

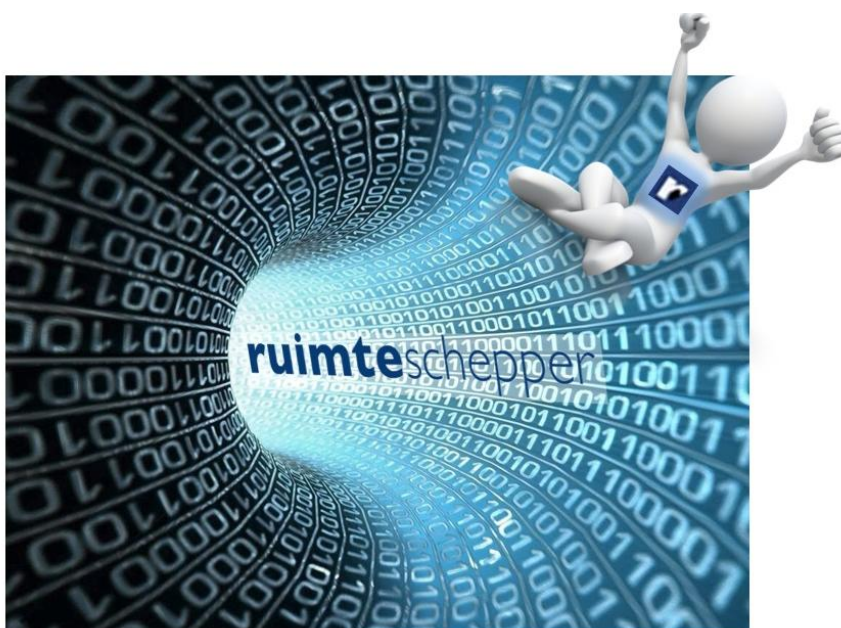
Deel 3

Het waarom van het benutten van externe open geodata is veelal geen vraag meer, het hoe is voor velen echter wel een aandachtspunt. Door de vele facetten ervan vraagt dit de nodige kennis....

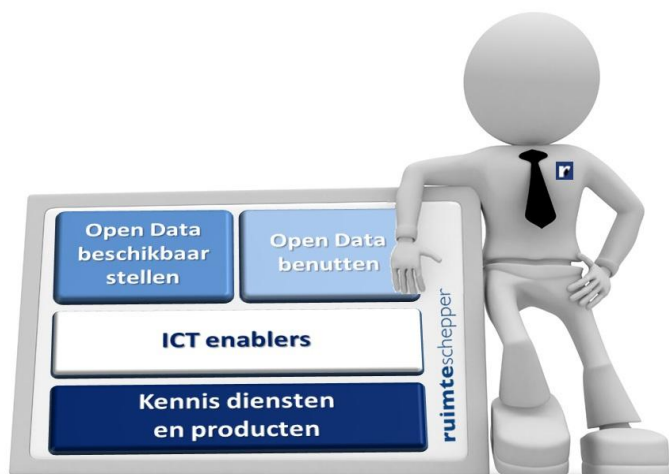
In de vorige artikelen over externe open geodata is ingegaan op de meerwaarde van het benutten van deze data en op WAT aspecten die bij de aanbieder van deze data, in veel gevallen de overheden, een rol spelen. De volgende artikelen richten zich meer op de HOE vraag. Wat komt erbij kijken om de externe open data daadwerkelijk beschikbaar te stellen en te benutten binnen de eigen bedrijfsvoering. Daarbij put ik graag uit de concrete ervaringen vanuit de activiteiten in dit kader binnen Ruimteschepper.

Open data kennisgedreven

Naarmate je je meer in de mogelijkheden van open data verdiept, kan het bijna niet anders dan dit je enthousiast maakt. Dat was in elk geval ook de algemene reactie van de deelnemers van Ruimteschepper na een verdiepingssessie over dit onderwerp. Samen met bekende namen uit het open data werkveld is naar de verschillende facetten van open data gekeken. Belangrijke conclusie van de bijeenkomst was dat er zoveel aspecten een rol spelen bij de implementatie van open data binnen organisaties, dat dit al snel in vragen naar kennis over het WAT en het HOE zal leiden. Zowel voor de aanbieders van de data als de potentiële gebruikers speelt dit een rol. De veelheid aan data, de bij behorende do's and don'ts en de benodigde ict-kennis over minder alledaagse zaken als webservices en dergelijke, kunnen al snel een drempel voor potentiële gebruikers opwerpen. Onderwerpen die zeker daar waar ze binnen het geo-domein passen binnen de competenties van de aangesloten deelnemers van Ruimteschepper vallen. Op basis van kennisdiensten en kennisproducten (sessies, workshops, publicaties) zien wij een rol weggelegd voor de Ruimteschepper geo-experts. We laten ons dan ook graag meesleuren in de hype rondom de externe open data.



Meegaan met open data hype



Gestructureerde ondersteuning implementatie open data

Structuur in ondersteuning

Om de ondersteuning van de implementatie van open data te ondersteunen zijn er een hoofdopdeling gemaakt met de volgende aandachtsgebieden:

- het beschikbaar stellen van open data
- het benutten van open data
- de ict-aspecten om dit mogelijk te maken
- de ondersteunende kennisdiensten en producten.

Deze artikelreeks volgt ook de lijnen van deze opdeling. In de rest van dit artikel ligt het accent op praktische aspecten die met het beschikbaar stellen van de data,

Vertrekpunt open data

In het vorige artikel zijn belangrijke bronnen als PDOK, de juridische aspecten en datakwaliteit genoemd. Hier willen we kijken naar meer inhoudelijke aspecten. Een logisch vertrekpunt daarvoor vormt de definitie die de overheid gemaakt voor het beschikbaar stellen van open data:

- Primair: 'brongegevens, niet geaggregeerd', daarmee de gebruiker de maximale mogelijkheid biedend de data voor de eigen toepassing te benutten;
- Compleet: 'alle openbare data en informatie van een thema beschikbaar stellen' waarmee de gebruikers de maximale mogelijkheid krijgt deze te benutten voor de eigen toepassingen;
- Tijdig: 'zo snel als kan beschikbaar stellen', om de actualiteit van het "nu" te behouden, verouderde informatie is veelal niet nuttig voor gebruikers en kan verwarrend werken;
- Toegankelijk: 'voor iedereen, voor elk legaal gebruik, zonder (financiële) drempels', waarmee de overheid garanderen dat de data daadwerkelijk open is en niet een schuil gaat achter een veelheid aan regels;
- Open licentie: 'geen beperking op basis van auteurswet, databankenwet, of geheimhouding', past onder de noemer toegankelijkheid, zie in dit kader ook de toelichting op de disclaimers en proclainers zoals deze internationaal zijn gedefinieerd in het vorig artikel;
- Machine leesbaar: 'voor automatische verwerking' is het digital beschikbaar stellen van de data een evidente eis aan de data;
- Open standaarden: om een zo breed mogelijke doelgroep te kunnen bedienen is 'het benutten van open standaarden' een belangrijk vertrekpunt, daarmee zorgt de overheid dat het format geen drempel is;

Deze definitie kent inmiddels een breed gebruik. Het vertalen van de op zich logisch gekozen vertrekpunten naar een concrete implementatie ervan laat evenwel nog genoeg vraagtekens en keuze mogelijkheden open.

Internationale definities

Een tweede belangrijke definitie in dit kader, is gedefinieerd door Tim Berners-Lee. Voor wie hem niet kennen, hij was één van de grondleggers van het World wide web waar hij in het begin van de jaren negentig aan heeft gewerkt. Hij is tot de dag van vandaag voorzitter van het World wide web consortium. In 2009 werd Berners-Lee door de Britse overheid ingehuurd om toezicht te houden op hun initiatief om een website te realiseren met daarin publiek beschikbare Britse data. Deze site (<http://data.gov.uk/>) is voor geïnteresseerden zeker een bezoek waard. Tim heeft ten aanzien van open data standaards een belangrijke bijdrage geleverd door het opstellen van het zogenaamde "5 star deployment scheme for Open Data". Daarin geeft hij in een gelaagd model de mogelijkheden om data open aan te bieden. Het model ziet er globaal als volgt uit:



★	Stel data beschikbaar, in welk format dan ook;
★★	Maak de data toegankelijk door deze gestructureerd aan te bieden, bijvoorbeeld als een Excel spreadsheet, tabel, of gescand image;
★★★	Gebruik niet leverancier afhankelijke formaten bijvoorbeeld CSV in plaats van een Excel spreadsheet)
★★★★	Maak het beschikbaar op het internet waardoor het via de url links door iedereen om (onderdelen van) de data te benaderen en ernaar te refereren; Koppel de data aan andere data zodat er een context voor de data ontstaat, de zogenaamde linked open data. Hierdoor vormt de data onderdeel van een semantisch netwerk aan datasets die gebruikers de maximale vrijheid geven de voor hen relevante informatie te achterhalen.
★★★★★	

Per stap schetst hij ook bijbehorende kosten en baten. Het moge duidelijk zijn dat de baten voor de eindgebruiker oplopen met de groei van het aantal sterren, maar dat omgekeerd de gevraagde inspanning van de publicerende organisatie, en daarmee de kosten, ook oplopen in diezelfde richting.

Linked open data

Linked Data is in feite de wijze waarop de data beschikbaar zou moeten komen. Deze manier maakt de samenhang tussen verschillende (open) databronnen zichtbaar. Door deze samenhang is het makkelijk om relevante informatie te vinden en realiseert toegevoegde waarde aan individuele databronnen door betekenisvolle relaties zichtbaar en klikbaar te maken. De data is in deze benadering zowel onderling verbonden, als open toegankelijk en deelbaar via het Semantische Web. Hoewel het idee van het koppelen van gegevens op deze manier niet nieuw is, is de huidige wijdverbreide praktijk van daadwerkelijk koppelen van datasets wel relatief nieuw. Daarom zijn de protocollen, standaarden, en licentiemogelijkheden voor gekoppelde open data nog steeds in ontwikkeling.

Het benutten van externe open data kent dermate veel facetten dat ook komende artikelen over dit onderwerp zullen gaan. Daarbij zal het volgend artikel weer wat nadrukkelijker de geodimensie oppakken, en kijken naar het benutten van de vele waardevolle open data bronnen in ons werkveld.

 j.roodzand@net4s.nl

 www.net4s.nl

 j.roodzand@ruimteschepper.nl

 www.ruimteschepper.nl

 @JanRoodzand

 06 200 133 65

Jan Roodzand heeft een ruime ervaring als ondernemer, adviseur en programmamanager in het implementeren van GIS-omgevingen. Kenmerkend daarbij is de integratie ervan met andere informatiesystemen en de aandacht voor datakwaliteitsaspecten. Hij is mede initiatiefnemer en partner in Net4s en Ruimteschepper

