



Q&A eindpresentatie innovatiesprint DiS Geo

DiS Online 14 juli 2020

Gestelde vragen naar aanleiding van de presentaties en demonstrators tijdens de DiS Online webinar 'Innoveren voor DiS – eindpresentatie innovatiesprint':

In hoeverre zijn gegevens gelinkt met BGT gegevens?

De BGT is niet expliciet verlinkt voor deze High 5. Echter is er al wel een experimentele set van de BGT (in Linked Data) beschikbaar van een eerdere innovatiesprint. Deze is wel aangesproken in combinatie met de gesloten IMGeo data van de Provincie. Zie ook: <https://labs.kadaster.nl/stories/high5-imbor/>).

Het is goed om te weten dat er een opdracht van BZK/DiS Geo ligt bij het Kadaster om de BGT te verlinken. De verwachting is dat deze in 2020 nog gerealiseerd zal worden.

Is er ook gekoppeld aan IHW Aquo data?

Nee, dit is niet gedaan. We hebben wel geo-gegevens van het Informatie Huis Water (IHW) beschikbaar gekregen, echter is het niet gelukt deze in de High 5 beet te pakken.

Hebben jullie de wijze van Linked Data modelleren expliciet vastgelegd? Wat is de link met de pas gelanceerde NTA8035, semantisch modelleren in de gebouwde omgeving? Is er link met Geonovum MIM?

In de High 5 werkten deelnemers mee die ook deelnemers zijn in de verschillende werkgroepen met betrekking tot de standaarden (zoals: Niels Hoffman, Provincie NH; Bart van Leeuwen, Netage; en Pano Maria, Kadaster).

Geonovum zal bij het modelleren het informatiemodel van de Samenhangende Objectenregistratie binnen DiS Geo uitgaan van MIM.

De WOZ ontbreekt in deze demonstrators, wordt de WOZ niet gezien als georegistratie?

Vanwege de gesloten natuur van de WOZ is deze niet volledig als Linked Data beschikbaar. Wel is er een experimentele set beschikbaar die de waarde van WOZ-objecten (Voor een adres) in een categorie indeelt (de chatbot op <https://labs.kadaster.nl/> gebruikt deze bijvoorbeeld).

We erkennen wel dat de WOZ een enorm belangrijk onderdeel is van een samenhangende registratie. WOZ neemt veel data over van gerelateerde basisregistraties (oa. BAG, BRK). Daarnaast is de WOZ uiteraard wél in scope bij de doorontwikkeling in samenhang van de geobasisregistraties, [DiS Geo en de ontwikkeling van de samenhangende objectenregistratie](#).

Zijn jullie tegen een les aan gelopen ten aanzien van de actualiteit van gegevens?

Er is zeker een risico aan het gebruiken van bronnen met een uiteenlopende actualiteit. Zo zien we bijvoorbeeld dat niet iedere linkset (indien noodzakelijk) up-to-date wordt



gehouden of beheert door een partij (i.e. geen governance). In de high 5 liepen we hier niet expliciet tegenaan, maar we nemen dit punt wel mee in ons verder werk voor DiS Geo. Dit punt is tevens terug te vinden in de *lessons learned*.

Tijdens de sessie is de waarde van metadata benadrukt op dit vlak. Het moet voor de gebruiker duidelijk zijn met welke actualiteit hij/zij werkt.

Stel we hebben meer uniforme afstemming (semantiek) over modelleren in Linked Data, wat wordt er dan nog meer mogelijk?

Er zijn een aantal partijen waaronder TNO en Geonovum die zich bezighouden met standaarden en manieren van modelleren. Alhoewel we er naar moeten streven dit zoveel mogelijk uniform te doen is er tijdens de sessie wel gewaarschuwd voor de ambitie om de hele wereld in één model / type modellering te vatten. Dat is vermoedelijk een brug te ver.

Wel is er tijdens de High 5 ook veel gesproken over een zekere uniformering vanuit de basisregistraties. (Weten dat je naar een basisregistratie linkt bijvoorbeeld.). Zeker een belangrijk thema om bij stil te (blijven) staan.

Hoe hebben jullie in de High 5 de beslissing genomen wie de linkset tussen twee registraties aandraagt?

Tijdens de High 5 zijn we hier zeer pragmatisch mee omgegaan, maar in het stelsel moet dit goed geregeld worden. Zowel tijdens deze High 5 als de vorige is het stuk Governance op het Snijvlak (i.e. de linkjes) als belangrijk risico / tekortkoming aan het huidige stelsel teruggekomen.

Wat moet er gebeuren om het stelsel van basisregistraties generiek in samenhang te kunnen gebruiken m.b.v. Linked Data? Dus zonder dat je van te voren weet welke vraag gesteld gaat worden.

Uitgangspunt is dat de data op een generieke manier beschikbaar wordt gesteld o.b.v. open standaarden, duidelijke metadata en verbindingen naar andere registraties. Daarna kan voor een *specifieke* vraag gebruik gemaakt worden van het *generieke* dataplatform. De aanpak is zoveel mogelijk tool agnostisch, waardoor ook een vendor lock in wordt voorkomen. Je kan dus inderdaad de data linken zonder dat je weet welke vragen er gesteld gaan worden.

Nu lijkt het veel handwerk om data te koppelen. Als je dat gestandaardiseerd kan krijgen wordt het dan generieker?

Het maken van 'linksetjes' is vaak redelijk triviaal en kan, zeker naarmate de kennis van Linked Data stijgt, steeds sneller en vrijwel volledig ge-automatiseert. De uitdaging is hierin vaak niet technisch, maar organisatorisch (wie maakt de links en wie beheert ze).

Hoe weet je welke data uit welke silo je nodig hebt om een antwoord te vinden op de vraag?

Het is een uitdaging om in de Knowledge Graph je weg te vinden naar de juiste data. Vaak vereist dit enige domeinkennis. Nu hebben we best wel moeten graven in de graph, maar zonder knowledge graph is het helemaal moeilijk.



De stelselcatalogus zou de rol zou kunnen spelen van wegwijzer in de data. Op dit moment verwijst nog geen enkele basisregistratie die als Linked Data beschikbaar is naar de stelselcatalogus.