaire **wat** er beschreven wordt.



Figuur 4.4 Klassen vs relaties

Er zijn op het moment circa 350 verschillende vocabulaires beschikbaar¹³ vanuit verschillende domeinen. Bij het tot stand brengen van nieuwe Linked Data is de kans redelijk aanwezig dat er een geschikte vocabulaire aanwezig is. Voorbeelden van veelgebruikte vocabulaires zijn OWL (Web Ontology Language), SKOS (Simple Knowledge Organization System), FOAF (Friend-of-a-friend) en RDFS (RDF Schema).

Een aantal voorbeelden van beschrijvende relaties (of predicates) vanuit deze vocabulaires zijn:

- owl:sameAs = is hetzelfde als..
- skos:narrower = Is een specialisatie van..
- foaf:homepage = verwijzing naar een homepage van een persoon of organisatie
- foaf:knows = kent de persoon..
- foaf:made = geproduceerd door..
- rdfs:seeAlso = heeft te maken met..

4.4.3 Triplestores en SPARQL

Triple Stores zijn opslagsystemen, zeg databases, voor (meta)data die in RDF zijn uitgedrukt. Om de RDF data te kunnen bevragen en analyseren is een taal nodig. Hiervoor is de querytaal SPARQL ontwikkeld. SPARQL is voor RDF wat SQL is voor relationale databases. Specifiek voor het gebruiken van geografische relaties (bijv ligt in, binnen een buffer van x m, etc) is vanuit het Open Geospatial Consortium (OGC) GEOSPARQL ontwikkeld.

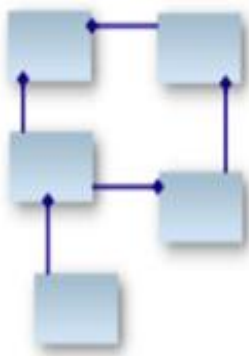
Een voorwaarde voor Linked Data is dat deze online beschikbaar is. RDF-bestanden in triplestores kunnen online worden gebracht via een zogenaamde SPARQL-endpoint. Dit is een webservice die wordt aangeboden door een organisatie of persoon, zodat de Linked Data via het web kan worden bevraagd en gebruikt. Er is verschillende software die de opslag in triplestores en/of de ontsluiting via een SPARQL-endpoint kunnen verzorgen. Bekende pakketten zijn:

- Allograph (triplestore)
- Virtuoso Universal Server (triplestore en SPARQL-endpoint)
- 4store (triplestore)
- SPARQLer (SPARQL-endpoint)

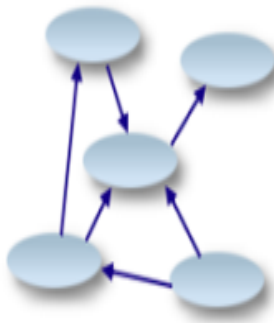
¹³ Bron: linked open vocabularies (<http://lov.okfn.org/dataset/lov/>)

4.5 Netwerk vs relationeel datamodel

Linked Data maakt gebruik van een zogenaamd netwerk datamodel. In de volgende paragraaf wordt het verschil uitgelegd met het conventionele relationeel datamodel;

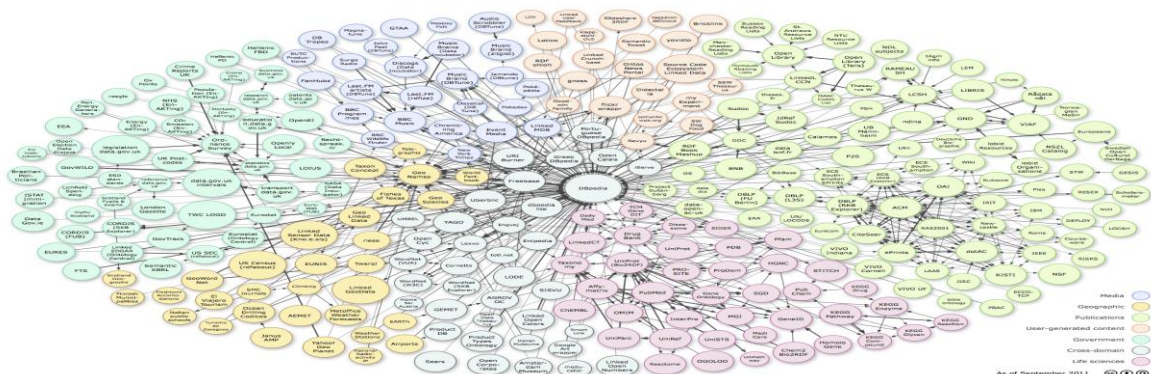


De meeste databases binnen bedrijfsomgevingen zijn ingericht volgens een **relationeel datamodel**. De gegevens worden hierbij opgeslagen in tabellen waarin de rijen de soortgelijke groepen informatie, de records vormen, en de kolommen de informatie die voor elk record moet worden opgeslagen. Verschillende tabellen kunnen met elkaar worden verbonden door een kolom toe te voegen waarin een verwijzing naar een record in een andere tabel (de zogenaamde primary key) wordt opgenomen.¹⁴ Voordelen van een relationele database zijn dat gegevens maar één keer en op één plek zijn opgeslagen (weinig redundantie) en dat het mogelijk is om zeer snel gegevens te zoeken en te presenteren. Een nadeel is dat de inrichting weinig flexibel is bij veranderingen en dat de data lastig uitwisselbaar is met andere databases met een andere indeling.



Linked Data maakt zoals gezegd gebruik van een **netwerk datamodel**. Binnen een netwerkmodel worden objecten met elkaar verbonden door middel van het beschrijven van hun onderlinge relatie. Het voordeel van een netwerk datamodel is dat nieuwe objecten, maar ook eigenschappen, kunnen worden toegevoegd zonder dat de structuur van het bestaande model hoeft te worden aangepast. Een nadeel kan zijn dat de structuur onoverzichtelijk en lastig beheersbaar wordt. Er ontstaat echter een datamodel dat flexibel is en waaraan eenvoudig iets kan worden toegevoegd of veranderd.¹⁵ Voor een web van data is deze flexibiliteit van essentieel belang.

In 2007 was Wikipedia één van de eerste databases die volgens het netwerk datamodel van Linked Data online werd gebracht. De Linked Data variant van Wikipedia heeft de naam DBpedia. Alle “dingen” op Wikipedia hebben dan een eigen URI waarnaar gelinkt kan worden. Het web van data (zie figuur 4.5) groeide sindsdien snel met DBpedia als centrum. In 2007 waren er slechts 12 databases met elkaar gelinkt, eind 2011 waren er 295. Deze 295 databases waren toen met elkaar verbonden met maar liefst 31 miljard triples.



Figuur 4.5 Visualisatie van het web van data eind 2011 (bron: www.w3.org)

¹⁴ Bron: Wikipedia

¹⁵ Bron: Gegevens delen via Linked Open Data (Alterra-rapport 2306, 2012)

4.6 Linked geodata

Binnen het geo-domein heeft de afgelopen jaren al een redelijke standaardisatie plaatsgevonden. De meest bekende data- en uitwisselformaten zijn:

- **ESRI shapefile:** is een formaat voor het opslaan van geografische informatie, ontwikkeld en beheerd door ESRI. Binnen een shapefile kunnen punten, lijnen en vlakken worden opgeslagen. Voor ieder object kunnen attributen worden opgeslagen, maar ook informatie over het coördinaten- en projectiestelsel. Hoewel het geen volledig open formaat is kunnen shapefiles worden ingelezen en gebruikt in diverse GIS-applicaties.
- **Web Map Service (WMS):** is een protocol voor het serveren van geogerefereerde kaartbeelden over het internet. Deze kaartbeelden bieden een representatie van vectordata aan en de onderliggende data kan worden bevraagd, maar niet gedownload. WMS kan dus worden beschouwd als een view-service. Een WMS service wordt gegenereerd door een webserver vanuit een GIS database. Het WMS protocol is ontwikkeld en wordt beheerd door het Open Geospatial Consortium (OGC).
- **Web Feature Service (WFS):** is een zusterprotocol van WMS in die zin dat deze gebruik maakt van dezelfde infrastructuur als WMS en eveneens vanuit het OGC is ontwikkeld en wordt beheerd. Echter waar met een WMS de data alleen kan worden bekeken kan via een WFS service de onderliggende data worden gedownload of gewijzigd (dit laatste als de rechten dit toelaten).
- **Web Coverage Service (WCS):** is een 3^e lid uit de familie van serviceprotocollen vanuit het OGC specifiek voor het uitwisselen van geografische rasterdata.
- **Geography Markup Language (GML):** is een op XML gebaseerde uitwisselstandaard waarmee verschillende soorten geoinformatie zoals geometrie, coverages of sensordata kunnen worden opgeslagen en uitgewisseld. GML is eveneens ontwikkeld en wordt beheerd door het OGC.
- **Keyhole Markup Language (KML):** is een variatie op GML specifiek gericht op het visualiseren van geografische data in Google Earth en Maps. KML's worden vaak gedistribueerd als **KMZ**. Een KMZ comprimeert de informatie in de KML, waardoor de omvang van het bestand wordt verkleind, maar ook kunnen er binnen een KMZ mediabestanden (zoals foto's) worden toegevoegd waar binnen de KML naar verwezen wordt.

Door de standaardisatie van geodata in de afgelopen jaren is het toepassen van de Linked Data principes, in vergelijking met andere domeinen, een relatief kleine stap. Vooral GML lijkt qua structuur erg op RDF. Hoewel GML vooral gericht is op het gebruik binnen GIS-applicaties biedt het ook mogelijkheden om links naar andere data op te nemen. Technisch gezien lijkt de impact niet heel groot.

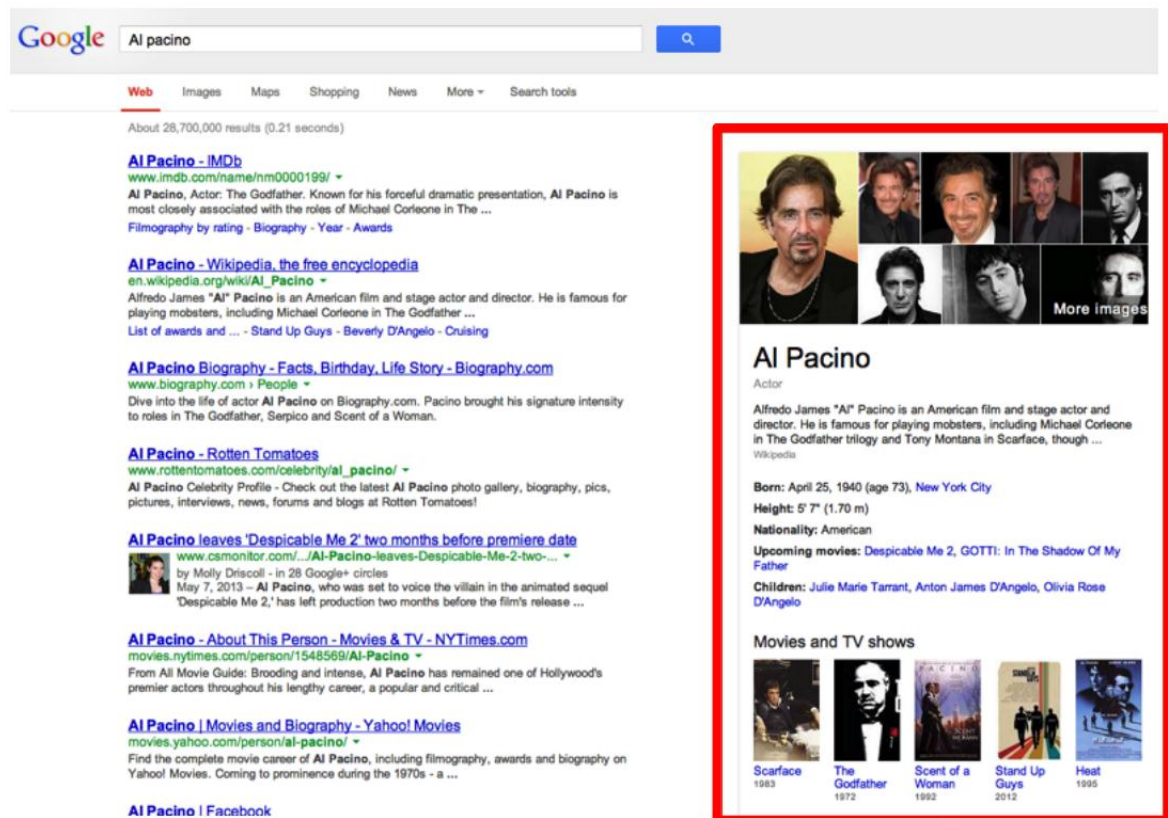
RDF is dan ook geen vervanger voor de huidige geo-standaarden, maar eerder een aanvulling. Binnen het geo-domein zal er dan ook niet direct een verdere verbetering optreden in de uitwisseling van informatie. Interessanter is om te kijken naar de uitwisseling en combinatie met niet-geodata. Doordat geodata en niet-geodata via een zelfde datamodel beschikbaar zijn en er onderlinge relaties worden gelegd zal de grens tussen ruimtelijke en reguliere informatie langzaam vervagen. Dit zal een enorme verrijking betekenen binnen de verschillende informatiewerkvelden, maar ook voor consumenten en onderling tussen aanbieders van data.

4.7 Voorbeeldtoepassingen Linked Data

4.7.1 Google Knowledge Graph; het nieuwe zoeken

Google loopt vaak voorop als het gaat om nieuwe technologie. Zeker als het gaat om het verbeteren van zoekresultaten. Op www.google.com (nog niet beschikbaar in de Nederlandse versie) wordt Linked Data inmiddels toegepast bij het tonen van relevante zoekresultaten. Google noemt deze toepassing van Linked Data de Google Knowledge Graph.

In figuur 4.6 is hiervan een voorbeeld te zien. Door als zoekterm Al Pacino in het voeren wordt, zoals we gewend zijn, aan de linkerzijde een lijst getoond van webpagina's waarin het woord "Al Pacino" voorkomt. Aan de rechterzijde is nu extra informatie te zien over de persoon Al Pacino, maar ook informatie over zijn nieuwe films, zijn familieleden en zijn geboorteplaats. Google hoeft deze informatie niet op te slaan in een eigen database, maar kan deze ophalen vanaf diverse locaties op het web waar deze als Linked Data beschikbaar is en een beschreven relatie heeft met de persoon "Al Pacino".

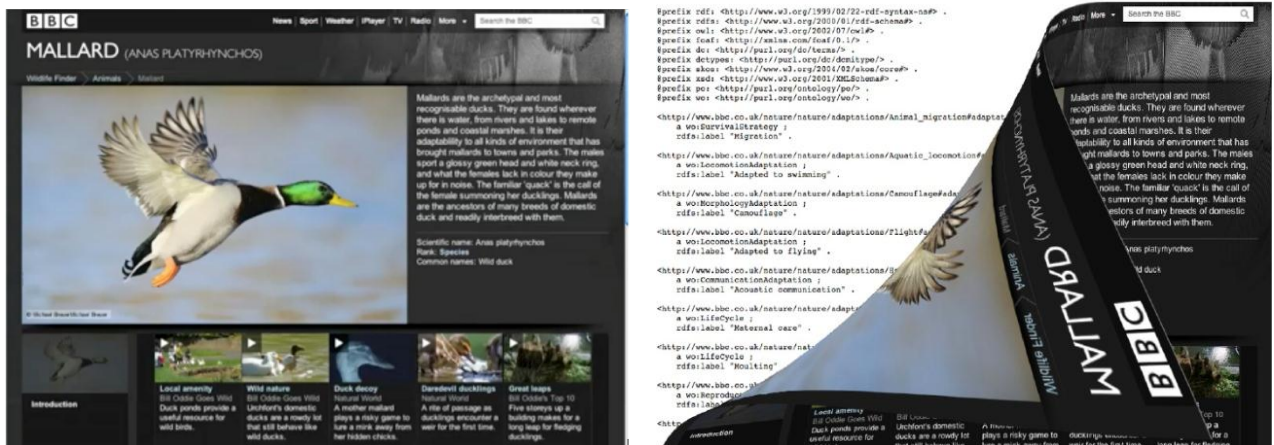


Figuur 4.6 Screenshot toepassing Google Knowledge Graph

4.7.2 BBC wild life; het web als CMS

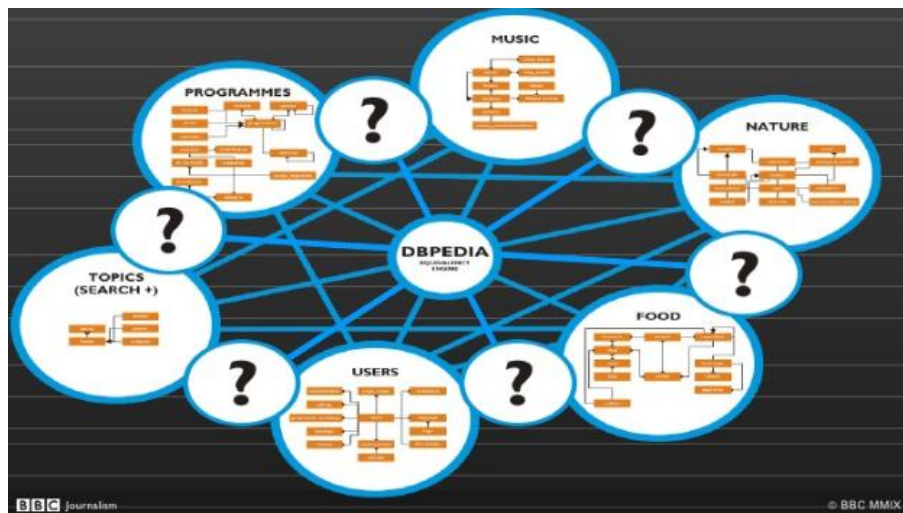
<http://www.bbc.co.uk/nature/life>

Ook de website van de BBC maakt onder de motorkap dankbaar gebruik van de Linked Data principes. Voor het vullen van de Wildlife Finder gebruikt de BBC Linked Data van onder meer Wikipedia, het Wereld Natuurfonds, het Animal Diversity Web en The IUCN's Red List of Threatened Species.



Figuur 4.7 Linked Data op Wild Life finder BBC (bron: BBC)

De BBC gebruikt Linked Data inmiddels ook op hun website bij informatie over programma's, muziek en lifestyle. DBpedia is een veel gelinkte bron (zie figuur 4.8), maar ook online databases zoals Music Brainz worden gebruikt om informatie op te halen. Een groot deel van de content van de BBC website wordt dus niet meer in een eigen content management systeem (CMS) beheerd, maar wordt verkregen vanaf het web.



Figuur 4.8 Visualisatie datalinks BBC (Bron: BBC)

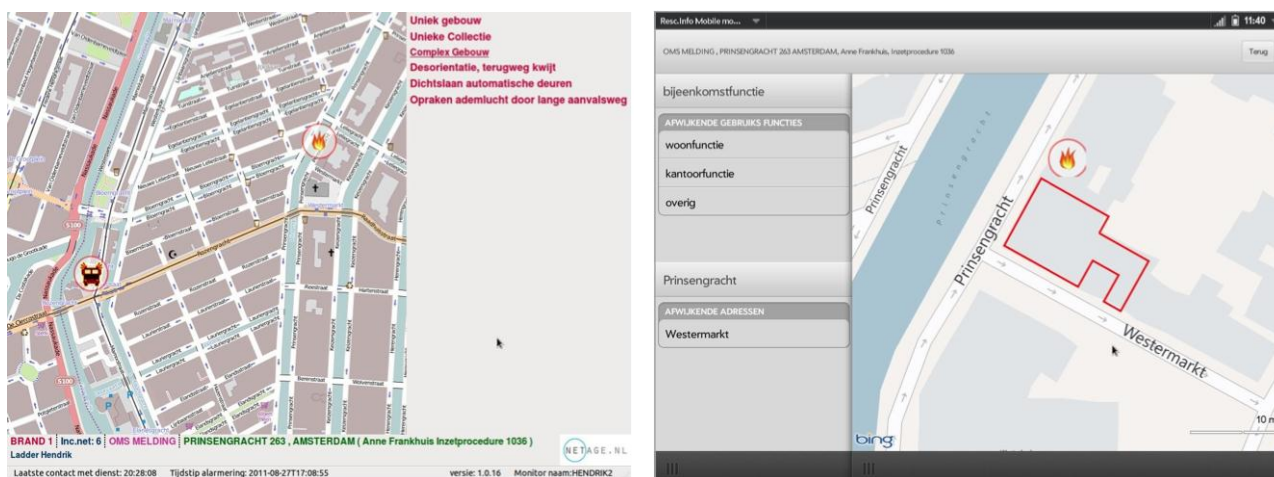
4.7.3 Brandweer Amsterdam-Amstelland; snel de juiste informatie

Iets dichterbij huis is de toepassing van de Brandweer Amsterdam-Amstelland¹⁶. Hier wordt Linked Data gebruikt om bij het uitrukken naar een brand snel relevante informatie te tonen. In een bluswagen zijn verschillende informatiesystemen aanwezig variërend van een portofoon, een tabletpc tot analoge ordners. Bij een melding moet de informatie snel worden geïnterpreteerd, zodat de juiste keuzes kunnen worden gemaakt voor de aanpak.



Figuur 4.9 Diversiteit aan systemen aanwezig in een bluswagen (bron: www.netage.nl)

In de Linked Data-toepassing bij Amsterdam-Amstelland wordt Linked Data uit onder meer het eigen protocollensysteem, van het historisch museum en de Basisregistratie Adressen en Gebouwen (BAG) met elkaar gecombineerd in één applicatie. In de voorbeeldschermen zoals te zien in figuur 4.9 is een melding voor het Anne Frankhuis te zien. Hierbij worden de juiste protocollen getoond naar aanleiding van de monumentenstatus zoals uit de Linked museumdata blijkt en de meerdere gebruikersfuncties van het gebouw vanuit de Linked BAG-data.



Figuur 4.10 Voorbeeldschermen Linked data applicatie bluswagen (bron: www.netage.nl)

¹⁶ Bron: B.van Leeuwen, Netage.nl

4.8 Nederlandse Pilot Linked Open Data

In Nederland waren er tot voor kort, naast de toepassing bij de Brandweer Amsterdam-Amstelland uit de vorige paragraaf, niet veel praktijkvoorbeelden te vinden op het gebied van Linked Open Data. In september 2012 is Geonovum gestart met het project “Linked Open Data Pilot”.

De ambitie van deze pilot is om op basis van praktijkervaring inzicht te krijgen in de mogelijkheden en wensen op het gebied van Linked Open Data. De verkregen inzichten zouden kunnen leiden tot algemene afspraken en afgesproken standaarden voor ontsluiting van open data op het web. Na de uitvoering van de LOD pilot heeft Nederland een beter inzicht in de LOD mogelijkheden en kan het gericht aan de slag.

De in het plan van aanpak geformuleerde doelstellingen van dit project zijn de volgende:

- Samenwerken met veel stakeholders voor de grondgebieden van de gemeente Nijmegen en Amersfoort waar toepassingen worden gerealiseerd met als doel om vast te kunnen stellen of LOD een techniek is om overheidsdata beschikbaar te stellen op het web
- Naast LOD de data in gangbare andere web-formaten beschikbaar te stellen om te komen tot een goed vergelijk
- Vaststellen van open data standaard(en) en referentiekader
- Het leggen van links tussen de ‘formele’ overheidsregistraties en de ‘informele’ registraties op het web
- Samenbrengen van kennis, technologie en behoefte
- Awareness en communicatie
- Vastlegging van de ervaringen en aanbevelingen als een open data referentiekader

De uitvoering van de LOD pilot is opgedeeld in drie deelsporen:

1. Aanbod
2. Toepassingen
3. Techniek & standaarden

Voor dit afstudeeronderzoek is aansluiting gezocht bij de LOD pilot, specifiek de actielijn Techniek & standaarden, om het concept van Linked Data en de conversie van reguliere data naar Linked Data te onderzoeken.

De opleverdatum van de Linked Open Data Pilot is augustus 2013.

5 CASESTUDIE LINKED WATERSCHAPSGEODATA

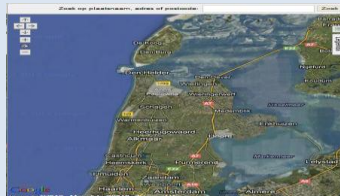
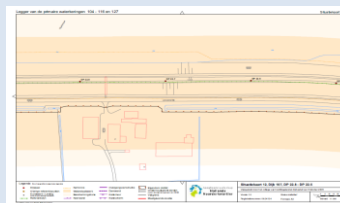
In de nu volgende casestudie zal binnen waterschap HHNK onderzoek worden gedaan naar de openheid van de huidige geodata, mogelijke toepassingen voor Linked waterschapsdata en een werkwijze om waterschapsdata als Linked Data beschikbaar te stellen.

5.1 Open geodata bij HHNK

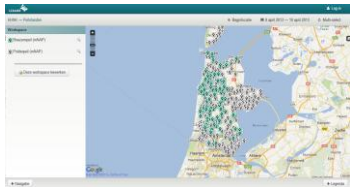
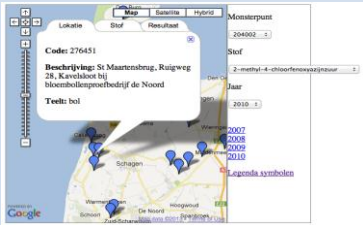
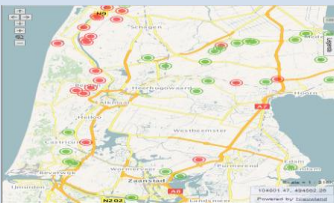
HHNK heeft nog geen duidelijke strategie voor open (geo)data. Data wordt enerzijds op de website gepubliceerd als dit wettelijk verplicht is, bijvoorbeeld bij het publiceren van een Legger¹⁷. Anderzijds via losse initiatieven.

Het eerste wat op de website (www.hhnk.nl) opvalt is dat het onduidelijk is of en waar er data door het waterschap beschikbaar wordt gesteld. De zoekfunctie of de link naar de digitale balie biedt hier geen duidelijkheid over. Naast een mogelijkheid tot een informatieverzoek in het kader van de Wet Openbaar Bestuur (WOB) is voor het opvragen van gegevens geen procedure te vinden.

Toch blijkt er op diverse locaties op de website geodata te worden aangeboden in verschillende vormen. In onderstaande tabel staan deze op een rij gezet en wordt ook de sterrenrating van Berners Lee op de data toegepast.

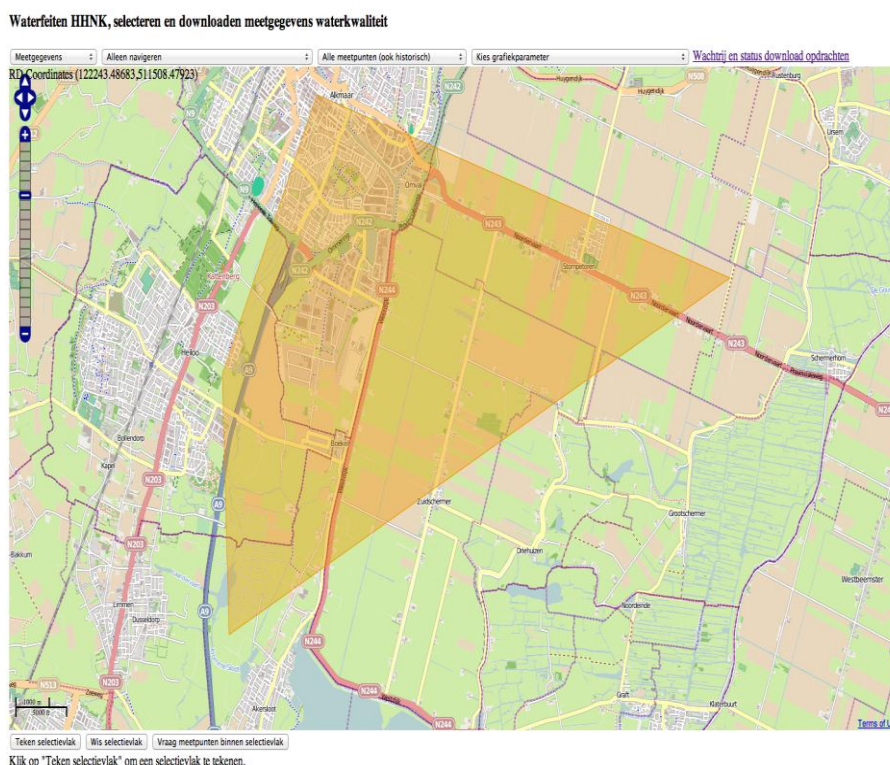
Onderwerp	Vorm	Representatie	Linked Open Data Sterren (1-5)
Legger wateren	Webviewer (WebGIS Publisher)		★ (Raadpleegbaar op het web, in welk formaat dan ook)
Legger zandige Kust	Webviewer (WebGIS Publisher)		★ (Raadpleegbaar op het web, in welk formaat dan ook)
Legger primaire waterkeringen	Kaarten als Pdf bestanden		★ (Raadpleegbaar op het web, in welk formaat dan ook)

¹⁷ Leggers zijn openbare registers waarin wordt aangegeven aan welke eisen de waterstaatswerken moeten voldoen voor wat betreft ligging, vorm, afmeting en constructie.

Onderwerp	Vorm	Representatie	Linked Open Data Sterren (1-5)
Wegen in beheer van HHNK	Kaart als Pdf bestand		★ (Raadpleegbaar op het web , in welk formaat dan ook)
Waterstanden	Webviewer (Lizard)		★ (Raadpleegbaar op het web , in welk formaat dan ook)
Metingen Waterkwaliteit	Webviewer (Open Layers) en verwijzing naar aparte website met downloadlinks voor ruwe data		★★★★ (Bruikbaar in vorm van niet-merkgebonden formats zoals CSV in plaats van Excel)
Werk bij u in de buurt	HTML pagina met informatie over projecten per gemeente		★ (Raadpleegbaar op het web , in welk formaat dan ook)
Waterbergingsprojecten	Webviewer (WebGIS Publisher)		★ (Raadpleegbaar op het web , in welk formaat dan ook)
Objecten en gebouwen	HTML pagina met informatie over verschillende objecten (bijv historische gebouwen, gemalen, molens, etc)		★ (Raadpleegbaar op het web , in welk formaat dan ook)

Wat allereerst opvalt is dat de informatie nogal verstopt zit binnen de website en op verschillende webpagina's moet worden gevonden. Daarnaast is de vorm van de aangeboden kaartinformatie divers, variërend van een pdf-kaart tot verschillende webviewers en een webpagina met informatie over HHNK-objecten en gebouwen. De onderliggende data zijn niet beschikbaar.

Het blijkt echter dat er vanuit het Onderzoekscuster van HHNK een aparte website (<http://hnk-water.nl>) beschikbaar is met gegevens over waterkwaliteit en –kwantiteit, maar ook neerslaggegevens en flora informatie. Deze data is in aantal gevallen downloadbaar in .txt, CSV- en KML-formaat, maar ook worden er WMS en WFS-services aangeboden. Hoewel het vinden van de website wat moeite kost en de website gericht lijkt op specialisten is hier wel degelijk sprake van 3-sterren Open Data.



Figuur 5.1 Screenshot selectie en download meetgegevens waterkwaliteit (<http://hnk-water.nl>)

Ondanks de Open Data op de onderzoekswebsite is de overall score van HHNK op de sterrenrating van Berners Lee slechts 1 ster.

Overall score HHNK op het gebied van Open Data



Raadpleegbaar op het Web (in welk formaat dan ook) onder een open licentie

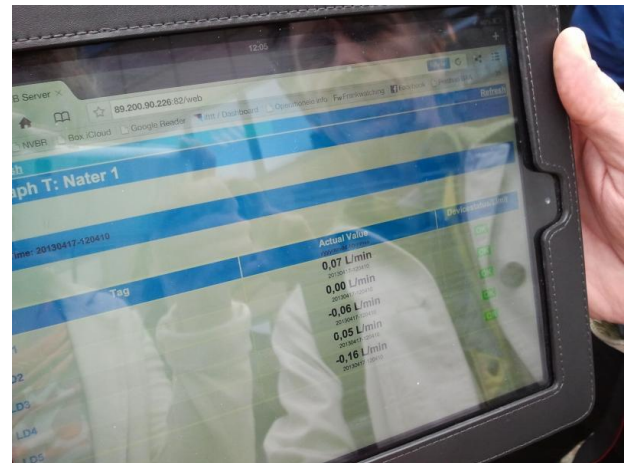
5.2 Toepassingsmogelijkheden Linked waterschapsdata

Er zijn verschillende toepassingen te bedenken, waarin de principes van Linked waterschapsdata een meerwaarde zullen betekenen. In deze paragraaf volgen twee praktische en actuele voorbeelden.

Linked Data Autospuut

Als bluswater uit oppervlaktewater getrokken wordt dan zal het waterschap door bediening van gemalen, stuwen en/of noodpompen de watergangen zo goed als mogelijk op peil proberen te houden mits het waterschap hierover tijdig is geïnformeerd. Daarnaast kan het voorkomen dat het bluswater moet worden afgevoerd om te worden gereinigd. In beide gevallen is het interessant voor het waterschap om te weten hoeveel water er opgenomen wordt.

Binnen het brandweercorps Amsterdam-Amstelland is gestart met een proef waarbij een bluswagen is uitgerust met een flow meter waarmee de hoeveelheid opgenomen water kan worden gemonitord. Deze informatie is daarbij als Linked Data beschikbaar. Door de data van de bluswagen en data van het waterpeil met elkaar te combineren kan een snelle analyse gedaan worden op welke acties er nodig zijn. Hiermee kan bijvoorbeeld onvoldoende watertoevoer of grondwaterverlaging worden voorkomen.



Figuur 5.2 Demonstratie Linked Data Autospuut (bron: Geonovum)

Huiskluis

De Huiskluis is één van de pilottoepassingen binnen de Linked Open Data pilot van Geonovum. Het is een internetomgeving waar bewoners/eigenaren een informatiedossier over hun eigen woning of bedrijfsgebouw kunnen aanleggen. Het dossier kan informatie van bevatten die geautomatiseerd wordt opgehaald vanuit verschillende organisaties. Hierbij valt te denken aan vergunningsaanvragen op het adres, het bestemmingsplan of de vuilnisophaal tijden. Ook kan de bewoner/eigenaar eigen informatie toevoegen zoals het energieverbruik of facturen/foto's van verbouwingen. De Huiskluis wordt hiermee een verzamelpunt van informatie over een huis en helpt met het verzamelen en delen ervan. Voor het geautomatiseerd verzamelen van gegevens maakt de Huiskluis gebruik van de principes van Linked Data.

In de context van waterschapsdata zou het interessant zijn voor een (toekomstig) huiseigenaar om te weten welke beperkingen er vanuit het waterschap zijn op het perceel. Naast het meer bekende bestemmingsplan van een gemeente heeft ook het waterschap regelgeving over het gebruik van grond vastgelegd in de eerder genoemde leggers. Een leggerzonering voor een waterkering geeft een gebied aan rondom een waterkering waarbinnen geen (graaf) werkzaamheden mogen plaatsvinden, omdat dit de stabiliteit van de waterkering kan beïnvloeden.

In figuur 5.3 is een situatie te zien waar er zich een aantal gebouwen bevinden binnen deze leggerzones.



Figuur 5.3 Situatieschets gebouwen binnen leggerzonerings

De (toekomstige) eigenaren van deze gebouwen moeten deze informatie op dit moment ophalen bij het waterschap, terwijl het bestaan van dit soort zones niet bij iedereen bekend is. In de Huiskluis verschijnt deze informatie, middels het Linked Data principe, geautomatiseerd tussen andere (niet-geo)informatie over de woning of het perceel.

5.3 Proof-of-concept Linked waterschapsgeodata

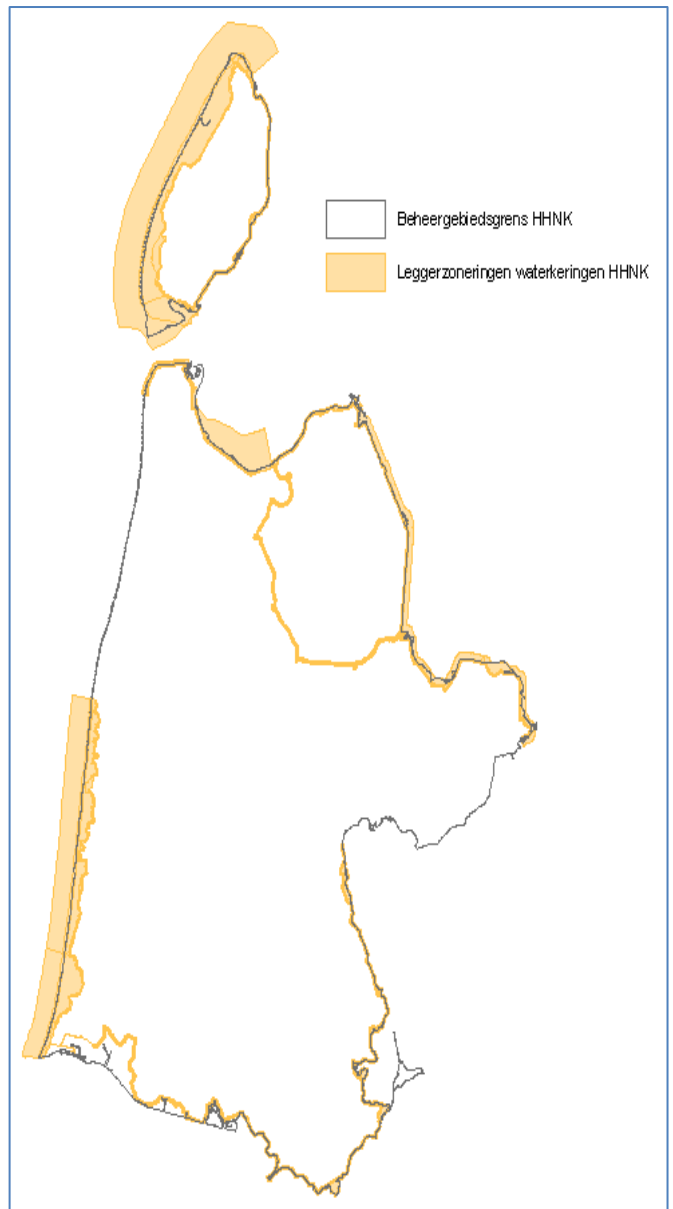
In dit hoofdstuk zullen de eerder genoemde leggerzoneringen, specifiek van waterschap HHNK, stap voor stap worden omgezet in Linked Data. Het doel hiervan is de data van de Leggerzoneringen online te publiceren in RDF-formaat.

5.3.1 De data

De dataset met leggerzoneringen is vastgelegd in de HHNK GIS-database onder de naam GW_GZN. De dataset GW_GZN bestaat uit 231 records (polygonen) en ziet er als volgt uit:

Attributen	
GZN_ID:	Automatische codering
GZNIDENT:	Identificatie zonering
GZNOMSCH:	Tekstuele omschrijving zonesoort (= niet gevuld)
GZNSOORT:	Soort zone

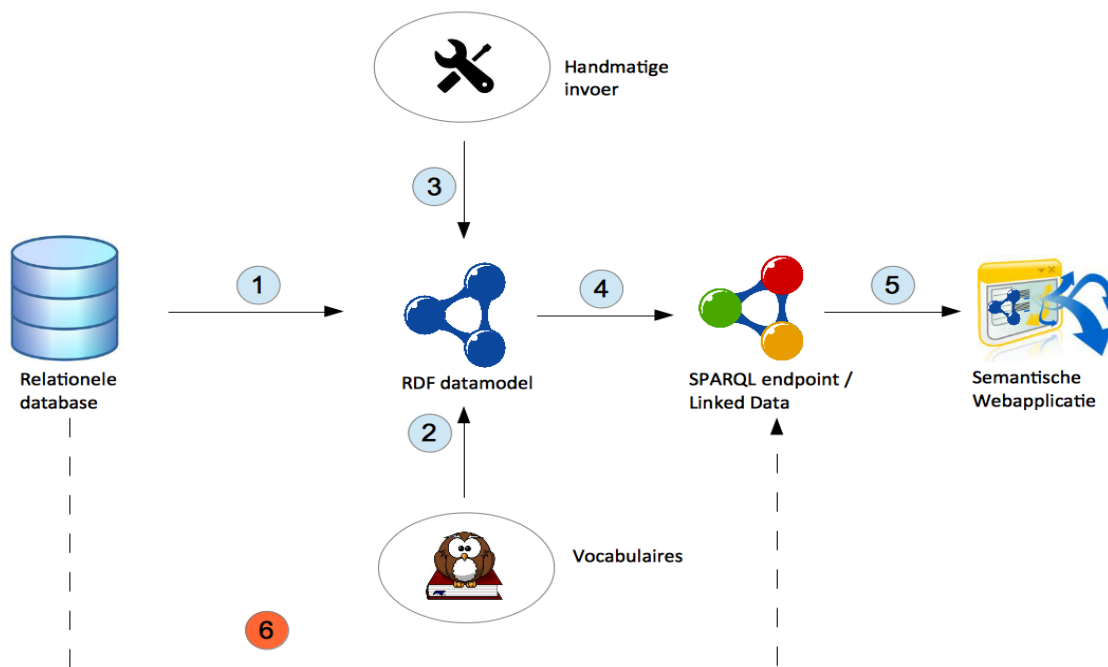
Metadata	
Omschrijving:	Een zone gelegen langs/op een primaire waterkering met een juridische status die uit de keur keur voortvloeit.
Bronhouder:	Hoogheemraadschap Hollands Noorderkwartier
Coördinatenstelsel:	Rijksdriehoekstelsel
Toepassingschaal:	1:1000
Inwinningswijze:	Gevolg van berekeningen
Gebruiksbeperking:	Open Data Commons Public Domain Dedication and License (PDDL)



Figuur 5.4 Dataspecificaties GW_GZN (Leggerzoneringen HHNK)

5.3.2 Omzetten naar Linked Data

In onderstaand figuur is het proces van het omzetten van de leggerzoneringen naar Linked Data geschematiseerd. Er zijn een aantal manieren om tot Linked Data te komen, dit stappenplan beschrijft er één.



Figuur 5.5 Schematisatie conversie naar Linked data

De stappen in het schema zijn:

1. Data verzamelen
2. Inventariseren toe te passen vocabulaires
3. Handmatige invoer
4. Valideren RDF
5. Publiceren Linked Data
6. Optioneel: rechtstreekse mapping vanuit Oracle database

Voor het uitvoeren van deze stappen zijn de volgende zaken nodig:

- (GIS) Data
- Een webserver
- Een PostGIS database
- Een XML-editor
- Virtuoso Universal Server (software voor het publiceren van Linked Data)

De stappen zien er als volgt uit:

1. Data verzamelen

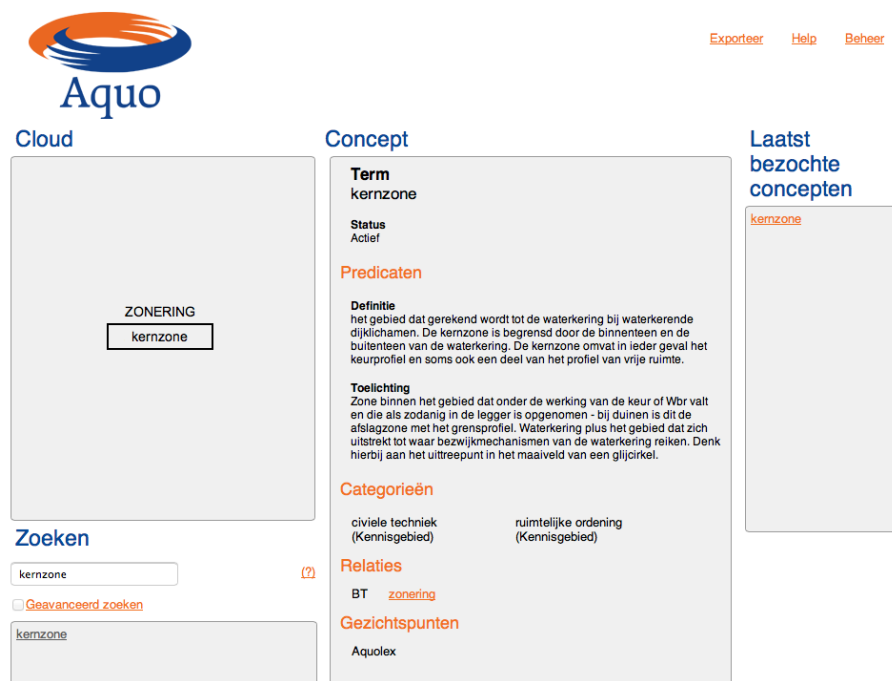
- Het bestand GW_GZN is geëxporteerd vanuit de HHNK GIS database middels een SQL script.
- De verschillende soorten zoneringen zijn in in het veld GZNSOORT slechts als domeinwaarde beschikbaar. In het veld GZN_omschrijving zijn, met behulp van de betreffende domeintabel, de tekstuele omschrijvingen van de zonesoorten ingevuld.
- Het exportbestand is vervolgens geïmporteerd in een PostGIS database.

2. Inventariseren toe te passen vocabulaires

De definities van concepten, maar ook die van relaties, zullen ergens moeten worden beschreven. Voor waterdata zoals leggerzoneringen blijkt dat er vanuit de Aquo-standaard een vocabulaire beschikbaar is. Hoewel niet alle typen zones zoals beschreven in het bestand GW_GZN beschikbaar zijn, kunnen we deze goed gebruiken.

Vanuit het Informatiehuis water is men gestart met een nieuw uiterlijk voor de Aquo-objectencatalogus. Deze is terug te vinden op: <http://aquo.gridwalker.nl/html5>.

Hierin kunnen alle beschikbare begrippen gemakkelijk worden opgezocht. Ook relaties van een object met andere objecten zijn in kaart gebracht.



Figuur 5.6 Screenshot HTML5 versie Aquo objectencatalogus

Van deze catalogus blijkt ook een 'linked' versie beschikbaar in SKOS (RDF variant). Deze is benaderbaar op <http://aquo.gridwalker.nl/items/skos>. Hiermee kunnen in een Linked Data bestand definities van concepten worden opgevraagd aan de hand van een ID-nummer. Bijvoorbeeld: Het concept kernzone heeft ID-nummer 30552, hier kunnen met de volgende URL de gegevens worden opgevraagd: <http://aquo.gridwalker.nl/items/skos?id=30552>.

Dit geeft in SKOS formaat de gegevens over het concept kernzone terug, zoals bijvoorbeeld de definitie:

“Het gebied dat gerekend wordt tot de waterkering bij waterkerende dijklichamen. De kernzone is begrensd door de binnenteen en de buitenteen van de waterkering. De kernzone omvat in ieder geval het keurprofiel en soms ook een deel van het profiel van vrije ruimte.”

Naast deze domein specifieke vocabulaire zullen er een aantal algemene vocabulaires worden gebruikt bijvoorbeeld voor het beschrijven van geodata zoals (GEOSPARQL) en relaties (FOAF).

3. Handmatige invoer

Als dataformaat is gekozen voor het ttl-formaat. Reden hiervoor is dat dit bestand overzichtelijk in opbouw is. Het ttl-bestand wordt in een XML-editor handmatig gevuld, gekozen is hier voor Notepad.

Het bestand is als volgt opgebouwd:

a) Beschrijving gebruikte vocabulaires

Allereerst wordt hier de URI voor het bestand beschreven.

Dit wordt <http://lod.geodan.nl/hhnk/zonerings>

Daarna worden de URI's van alle gebruikte vocabulaires beschreven.

```
@prefix : <http://lod.geodan.nl/hhnk/zonerings/> .
@prefix geosparql: <http://www.opengis.net/ont/geosparql#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix void: <http://rdfs.org/ns/void#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
```

b) Beschrijving metadata

In dit deel worden de eigenschappen van het bestand beschreven. Naast een aantal zaken zoals de licentie, geldigheid en extent van de data worden hierin ook een aantal verwijzingen opgenomen.

Zo wordt naast een verwijzing naar de HHNK website ook een link gemaakt ("same as") met informatie van HHNK op DBpedia. Op deze manier kan eenvoudig informatie van HHNK worden opgehaald die op Wikipedia is beschreven. Ook voor Geodan als publisher wordt een verwijzing naar een URI gemaakt.

```
# begin metadata
<http://lod.geodan.nl/hhnk/zonerings/>
a void:DataSet;
rdfs:label "Leggerzonerings van Hoogheemraadschap Hollands Noorderkwartier"@nl;
dct:license <http://www.opendatacommons.org/odc-public-domain-dedication-and-licence/>;
dct:creator :hhnk;
dct:publisher <http://lod.geodan.nl/org/geodan>;
void:sparqlEndpoint <http://lod.geodan.nl/sparql>;
dct:temporal "start=2008; end=2014;"^^<http://purl.org/dc/terms/Period>;
dct:spatial <http://nl.dbpedia.org/resource/Hoogheemraadschap_Hollands_Noorderkwartier>;
dct:spatial :spatialextent;
.
:hhnk
a foaf:Organization;
foaf:Homepage <http://www.hhnk.nl>;
rdfs:label "Hoogheemraadschap Hollands Noorderkwartier"@nl;
owl:sameAs <http://nl.dbpedia.org/resource/Hoogheemraadschap_Hollands_Noorderkwartier>;
.
:spatialextent
a geosparql:Geometry;
geosparql:asWKT "<http://www.opengis.net/def/crs/EPSG/0/28992> POLYGON((97372 487152,97372 580407,149636 580407,149636 487152,97372 487152))"^^geosparql:wktLiteral;
```

c) Beschrijving definities

Allereerst wordt er verwezen (“seeAlso”) naar de definities in de SKOS versie van de Aquo objectencatalogus: <http://aquo.gridwalker.nl/items/skos> . Voor de concepten zonering, kernzone, beschermingszone, buitenbeschermingszone en waterstaatswerk blijken definities beschikbaar. Hiernaar zal worden verwezen.

Bijvoorbeeld in het geval van de kernzone gebeurt dit als volgt:

rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30552>

Omdat de zonetypen niet exact overeenkomen met de definities uit Aquo wordt er daarnaast een eigen aanvullende beschrijving gemaakt. Bijvoorbeeld:

Buitenbeschermingszone_buitenkant: “Buitenbeschermingszone aan de waterzijde van een waterkering”

```
# begin definities
:zonering
  a skos:Concept;
  rdfs:seeAlso <http://aquo.gridwalker.nl/items/skos?id=30574>;
  skos:prefLabel "zonering"@nl;
  skos:definition "Een zone gelegen langs/op een primaire waterkering met een juridische status die uit de keur voortvloeit.
  Neem contact op met het waterschap (via http://hhnk.nl) bij voorgenomen graafwerkzaamheden."@nl;
.
:buitenbeschermingszone_buitenkant
  a skos:Concept;
  rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30537>;
  skos:prefLabel "buitenbeschermingszone buitenkant"@nl;
  skos:definition "Buitenbeschermingszone aan de waterzijde van een waterkering"@nl;
  skos:broader :zonering;
.
:kernzone_waterkering
  a skos:Concept;
  rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30552>;
  skos:prefLabel "kernzone waterkering"@nl;
  skos:definition "Kernzone van de waterkering"@nl;
  skos:broader :zonering;
.
:waterstaatswerk
  a skos:Concept;
  rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30836>;
  skos:prefLabel "waterstaatswerk"@nl;
  skos:definition "Waterstaatswerk"@nl;
  skos:broader :zonering;
.
:beschermingszone_binnenkant
  a skos:Concept;
  rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30873>;
  skos:prefLabel "beschermingszone binnenkant"@nl;
  skos:definition "Beschermingszone aan de landzijde van een waterkering."@nl;
  skos:broader :zonering;
.
:beschermingszone_buitenkant
  a skos:Concept;
  rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30873>;
  skos:prefLabel "beschermingszone buitenkant"@nl;
  skos:definition "Beschermingszone aan de waterzijde van een waterkering."@nl;
  skos:broader :zonering;
.
:buitenbeschermingszone_binnenkant
  a skos:Concept;
  rdfs:seeAlso <http://ihw.gridline.nl/items/skos?id=30537>;
  skos:prefLabel "buitenbeschermingszone binnenkant"@nl;
  skos:definition "Buitenbeschermingszone aan de landzijde van een waterkering."@nl;
  skos:broader :zonering;
.
# einde definities
```

d) Beschrijving geometrie en features

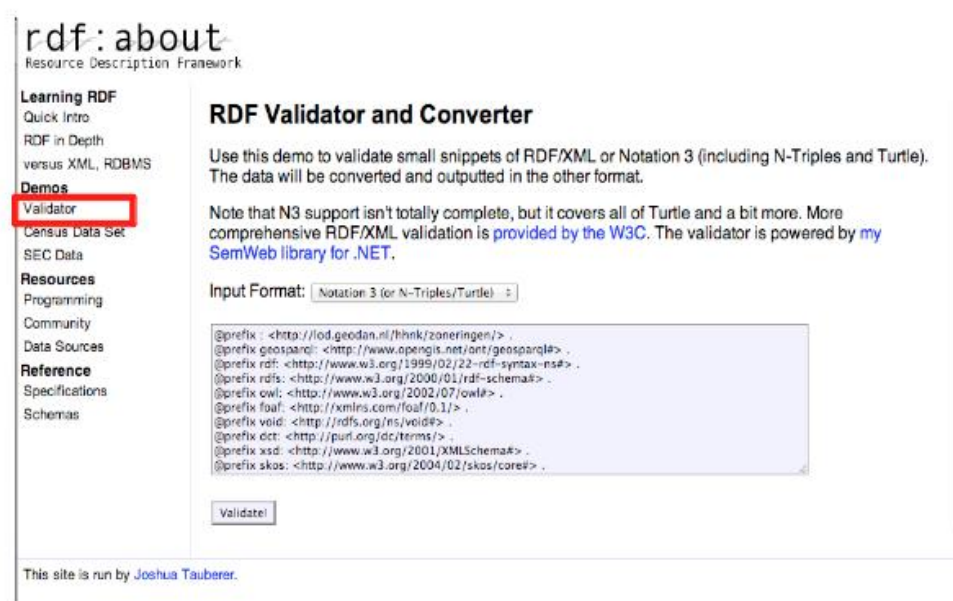
De geometrie en attributen worden met behulp van een script vanuit de PostGIS database in het ttl-bestand beschreven. Iedere feature krijgt hierbij zometeen een eigen URI bestaande uit de namespace van de dataset en de codering van de zonering. Dus bijvoorbeeld:

http://lod.geodan.nl/hhnk/zoneringen/G_278

```
# begin zoneringsdata
:G_278
void:inDataset <http://lod.geodan.nl/hhnk/zoneringen/>;
a geosparql:Geometry;
geosparql:asWKT "<http://www.opengis.net/def/crs/EPSSG/0/28992> POLYGON((110310.77 551851.033,110311.682 551851.601,110312.
rdfs:label "G_278"^^xsd:string;
dct:subject :beschermingszone_binnenkant;
```

4. Valideren RDF

Een handige manier om de syntax van het ttl-bestand te controleren op fouten of onvolkomenheden is de validatietool op de website www.rdfabout.com. Het betreft op dit moment overigens een demo-versie.



Figuur 5.7 Validatietool op rdfabout.com



Figuur 5.8 Validatiescore

5. Publiceren Linked Data

Nu het bestand klaar is zal deze online moeten worden gebracht, zodat anderen de data kunnen gebruiken. De software Virtuoso Universal Server helpt hierbij.



Virtuoso heeft diverse mogelijkheden. Naast functies voor databasemanagement en een webserver heeft Virtuoso de functionaliteit om Linked Data te serveren op het internet. Het ttl-bestand bevat alle data: metadata, de zonedata en ook nog een stukje semantiek. Deze data wordt allereerst opgeslagen in de triplestore binnen Virtuoso. Om de data vanaf internet te kunnen bevragen wordt er in Virtuoso een SPARQL endpoint gemaakt. Een SPARQL endpoint is een webservice die over het internet kan worden benaderd om de Linked Data van een organisatie op te vragen en hier met behulp van statements queries en analyses mee te doen. Het adres van onze SPARQL endpoint is:

<http://lod.geodan.nl/sparql>

Wanneer deze url via een browser wordt benaderd wordt er een HTML-pagina getoond waarop een SPARQL query kan worden ingegeven. Een SPARQL query die onze dataset zou kunnen oproepen ziet er als volgt uit:

```
select ?p ?o
from <http://lod.geodan.nl/hhnk/zoneringen/>
where {
  <http://lod.geodan.nl/hhnk/zoneringen/> ?p ?o.
}
```

SPARQL queries worden in de praktijk echter meestal van machine naar machine uitgevoerd, zonder tussenkomst van een dergelijke HTML-pagina.

Figuur 5.9 HTML-pagina met daarop de SPARQL endpoint Query Editor

<http://lod.geodan.nl/sparql>

Ook vanuit een applicatie kan het ttl-bestand via het SPARQL endpoint worden opgevraagd. De dataset, zou, zoals we gewend zijn, gevisualiseerd kunnen worden op een kaart. In onderstaande Open Layers viewer is het bestand opgeroepen vanaf het SPARQL endpoint.

De echte meerwaarde van onze nieuwe Linked Data is uiteraard als deze in een toepassing wordt gecombineerd en verrijkt met andere (niet-geo)data.



Figuur 5.10 Voorbeeld visualisatie Linked Data in Open Layers viewer

<http://research.geodan.nl/sites/frans/hhnhk.html>

6. Rechtstreekse mapping vanuit Oracle database

Een andere mogelijkheid is om het proces van converteren en publiceren van Linked Data te automatiseren door deze rechtstreeks op een relationele database uit te voeren. De stappen 1 t/m 5 worden dan in een script beschreven en periodiek uitgevoerd. Vooral bij dynamische data zal dit een toegevoegde waarde zijn, omdat het proces van omzetten naar RDF niet steeds opnieuw hoeft te worden uitgevoerd en de data in de meest actuele versie kan worden aangeboden.

Deze stap wordt in deze casestudie niet verder beschreven.

6 CONCLUSIE EN AANBEVELINGEN

6.1 Conclusie

Het zal niet lang duren voordat Open Data ook in Nederland een logisch gegeven zal zijn. Overheden, maar ook andere organisaties zullen hun data steeds meer open stellen. De principes van Linked Open Data helpen om deze data gecombineerd te gebruiken.

Op dit moment is het proces van aanbieden van Linked Data voor bronhouders echter nog niet erg laagdrempelig. Er zijn verschillende tools beschikbaar waarmee de conversie naar Linked Data kan worden gemaakt en de semantische wereld van datamodellen en vocabulaires vereist een zekere basiskennis, voordat men hiermee aan de slag kan. Echter zodra de tools laagdrempeliger worden en het aanbod van praktijkgerichte trainingen op het vlak van semantiek en Linked Data toeneemt zal ook in Nederland de toepassing van Linked Open Data zeer waarschijnlijk een vlucht nemen. Het lijkt dan ook niet de vraag of de Nederlandse overheden ermee aan de slag gaan, maar eerder wanneer dit gebeurt. De eerste initiatieven zijn al gestart!

Linked Open Data zal in de Nederlandse overheidswereld zorgen voor verdere transparantie, maar ook voor kostenbesparingen, doordat integratie van informatie met andere overheden veel minder hoeft te worden gezocht in het koppelen van systemen of het overdragen van data. Er ontstaat een web van data waar optimaal gebruik kan worden gemaakt van elkaars informatie. Daarnaast zal dit web van data nieuwe, nu nog niet voor te stellen, toepassingen voortbrengen. Deze innovatie zal de komende jaren een stimulans zijn voor de economische groei binnen Nederland, maar ook binnen de rest van Europa.

Voor ruimtelijke informatie in het algemeen betekent Linked Open Data dat de grens met reguliere informatie langzaam vervaagt. De betekenis van data wordt bepaald door de context en de geografische context zal een belangrijke meerwaarde creëren voor reguliere informatie, evenals andersom. Een algemeen aandachtspunt bij het combineren van diverse data is wel dat de dwarsverbanden beelden kunnen opleveren die in het kader van privacywetgeving niet wenselijk zijn.

Linked Open Data lijkt meer te zijn dan een modewoord. In de ontwikkelingen van steeds meer en grotere hoeveelheden Open Data lijken de Linked Data principes niet alleen een toegevoegde waarde, maar een voorwaarde om de kracht van Open Data optimaal te kunnen benutten.

Meerwaarde voor ruimtelijke waterschapsdata

Voor de interne bedrijfsvoering van een waterschap lijkt Linked Open data niet direct een meerwaarde te bieden. Wel is het zo dat het bedenken van innovatieve toepassingen niet de corebusiness is van een waterschap. Door de data die is verzameld voor het uitvoeren van de kerntaken beschikbaar te maken wordt innovatie in de markt gestimuleerd. Door deze data aanbieden als Linked Open Data ontstaan slimme toepassingen, zonder dat het waterschap hiervoor grote investeringen hoeft te doen.

Door het gebruik van Linked open waterschapsdata binnen toepassingen zoals de Huiskluis komt de informatie van een waterschap dichterbij de burger. Ruimtelijke waterdata krijgt context, doordat deze wordt gecombineerd met andere vormen van data. Leggerinformatie bijvoorbeeld wordt voor een burger een stuk relevanter in de context van andere informatie die van een woning of perceel beschikbaar is.

Tot slot spreekt het hopelijk voor zich dat (Linked) Open Data geen doel op zich mag worden. Een aanbod moet immers voorzien in een behoefte. Het is daarom van belang om in contact te blijven met de omgeving om samen te bepalen welke data moet worden opengesteld.

6.2 Aanbevelingen voor HHNK

Het bewust zijn van de meerwaarde van Open waterschapsdata voor de maatschappij is een eerste stap. Dit rapport kan hierbij helpen. Op dit moment wordt er binnen de HHNK-website op verschillende plekken en in verschillende vormen ruimtelijke informatie aangeboden. De onderliggende data zijn echter vaak niet beschikbaar en er is geen sprake van een samenhangend geheel. Daarom de volgende aanbeveling:

Bepaal een (Linked) Open Data strategie voor geodata op internet

Adopteer de 5 sterren rating van Berners Lee om een HHNK-strategie voor Open Data te bepalen. Start met het ordenen van de verschillende aangeboden data en benader dit vanuit het oogpunt van de websitebezoeker. Wat zou deze graag willen zien en op welke manier?

5 sterren rating	Omschrijving
★	Raadpleegbaar op het Web (in welk formaat dan ook) onder een open licentie
★★	Raadpleegbaar als gestructureerde data (zoals Excel in plaats van een plaatje van een tabel)
★★★	Bruikbaar in vorm van niet-merkgebonden formats (zoals CSV in plaats van Excel)
★★★★	Bruikbaar via URI's om dingen te identificeren, zodat mensen naar de data kunnen verwijzen
★★★★★	Gelinkt aan andere data om context te verstrekken

Hierbij een voorzet voor een HHNK-strategie voor de komende 3 jaar om geodata beschikbaar te stellen als (Linked) Open Data:

2013-2014: ★ "Harmoniseren"

- Breng de behoefte aan open waterdata in beeld door in gesprek te gaan met de omgeving
- Investeer in kennis en ervaring op het gebied van (Linked) Open Data
- Breng alle huidige online geodata samen op 1 locatie op de website, zodat de bezoeker snel en gemakkelijk kan zien welke informatie het waterschap op kaart aanbiedt. Op korte termijn wordt de website van de HHNK vervangen. Het is een kans om deze wijziging bij de opbouw van de nieuwe website mee te nemen.
- Harmoniseer zoveel mogelijk de bestaande verschijningsvormen (bijvoorbeeld 1 type webviewer).

2014: ★★★ "Webservices en ruwe data"

- Bied webservices (WMS/WFS) aan van de data die in de viewers getoond worden en registreer deze in het NGR. Op deze manier wordt de data vindbaar en kan de geodata in GIS software buiten HHNK als service worden gebruikt. De benodigde infrastructuur (hardware/software) is hiervoor reeds bij HHNK aanwezig. Echter is deze nog niet operationeel.
- Organiseer, bijvoorbeeld met andere waterschappen, een competitie voor ontwikkelaars (een zgn Hackaton) om applicaties te ontwikkelen met de Open Data vanuit het waterschap.
- Bied enige geodata aan als Linked Data om het effect ervan te onderzoeken.

2015-2016: ★★★★★(★) "URI's en linken"

- Definieer URI's voor de data en gebruik vocabulaires van bijvoorbeeld het Informatiehuis Water om de Open geodata als Linked Data aan te bieden.

7 REFERENCES

Hoofdstuk 1

- <http://policybythenumbers.blogspot.nl/2012/03/historical-anecdote-about-open-data.html>
- <http://www.nu.nl/internet/3394340/eu-wil-meer-overheidsdata-openbaar-maken.html> (10 april 2013)
- “Commission welcomes Member States' endorsement of EU Open Data rules”
http://europa.eu/rapid/press-release_IP-13-316_en.htm (10 april 2013)
- www.hetwaterschapshuis.nl
- www.informatiehuiswater.nl

Hoofdstuk 3

- www.data.overheid.nl.
- Website HHNK, www.hhnk.nl
- Website Hoogheemraadschap Stichtse Rijnlanden, <http://www.hdsr.nl>
- “Gevoorziening-klaar”
<http://www.hetwaterschapshuis.nl/nieuws/@214755/gevoorziening-klaar/>
- <http://wetten.overheid.nl>
- <http://www.geonovum.nl/dossiers/inspire>
- Notitie-verslag INSPIRE Annex III_ 20121129 (K. Van Gijssel, 2013)
- Waterschappen en Open Data beleid: een afwegingskader (Informatiehuis Water, 22 januari 2013)
- www.pdok.nl

Hoofdstuk 4

- Linked Data. Design Issues for the World Wide Web - W3C, (Tim Berners-Lee, 2006, <http://www.w3.org/DesignIssues/LinkedData.html>)
- Linked Open Data: The Essentials - A Quick Start Guide for Decision Makers (F. Bauer, M. Kaltenbock, 2012)
- RDF Schema, Presentatie VU Amsterdam (Eyal Oren, 2009, <http://webkr.cs.vu.nl/slides/rdfs.pdf>)
- Linked Open Vocabularies LOV (<http://lov.okfn.org/dataset/lov/index.html>)
- Gegevens delen via Linked Open Data (Alterra-rapport 2306, 2012)
- Linked Open Data Around-The-Clock (www.latc-project.eu)
- <http://techcrunch.com/2012/05/16/google-just-got-a-whole-lot-smarter-launches-its-knowledge-graph/>
- Plan van aanpak Pilot LOD (Geonovum, 2012)

Hoofdstuk 5

- Website HHNK, www.hhnk.nl
- <http://hnk-water.nl>
- Pilot Linked Open Data (Een project van Geonovum, 2012-2013)
<http://www.geonovum.nl/dossiers/linkedopendata>
- Waterschap en bluswatervoorziening (www.brandweerkennisnet.nl)
- <http://linkeddatabook.com>
- <http://www.rdfabout.com>
- <http://aquo.gridwalker.nl/html5/>
- Trentino government linked open geo-data: a case study (P. Shvaiko, F. Farazi, V. Maltese, A. Ivanyukovich, V. Rizzi, D. Ferrari, G. Ucelli, 2012)

BIJLAGE 1- RESULTATEN INTERVIEWS

Vragenlijst afstudeeronderzoek “Linked Open Data”

Datum: Diverse (zonder script)

Naam: Bart van Leeuwen

Bedrijf: Brandweer Amsterdam-Amstelland / Netage.nl

1. Wat is je rol binnen HHNK?

Ik ben brandweerman bij het korps Amsterdam-Amstelland en daarnaast eigenaar van Netage.nl. Vanaf 2007 houd ik mij vanuit Netage.nl bezig met Semantic Web Technology. Het uiteindelijke doel van Netage.nl is om de informatievoorziening aan gebruikers te verbeteren door het benutten van semantiek.

2. Wat is je standpunt op het openbaar beschikbaar stellen van data (geodata en van andere soorten) door de overheid? En specifiek vanuit de waterschappen?

Ik vind het niet meer dan logisch dat overheden hun data opstellen. Er is reeds voor betaald door de burger. Toch zie je dat data nog veel blijft opgesloten in eigen systemen. Dat is erg jammer. In een crisissituatie bijvoorbeeld kan de juiste informatie het verschil maken.

3. Welke knelpunten zie je?

Innovatie blijft vaak hangen binnen organisaties, doordat kennis niet wordt gedeeld. Zo ook bij Open Data en dat is zonde. Het zou goed zijn als er meer kruisbestuiving zou plaatsvinden, zodat het wiel niet steeds opnieuw hoeft te worden uitgevonden.

4. Welke meerwaarde en kansen zie je bij het bevorderen van het gebruik van open data, bijvoorbeeld door het principe van Linked Open Data?

Vanuit de waterschappen bekeken zou het gebruik van Linked Data een toegevoegde waarde zijn bij bijvoorbeeld calamiteiten. Een voorbeeld vanuit mijn eigen praktijk is bijvoorbeeld dat 1 van onze bluswagens op dit moment is uitgerust met telemetrie, waarbij het waterverbruik wordt gemonitord en de data als Linked Data beschikbaar is. Het zou interessant zijn om deze data te combineren met data over waterstanden vanuit het waterschap. Ook andere samenwerkingsverbanden bij calamiteiten vragen om snelle en efficiënte informatie-uitwisseling. Linked Data is universeel en licht in gebruik, maar is ook flexibel.

5. Welke verdere ambities heb je op het vlak van Linked Open (geo)data?

Niet besproken

6. Op welke manier zul je hier de komende 3 jaar invulling aan geven?

Niet besproken

7. Heb je nog andere vragen of opmerkingen?

Nee

Vragenlijst afstudeeronderzoek “Linked Open Data”

Datum: 25 april 2013

Naam: Reind van Olst

Bedrijf: 2CoolMonkeys

1. Kun je kort iets vertellen over 2CoolMonkeys en wat is jouw rol?

2CoolMonkeys werkt sinds 2008 als producent van apps voor alle platforms, maar ook mobiele websites, webapps, etc. Dit noemen wij mobiele media producties. Wij hebben inmiddels zo’n 40 apps op onze naam staan voor diverse opdrachtgevers zoals Rijkswaterstaat, het Kadaster en Royal Haskoning DHV. Wij zijn gespecialiseerd in apps met ruimtelijke informatie en zijn tevens (silver) partner van ESRI.

Ik ben Open Data Specialist en verantwoordelijk voor de commercie rondom onze apps en het bedenken van nieuwe concepten.

2. Wat is je standpunt op het openbaar beschikbaar stellen van data (geodata en van andere soorten) door de overheid? En specifiek vanuit de waterschappen?

Door als overheidsorganisatie data beschikbaar te maken vergroot je de betrokkenheid van de burger, doordat deze beter geïnformeerd is en zich een mening kan vormen over de wijze waarop de overheidstaak wordt uitgevoerd. Daarbij stel je de markt in staat om nieuwe toepassingen voor de data te bedenken. Die kan dit meestal beter als de overheid en nuttige, maar ook onverwachte toepassingen ontstaan.

Vanuit de waterschappen denk ik aan bijvoorbeeld (grond) waterstanden. Dit zal weliswaar niet direct een “killer-app” voor het grote publiek opleveren, maar voor specifieke doelgroepen zoals agro-adviseurs of boeren (wanneer moet ik beregenen?) kan dit erg interessant zijn.

3. Welke belemmeringen zie je tbv Open Data?

Vooroordelen en te weinig kennis betreft Open Data bij overheidsorganisaties zorgt ervoor dat men vaak nog terughoudend is. Men kan ook bang zijn voor reacties van burgers over de eventuele onjuistheid van de gegevens. Dit terwijl het publiek juist kan bijdragen aan het verbeteren van de kwaliteit. Een voorbeeld hiervan is één van onze nieuwe apps waarbij inwoners van een gemeente de kaart kunnen verbeteren, maar ook meldingen kunnen doen over de openbare ruimte.

Je kunt de koudwatervrees wegnemen door bijvoorbeeld ontwikkelaars uit te nodigen om voorbeelden te laten zien of het organiseren van een zgn hackathon waarbij je een prijs uitreikt aan de beste ingezonden app.

4. Welke meerwaarde en kansen zie je bij het bevorderen van het gebruik van open data, bijvoorbeeld door het principe van Linked Open Data?

De werkelijke meerwaarde van Open Data en zeker Linked Open data is het koppelen van afzonderlijke data, zodat nieuwe informatie en inzicht ontstaat. Het gebruik van de data wordt zo steeds gemakkelijker, terwijl de data bij de bron blijft.

Een aandachtspunt kan echter privacy zijn. Door het combineren van verschillende data kunnen nieuwe beelden ontstaan die niet altijd even wenselijk zijn in het kader van de privacywetgeving.

5. Welk ambities heeft 2CoolMonkeys op het vlak van Open (geo)data?

Wij zullen apps blijven bouwen en ons daarnaast richten op het klant-en-klaar aanbieden van data, als service, voor andere app-ontwikkelaars.

6. Heb je nog andere vragen of opmerkingen?

Nee

Vragenlijst afstudeeronderzoek “Linked Open Data”

Datum: 21 mei 2013

Naam: Bart Bos

Bedrijf: HHNK

1. Kun je kort iets vertellen over 2CoolMonkeys en wat is jouw rol?

Ik ben Informatie-analist binnen HHNK bij het cluster Onderzoek, afdeling Ingenieursbureau. Ik houd mij onder meer bezig met waterkwaliteitsgegevens en de Kader Richtlijn Water (oa ontwikkeling van het KRW Volg en Stuursysteem). Daarnaast zit ik in een werkgroep bij het Informatiehuis Water voor de landelijke standaardisatie van waterkwaliteitsinformatie.

2. Wat is je standpunt op het openbaar beschikbaar stellen van data (geodata en van andere soorten) door de overheid? En specifiek vanuit de waterschappen?

Alles is open tenzij! Alle overheden die worden gefinancierd worden door de burger moeten de informatie die zij verzamelen openbaar stellen, tenzij dit gevoelige of persoonsgebonden informatie betreft.

3. Welke belemmeringen zie je tbv Open Data?

Nu is zeker bij de waterschappen alles gesloten tenzij. Reden hiervoor zou angst kunnen zijn vanuit beveiligingsoogpunt bij ICT (verbindingen van buiten naar binnen). Binnen de organisatie heeft men vooral oog voor het uitvoeren van de watertaak, maar minder op de interactie met de maatschappij, bijvoorbeeld door het aanbieden van data. Om als organisatie met Open Data aan de slag te gaan is het echter nodig hier inzet en wellicht ook budget voor te reserveren. Innovatie op dit gebied lijkt, afgezet tegen de watertaken, minder prioriteit te krijgen bij het toebedelen van uren en geld.

Op landelijk niveau zitten de waterschappen erg op “hun” data. Er is onder andere angst voor ongecontroleerd hergebruik. Dit terwijl ruwe data in mijn oogpunt geen risico met zich meebrengt. Pas als het waterschap hier conclusies aan verbindt in bijvoorbeeld rapporten dan zijn wij aansprakelijk. Conclusies die worden gemaakt door andere partijen kunnen niet worden “verhaald” op het waterschap.

4. Welke meerwaarde en kansen zie je bij het bevorderen van het gebruik van open data, bijvoorbeeld door het principe van Linked Open Data?

Wat ik begrijp van het concept Linked Data is dat verschillende soorten aangeboden data snel met elkaar kunnen worden gecombineerd. Voor het waterschap zou dit kansen betekenen bij het samenwerken met partners bijvoorbeeld op het vlak van een landelijk beeld van de toetsing van waterkeringen. Hier zijn verschillende informatie- typen die moeten worden gecombineerd. Naar de burger toe zou de schouwplicht van wateren gericht naar de burger kunnen worden gecommuniceerd.

5. Welk ambities heeft jouw afdeling op het vlak van Open (geo)data?

Ik denk dat allereerst dat Open Data voor HHNK geen optie is, maar een plicht. We zouden het positieve effect van Open Data kunnen laten zien door toepassingen te (laten/helpen) ontwikkelen. We zouden binnen HHNK dan ook meer moeten investeren in innovatie op dit vlak.

Vanuit cluster Onderzoek vinden wij innovatie belangrijk, ook op het vlak van data. Zo zullen we in 2013 biologische en chemische datareeksen (van circa 2000 locaties) online beschikbaar stellen, in open formaten en volgens het Aquo uitwisselingsmodel.

6. Heb je nog andere vragen of opmerkingen?

Ik denk dat (Linked) Open Data in het algemeen een mogelijke keerzijde heeft waarvan men zich bewust moet zijn. Het combineren van data kunnen beelden opleveren die op het randje van de privacywetgeving balanceren. Bedrijven als Facebook en Google zullen hun voordeel doen met deze nieuwe technieken. Daarbij kunnen andere partijen op hun beurt waterdata voor commerciële doeleinden inzetten. De vraag is of dat laatste erg is.

Echter vind ik de voordelen zwaarder wegen dan deze keerzijde!

Vragenlijst afstudeeronderzoek "Linked Open Data"

Datum: 27 mei 2013

Naam: Ton Hugens

Bedrijf: HHNK

1. Wat is je rol binnen HHNK?

Ik ben directeur van de directie Bedrijfsvoering binnen waterschap Hoogheemraadschap Hollands Noorderkwartier.

2. Wat is je standpunt op het openbaar beschikbaar stellen van data (geodata en van andere soorten) door de overheid? En specifiek vanuit de waterschappen?

Ik ben een voorstander van zowel Open Data als Open Source. Een belangrijk criteria vind ik echter dat de (kritische) bedrijfsvoeringsactiviteiten gewaarborgd blijven van risico's. Een discussie over de kosten die op dit moment vaak gepaard gaan met het gebruik van software vind ik echter gezond.

Door data open te stellen kan de overheid rekenschap af leggen, is deze controleerbaar en transparant. De data is al betaald door de burger, dus het lijkt logisch deze aan te bieden.

3. Welke knelpunten zie je?

Er is veel koudwatervrees bij overheden, over hergebruik en door twijfels over de kwaliteit van de data. Voor het openstellen van data is een andere mindset nodig. Van "Kennis is macht" naar "Kennis is voor iedereen". De overheid heeft geen voorsprong meer op het vlak van informatie.

Door toch te moeten voldoen aan de wettelijke eisen ontstaat mogelijk "Open data in name only"; men stelt wel data beschikbaar, maar veel gegevens worden geclassificeerd als zijnde bedrijfskritisch om deze toch niet te hoeven open stellen.

4. Welke meerwaarde en kansen zie je bij het bevorderen van het gebruik van open data, bijvoorbeeld door het principe van Linked Open Data?

Voor de eigen bedrijfsvoering is de meerwaarde denk ik niet direct aanwezig. Wel zie ik voordelen in de samenwerking met partners. Of bij het voeren van een bredere discussie met de omgeving over de taken van het waterschap. Alle belanghebbenden kunnen beschikken over dezelfde informatie, waardoor in alle redelijkheid tot het beste besluit kan worden gekomen.

5. Welk ambities heeft HHNK op het vlak van Open (geo)data?

Bij het bepalen van prioriteiten binnen de directie van HHNK staat specifiek Open Data op dit moment niet hoog op de agenda. Ik vind het echter belangrijk om de komende jaren te investeren in ICT, waarbij gegevensuitwisseling een thema kan zijn.

6. Op welke manier zal HHNK hier de komende 3 jaar invulling aan geven?

NVT

7. Heb je nog andere vragen of opmerkingen?

Nee

Titel

Linked Open Data

De meerwaarde van Linked Open Data bij de uitwisseling van ruimtelijke waterschapsinformatie

Afstudeerscriptie van Elvin Martha binnen de duale opleiding Geodesie / Geoinformatica aan de Hogeschool Utrecht (2013)
