

Bedrijven kunnen profiteren van open data, maar weten nog niet goed hoe

Markt voor verwerken van vrij beschikbare data staat nog in de kinderschoenen

Hanne Dijkstra
Amsterdam

'Door effectief gebruik te maken van open data, kan beleid efficiënter en duurzamer worden', zegt Laurens Ruster van Data Kitchen. Vanuit zijn in januari opgericht start-up helpt hij bedrijven beter beleid te voeren door open data te combineren. Zo heeft hij op basis van vrij toegankelijke open data voor een incassobureau verschillende debiteuren-profielen opgesteld. 'Bedrijven hebben vaak geen idee tot hoeveel informatie zij toegang hebben. Wij halen er meer uit door de hele pot data te gebruiken.'

Het begrip 'open data' is de laatste jaren een hype, het wordt zelfs 'het nieuwe goud' genoemd. In de puurste vorm zijn open data alle data die vrij toegankelijk zijn en zonder beperkingen gedeeld kunnen worden. Ook worden ze in een computer-vriendelijke dataset aangeboden, zodat ze gemakkelijk met andere gegevens gecombineerd kunnen worden. In Nederland is zo'n 70% van de beschikbare data afkomstig van de overheid. Maar ook vanuit de wetenschap en het bedrijfsleven groeit het aanbod. De grootte van een dataset kan variëren van een paar gegevens tot een paar miljoen.

Wereldwijd staat Nederland op de vijfde plaats in de Open Data Index. Deze index meet de kwaliteit en de beschikbaarheid van de open data die door de overheid beschikbaar worden gesteld. De Nederlandse database bevat op dit moment 5711 data-

sets. Ter vergelijking: de Britse en Amerikaanse databases hebben op het moment respectievelijk 18.447 en 105.593 sets beschikbaar. Eind deze maand stelt het Centraal Bureau voor de Statistiek (CBS) 3300 nieuwe datasets beschikbaar. Dat zijn de gegevens uit het statistiekprogramma Statline die als open data worden aangeboden. Het CBS heeft zo'n € 200 per dataset geïnvesteerd en de sets zijn gratis toegankelijk.

Doordat steeds meer beschikbare datasets gecombineerd kunnen worden, ontstaat een groot informatiepotentieel. Laurens Ruster en zijn collega's van Data Kitchen hebben bijvoorbeeld de gegevens van onder andere Basisregistraties Adressen en Gebouwen (BAG), het CBS en de verbruiksdata van netbeheerder Alliander gecombineerd om per wijk het inkomen, energieverbruik en grootte van percelen in kaart te brengen.

Op basis van die gegevens hebben zij debiteurenprofielen voor het incassobureau vastgesteld, waarop de aanpak per debiteur kan worden afgestemd. Ruster: 'Er is een groot verschil tussen iemand die niet kan betalen of niet wil betalen. Als de oorzaak onmacht door werkloosheid is, is het verstandig een betalingsregeling aan te bieden. Maar als wij kunnen zien dat iemand een groot huis bewoont en veel energie verbruikt, kan deze persoon harder aangepakt worden.'

Hoewel steeds meer open data beschikbaar komen, blijken veel bedrijven er nog weinig mee te kunnen. Grote, ongestructureerde 'big data'-bestanden zijn lastig te verwerken. Met de komst van open data is dus ook een nieuwe markt van data ordenen en verwerken ontstaan. En het zijn innovatieve nieuwe start-ups, zoals Data Kitchen, die in dit kennisgat springen.

Open data staan nog in de kinderschoenen en weinig bedrijven weten er al winst uit te halen. Toch zijn er bedrijven die wel de stap naar open data hebben gezet. KPN, netbeheerder Enexis en TNO ontwikkelen in het project Toegankelijke Energie Informatie een ICT-infrastructuur die het voor bedrijven makkelijker moet maken datasets aan te bieden.

Esri Nederland, expert in standaard Geografisch Informatie Systeem (GIS), heeft zich achter de schermen tot een belangrijke drijvende kracht ontwikkeld. Door technische ondersteuning en open data-modellen aan te bieden, geeft het start-ups een platform om uit te groeien. Visualisering van data is hun specialisatie en door de broncode van de programma's openbaar te maken, kunnen app-ontwikkelaars er ook zelf mee aan de slag.



Een 3D-model van Valkenburg, op basis van open data, gemaakt door het Kadaster en Esri Nederland om de landschapsvervuiling door windmolens te laten zien.

FOTO'S: KADASTER

Door datasets te combineren ontstaat informatiepotentieel

Ondanks een toename aan beschikbare datasets, is de impact van open data tot nu toe beperkt. Gemeenten bijvoorbeeld leveren vaak sporadisch datasets aan, omdat een onderwerp op een bepaald moment een 'hype' is. Maar de aandachtsgolven zakt daarna vaak weer snel, waardoor de gegevens maar kort actueel zijn en maar tijdelijk online staan.

Versillende start-ups werken nu met gemeenten samen om het aanbod voor bedrijven aantrekkelijker en betrouwbaarder te maken.

Privacy in gevaar door opkomst van open data

Van onze redacteur
Amsterdam

Met de komst van grote datasets wordt het steeds gemakkelijker informatie te combineren. Met name persoonsgegevens worden daardoor kwetsbaar. Volgens Axel Armbak, onderzoeker Cybersecurity en Informatierecht bij de Universiteit van Amsterdam, schiet de naleving van de Wet Bescherming Persoonsgegevens ernstig tekort en wordt de huidige privacywetgeving massaal overtreden.

De privacywetgeving, die stamt uit 1992, is volgens Armbak door de technologie ingehaald. Overheden en bedrijven bieden hun open data geanonimiseerd aan, maar dat is volgens de onderzoeker allang niet meer toereikend. 'Het concept anonimiseren staat uit de jaren 90, toen datasets veel simpeler waren. Nu kun je met een paar datapunten vaak al iemand achterhalen.'

De kans dat persoonsgegevens uit data achterhaald worden, wordt dus steeds groter. Dit effect wordt bovendien versterkt door de mogelijkheid verschillende datasets te combineren. Hoewel het volgens Armbak mogelijk is te anonimiseren, hebben de gegevens volgens hem in zo'n geval nog maar weinig commerciële waarde.

In het Europees Parlement wordt momenteel gedebatteerd over nieuwe privacywetgeving. Het einde is nog niet in zicht, maar nu al is een recordhoeveelheid aan amendementen ingediend. Voor dataleveranciers kan het debat grote gevolgen hebben, want de aansprakelijkheid voor de bescherming van de persoonsgegevens ligt bij hen.

Ondanks de mogelijke gevaren, ziet Armbak ook positieve toepassingen van open data. 'De website 9292.nl is ook op open data gebaseerd. En het is heel fijn om per bushalte de real-time vertrektijden te kunnen vinden.'

1992

De huidige privacywetgeving is na 22 jaar grotendeels achterhaald.

Bedrijven hebben vaak geen idee tot hoeveel informatie zij toegang hebben'



Laurens Ruster van Data Kitchen

TOEGANKELIJK

De intermediair

Bedrijf: Civity

Sinds: Juni 2013

Arlen Hof fungeert met zijn bedrijf Civity als intermediair tussen vraag en aanbod. In zijn 20 jaar bij de overheid heeft hij de opkomst van open data nauwlettend gevolgd en van dichtbij de moeilijkheden voor dataleveranciers gezien.

Hij constateerde dat datasets, met name afkomstig van gemeenten, vaak slechts gedeeltelijk online komen en nauwelijks worden bijgehouden, met als gevolg dat de data minder aantrekkelijk zijn voor bedrijven.

Civity biedt meerdere producten en diensten aan die het openstellen van data vergemakkelijken. Dataleveranciers kunnen ervoor kiezen de data op hun eigen website aan te bieden.

Civity helpt om de gegevens op te schonen en openbaar te maken. Door een platform te bieden voor een uniforme toegang tot de datasets, worden de data voor app-bouwers toegankelijker.

Daarnaast helpt Civity ook bij het veilig aanbieden van datasets. Omdat dataleveranciers bang zijn aansprakelijk te worden gesteld voor het vrijgeven van privacygevoelige gegevens, helpt Civity de leveranciers door toezicht te houden en door toegangsrechten voor bezoekers van de website op te zetten.

Aan de vraagkant probeert Civity open data aantrekkelijker te maken, door beschikbare open data van overheden met elkaar te verbinden en

aan te bieden. Met deze 'linked open data' ontstaat een netwerk van datasets. Ook helpt het bedrijf geïnteresseerden de benodigde data te vinden of, zo nodig, aan te vragen.



OPTIMALE LOCATIES

De vastgoedadviseur

Bedrijf: Hemm

Sinds: september 2013

Ellen-Jane van Straten heeft met haar partner Martijn Kogelman ruim twintig jaar ervaring in de advieswereld. Nu adviseren ze gemeenten en andere maatschappelijke instellingen over hun vastgoedportfolio vanuit hun eigen bedrijf Hemm. Door het bestaande vastgoed te inventariseren en in een model te combineren met beschikbare open data, willen ze instellingen helpen keuzes voor de toekomst te maken.

In het model Citable wordt alle vastgoed op een kaart weergegeven. Gebruikers kunnen zelf lagen met informatie toevoegen of verwijderen. Deze informatie is zowel afkomstig uit open data als uit gegevens van de gemeenten

zelf. Volgens Van Straten weten bedrijven niet welke schat aan informatie ze in huis hebben. Als een gemeente nu een school wil bouwen, kan ze op basis van eigen gegevens en die van het Centraal Bureau voor de Statistiek in het model kijken waar de meeste kinderen wonen.

Het model wordt continu bijgewerkt, waardoor het niet nodig is elke vier jaar een nieuw plan te maken. Samen met de gemeenten optimaliseert Hemm het vastgoedportfolio en brengt advies uit over de beste toekomststrategie.

De start-up wil snel uitbreiden, want ook voor woningcorporaties of andere houders van grootvastgoed biedt het model uitkomst. Zo kan op basis van filters in een oogwenk de optimale locatie van een nieuw gebouw worden bepaald of naar alternatieven worden gezocht.



voor maatschappelijk vastgoed

INZICHT GEVEN

De ontwikkelaar

Bedrijf: 2CoolMonkeys

Sinds: 2008, open data sinds 2011

De app-bouwers van 2CoolMonkeys hebben al veel apps gemaakt, maar met de komst van open data in 2011 begon het balletje volgens consultant Reind van Olst echt te rollen. 'De eerste app die we met open data maakten, was voor de Rijksoverheid. De burger moest daarmee beter inzicht krijgen in de taken en middelen van de overheid. Maar uiteindelijk waren het juist de medewerkers van de overheid die de app gebruikten omdat die veel overzichtelijker was dan het systeem zelf. Toen kregen we ineens meer opdrachten om apps te maken.'

Langzamerhand bouwen Van Olst en zijn collega's een stevige verzameling apps. De achterlig-

gende open data die ze hiervoor gebruiken, publiceren ze via hun eigen platform Robin. Nu, drie jaar later, hebben ze dit platform opengesteld voor andere app-bouwers. Door de data toegankelijk aan te bieden, willen ook derden stimuleren om apps te ontwikkelen.

Een belangrijke vraag vanuit de markt op dit moment is goede kope toegang tot de gegevens van het Kadaster. In samenwerking met twee andere bedrijven wil 2CoolMonkeys de gegevens direct op platform Robin publiceren, in combinatie met open data. Voor een klein bedrag krijgen bedrijven toegang tot de gegevens en blijft het Kadaster aan de

data verdienen. Het kost nu naar schatting zo'n € 4,8 mln om toegang tot alle gegevens te krijgen. Met deze toepassing neemt 2CoolMonkeys dus een grote belemmering weg.



ROBIN