

De waarde van open data

Keuzes en effecten van open-datastrategieën voor publieke organisaties

drs. Tom Kronenburg
ir. Ton Monasso
Ellen Boschker MSc
prof. dr. Marcel Thaens

De waarde van open data

Keuzes en effecten van open-datastrategieën voor publieke organisaties

In opdracht van:
Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Auteurs:
drs. Tom Kronenburg
ir. Ton Monasso (projectleider)
Ellen Boschker MSc
prof. dr. Marcel Thaens

26 april 2012

© Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Het ministerie stelt dit werk beschikbaar onder een Creative Commons Naamsvermelding 3.0 Nederland-licentie. Bezoek <http://creativecommons.org/licenses/by/3.0/nl/> om een kopie te zien van de licentie.



Aangemeld bij de Nationale Onderzoeksdatabank onder nummer OND1345708.

Leeswijzer

Het rapport dat u nu in handen heeft, bestaat uit verschillende delen. Natuurlijk raden wij u aan om het rapport in zijn geheel te lezen. Maar, afhankelijk van uw achtergrond en interesse kunnen specifieke hoofdstukken uw aandacht te hebben.

- Wilt u of uw organisatie zelf aan de slag met open data? In hoofdstuk 4 leest u uit welke onderdelen een open-datastrategie kan bestaan en welke effecten daarmee bereikt kunnen worden; in hoofdstuk 6 ziet u welke businessmodellen voor open data u kunt gebruiken en in de eerste bijlage kunt u nalezen hoe andere organisaties met open data aan de slag zijn gegaan.
- Wilt u anderen stimuleren aan de slag te gaan met open data? Lees dan ook hoofdstuk 7 (discussie en aanbevelingen) extra zorgvuldig.
- Bent u benieuwd hoe dit onderzoek is uitgevoerd? U leest de onderzoeksaanpak terug in hoofdstuk 1 en vindt de belangrijkste analytische tussenstappen in bijlage IV.

Bij dit onderzoek hoort ook een poster. Hierin zijn de belangrijkste uitkomsten van dit onderzoek op aantrekkelijke wijze gepresenteerd.

Inhoudsopgave

Leeswijzer	3
Inhoudsopgave.....	4
1 Inleiding.....	6
1.1 Aanleiding.....	6
1.2 Onderzoeksopdracht.....	7
1.3 Inkadering onderzoek.....	8
1.4 Onderzoeksaanpak.....	9
1.5 Opbouw van het rapport.....	10
2 Eerder onderzoek.....	11
2.1 Macro-economisch onderzoek.....	11
2.2 Onderzoek naar prijsstrategieën.....	11
2.3 Marktversturende effecten.....	12
2.4 Implicaties van eerder onderzoek.....	14
3 Introductie op de cases.....	15
3.1 Het BIG-register van het CIBG.....	15
3.2 Erfgoed informatie van de RCE.....	15
3.3 Luchtkwaliteit- en emissiedata bij het RIVM.....	16
3.4 De wettenbank van het ministerie van BZK.....	16
3.5 Diverse datasets van de gemeente Rotterdam.....	16
3.6 Geo-informatie van de provincie Noord-Brabant.....	17
3.7 Scholeninformatie van DUO.....	17
4 Typen keuzes en effecten.....	19
4.1 Open-datastrategie.....	20
4.2 Kenmerken organisatie en omgeving.....	22
4.3 Kenmerken data.....	22
4.4 Effecten.....	22
4.4.1 Effecten voor de individuele organisatie.....	23
4.4.2 Effecten voor de overheid als geheel.....	23
4.4.3 Effecten voor de samenleving.....	23
5 Keuzes en effecten in de praktijk.....	25
5.1 Elementen en relaties open-datastrategie en kenmerken organisatie, omgeving en data.....	25
5.1.1 Aanleiding, doel en investeringsbereidheid.....	25
5.1.2 Gebruik en gebruikers.....	26
5.1.3 Vorm en presentatie van de data.....	27
5.2 Effecten van de open-datastrategie.....	29
5.2.1 Financiële effecten.....	29
5.2.2 Niet-financiële effecten.....	34
6 Conclusie: businessmodellen voor open data.....	36
6.1 Radicale innovatie: nieuwe mogelijkheden.....	36
6.2 Incrementele innovatie: verbetering van het bestaande.....	37
6.3 Bouwen op bestaande data.....	38
6.4 Hergebruik naast eigen aanbod.....	38

7	Discussie en aanbevelingen	40
7.1	Implicaties.....	40
7.2	Reflectie.....	41
	Literatuurlijst.....	44
	Bijlage I. Casebeschrijvingen	46
	Het BIG-register van het CIBG	46
	Erfgoed informatie van de RCE	51
	Luchtkwaliteit- en emissiedata bij het RIVM	58
	De wettenbank van het ministerie van BZK.....	62
	Scholeninformatie van DUO	66
	Diverse datasets van de gemeente Rotterdam	75
	Geo-informatie van de provincie Noord-Brabant	82
	Bijlage II. Interviewgide	88
	Bijlage III. Begeleiding van het onderzoek	91
	Bijlage IV. Elementsgewijze analyse van de cases	92
	Open-datastrategie	92
	Kenmerken organisatie en omgeving	100
	Kenmerken van de data	103
	Effecten	106

1 Inleiding

1.1 Aanleiding

De laatste jaren is er in Europa en Noord-Amerika toenemende aandacht voor open data. Ook in Nederland ontplooiën steeds meer overheidsorganisaties open-data-initiatieven en stellen actief data beschikbaar voor hergebruik. De toenemende interesse is ontstaan door de voordelen van beschikbaarheid van open data. Bedrijven kunnen met het ontwikkelen van nieuwe producten bijdragen aan economische groei (innovatie, economische groei). Burgers krijgen nieuwe mogelijkheden om de overheid te controleren (transparantie). En innovatieve bedrijven kunnen een deel van de overheidsdienstverlening vervangen door een beter aanbod (service-innovatie).

Het kabinet moedigt het beschikbaar stellen van open data aan. In de Kamerbrief over hergebruik en open data van 30 mei jl. geeft de minister van Binnenlandse Zaken en Koninkrijksrelaties aan dat overheden moeten worden aangespoord om de door hen beheerde data ter beschikking te stellen (Minister van BZK, 2011). Volgens de minister kan onderzoek naar de kosten en baten van beleid inzake hergebruik op het niveau van individuele overheidsorganisaties daarbij helpen. Dergelijk onderzoek moet overheidsorganisaties inzicht verschaffen in de vraag in hoeverre en onder welke voorwaarden het loont om te investeren in een open-datastrategie. Omdat enkele overheidsorganisaties de laatste jaren al flinke stappen hebben gezet in het openen van hun data, is er een interessante empirische basis voor een onderzoek beschikbaar. Het voorliggende rapport is het resultaat van het onderzoek dat de minister in de Kamerbrief voorstelt.

De Kamerbrief beschrijft open data als bronnen van onbewerkte overheidsinformatie

- die openbaar zijn;
- waar geen auteursrechten of andere rechten van derden op berusten;
- die bekostigd zijn uit publieke middelen, beschikbaar gesteld voor de uitvoering van die taak;
- die bij voorkeur voldoen aan 'open standaarden' (geen barrières voor het gebruik door ICT-gebruikers of door ICT-aanbieders) en
- die bij voorkeur computerleesbaar zijn, zodat zoekmachines informatie in documenten kunnen vinden.

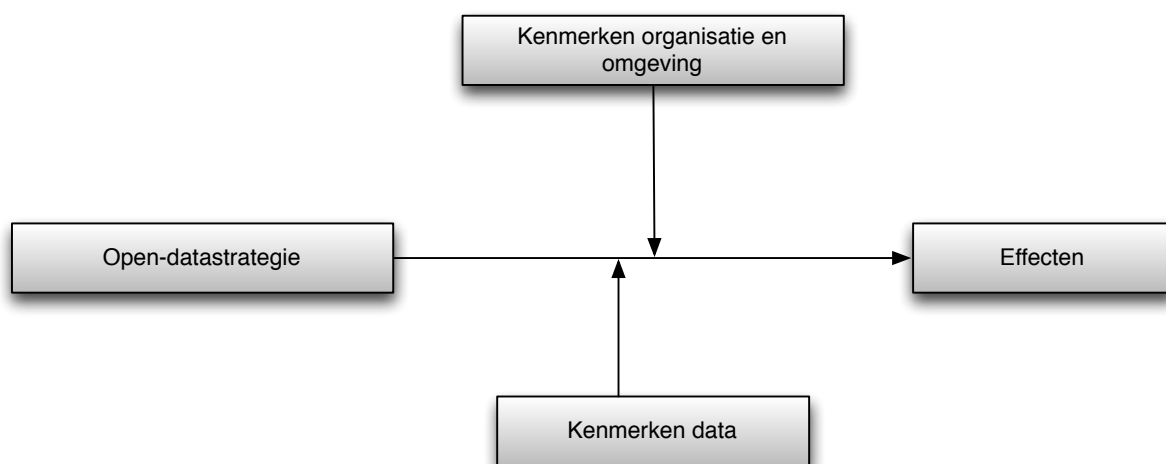
In dit rapport hanteren we dezelfde definitie.

1.2 Onderzoeksopdracht

Het doel van dit onderzoek is tweeledig. We willen ten eerste inzicht verkrijgen in de keuzes die overheidsorganisaties hebben wanneer zij open data beschikbaar (willen) stellen en de effecten van deze keuzes. De keuzes die organisaties maken, zijn onderdelen van een (al dan niet expliciete) open-datastrategie. Er kunnen zowel financiële als niet-financiële effecten worden onderscheiden. En, de effecten kunnen de individuele organisatie betreffen, maar ook de overheid als geheel en de maatschappij. We spreken ook wel over een open-databusinessmodel: een intern consistente set van strategische keuzes ten aanzien van open data en de wijze waarop deze binnen de organisatie leidt tot waardecreatie. Met dit inzicht wordt ook het tweede doel van het onderzoek bereikt. We willen organisaties laten zien in hoeverre het loont om met open data aan de slag te gaan en welke mogelijkheden zij daartoe hebben.

Tijdens het onderzoek is gebleken dat het moeilijk is om de kosten en baten van open data te kwantificeren en tussen organisaties vergelijkbaar te maken. De belangrijkste reden daarvoor is dat open data zelden als een geïsoleerde en geëxpliciteerde investeringsbeslissing wordt bekeken, maar vaak tegelijk met andere ontwikkelingen wordt opgepakt. Daarmee is het een onderdeel of een uitvloeisel van een bredere organisatie- en informatiestrategie.

Hieronder hebben we de relatie tussen de open-datastrategie en de effecten daarvan in een model weergegeven. Verschillende variabelen beïnvloeden deze relatie. Die variabelen hebben te maken met de organisatie en de omgeving van de organisatie waarin de open-datastrategie wordt uitgevoerd. Denk bijvoorbeeld aan de discretionaire ruimte. Als een organisatie veel discretionaire ruimte heeft, is er meer vrijheid voor de open-datastrategie dan wanneer de discretionaire ruimte beperkt is. Ook zal de data op zichzelf invloed hebben op de open-datastrategie en zijn effecten. Als de data bijvoorbeeld veel privacygevoelige informatie bevatten, dan kunnen deze niet zomaar naar buiten worden gebracht.



Dit onderzoek beantwoordt de volgende hoofd- en deelvragen:

Hoofdvraag

Wat zijn businessmodellen voor open data, die individuele overheidsorganisaties kunnen gebruiken voor het beschikbaar stellen ervan?

Deelvragen

1. Op welke aspecten maken individuele overheidsorganisaties keuzes als zij hun data open beschikbaar stellen (open-datastrategie)?
2. Welke effecten van het open beschikbaar stellen van data kunnen worden onderscheiden?
3. Welke kenmerken van de organisatie, haar omgeving en van de data zijn van invloed op deze keuzes en effecten?
4. Wat zijn de keuzes, de kenmerken van de organisatie en omgeving, de kenmerken van de data en de effecten in de onderzochte cases?
5. Wat zijn de relaties tussen de keuzes, de kenmerken van de organisatie en omgeving, de kenmerken van de data en de effecten (zowel kwantitatief als kwalitatief en zowel op het niveau van de individuele organisatie, het niveau van de overheid als geheel als op het niveau van de maatschappij)?

1.3 Inkadering onderzoek

Het onderzoek stelt individuele organisaties centraal...

Het onderzoek richt zich op individuele overheidsorganisaties. Daarmee onderscheidt het zich van studies die gericht zijn op macro-economische en andere maatschappelijke effecten van open data.

...die niet 'leven' van open data...

Het onderzoek richt zich op organisaties die niet voor een (substantieel) deel van hun inkomsten afhankelijk zijn van de omzet door verkopen van informatie. Tot nu toe is het meeste onderzoek en aandacht gericht geweest op organisaties waar dat wel het geval is.

...en mondt uit in businessmodellen die door organisaties gebruikt kunnen worden

In dit onderzoek wordt gesproken over businessmodellen en niet over businesscases. De doelstelling van een businesscase is het onderbouwen of legitimeren van een investeringsbeslissing. Een businessmodel is een beschrijving van de wijze waarop een organisatie waarde genereert, en beschrijft de combinatie van onder andere de propositie, de interactie met de omgeving en het financieringsmodel. Een businessmodel is veelal ook geen expliciet en vooraf opgesteld product, zoals een businesscase vaak is, maar eerder een impliciet model.

1.4 Onderzoeksaanpak

Om de hierboven gestelde onderzoeksvragen te beantwoorden, hebben we de open-datastrategie van zeven verschillende Nederlandse publieke organisaties onderzocht. De selectie hiervan heeft plaatsgevonden in overleg met de begeleidingscommissie van het onderzoek. De zeven cases zijn:

- Het BIG-register van het Centraal Informatiepunt Beroepen Gezondheidszorg (CIBG);
- Erfgoed informatie van de Rijksdienst voor het Cultureel Erfgoed (RCE);
- Luchtkwaliteit- en emissiedata bij het Rijksinstituut voor Volksgezondheid en Milieu (RIVM);
- De wettenbank van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties;
- Diverse gemeentelijke datasets van de gemeente Rotterdam;
- Geo-informatie van de provincie Noord-Brabant;
- Scholeninformatie van de Dienst Uitvoering Onderwijs (DUO).

Elke case bestaat uit de combinatie van een organisatie en een of meerdere van haar datasets. Bij de caseselectie en de eerste contacten bleek dat het vaak niet mogelijk is om van de open-datastrategie van de organisatie als geheel te spreken, maar dat de strategie per dataset verschilt. Een organisatie kan meerdere datasets beheren, maar evengoed is het mogelijk dat een dataset de organisatiegrens overschrijdt. Dat is bijvoorbeeld het geval bij de emissieregistratie van het RIVM, die de gegevens van verschillende organisaties combineert en ontsluit.

Van deze zeven cases is een casebeschrijving gemaakt, te vinden in bijlage I. De casebeschrijvingen zijn gemaakt op basis van interviews met het management (CIO), uitvoerende medewerkers (functioneel beheer, beleid) en afnemers van de data (hergebruikers). Daarnaast zijn ook projectplannen, begrotingen en andere relevante documentatie onderzocht. Alle casebeschrijvingen zijn gevalideerd door de geïnterviewde personen.

Er is nauwelijks eerder onderzoek gedaan naar de open-datastrategie van individuele organisaties. Daarom bleek het niet mogelijk om een bestaande theorie toe te passen en te toetsen. We hebben gekozen voor een inductieve werkwijze, waarbij we vanuit de bijzondere situaties (de cases) algemene inzichten (de businessmodellen) willen afleiden. Bestaand onderzoek hebben we gebruikt om relevante variabelen te onderscheiden, die wij in de casestudy's hebben onderzocht. In drie stappen hebben we generieke inzichten afgeleid:

- We hebben de beschrijvingen aan de hand van een lijst van elementen 'gecodeerd' (zie bijlage IV). Daardoor konden we de cases eenvoudiger vergelijken.
- We hebben geanalyseerd welke elementen met elkaar samenhangen en welke lessen hieruit te trekken zijn.
- Ten slotte hebben we hieruit twee generieke businessmodellen afgeleid.

Over de caseselectie, relevante onderwerpen voor de interviews (zie bijlage II voor de gebruikte interviewguide) en een eerste duiding van het casemateriaal is in twee bijeenkomsten gesproken met experts (zie bijlage III voor een overzicht van de geraadpleegde experts en de samenstelling van de begeleidingscommissie). Ook is er regelmatig contact geweest met de begeleidingscommissie over de opzet en de rapportage van het onderzoek.

1.5 Opbouw van het rapport

In hoofdstuk 2 bespreken we relevante inzichten uit eerder onderzoek. In hoofdstuk 3 introduceren we de onderzochte cases. Dit dient als achtergrond voor de hoofdstukken 4 en 5 waarin we antwoord geven op de deelvragen van het onderzoek. Hoofdstuk 6 bevat de conclusies en de beantwoording van de hoofdvraag van het onderzoek. In hoofdstuk 7 bespreken we de implicaties, reflecties en aanbevelingen van het onderzoek.

2 Eerder onderzoek

Hergebruik van digitale overheidsinformatie is vrij recent op de politieke agenda gekomen. Sinds eind jaren negentig is er dan ook pas academisch onderzoek naar gedaan. De grootste en bekendste studies op dit gebied behandelen de macro-economische impact van hergebruik van open data. Een tweede populair onderzoeksthema zijn de prijsstrategieën die door individuele instellingen gehanteerd kunnen worden. Daarnaast is er ook aandacht voor de versturende effecten van open data op bestaande markten.

2.1 Macro-economisch onderzoek

Onderzoek door PIRA (2000) gaf voor het eerst aanleiding om te veronderstellen dat hergebruik van de grondstof overheidsinformatie, net als in de Verenigde Staten, ook in de Europese Unie (EU) een grote economische waarde kan hebben. Het MEPSIR-onderzoek (Dekkers et al., 2006) bevestigde dat de waarde van 'Public Sector Information' (PSI)-hergebruik in de hele EU rond de € 27 miljard op jaarbasis zou bedragen. Te Velde (2009) beargumenteerde dat dit bedrag vijf tot negen keer te hoog zou zijn, terwijl een recent comparatief onderzoek van Vickery (2011) de uiteindelijke effecten van open data op de economie van de EU27 tussen de € 140 en 200 miljard raamt. De verschillen in gevonden opbrengsten worden vooral verklaard door verschillende meetmethodieken. De definitie van het begrip waarde (omzet, winst, impact op het BNP, etc.) verschilt wel tussen de studies. Al deze macro-economische studies hebben wel met elkaar gemeen dat zij aangeven dat hergebruik van overheidsinformatie een positieve impact kan hebben op de economie van de lidstaten van de EU.

2.2 Onderzoek naar prijsstrategieën

In een aantal zogenaamde 'pricing-onderzoeken' is onderzocht op welke manier individuele overheidsorganisaties ervoor kunnen zorgen dat de economische impact van hergebruik van de door hen beheerde data zo groot mogelijk is. Het belangrijkste element van onderzoek is daarbij de prijs die voor de data gevraagd wordt door een overheidsorganisatie. Een aantal bekende onderzoeken met deze thematiek zijn van Pollock (2009), Uhler (2009) en het recente POPSIS-onderzoek van Deloitte (Vries et al., 2011). De conclusie van al deze onderzoeken is dat de economische effecten het grootst zijn wanneer informatie gratis ('zero costing') of alleen tegen verstrekingskosten ('marginal costing') wordt aangeboden.

Er bestaat echter een obstakel voor het gratis of tegen marginale kosten verstrekken van informatie. Namelijk, wanneer publieke organisaties het geld dat binnenkomt met het ter beschikking stellen van data nodig hebben om het verzamelen en het beheren van de data te subsidiëren. Het hanteren van een 'zero costing'- of 'marginal costing'-beleid is voor deze

organisaties geen optie, tenzij er andere bronnen van inkomsten kunnen worden aangeboord. POPSIS zegt hierover:

“A large majority of Public Sector Bodies (PSBs) interviewed do not seem to have fundamental objections to lowering charges. Yet, PSBs that rely on sales revenues from PSI and their own value-added products appear to be stuck in a situation of deadlock: although they are sympathetic to lowering charges and allowing more data re-use, their dependency on sales revenues compels them to protect their current revenue streams when there is no other sustainable alternative income stream available.

Such an alternative income stream can often only be provided by the Treasury, since the benefits from lowered charges are often concentrated in the form of increased tax gains. Thus, the power to enable change does not necessarily lie with the Ministry willing to support the move, let alone with the PSB concerned.” (Vries et al., 2011. p. 6-7)

Het collectieve belang (economische groei) weegt voor een individuele overheidsorganisatie, die voor een belangrijk deel van zijn inkomsten afhankelijk is van dataverstrekking, dus vaak niet op tegen de financiële nadelen van gratis of marginale verstrekking.

Vrij recent is er ook in Australië onderzoek gedaan naar de open-databusinesscase voor individuele organisaties. Houghton (2011) schrijft onder andere dat het moeilijk is om resultaten van onderzoek bij een individuele organisatie te generaliseren. De omstandigheden van een organisatie zijn in veel gevallen te specifiek. Hij concludeert echter wel:

“It is clear from the case studies presented that even the subset of benefits that can be measured outweigh the costs of making PSI more freely and openly available. It is also clear that it is not simply about access prices, but also about the transaction costs involved. Standardised and unrestrictive licensing, such as Creative Commons, and data standards are crucial in enabling access that is truly open (i.e. free, immediate and unrestricted).” (Houghton, 2011, p. 4)

Naast deze ‘pricing-onderzoeken’ is er weinig onderzoek gedaan naar individuele beweegredenen van organisaties om open data beschikbaar te stellen.

2.3 Marktversturende effecten

Het beschikbaar stellen van informatie heeft impact op de economie als geheel. Maar individuele (private) bedrijven of de markten waarin zij handelen, kunnen ook positieve en

negatieve effecten ondervinden van de beschikbaarheid van open data. Voor dit onderzoek zijn twee typen effecten relevant (afgeleid uit Dekkers et al., 2006, paragraaf 13.3):

- Overheden concurreren soms met bedrijven wanneer zij niet alleen de informatie zelf ter beschikking stellen, maar die informatie ook op een consumentvriendelijke wijze presenteren. Het bezwaar hierbij is dat de overheid het marktaandeel van commerciële partijen verkleint, door een met belastinggeld gefinancierd alternatief. De RDW werd bijvoorbeeld bekritiseerd om het beschikbaar stellen van zijn gratis kentekencheck-app. De kritiek kwam van leveranciers van gratis (in combinatie met advertenties) en betaalde kenteken-apps op websites, via sms en voor iPhone- en Android-smartphones. De RDW vond het publieksvriendelijk verstrekken van gratis kentekeninformatie echter een belangrijke taak en heeft de functie op zijn website gehandhaafd.
- De overheid kan de waardeketen ook veranderen door data vrij te geven, die ook al (in vergelijkbare vorm) door commerciële partijen verzameld en verkocht wordt. De marge, prijs en waardeketen van de bestaande markt kan daardoor sterk veranderen. De recente rechtszaak van Falkplan over het vrijgeven van bijvoorbeeld het Nationaal Wegenbestand (NWB) illustreert dit. Falkplan constateert dat haar klanten weglopen omdat ze verwachten dat ze ofwel de NWB-dataset direct kunnen gebruiken, ofwel dat ze dezelfde informatie die ze eerder van Falkplan kochten, elders veel goedkoper kunnen kopen (Rechtspraak, 2012). Ook wanneer overheden gewend waren om in samenwerking met één of enkele bedrijven informatie te verspreiden (monopolie), heeft openbaar beschikbaar stellen van diezelfde data soms een verstoring effect. Het beschikbaar stellen van de wettenbank heeft bijvoorbeeld ingrijpende effecten gehad voor de verkoop van wettenbundels e.d. door de grote juridische uitgevers (Boom, Kluwer, Sdu).

Het begrip 'publieke taak' staat vaak centraal in de discussie over markteffecten. Welke taken heeft de overheid (als geheel of als individuele publieke instelling) en wanneer voert zij een markttaak uit? Als een overheidsinstelling niet alleen informatie ter beschikking stelt, maar zelf ook een website of een applicatie voor smartphones maakt, is dit dan haar taak? Of concurreert ze op oneigenlijke gronden (want gesubsidieerd met belastinggeld) met ondernemers die dezelfde dienst aanbieden? Lundqvist (2011) geeft voor deze vragen zeer uitgebreide richtlijnen, gebaseerd op de PSI Directive en de Zweedse wet. Een interessant voorbeeld hiervan is dat organisaties, die hun eigen data hergebruiken, zichzelf niet mogen bevoordelen. Met andere woorden, de gebruiksvoorwaarden voor de eigen organisatie moeten hetzelfde zijn als die voor andere organisaties. De bevindingen van Lundqvist zijn overigens ook goed bruikbaar in de Nederlandse context, waar de regels voor concurrentie tussen overheden en marktpartijen in de wet 'markt en overheid' zijn vastgelegd.

2.4 Implicaties van eerder onderzoek

Dit onderzoek vult het bestaand onderzoek aan, en is geen herhaling van dezelfde thematiek in een andere sector. Het eerdere onderzoek maakt duidelijk dat de scheidslijn tussen marktactiviteiten en publieke taak moet worden meegenomen in de analyse voor dit onderzoek. Het gaat dan vooral om de vraag of publieke organisaties eigen presentatievoorzieningen ontwikkelen ('value-added services'), of dergelijke toepassingen aan private partijen overlaten, zodra de markt in staat is een volwaardig alternatief te bieden. Verwacht kan worden dat het al dan niet ontwikkelen van presentatievoorzieningen de financiële effecten van de open-datastrategie zal beïnvloeden.

3 Introductie op de cases

In dit hoofdstuk introduceren we kort de zeven cases, als achtergrond voor de volgende hoofdstukken.

3.1 Het BIG-register van het CIBG

Het CIBG beheert namens het ministerie van Volksgezondheid, Welzijn en Sport (VWS) meerdere registers. Eén van deze registers is het register Beroepen in de Individuele Gezondheidszorg (BIG-register). Dit houdt bij welke zorgverleners bevoegd zijn een medisch beroep uit te oefenen. Een BIG-registratie is bij wet verplicht voor een zorgverlener om in de zorg te mogen werken. Elke ingeschreven zorgverlener heeft een eigen BIG-nummer. Het CIBG stelt het register via een website en een webservice beschikbaar. Burgers, patiëntenverenigingen en bedrijven kunnen deze informatie gebruiken. Burgers om te controleren of de zorgverlener staat ingeschreven, patiëntenverenigingen om patiënten voor te lichten en bedrijven om ingeschrevenen bijvoorbeeld korting te verstrekken op boeken of zorgverzekeringen. Het CIBG houdt - naast een overzicht van BIG-geregistreerden - ook een 'zwarte lijst' bij van zorgverleners die (tijdelijk) niet mogen werken in de zorg. Deze lijst is op de website te vinden.

3.2 Erfgoedinformatie van de RCE

De RCE valt onder het ministerie van Onderwijs, Cultuur en Wetenschap (OCW). Deze rijksdienst is belast met de bescherming van het Nederlands cultureel erfgoed, bestaande uit archeologie, monumenten, landschap en roerend goed. De RCE bewaakt in een netwerk samen met andere organisaties het beheer van het Nederlands erfgoed. Andere organisaties zijn bijvoorbeeld gemeenten en provincies, restauratiebedrijven en eigenaren van erfgoed zoals kasteel eigenaren.

De RCE heeft veel informatie over het erfgoed, variërend van boeken, tijdschriften, foto's, tekeningen en registers. Vanuit de rol als facilitator van het netwerk heeft de RCE ervoor gekozen de ontsluiting van data te verbreden. Onder andere door het ontsluiten van halffabricaten, de thesaurus en meta-datastandaarden vergroot de RCE zijn impact op de gehele erfgoedsector. De RCE ontsluit dus ook meer dan alleen de data waartoe de organisatie wettelijk verplicht is. In tegenstelling tot de in de open-datawereld gebruikelijke filosofie dat het beste 'ruwe en volledig onbewerkte data' geleverd kan worden, en passend bij de bedrijfsstrategie biedt de RCE geen 'ruwe' data aan. De ruwe data worden door middel van een informatie-verbindingsslaag ontsloten. Applicaties en gebruikers krijgen daardoor alleen herkenbare eenheden van gegevens gerepresenteerd. Gebruikers krijgen bijvoorbeeld de complete informatieset van een monument (inclusief locatiegegevens, redengevende omschrijving, foto's, etc.).

Hergebruikers en ketenpartners vinden de data van de RCE onder meer via de beeldbank en Wikipedia. Hergebruikers zijn eigenaren van erfgoed, intermediairs en andere erfgoedbeherende organisaties.

3.3 Luchtkwaliteit- en emissiedata bij het RIVM

Het RIVM valt onder het ministerie van VWS. Het ondersteunt overheden op het gebied van volksgezondheid en milieu. Hiervoor genereert en behoudt het kennis over gezondheid en milieu. Via meetstations verzamelt het RIVM data over onder andere luchtkwaliteit en emissies. De gegenereerde data worden automatisch doorgestuurd naar de database van het RIVM. Beide datasets worden via websites ter beschikking gesteld aan overheden, professionele gebruikers en burgers. Overheden gebruiken de milieudata om beleid te maken, te monitoren of te evalueren. Het RIVM rapporteert ook aan de Europese Commissie zodat vastgesteld kan worden in hoeverre Europese normen worden overschreden. Ingenieursbureaus en milieu- en natuurorganisaties zijn de belangrijkste professionele gebruikers. Ook burgers kunnen de data gebruiken, maar hebben vaak minder kennis om de data te duiden dan professionele gebruikers.

3.4 De wettenbank van het ministerie van BZK

De wettenbank bevat een dataset met geconsolideerde teksten van nationale wet- en regelgeving. Het gaat bijvoorbeeld om wetten, Algemene Maatregelen van Bestuur en ministeriële regelingen. Deze dataset wordt via een website en een webservice ter beschikking gesteld. Voor het ministerie van BZK - de opdrachtgever - is de wettenbank om twee redenen van belang. De wettenbank maakt nationale wet- en regelgeving transparant voor burgers en borgt op die manier democratische waarden als transparantie in het openbaar bestuur. Daarnaast bevordert de wettenbank samenwerking binnen de Rijksdienst. Elk ministerie is afnemer van dezelfde centrale wettenbank en hoeft daarvoor geen aparte abonnementen met particuliere uitgevers af te sluiten. Dit verhoogt de interne efficiëntie Rijksoverheid. Burgers, ambtenaren en juridische professionals zijn gebruikers van de wettenbank.

3.5 Diverse datasets van de gemeente Rotterdam

De gemeente Rotterdam stelt verschillende gemeentelijke datasets voor hergebruik beschikbaar. De achtergrond hiervan ligt in een pilotproject geïnitieerd door de Hogeschool Rotterdam en enkele Rotterdamse bedrijven. Zeven verschillende gemeentelijke diensten hebben daarvoor data vrijgegeven, onder andere de data van de bibliotheek, het gemeentearchief en de GGD. Deze data worden via een website en een webservice van de Rotterdam Open Data Store (RODS) ontsloten. De data worden voornamelijk gebruikt door studenten van de hogeschool en door bedrijven. In de toekomst zal meer data van de dienst Stadsontwikkeling worden ontsloten, zo heeft het College van B&W in september 2011

besloten. Alle data die door de dienst Stadsontwikkeling worden beheerd en die geen persoonsgegevens bevat, worden zonder gebruiksbependingen openbaar gemaakt. De gemeente wil hiermee innovatie en economie in de stad stimuleren. Eind november 2011 waren er nog geen nieuw toegevoegde datasets van de dienst Stadsontwikkeling beschikbaar in de RODS.

3.6 Geo-informatie van de provincie Noord-Brabant

De provincie Noord-Brabant werkt met andere Nederlandse provincies samen in het Provinciaal Georegister (PGR). Hierin worden minimaal tien datasets ontsloten die de INSPIRE-richtlijn van de EU vereist. De sets zijn beschikbaar via de website en een webservice. De provincie Noord-Brabant zelf ontsluit ongeveer 50 datasets, onder meer ten aanzien van ruimtelijk beleid, cultuur en waterhuishouding. De provincie ontsluit alleen informatie waarvan de provincie zelf bronhouder is. Noord-Brabant heeft ook een eigen website met een presentatievoorziening voor de data. Via deze site kunnen geen gegevens worden gedownload, dat gebeurt via het PGR.

Met het openstellen van data wil de provincie transparantie en openheid bevorderen, efficiëntie vergroten en innovatie stimuleren. De efficiëntie wordt vergroot doordat data op een plaats wordt aangeboden en daar (vaker) gedownload kan worden. De provincie wil niet zelf ontwikkelen, maar ontwikkeling en innovatie in de markt stimuleren door het openstellen van data. Hergebruik vindt tot nu toe plaats door andere provincies en organisaties als het RIVM en het Planbureau voor de Leefomgeving.

3.7 Scholeninformatie van DUO

De Dienst Uitvoering Onderwijs (DUO) is een uitvoeringsorganisatie van het ministerie van OCW. DUO beheert gegevens over het primair onderwijs, het voortgezet onderwijs, beroepsonderwijs en volwasseneducatie en het hoger onderwijs. De gegevens omvatten verschillende categorieën of gegevensdomeinen: onderwijsdeelname en -resultaten, instellingen en opleidingen, onderwijspersoneel en financiële gegevens. De datasets die DUO open beschikbaar stelt, zijn veelal geaggregeerd en ontdaan van identificerende gegevens als het BSN. De databestanden worden gepubliceerd op data.duo.nl. Deze site is de opvolger van www.onderwijsincijfers.nl, cijfers.minocw.nl en een aantal andere vergelijkbare sites.

Met het openstellen van de data willen het ministerie en DUO openbaarheid bevorderen. De stelregel voor het openbaar maken van data is ook 'alles is openbaar, tenzij'. Hiernaast speelden interne efficiëntie-overwegingen een rol in het actief openbaar maken van de gegevens. De standaardinformatieproducten van DUO vroegen veel tijd en capaciteit. Het bieden van data in plaats van informatie was een efficiëntere manier om aan de vraag te voldoen. Die vraag komt vanuit de sectorraden van de verschillende onderwijssectoren,

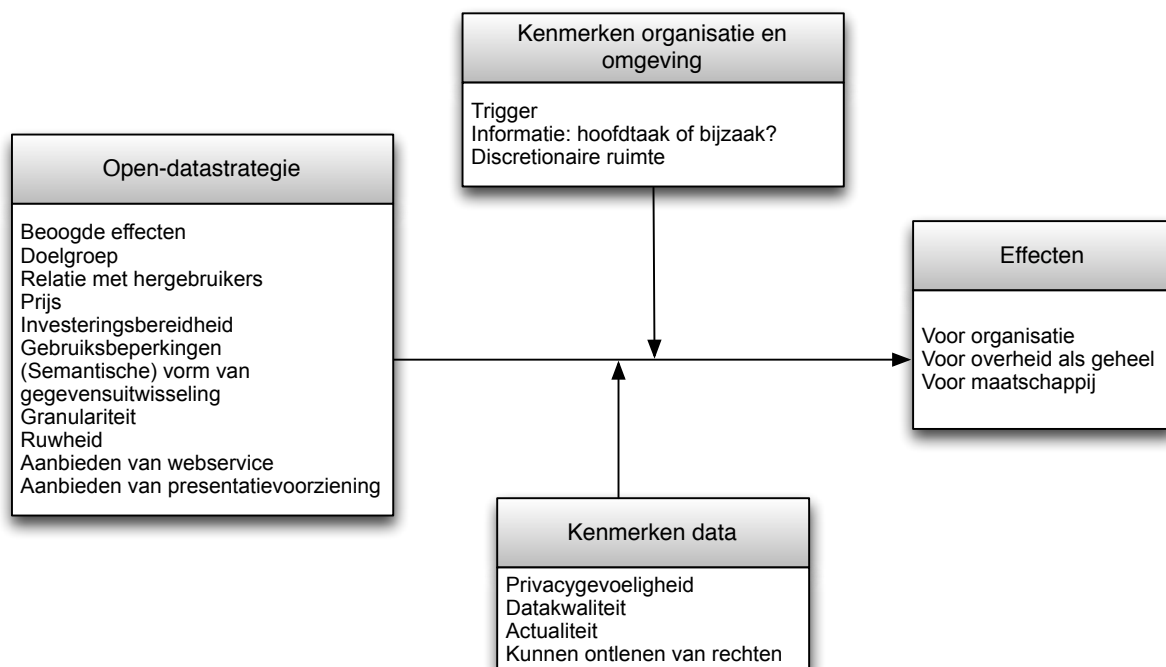
zoals de sectorraad primair onderwijs en de sectorraad voortgezet onderwijs. Ook maken gemeenten, onderzoeksbureaus, adviesbureaus en burgers gebruik van de informatie. Naast deze (her)gebruikers zijn er verschillende bedrijven die de DUO-data - veelal gecombineerd met andere datasets - via een website of via een ander medium aanbieden.

4 Typen keuzes en effecten

In dit hoofdstuk gaan we in op het eerste deel van de onderzoeksresultaten. We geven een overzicht van de keuzes die overheidsorganisaties in hun open-datastrategie maken en de effecten hiervan. Ook geven we aan welke kenmerken van de organisatie, de omgeving en de data deze keuzes en de effecten van de open-datastrategie beïnvloeden. In hoofdstuk 5 kijken we hoe dit zich in de praktijk van de verschillende cases vertaalt.

1. Op welke aspecten maken individuele overheidsorganisaties keuzes als zij hun data open beschikbaar stellen (open-datastrategie)?
2. Welke effecten van het open beschikbaar stellen van data kunnen worden onderscheiden?
3. Welke kenmerken van de organisatie, haar omgeving en van de data zijn van invloed op deze keuzes en de effecten?

Het onderstaande model is het resultaat van de interviews en de documentstudie voor de zeven cases. In dit model zijn de antwoorden van de drie onderzoeksvragen gezamenlijk neergezet. In de volgende paragrafen beschrijven we de verschillende elementen per onderdeel.



4.1 Open-datastrategie

We hebben elf elementen gevonden voor een open-datastrategie:

Element	Beschrijving
Beoogde effecten	De doelen die een organisatie met open data wil bereiken. Deze kunnen betrekking hebben op meerdere niveaus: organisatie, overheid als geheel en maatschappij (zie omschrijving 'blok' effecten).
Doelgroep	De (groep van) hergebruikers waarop de organisatie zich richt. Een organisatie kan data beschikbaar stellen aan eindgebruikers (burgers, bedrijven, journalisten, andere overheden, etc.) en/of aan intermediairs. Een intermediair levert de data (bijvoorbeeld via een applicatie of een website) weer door aan een (eind)gebruiker of een volgende intermediair.
Relatie met hergebruikers	De mate waarin en de wijze waarop een organisatie contact heeft met hergebruikers.
Prijs	De prijs die hergebruikers voor de data moeten betalen. Data kunnen gratis ('zero costing'), tegen marginale verstrekingskosten ('marginal costing'), kostendekkend ('full cost recovery') of tegen een winstmaximalisatieprijs ('commercial costing') aan hergebruikers aan worden geboden.
Investeringsbereidheid	De mate waarin een organisatie wil investeren in het beschikbaar stellen van open data.
Gebruiksbeperkingen	De beperkingen die een organisatie hergebruikers aan het gebruik van data oplegt. Deze beperkingen zijn veelal in de vorm van licenties vastgelegd.

(Semantische) vorm van gegevensuitwisseling	<p>Data kunnen volgens een semantische structuur of standaard worden vrijgegeven.</p> <p>Data die volgens een semantische structuur worden vrijgegeven, zijn in het simpelste geval een .csv, een database of een .xls. In de meest vergaande vorm is de data zodanig gestructureerd dat die geautomatiseerd te relateren is aan andere datasets. Een standaard is een formeel vastgestelde structuur en wordt in het algemeen ook door andere organisaties herkend en gebruikt.</p>
Granulariteit	<p>De mate waarin de data zijn geaggregeerd en tot individuele objecten zijn te herleiden.</p> <p>De data zijn zeer granulaair wanneer elk record in een dataset te herleiden is tot een individueel object. Is de data alleen gegroepeerd weergegeven dan wordt deze minder granulaair. Het minst granulaair is een dataset die alleen statistieken over een gehele populatie weergeeft.</p>
Ruwheid	<p>De mate waarin data zijn bewerkt voor ontsluiting.</p> <p>Ruwe data zijn een één-op-één-kopie van de data die in een systeem aanwezig zijn, eventueel ontdaan van identificerende gegevens als burgerservicenummer, naam en adres. Voor bewerkte data heeft de organisatie moeite gedaan om ze op een andere manier te presenteren dan waarop deze in de (bron)systemen worden gebruikt.</p>
Aanbieden van webservice	<p>De vraag of een organisatie wel of geen webservice gebruikt voor de ontsluiting van data.</p> <p>Een webservice is een algemeen geaccepteerde techniek om informatie tussen twee geautomatiseerde informatiesystemen uit te wisselen.</p>
Aanbieden van presentatievoorziening	<p>De vraag of een organisatie zelf een presentatievoorziening zoals een website of een mobiele applicatie aanbiedt.</p>

4.2 Kenmerken organisatie en omgeving

Kenmerken van de organisatie en haar omgeving beïnvloeden de open-datastrategie en de effecten van die strategie. Verschillende elementen blijken in de praktijk van invloed op de open-datastrategie en de effecten daarvan.

Element	Beschrijving
Trigger	De aanleiding of motivatie voor een organisatie om over te gaan op open data. Deze motivatie kan extern gedreven zijn, zoals wetgeving, maar ook intern, bijvoorbeeld bezuinigingen.
Informatie: hoofdtaak of bijzaak?	Het verzamelen en beheren van informatie kan een hoofdtaak of een bijzaak voor organisaties zijn. Organisaties met informatiebeheer als hoofdtaak (bijvoorbeeld een registerhouder) hebben een andere taak dan organisaties die informatie beheren als ondersteuning van hun eigen processen.
Discretionaire ruimte	De speelruimte of bewegingsvrijheid die een organisatie heeft om zelfstandig de uitvoering van haar taken in te vullen.

4.3 Kenmerken data

Ook kenmerken van de data hebben invloed op de open-datastrategie en de effecten:

Element	Beschrijving
Privacygevoeligheid	De mate waarin data herleidbaar zijn tot (natuurlijke) personen.
Datakwaliteit	De foutgevoeligheid van de data en de mate waarin controlemechanismen zijn ingebouwd om fouten te voorkomen.
Actualiteit	De mate waarin de data, op het moment dat ze beschikbaar komt, overeenstemmen met de data in de werkelijkheid.
Kunnen ontlene van rechten	De vraag of hergebruikers rechten kunnen ontlene op basis van de data.

4.4 Effecten

Effecten kunnen de organisatie zelf, de overheid als geheel of de maatschappij betreffen. In dit onderzoek staan individuele overheidsorganisaties centraal, daarom zijn de effecten op de organisatie zelf in deze analyse de belangrijkste.

4.4.1 Effecten voor de individuele organisatie

Element	Beschrijving
Financiële besparing	Besparingen, die ontstaan als andere instrumenten niet meer ingezet hoeven te worden of als er andere financiële voordelen ontstaan.
Verhoging datakwaliteit	De datakwaliteit kan omhoog gaan doordat de organisatie gedwongen wordt om de eigen datakwaliteit te verhogen alvorens data naar buiten te brengen of doordat er feedbackmechanismen worden ingericht.
Versterking van de legitimiteit	De legitimiteit van de organisatie wordt versterkt doordat de informatie die de organisatie in beheer heeft, vaker wordt gebruikt. Het effect van haar bestaan heeft vaker impact op de samenleving.
Vergroting van de externe oriëntatie van de organisatie	Door voor een open-datastrategie te kiezen, kan een organisatie gedwongen worden om naar buiten te kijken en meer samen te werken met ketenpartners.

4.4.2 Effecten voor de overheid als geheel

Element	Beschrijving
Financiële besparing	Doordat een individuele publieke organisatie met een open-data-aanpak informatie-uitwisselingen anders organiseert, kunnen ook andere publieke instellingen daarvan de (financiële) vruchten plukken.
Standaardisatievoordelen	Wanneer meer data volgens (open) standaarden ontsloten worden, kunnen andere overheidspartijen deze data ook eenvoudiger hergebruiken. De interoperabiliteit van de overheid wordt hierdoor vergroot.

4.4.3 Effecten voor de samenleving

Element	Beschrijving
Economische waardecreatie	Door data voor hergebruik beschikbaar te stellen, kunnen overheidsorganisaties innovatie en nieuwe bedrijvigheid stimuleren.

Vergroting transparantie	Door data voor het publiek (burgers en journalisten) beschikbaar te maken, wordt de openheid van overheidsorganisaties (en indirect daarmee ook de legitimiteit van de overheid) vergroot.
Betere dienstverlening aan burgers en bedrijven	Op basis van open data kunnen toepassingen worden ontwikkeld die een uitbreiding en verbetering zijn ten opzichte van de bestaande dienstverlening van de overheid. Dit kan burgers en bedrijven zowel een besparing in tijd als in geld opleveren.

5 Keuzes en effecten in de praktijk

In het vorige hoofdstuk schetsten we een model met daarin verschillende elementen van een open-datastrategie en de effecten daarvan. Wat hebben wij daarvan in de praktijk van de zeven cases aangetroffen? In dit hoofdstuk geven we antwoord op deelvragen 4 en 5 van het onderzoek. De daarvoor noodzakelijke tussenstap, het per element vergelijken van de cases, is uitgewerkt in bijlage IV.

4. Wat zijn de keuzes, de kenmerken van de organisatie en haar omgeving, de kenmerken van de data en de effecten in de onderzochte cases?
5. Wat zijn de relaties tussen de keuzes, de kenmerken van de organisatie en omgeving, de kenmerken van de data en de effecten (zowel kwantitatief als kwalitatief en zowel op het niveau van de individuele organisatie, het niveau van de overheid als geheel als op het niveau van de maatschappij)?

We beantwoorden deze deelvragen gezamenlijk. Allereerst gaan we in op het eerste gedeelte van het model: de elementen van de open-datastrategie en de wijze waarop kenmerken van de organisatie, omgeving en data deze beïnvloeden. Daarna geven we aan wat de effecten in de praktijk van de zeven cases zijn. De verschillende elementen zijn ter verduidelijking **vetgedrukt**.

5.1 Elementen en relaties open-datastrategie en kenmerken organisatie, omgeving en data

We bespreken de keuzes van de onderzochte organisaties rond hun open-datastrategie en de kenmerken die daarop inwerken van de organisatie, de omgeving en de data.

5.1.1 Aanleiding, doel en investeringsbereidheid

Van de **beoogde effecten** die organisaties met open data nastreven, is het realiseren van financiële besparingen een van de belangrijkste. Dit geldt voor vijf van de zeven cases. Daarnaast is transparantie (als uitgangspunt van goed bestuur) een vaak genoemd doel van de open-datastrategie. Transparantie lijkt echter eerder een nevensdoel dan een hoofddoel. Het is niet altijd mogelijk de verschillende doelen uit elkaar te houden. Een open-datastrategie is vaak gericht op meerdere doelen. Zo is de wettenbank gericht op transparantie (voor de burger) en op financiële voordelen (voor de overheid als geheel). De twee doelen staan nauwelijks op gespannen voet met elkaar.

De **trigger** om met open data aan de slag te gaan is, net als de beoogde effecten, vaak intern (binnen de organisatie) en maar in beperkte mate extern (buiten de organisatie)

gemotiveerd. Dit is heel sterk bij het CIBG terug te zien. Bij het CIBG kiest de organisatie bewust voor het goedkoopste dienstverleningskanaal (internet). Open data helpen de kosten te reduceren van duurdere dienstverleningskanalen (telefoon, balie). Vergelijkbare overwegingen spelen bij DUO. Het leveren van maatwerkrapporten en analyses legde een te groot beslag op de aanwezige capaciteit bij de afdeling informatieproducten. Om deze reden ging DUO open data leveren in plaats van informatie en maatwerkproducten. Het RIVM en de provincie Noord-Brabant hadden een sterke externe motivatie om te starten met open data. De organisaties hebben wettelijke taken om informatie te verspreiden. Echter, door interne besparingen was het wel zo dat vooral bij de provincie Noord-Brabant en de wettenbank de projecten voortvarender en uitgebreider zijn opgepakt dan anders wellicht het geval zou zijn geweest.

In de cases kwamen ook andere doelstellingen naar voren, zoals:

- Het verhogen van de legitimiteit (de samenleving maakt vaker/meer gebruik van de diensten van de overheidsinstelling);
- Het verhogen van datakwaliteit, door het publiek feedback en aanvullingen te laten geven op de data;
- Het vergroten van de externe oriëntatie van de organisatie.

Hoewel dit in sommige cases weliswaar belangrijke doelstellingen zijn, bleken ze voor de organisaties in kwestie echter niet doorslaggevend te zijn om tot actie over te gaan. Het realiseren van financieel voordeel blijkt vrijwel altijd een voorwaarde te zijn voor het overgaan tot het uitvoeren van een open-datastrategie.

Ook zien we dat het realiseren van macro-economische welvaartseffecten (economische groei) nauwelijks van invloed is op het beslisproces om te starten met open data. Dit is geheel in overstemming met de resultaten van POPSIS (Vries et al., 2011).

Kijken we naar de mate van **investeringsbereidheid**, dan zien we dat alleen bij de RCE sprake is geweest van grote investeringen. Dit hangt nauw samen met de **discretionaire ruimte**. De RCE heeft relatief veel discretionaire ruimte. Daarom kon de RCE eenvoudiger grote investeringen doen voor zijn open-datastrategie, dan organisaties zoals het CIBG en het RIVM. Dergelijke projectgefinancierde organisaties moeten veel meer de boer op om hun plannen aan hun opdrachtgever te verkopen.

5.1.2 Gebruik en gebruikers

De **doelgroep** voor open data is bij geen van de onderzochte cases gespecificeerd. De data zijn in bijna alle gevallen voor iedereen op een begrijpelijke wijze inzichtelijk gemaakt (bijvoorbeeld in de vorm van een webpagina of een Excel-bestand). Daarnaast bieden de meeste organisaties ook een webservice aan, waarmee de overheidsorganisaties ook professioneel hergebruik mogelijk maken. Data zijn dus voor ieder (van burger tot journalist en van onderzoeker tot informatie-ondernemer) herbruikbaar. De onderzochte organisaties

leggen hergebruikers van hun informatie daarbij ook niet of nauwelijks (juridische) **gebruiksbeperkingen** op. Organisaties willen (maximaal) hergebruik stimuleren, en gebruiksbeperkingen zijn daarmee niet in overeenstemming. Organisaties stellen ook geen richtlijnen op voor het type gebruik. Eenieder heeft dus gelijke rechten en mogelijkheden om data te hergebruiken.

Een ander element van de open-datastrategie is de **relatie** die organisaties **met hergebruikers** aangaan. De meeste organisaties kennen (een aantal van) hun hergebruikers en hebben meer of minder frequent contact met hen. In twee gevallen (PGR en wettenbank) bestaat er zelfs een formele gebruikersraad. In andere gevallen, zoals bij DUO en de RCE, is het contact met de hergebruikers vooral incidenteel. Het CIBG vormt een uitzondering. Zij kent haar hergebruikers in principe niet. Het CIBG vindt dit haar taak ook niet. Zij stellen data ter beschikking, zodat hergebruikers ermee aan de slag kunnen en daar houdt de taak van het CIBG op.

In de cases zien we dat contact met hergebruikers de organisaties helpt in het afstemmen van vernieuwingen, het verhogen van de datakwaliteit (DUO en RCE), het bepalen van de prioriteiten (PGR) en het bepalen van (nieuwe) functionele eisen (wettenbank). Opvallend is dat de RCE op dit moment bezig is om feedbackmechanismen, bijvoorbeeld voor het verwerken van feedback vanuit Wikipedia, in te regelen. Daarbij wordt niet het contact gezocht met specifieke hergebruikers, maar de organisatie staat wel open voor tweerichtingscommunicatie op dataniveau. Wanneer de organisatie hergebruik van haar data actief wil stimuleren, bijvoorbeeld om haar legitimiteit te verhogen, dan is contact met en assistentie van vooral professionele hergebruikers een van de meest voor de hand liggende opties. Het gaat daarbij dan bijvoorbeeld om het geven van uitleg over de datastructuur, eventueel onderliggende verbanden in de data en het geven van duidelijkheid over de update frequentie.

Hergebruikers hoeven in geen van de case een **prijs** voor de data te betalen. De data worden gratis ter beschikking gesteld. Alle organisaties hebben ook de intentie de data gratis te blijven aanbieden. Zolang dat tenminste geen onevenredig beslag legt op de (capaciteit van de) organisatie. Zowel Noord-Brabant, DUO/OCW als de RCE zouden, als er door excessief hergebruik hoge kosten ontstaan voor de organisatie, zich wel kunnen voorstellen dat hergebruikers toch moeten betalen. Dat is nu echter niet aan de orde. Voor de wettenbank is het gratis beschikbaar stellen van informatie een bewuste beleidskeuze vanuit transparantie voor de burger. Op dit moment is het dan ook de verwachting dat er ook in de toekomst niet voor de informatie betaald hoeft te worden.

5.1.3 Vorm en presentatie van de data

Alle data worden in de vorm van een **structuur** aangeboden, ten minste in een (machineleesbare) tabelvorm. Over het algemeen wordt geen extra **bewerkingslag** op de

data gemaakt. Er bestaan wel verschillen in de mate van detail van de beschrijving van de structuur. Het RIVM, DUO/OCW en het CIBG geven aan welke informatie geleverd wordt en welke betekenis deze heeft. Bij de RCE zijn een onderliggende thesaurus en objectmodel beschikbaar. De wettenbank maakt zelfs gebruik van een **standaard** voor de data. Deze is ontstaan vanuit hergebruik en mede ontwikkeld door de hergebruikers van de data.

Ten aanzien van de **granulariteit** van de data kunnen we uit de cases de conclusie trekken dat er alleen sprake is van aggregatie bij privacygevoelige informatie. DUO aggregiert bijvoorbeeld de data om herleidbaarheid tot individuele docenten of leerlingen te voorkomen.

RIVM, RCE en DUO bewerken de data die zij naar buiten brengen. Dit heeft vooral met de **datakwaliteit** van doen. In het algemeen lijkt de volgende redenering te worden gehanteerd: 'data bewerk je als het om kwaliteitsredenen moet, **ruwe** data stel je ter beschikking als het kan'. Zo controleert het RIVM haar ruwe data en past zij deze aan. Ook DUO maakt veel interne controleslagen voordat zij haar data beschikbaar stelt, en haalt onder meer **privacygevoelige** informatie eruit. De RCE probeert ook zo weinig mogelijk privacygevoelige informatie open te maken, maar aggregiert de data daarnaast ook met het oog op de toegankelijkheid.

De taakopvatting van de overheid bepaalt in belangrijke mate het wel of niet bieden van een **presentatievoorziening**. Een gebruikelijk open-dataprincipe is dat de overheid de data levert en de samenleving de apps (presentatievoorzieningen) bouwt. Bij drie cases is dit ook het geval. Bij DUO wordt alleen data aangeboden, die vervolgens door 10.000scholen.nl en anderen wordt hergebruikt. Bij de provincie Noord-Brabant wordt alleen data aangeboden, in de vorm van een **webservice**. De gemeente Rotterdam biedt data aan zonder presentatievoorziening. Studenten hebben hierop een aantal apps ontwikkeld. Bij de wettenbank, de RCE, CIBG en het RIVM wordt data zowel als ruwe data aangeboden (via webservices of downloads van bestanden) als via een presentatievoorziening (bijvoorbeeld wetten.nl en beeldbank.cultureelerfgoed.nl).

Sommige van de organisaties zien het als hun taak zelf presentatievoorzieningen te beheren of te bouwen. Zij vinden een eigen presentatievoorziening noodzakelijk, waarmee de 'gewone' burger de informatie kan analyseren en gebruiken. In veel gevallen is dit een website, maar een aantal organisaties vindt ook een mobiele applicatie denkbaar. Het maakt deze organisaties in beginsel niet zoveel uit of er alternatieven komen, die dezelfde informatie op andere manieren ontsluiten. Zij zullen ondanks concurrentie van anderen hun eigen kanaal in stand houden. BZK vindt bijvoorbeeld dat wetten.nl altijd moet blijven bestaan, hoeveel commerciële alternatieven er ook komen. Het CIBG staat hier anders in. Wanneer er een gezonde markt van commerciële aanbieders van presentatievoorzieningen voor de betreffende dataset ontstaat, kan deze taak voor het CIBG vervallen. Het beheer van het register zelf wordt vanzelfsprekend beschouwd als een overheidstaak.

5.2 Effecten van de open-datastrategie

In deze paragraaf geven we aan wat open data de verschillende onderzochte organisaties oplevert. We onderscheiden financiële en niet-financiële effecten.

5.2.1 Financiële effecten

In elke onderzochte case zijn door de geïnterviewde personen van de bronhoudende organisaties vermoedens of bewijzen van de financiële voordelen van open data benoemd. Open data leiden volgens hen vooral tot een positief financieel resultaat door de mogelijkheid van besparingen binnen de eigen organisatie, door bijvoorbeeld de kanalenmix aan te passen in de richting van minder dure kanalen of door minder handmatige handelingen te hoeven verrichten. DUO en het PGR besparen bijvoorbeeld op het aantal uren dat medewerkers besteden aan het vervaardigen van informatieproducten. De initiële jaarlijkse kosten bij DUO en de provincie zijn hoger door informatie pro-actief aan te bieden, maar de transactiekosten per bevraging zijn fors lager. Dit kan flinke besparingen opleveren, omdat het aantal bevestigingen in beide gevallen hoog is. De provincie Noord-Brabant bespaart 200 tot 500 uur werk per jaar aan uitvragen, hoewel zij aan de andere kant ook een tijdsinvestering heeft moeten doen in het PGR. Bij DUO gaat het om 0,5 fte binnen de afdeling informatieproducten. Bij het CIBG is het aantal uitvragen in de afgelopen jaren met 80.000 per maand toegenomen, zonder extra personeel te hoeven inzetten. Hier wordt dus meer werk verzet met ongeveer hetzelfde aantal mensen.

Een andere categorie financiële voordelen zijn de 'inkoopvoordelen'. Die zien we vooral bij het ministerie van BZK. Centraal inkopen en financieren van wetteksten geeft de overheid meer marktmacht om goedkoper in te kopen. Bijkomend voordeel is dat alle departementen, gemeenten en andere overheidsorganisaties ook geen of veel minder rekeningen meer (hoeven te) ontvangen voor wetsteksten. Binnen de (Rijks)overheid hoeft geen verrekening of vereffening (meer) uitgevoerd te worden omdat de wettenbank gefinancierd is uit de rijksbegroting. Ook dat wordt gezien als een groot voordeel.

Deze kosten verbonden aan het uitvoeren van de open-datastrategie vallen vaak onder een van de volgende categorieën:

- Kosten om data geschikt te maken voor ontsluiting
 - Kwaliteitsverbetering en aanbrengen structuur
 - Digitalisering

- Kosten van het open-datakanaal
 - Bouw en beheer website
 - Bouw en beheer webservice
 - Down- en uploadcapaciteit
 - Servers
 - Bouw en beheer apps voor 'mobile devices'
- Operationele kosten van open-data-activiteiten
 - Beheer data
 - Beheer klantrelatie
 - Verwerking financiële transacties
 - Publiciteit
 - Helpdeskfunctie

Kosten om de data geschikt te maken voor ontsluiting zijn kosten die soms toegerekend worden aan de open-datastrategie en soms aan de informatiestrategie. Het op orde krijgen van de datakwaliteit en de structuur van de data hangt sterk samen met gemaakte keuzes rondom **granulariteit** en **ruwheid** van de data, alsmede de **(semantische) vorm van de gegevensuitwisseling**. Aan data van hoge kwaliteit kunnen vaak wel meer **rechten ontleend** worden dan aan data van mindere kwaliteit. De dominante opvatting over deze kosten is overigens dat het helemaal niet nodig is om dit type kosten te maken, omdat data zo ruw mogelijk aangeboden moeten worden. Organisaties hebben soms toch redenen om wel te investeren in de kwaliteit van de data. Bij het PGR bijvoorbeeld is een standaardisatieslag doorgevoerd om de bruikbaarheid van de data van meerdere provincies tegelijkertijd te verhogen.

De kosten voor het open-datakanaal zijn het meest tastbaar. Het bouwen van een website worden door veel organisaties als noodzakelijke kosten ervaren. In Rotterdam, daarentegen, bieden de gemeentelijke diensten de data eigenlijk alleen als te downloaden bestand of webservice aan. Een website is dus niet perse noodzakelijk. Wanneer een website niet alleen de data beschikbaar stelt, maar ook mogelijkheden biedt om de data te bekijken of te analyseren, is er sprake van een **presentatievoorziening**. De kosten van een presentatievoorziening zijn in de regel veel hoger dan de kosten van een webservice of website. Het bouwen van een **webservice** wordt in de regel alleen gedaan wanneer er een presentatievoorziening gebouwd wordt die gebruik maakt van een webservice voor de communicatie met de databases, of er al een service-architectuur beschikbaar is, waarop deze eenvoudig kan worden gebouwd. Kosten voor servers, down- en uploadcapaciteit zijn bij de door ons onderzochte cases beperkt. Veelal kan hier gebruik worden gemaakt van bestaande infrastructuur.

Ten slotte de operationele kosten van de open-datastrategie. Het beheren van de open data is vrijwel nooit een open-data-activiteit, maar vaak een lijnactiviteit. Dit wordt door bijna geen

enkele organisatie gezien als open-datakosten. De kosten voor het beheren van een **website** en of **presentatievoorziening** worden in de regel wel als open-data-activiteiten gezien. Voor het beheren van de relatie met hergebruikers (al dan niet in een gebruikersoverleg) en het 'adverteren' van de beschikbaarheid van open data (bijvoorbeeld in het nationaal dataportaal of op blogs) wordt door de onderzochte organisaties vaak wel capaciteit beschikbaar gesteld. Hierbij hangt veel af van de intensiteit van de **relatie met hergebruikers** die de organisatie voorstaat. De kosten van een helpdesk zijn vaak minimaal, omdat de open-datastrategie in organisaties waar zo'n helpdesk is ingericht, helpt met het naar beneden brengen van het aantal telefoontjes of mailtjes. Mochten er dan toch vragen binnenkomen bij de helpdesk, dan zijn dit maar een fractie van de kosten die eerder gemaakt werden. Organisaties richten over het algemeen geen helpdesk speciaal voor hergebruikers in.

Toerekenen van kosten en baten

Kosten of baten blijken in de praktijk nauwelijks toe te rekenen aan een open-datastrategie. Veelal worden kosten of baten toch al gerealiseerd. "De helpdesk bestond vroeger ook al, ze krijgen nauwelijks meer vragen dan vroeger" of "we moesten toch al aan de slag met de website" zijn veelgehoorde opmerkingen.

Het toerekenen van kosten en baten is afhankelijk van de manier waarop een organisatie zelf de scope van een open-dataproject definieert. Het CIBG, bijvoorbeeld, heeft aan het begin van deze eeuw haar hele informatievoorziening vernieuwd en volledig volgens de principes van de service-oriented architecture (SOA) ingericht. De kern van SOA is dat softwarefunctionaliteit in herbruikbare componenten ('bouwsteentjes') is opgedeeld, die onderling makkelijk kunnen worden verbonden.

Toen dit project klaar was, bleken de meerkosten van het ontwikkelen van een webservice voor het verstrekken van BIG-gegevens eenmalig slechts € 15.000,- te bedragen. Omdat de technische infrastructuur nauwelijks wordt belast, vindt het CIBG dat zij haar open-datastrategie voor € 15.000,- heeft gerealiseerd. In contrast daarmee staat de RCE, die € 7,5 miljoen investeert in het digitaliseren van informatie en verbeteren van de informatievoorziening en dit in zijn geheel beschouwt als open-dataproject. Het doel van de RCE is tenslotte al zijn beschikbare informatie ook openbaar te maken. Het beperken van onder andere archiefruimte en het besparen op receptionisten ziet zij dan ook als baten van de open-data-activiteiten. Het is echter ook mogelijk te beargumenteren dat het grootste gedeelte van de kosten en baten van deze projecten vooral binnen de organisatie liggen. Om effectief en efficiënt te opereren moet de informatievoorziening op orde zijn. Digitalisering van informatie zou daarom ook gezien kunnen worden als een activiteit die noodzakelijk is voor het opereren als moderne archief beherende organisatie.

De vraag is dus waar de grens ligt, of moet liggen, tussen activiteiten om de eigen informatievoorziening op orde te krijgen en om open data beschikbaar te stellen. Deze vraag

wordt door elke organisatie anders beantwoord. De cases kunnen daarom op dit punt niet kwantitatief met elkaar worden vergeleken. We bespreken om deze reden voor elke case de kosten en baten en de betreffende inkadering.

Financiële kosten en baten per case

Bij het CIBG is voor € 15.000,- (afgeschreven in 6 jaar) een extra webservice gebouwd. Het aantal bevestigingen van het BIG-register via de gecombineerde website en webservice is verdubbeld ten opzichte van de situatie waarin alleen de website als kanaal beschikbaar was. Hoewel er hier geen sprake is van een positieve businesscase op basis van inkomsten en uitgaven, zijn de kosten zo laag dat ze bijna verwaarloosbaar zijn. Het CIBG had voordat de webservice gebouwd werd al fors geïnvesteerd in het op orde brengen van de interne informatievoorziening.

De situatie in Rotterdam is vergelijkbaar. De gemeente heeft voor ongeveer € 100.000,- kosten gemaakt. Een gedeelte daarvan zijn 'sunk costs', kosten die niet vermijdbaar waren maar nu ingezet worden ten behoeve van Rotterdam Open Data. Deze kosten zijn vooral gemaakt voor het faciliteren van het project van de Hogeschool Rotterdam (HR): beschikbaar stellen van personeel en ruimte. De HR en de verschillende bedrijven verenigd in het 'Rotterdam Open Data'-platform hebben voor € 565.000,- aan financiering van onderzoeksprojecten binnengehaald. Vergelijkbaar met het CIBG, is er dus een investering gedaan die nog niet rendeert voor de gemeente zelf. Aan de andere kant is dit geen enorm bedrag, afgezet tegen de gehele gemeentebegroting. Wat ook vergelijkbaar is met het CIBG, is dat Rotterdam nog veel baten verwacht van open data. In de forse bezuinigingsopdracht die de stad nu heeft, is open data een van de mogelijke innovaties die ervoor kunnen zorgen dat de stad met minder personeel en minder geld toch een betere leefomgeving kan verzorgen. De investering moet dus gezien worden tegen de achtergrond dat er in de toekomst veel projecten in samenwerking met de ondernemers en burgers uit de stad plaats moeten vinden en open data maken een belangrijk deel van die innovatie mogelijk. De besparingen die dan moeten plaatsvinden komen voornamelijk uit het afstoten of 'crowdsourcen' van taken naar het publiek en het efficiënter kunnen uitvoeren van andere taken door het publiek te betrekken in de uitvoering van de taak.

Bij BZK is men ervan overtuigd dat de kosten voor de wettenbank veel lager zijn dan de besparingen op rijksniveau. Het beheer van de software en het uitbreiden en actualiseren van de content kosten jaarlijks € 2,2 miljoen. Om dat te financieren heeft een herverdeling van budgetten plaatsgevonden vanuit de budgetten van de andere ministeries. Gecorrigeerd voor prijsinflatie bedragen de kosten die vroeger werden uitgegeven aan de Algemene Databank Wet- en Regelgeving circa € 1,9 miljoen per jaar. Daarnaast is voor een onbekend bedrag bespaard op de individuele informatieproducten en redactie bij ministeries en uitvoeringsorganisaties. Er werden door deze individuele partijen namelijk ook andere wetsinformatieproducten aangeschaft (zoals boeken en wettenbundels) en dat hoeven zij nu

niet meer te doen, omdat het op wetten.nl vindbaar is. Hoewel er daarvoor geen centrale inkoopprocedure was (en dus lastig in te schatten is wat er bespaard wordt) is de algemene opinie wel dat de besparing die dit oplevert groter is dan de € 300.000,- verschil tussen de kosten vroeger en de huidige kosten. Daarnaast is de wettenbank voor de wetgevingsjuristen een veel geschikter instrument dan de oude databank Wet- en Regelgeving. Samengevat kan gesteld worden dat, hoewel BZK meer geld investeert in de wettenbank, de businesscase op rijksniveau waarschijnlijk positief is.

De RCE heeft een ambitieus programma van organisatieverandering verbonden met de transitie naar een open-data-organisatie. Zij verdient in ongeveer 10 jaar de volledige investering (ter grootte van € 7,5 miljoen) terug door te besparen op balie- en archiefmedewerkers. Doordat burgers veel minder bij het archief langs komen, maar gegevens online inzien is er minder bezoek. Ook wordt de postverwerking minder arbeidsintensief, doordat deze grotendeels gedigitaliseerd en geautomatiseerd is. Daardoor kan dus vooral worden gesneden in de ondersteuning van de primaire taken, die zelf ook makkelijker uit te voeren worden. Het grootste gedeelte van de investering wordt gebruikt voor het verbeteren van de informatiearchitectuur, de ontsluiting van diverse databronnen binnen die architectuur en de digitalisering van een gedeelte van het archief. Een groot gedeelte van de investering door de RCE is overigens binnengehaald met een rijkssubsidie en een substantieel gedeelte van de besparing zal (naar verwachting) in de vorm van een korting op het budget weer terugvloeien naar de rijksbegroting.

Het RIVM krijgt een vast budget van haar opdrachtgever voor een aantal vastgestelde taken rondom de emissie- en luchtkwaliteitsregistraties. Een van de taken is het beschikbaar stellen van de data. Daarmee is er geen geïsoleerd financieel effect voor het RIVM te onderkennen. Met de overgang van papieren naar digitale en in toenemende mate machineleesbare ontsluiting van de data zijn de kosten vooral verschoven tussen de verschillende publicatiekanalen, maar in omvang niet veranderd.

Bij OCW is € 500.000,- uitgegeven aan het project 'website openbare onderwijsinformatie'. Op basis van het projectplan zullen op termijn 4,5 fte formatie voor de afdeling informatieproducten verdwijnen. Na de terugverdientijd van iets meer dan 2 jaar wordt er door DUO dus structureel bespaard.

De Nederlandse provincies hebben gezamenlijk € 450.000,- bijgedragen aan het bouwen van het PGR. Dit portaal moest gebouwd worden om te kunnen voldoen aan nieuwe wetgeving rondom geo-informatie. De totale jaarlijkse besparing in uren van alle provincies samen wordt door een geïnterviewde op 100 tot 500 uur per jaar per provincie geschat, zonder rekening te houden met de in het PGR geïnvesteerde tijd (circa 10 tot 70 uur per jaar per provincie).

5.2.2 Niet-financiële effecten

Voor de organisatie

Het **versterken van de legitimiteit** van de organisatie is het voornaamste niet-financiële effect van open data. Dit resultaat zien we bij vijf cases terug. Bijvoorbeeld bij het CIBG, waar het gebruik van data fors is gestegen door openstelling van een webservice voor BIG-bevragingen. Het aantal bevragingen van het BIG-register dat niet per telefoon verloopt, is na de invoering van de webservice verdubbeld (van 80.000 bevragingen via de website naar 80.000 bevragingen via website en 80.000 via webservice). Ditzelfde effect is ook waarneembaar bij de RCE, DUO, het PGR en het RIVM, waar de data steeds vaker gebruikt worden.

Bij de gemeente Rotterdam en bij de RCE is sprake van een **vergroting van de externe oriëntatie**. Beide hebben hun netwerk en hun faciliterende rol daarin uitgebouwd. Een andere uitkomst voor de RCE is, dat de **datakwaliteit verbeterd** wordt door feedback van derden over (mogelijke) fouten in de data.

Voor de overheid als geheel

Effecten voor de overheid als geheel zijn zichtbaar bij het PGR en de wettenbank. Het PGR zorgt voor **standaardisatie** en mede daardoor verlaagde transactiekosten voor zowel de datahouders (provincies) als de afnemers van de open data (veelal andere overheidsorganisaties). De standaardisatie betreft de datastructuur (bijvoorbeeld definities) en de wijze van ontsluiting (bijvoorbeeld de gebruiksvoorwaarden). Ook de wettenbank heeft tot een grotere standaardisatie geleid, die het onder andere mogelijk maakt om relaties te leggen tussen verschillende soorten wetgeving, publicaties in het Staatsblad en parlementaire documenten te leggen. Dat maakt het vinden van officiële documenten gemakkelijker en eenvoudiger, en bevordert de samenwerking tussen onderdelen van de overheid.

Voor de samenleving

Op het niveau van de samenleving zien we drie effecten in de praktijk terug: betere dienstverlening aan burgers en bedrijven, economische waardecreatie en vergroting van transparantie. **Economische waardecreatie** zien we in bijna alle cases terug. De data van alle organisaties worden ook hergebruikt. Bij het CIBG maken bijvoorbeeld uitgeverij als Springer Verlag gebruik van de data. De data van de RCE worden gebruikt in erfgoed.mobi waarmee AB-C media hun meerwaarde kunnen tonen aan potentiële klanten. De data van het RIVM en het PGR wordt door advies- en ingenieursbureaus gebruikt.

Het **verbeteren van de dienstverlening** aan burgers is in een aantal gevallen terug te zien. Bij de RCE hoeven burgers niet meer langs de balie voor een foto, maar kunnen ze die via internet verkrijgen. Bij het CIBG kunnen ook anderen gebruik maken van de BIG-data,

waardoor hun dienstverlening verbeterd. De Vereniging Verpleegkundigen Nederland doet dit bijvoorbeeld.

De **transparantie** van een overheidsorganisatie wordt impliciet verbeterd wanneer open data beschikbaar worden gesteld. Maar toch kunnen hier enkele kanttekeningen bij geplaatst worden. In alle cases krijgt de burger (in welke rol dan ook) toegang tot meer informatie dan voorheen. Dit betekent meer transparantie. Aan de andere kant zien we slechts bij een aantal organisaties meer mogelijkheden om het functioneren van de overheid beter te beoordelen. De RIVM-emissiedata zijn bijvoorbeeld een bron om (de effecten van) het Nederlandse milieubeleid beter te beoordelen. Maar, met de data van de RCE kan het (effect van het) Nederlandse erfgoedbeleid niet direct worden beoordeeld.

6 Conclusie: businessmodellen voor open data

In dit hoofdstuk geven we antwoord op de hoofdvraag van het onderzoek:

Wat zijn businessmodellen voor open data, die individuele overheidsorganisaties kunnen gebruiken wanneer zij open data beschikbaar (willen) stellen?

Uit het casemateriaal en de analyse in het vorige hoofdstuk kunnen we wij twee businessmodellen afleiden. We onderscheiden allereerst het **radicale open-datamodel**. Open data gaan samen met een koerswijziging van de organisatie, waardoor ook nieuwe doelstellingen gerealiseerd kunnen worden. Daarnaast komt veel vaker het **incrementele open-datamodel** voor. Open data worden dan onderdeel van een al bestaande organisatie- en informatiestrategie en dragen bij aan al bestaande doelen.

6.1 Radicale innovatie: nieuwe mogelijkheden

Wanneer open data door een organisatie worden ingezet om nieuwe doelen te realiseren, zijn de keuzes rondom open data verweven met een gewijzigde organisatie- en informatiestrategie. Open data zijn het middel dat kan bijdragen om het businessmodel van de organisatie te veranderen. Het verstrekken van open data wordt één van de belangrijkste manieren om aan de legitimiteit van de organisatie te werken. Het model past dus bij organisaties die het **beheren van informatie tot hun hoofdtaken** rekenen.

In het radicale open-datamodel is de totstandkoming van de **open-datastrategie** tamelijk expliciet. De omvang van de beoogde effecten en daarmee het belang voor de organisatie zijn dermate groot dat de keuze voor open data in de top van de organisatie wordt gemaakt. Als de organisatie over voldoende **discretionaire ruimte** beschikt of haar financiers kan overtuigen, is het ook mogelijk om eenmalige financiële investeringen te doen die pas later worden terugverdiend. Een expliciete businesscase is daarvoor vereist. De kosten en opbrengsten van open data zijn daarin echter moeilijk te isoleren. Dit komt door de verstrengeling met andere wijzigingen door de nieuwe organisatie- en informatiestrategie.

De **granulariteit, ruwheid** en **kwaliteit van de** (bestaande) **data** hebben nauwelijks invloed op de keuze om voor het radicale open-datamodel te kiezen. Deze kenmerken van de strategie worden eerder gekozen aan de hand van de doelen die de organisatie zichzelf met de innovatie stelt. De kans is groot dat de organisatie die kiest voor het radicale open data model kiest voor een geïntegreerde levering, dat wil zeggen dat de **semantische vorm van de data** in verregaande mate gestructureerd en gestandaardiseerd zal zijn. De organisatie

wil immers haar **legitimiteit** ophangen aan de levering van open data en dan is het bijvoorbeeld belangrijk dat informatie niet met elkaar in tegenspraak is en dat de definities helder zijn.

De relatie met klanten is voor de meeste organisaties al belangrijk, maar wanneer je jezelf opnieuw uitvindt is dat zeker het geval. De **relatie met hergebruikers** is bij het radicale open-datamodel relatief goed ingericht. Op zijn minst is er op dit punt een expliciete afweging gemaakt. Het slagen of falen van de transformatie heeft tenslotte veel te maken met de manier waarop afnemers met informatieproducten omgaan en op welke manier zij de nieuwe open data gaan inzetten.

In dit onderzoek vinden we dit alternatieve businessmodel bij de RCE terug. De keuze om data open beschikbaar te stellen is nauw verbonden met de nieuwe taakopvatting, waarbij de RCE toegroeit naar een netwerkorganisatie en informatie anders inzet. Daartoe is de gegevensstructuur omgevormd en worden data ontsloten via partners als bijvoorbeeld Wikipedia. De data worden verregaand gestructureerd ontsloten en op het niveau van granulariteit, ruwheid en datakwaliteit zijn bewuste keuzen gemaakt.

6.2 Incrementele innovatie: verbetering van het bestaande

In zes cases worden open data vooral ingezet om - binnen de kaders van de bestaande organisatiestrategie - een aantal activiteiten of producten goedkoper of efficiënter samen te stellen. Vaak betekent dit dat de kanalenmix van balie, telefoon, website uitgebreid wordt met een nieuw gegevenskanaal genaamd 'open data'. In dit model worden financiële besparingen verwacht door het ontlasten van duurdere communicatiekanalen. Ook zien we wel dat helpdesks minder tijd besteden aan individuele hulpvragen rondom de informatie die zij beheren, maar meer verwijzen naar de datasets die al online staan.

De impact, voor de organisatie van toepassing, van het incrementele open-datamodel is meestal beperkt, omdat er geen structurele wijzigingen in productaanbod of takenpakket plaatsvinden. Ook hoeft de **investeringsbereidheid** dus niet heel groot te zijn. Het is vaak niet nodig om grote investeringen te doen om data op een bruikbare wijze beschikbaar te stellen. Door de beperkte omvang van de investeringen ontbreekt vaak ook een expliciete businesscase. In tegenstelling tot het radicale open-datamodel kan de investering in open data wel beter geïsoleerd worden van andere strategische beslissingen.

De data worden in het algemeen gratis beschikbaar gesteld. De overweging om eventueel **verstrekkingkosten** in rekening te brengen, wordt pas actueel als de omvang van het gebruik daar aanleiding toe geeft. Zolang er geen grote kosten worden gemaakt, is het voor de organisatie ook niet nodig om geld te vragen voor de open-datadiensten. Het verkopen van data is immers geen primair of bekend businessmodel van de organisatie.

Ook zien we dat niet altijd expliciet gestuurd wordt op het bereiken van de doelen van de open-datastrategie. De meeste organisaties hanteren een 'lerende' strategie, waarbij geëxperimenteerd wordt met open data zonder van tevoren harde doelen te definiëren. Vaak bestaat de wens om 'iets' aan open data te doen, omdat men de abstracte voordelen ervan inziet, maar wel op een zodanige manier dat het ook iets oplevert voor de organisatie zelf.

Door het beschikbaar stellen en het ontdekken van de effecten komen nieuwe vraagstukken op. Niet alle strategische keuzes worden dan ook op hetzelfde moment gemaakt. Organisaties die meer ervaring hebben met open data zullen een groter deel van hun keuzes kunnen expliciteren. Ook zullen zij intensiever **contact met hergebruikers** hebben. In de loop van de tijd tekent zich namelijk duidelijker af wie de hergebruikers zijn. Er ontstaan gezamenlijke ervaringen met de dataset en het wordt duidelijk dat er ook gezamenlijke belangen kunnen zijn.

6.3 Bouwen op bestaande data

In beide businessmodellen hangt de technische wijze waarop de data beschikbaar wordt gesteld vooral af van de kenmerken van de data die in het primaire proces van de organisatie zijn ontstaan. Zo komen de actualiteit, de structuur en de granulariteit van de data vaak overeen met de oorspronkelijke dataset. Als de data worden bewerkt voor openstelling, is dat veelal om privacygevoelige gegevens te verwijderen of om de data toegankelijker te representeren voor hergebruikers. Het is ook niet noodzakelijk om een foutloze dataset te bieden aan hergebruikers, als maar wel duidelijk is wat de kwaliteit van de data is en welke rechten eraan ontleend kunnen worden. Hergebruik van de data kan overigens een positieve impuls geven aan de datakwaliteit door feedback van hergebruikers. Ook gebruiksrechten ontstaan niet doordat data open beschikbaar wordt gesteld. Meestal was de dataset daar al op ingericht.

6.4 Hergebruik naast eigen aanbod

Organisaties kiezen er in beide businessmodellen vaak voor om een eigen presentatievoorziening in stand te houden. Daarmee wordt een kanaal in stand gehouden om eindgebruikers te bedienen. Dat kan bijdragen aan de legitimiteit van de organisatie en haar taken. Het is immers heel zichtbaar dat de data afkomstig zijn van de organisatie. Ook biedt een eigen kanaal de mogelijkheid om gevoel te krijgen voor het hergebruik. Het surfgedrag van bezoekers op wetten.nl geeft inzicht in de interesse in verschillende wetten. Gebruikers geven ook soms opmerkingen door. Die informatie zou ontbreken als er slechts een webservice voor intermediaire en professionele hergebruikers zou bestaan. De presentatievoorziening kan wel kostbaar zijn. In enkele gevallen zou de eigen voorziening geheel kunnen verdwijnen als intermediaire partijen voorzieningen aanbieden op basis van de beschikbaar gestelde open data. Dat zou (flinke) besparingen op kunnen leveren voor de

bronhouder, die dat dan moet afwegen tegen het belang van de legitimiteitsversterking door het aanbieden van een herkenbare eigen voorziening.

7 Discussie en aanbevelingen

In dit hoofdstuk bespreken we de implicaties van de conclusies, zowel voor organisaties die met open data aan de slag willen gaan, als voor het stimuleringsbeleid van de Nederlandse overheid voor open data. Ook reflecteren we op de uitkomsten en de aanpak van het onderzoek. De aanbevelingen die hieruit volgen hebben we in kaders weergegeven.

7.1 Implicaties

De afwegingen van organisaties rondom open data zijn zelden geïsoleerd. Ze zijn sterk verbonden met de aanwezige of beoogde organisatie- en informatiestrategie. Organisaties herkennen het label 'open data' ook niet altijd, terwijl zij wel data beschikbaar stellen volgens de definitie van open data. Bij het incrementele businessmodel zijn open data soms meer het resultaat van de overtuigingen en activiteiten van medewerkers op tactisch en operationeel niveau, dan van afwegingen in de bestuurskamer. Daarmee is duidelijk dat zowel het enthousiasme over als de daadwerkelijke inzet op open-datagebied voor een belangrijk deel afhangt van medewerkers lager in de organisatie. Zij zijn dan ook een belangrijke doelgroep bij het stimuleren van het hergebruik van overheidsinformatie.

Aanbeveling: Onderken het belang van enthousiaste medewerkers, die het open-datagedachtegoed begrijpen en tegelijkertijd de verbinding met de organisatie kunnen leggen.

De afweging van organisaties om open data beschikbaar te stellen kan sterk worden beïnvloed door opdrachtgevende of financierende partijen. Als bijvoorbeeld opdrachtgevers van onderzoek eisen dat de data open worden gesteld, is dat richtinggevend voor de opdrachtnemende publieke organisatie. Ook is het soms nodig een voorinvestering toe te staan die niet binnen hetzelfde boekjaar wordt terugverdiend. Ook wetgeving kan organisaties bewegen data te openen, als die verplichtingen bevat om (specifieke) data openbaar te maken.

Aanbeveling: Maak gebruik van de opdrachtgever-opdrachtnemerrelatie om het beschikbaar stellen van open data te bevorderen. Bronhoudende organisaties kunnen opdrachtgevers vragen de juiste randvoorwaarden te scheppen, en omgekeerd kunnen opdrachtgevers verzoeken data open te stellen.

Zowel voor organisaties als voor beleidsmakers is het van belang om te blijven beseffen dat open data een middel is om bij te dragen aan doelen binnen of buiten de organisatie. Die doelen motiveren meer dan het middel op zichzelf. Organisaties zullen gemotiveerd zijn om

data te openen als zij verwachten financiële voordelen te bereiken, hun legitimiteit te versterken (door een groter gebruik van data van de organisatie versterkt deze haar bestaansrecht) of een bijdrage te leveren aan de transparantie in hun sector. De noodzaak om tot een expliciete positieve businesscase te komen is vooral van belang als er grote investeringen nodig zijn. Veelal blijken de kosten voor het openen van bestaande datasets echter beperkt en is geen expliciete financiële analyse opgesteld. Het loont daarom de aanbeveling om niet teveel te focussen op het wegnemen van financiële drempels bij organisaties die data-inkomsten niet nodig hebben voor hun voortbestaan. Dit onderzoek heeft zich dan ook niet gericht op de businesscase als onderzoeksobject, maar op het businessmodel dat organisaties im- of expliciet hanteren.

Veel eerder onderzoek (zie hoofdstuk 2) richt zich op organisaties die een omvangrijke of economisch belangrijke registratie beheren en de inkomsten uit de data nodig hebben om die registratie in stand te kunnen houden. In dit onderzoek hebben we veel voorbeelden gevonden van organisaties voor wie de beschikbaarstelling van de data niet tot het primaire proces behoort of behoorde, of die al op een andere manier gefinancierd werden. Zij staan symbool voor een groot aantal publieke organisaties, waarbij 'pricing-vraagstukken' nauwelijks een rol spelen bij beslissingen rondom open data.

Aanbeveling: Focus niet teveel op het wegnemen van financiële drempels bij organisaties die data-inkomsten niet nodig hebben voor hun voortbestaan. Voor de meeste organisaties zijn pricing-vraagstukken nauwelijks aan de orde en kan open data juist geld opleveren.

In beide beschreven businessmodellen blijkt het betrekken van hergebruikers veel baten op te leveren. Als hergebruik bijdraagt aan de doelstellingen van de organisatie, is het ook van belang om het hergebruik effectief te maken. Door de dialoog met hergebruikers kan zicht worden verkregen op hun behoeften. Ook kan het potentieel van anderen om bij te dragen aan de verspreiding van de data en het terugkoppelen van afwijkingen of aanvullingen worden benut.

Aanbeveling: Zoek de dialoog met hergebruikers op. De bronhoudende organisatie en hergebruikers hebben verschillende rollen, maar kunnen gemeenschappelijke belangen hebben.

7.2 Reflectie

De inductieve opzet van het onderzoek maakt dat tamelijk generieke en brede onderzoeksvragen zijn beantwoord met specifieke casestudy's. Het aantal casestudy's is met zeven wat beperkt, en ook verschillen de onderzochte organisaties sterk qua sector, overheidslaag, omvang en belang van informatie voor de organisatie als geheel. Dat maakt

dat de conclusies niet zonder meer generaliseerbaar zijn. In het onderzoek is veel ruimte geweest voor reflectie, met de onderzochte organisaties, met experts, met de opdrachtgever en binnen het onderzoeksteam. Daarbij zijn ook ervaringen en expertise ingebracht die betrekking hebben op andere organisaties. De gevonden businessmodellen kunnen niet op een harde wijze worden gerelateerd aan de onderzoeksdata, maar zijn meer de uitkomst van een puzzel die zich richtte op een herkenbare en bruikbare veralgemenisering van de resultaten.

Het zoeken naar businessmodellen en het in beeld brengen van strategische overwegingen van organisaties gaat uit van een rationeel perspectief op strategievorming. Veel organisaties volgen een rationale of planmatige route voor de besluitvorming rond belangrijke beslissingen, maar er spelen ook andere overwegingen mee. Het feit dat open data 'in de mode' is en andere organisaties het omarmen, kan een belangrijke trigger zijn voor een organisatie. Ook ontstaan sommige inzichten als bij-effect van een impliciete of experimentele strategie; pas op een later moment komt open data dan als thema op de bestuurstafel terecht.

De bevindingen in dit rapport zijn gebaseerd op de situatie ten tijde van het onderzoek. De beslissingen rond open data van de onderzochte organisaties vinden hun wortels meestal in het verleden. Dat kan betekenen dat keuzes anders zouden zijn gemaakt, met de kennis van nu of als de organisaties nu pas tot open data zouden overgaan. Een voorbeeld kan de consequentie daarvan illustreren. De aandacht voor open data is de laatste tijd toegenomen. Er heerst een stimulerend discours, wat eventuele koudwatervrees kan wegnemen en drempels voor het beschikbaar stellen van data verder kan verlagen. Een andere stimulerende prikkel in de omgeving van overheidsorganisaties is de noodzaak tot bezuinigen, die bijna iedereen raakt. Veel van de onderzochte organisaties zien financiële besparing als een belangrijk doel van hun open-datastrategie. Daarmee kan financiële noodzaak een trigger vormen om tot het openen van datasets over te gaan. Het is denkbaar dat organisaties eerder voor het radicale open-datamodel zullen kiezen dan wat we in dit onderzoek hebben gezien, omdat zij genooddaakt worden hun organisatiestrategie te herzien. Open data kunnen daarin een interessante en profijtelijke rol spelen.

Bij de selectie van cases bleek het moeilijk om goede voorbeelden bij decentrale overheden te vinden. De provincies zijn zeer actief met geodata, maar het hergebruik van andere datasets is nog niet gevorderd. Het bleek ook moeilijk om gemeenten te vinden die ruime ervaring hebben met het openen van meerdere datasets. Er zijn weliswaar veel provincies en gemeenten die voornemens zijn om datasets te openen, maar de sets zijn vaak nog niet daadwerkelijk beschikbaar gesteld.

Bij de selectie van de cases is allereerst gekeken naar organisaties, en zijn daarbinnen vervolgens datasets geselecteerd. Dat bleek nodig omdat de open-datastrategie per dataset

kan verschillen. Het belang van een dataset voor de strategie van een organisatie of voor derden kan sterk verschillen per set. Zo zijn bijvoorbeeld de data over de verspreiding van infectieziekten van het RIVM van een andere orde dan de luchtkwaliteitsmetingen; de doelgroepen, hergebruiksdoelen, privacygevoeligheid en de huidige financieringswijze verschillen sterk. Overigens kunnen datasets van verschillende organisaties ook naar elkaar toegroeien. Dit kan leiden tot een organisatie-overstijgende open-datastrategie voor die datasets. Bij het Provinciaal Georegister is zichtbaar dat de provincies hun geodata via gezamenlijke afspraken en voorzieningen ontsluiten. Verdergaande harmonisatie ligt in de lijn der verwachting. Dat betekent ook steeds meer dat keuzes rondom open data niet meer binnen een enkele organisatie worden gemaakt.

Literatuurlijst

- Broek, T.A. van den, et al. (2010). *Open Overheid - Internationale Beleidsanalyse en aanbevelingen voor Nederlands beleid*. TNO.
[http://www.tno.nl/content.cfm?context=thema&content=prop_publicatie&laag1=897&laag2=918&laag3=113&item_id=784]
- Consumentenbond (2011). *Wet BIG beschermt patiënt slecht*.
[<http://www.consumentenbond.nl/campagnes/zorg-campagne/wet-big-bescherm-patient-slecht/>]
- Dekkers, M. et al. (2006). *Measuring European Public Sector Information Resources*.
[http://ec.europa.eu/information_society/policy/psi/docs/pdfs/mepsir/final_report.pdf]
- Gemeente Rotterdam (2010). [<http://emerging-media.rotterdamopendata.nl/>]
- Gemeente Rotterdam (2011a). B en W Bericht 2011-35.
[http://www.bds.rotterdam.nl/Bestuurlijke_Informatie:7/College_van_B_en_W/B_en_W_berichten/2011/B_en_W_bericht_2011_35_collegevergadering_van_20_september_2011]
- Gemeente Rotterdam (2011b) [http://www.rotterdam.nl/rotterdamse_innovatie_agenda]
- Gemeente Rotterdam (2011c). De Rotterdamse Innovatie Agenda
[<http://www.rotterdam.nl/Stadsontwikkeling/Document/Economie%20en%20Arbeidsmarkt/Rotterdamse%20Innovatie%20Agenda%20DEF.pdf>]
- Houghton (2011). *Cost and Benefit of Data Provision*.
[<http://ands.org.au/resource/houghton-cost-benefit-study.pdf>]
- Lundqvist, B. et al (2011). *Business Activity and Exclusive Right in the Swedish PSI Act*.
[http://www.konkurrensverket.se/upload/Filer/Trycksaker/Rapporter/uppdraagsforskning/forskrap_2011-2.pdf]
- Minister van BZK (2011). Kamerstukken 32802 nr. 2, *Hergebruik en Open Data: naar betere vindbaarheid en herbruikbaarheid van overheidsinformatie*.
[<https://zoek.officielebekendmakingen.nl/kst-32802-2.html>]
- Ministerie van Onderwijs, Cultuur en Wetenschappen (z.d.), *Agendapunt 6: Juridisch kader t.a.v. leveren en verwerken gegevens in onderwijsveld*, interne notitie, opvraagbaar.
- Ministerie van Onderwijs, Cultuur en Wetenschappen / DUO, *Business case project portaal openbare onderwijsinformatie*, interne notitie, 4 november 2011, opvraagbaar.
- PIRA (2000). *Commercial exploitation of Europe's Public Sector Information*.
[http://ec.europa.eu/information_society/policy/psi/docs/pdfs/pira_study/commercial_final_report.pdf]
- Pollock, R. (2009). *The Economics of Public Sector Information*.
[http://rufuspollock.org/economics/papers/economics_of_psi.pdf]
- RCE (2011). [<http://beeldbank.cultureelerfgoed.nl/rcc>]
- Stiglitz, J.E., (2000). *The role of government in a digital age*.
[http://archive.epinet.org/real_media/010111/materials/stiglitz.pdf]
- Uhlir, P.F. (ed.) (2009). *The Socioeconomic Effects of Public Sector Information on Digital Networks*
[http://www.nap.edu/catalog.php?record_id=12687]

- Velde, R.A. te (2009). *Public Sector Information: Why Bother?* in Uhler, P.F. (ed.) (2009). *The Socioeconomic Effects of Public Sector Information on Digital Networks*.
[\[http://www.nap.edu/catalog.php?record_id=12687\]](http://www.nap.edu/catalog.php?record_id=12687)
- Vickery, G. (2011). *Review of recent studies on PSI re-use and related market developments*.
[\[http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/psi_final_version_formatted.docx\]](http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/psi_final_version_formatted.docx)
- Voorzieningenrechter Rechtbank 's-Hertogenbosch (2011). Uitspraak sector Bestuursrecht, 4 december 2011. *LJN-nr. BU8010*. [\[http://zoeken.rechtspraak.nl/detailpage.aspx?ljn=BU8010\]](http://zoeken.rechtspraak.nl/detailpage.aspx?ljn=BU8010)
- Vries, M. de, et al. (2011). *Pricing Of Public Sector Information Study (POPSIS)*.
[\[http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/11_2012/models.pdf\]](http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/11_2012/models.pdf)

Bijlage I. Casebeschrijvingen

Het BIG-register van het CIBG

Gesproken met:

- Gerrit Arkesteijn - directeur CIBG
- Michel Kuipers - hoofd ontwikkeling CIBG
- Francis Bolle - woordvoerder en senior adviseur public affairs Verpleegkundigen en Verzorgden Nederland (hergebruiker)

Het CIBG en het BIG-register

Het CIBG is een uitvoeringsorganisatie van het ministerie van Volksgezondheid, Welzijn en Sport. Het CIBG heeft 25 verschillende taken. Het beheert registers waaronder enkele met gecertificeerde gegevens, draagt zorg voor veilige gegevensoverdracht, verstrekt vergunningen en ondersteunt toetsingscommissies. Het merendeel van de taken betreft de zorg, maar het CIBG is ook verantwoordelijk voor het register voor diergeneeskunde (in opdracht van het ministerie van Economie, Landbouw en Innovatie) en het lerarenregister (in opdracht van het ministerie van Onderwijs, Cultuur en Wetenschap).

Het CIBG beheert onder andere de volgende registers:

- Register van eicel- en spermadonoren, het KIDsS (Kunstmatige Inseminatie donorensysteem)
- Het donorregister
- Het BIG-register (Beroepen in de Individuele Gezondheidszorg)
- Register voor diergeneeskunde
- Lerarenregister
- Register van de late zwangerschapsafbrekingen en levensbeëindiging van pasgeborenen
- Het UZI-register
- Het register voor apothekhoudende geneesheren
- Het zorgregister (ten behoeve van privéklinieken en privé verpleeg- en verzorgingstehuizen)

De gegevens in het BIG-register worden via een website en webservice beschikbaar gesteld. Het is ook de bedoeling dat de gegevens in het lerarenregister per 1 februari 2012 via een webservice beschikbaar worden gesteld. Deze case gaat over het BIG-register. In het BIG-register kunnen zorgverleners zich registreren. Deze registratie maakt de bevoegdheid van een zorgverlener duidelijk. Een BIG-registratie is noodzakelijk om in de zorg als zorgverlener te kunnen werken. Ook worden verklaringen van vakbekwaamheid

afgegeven aan zorgverleners met een buitenlands diploma die in Nederland willen werken, om in het BIG-register te kunnen worden ingeschreven. Op de website bestaat ook een 'zwarte lijst' van hulpverleners die (tijdelijk) hun werk in de zorg niet meer mogen doen.

Kenmerken van de data

Het BIG-register komt voort uit de Wet op de Beroepen in de Individuele Gezondheidszorg. Volgens de Wet BIG worden de volgende beroepen geregistreerd in het BIG-register: apotheker, arts, fysiotherapeut, gezondheidszorgpsycholoog, psychotherapeut, tandarts, verloskundige en verpleegkundige. Deze beroepsbeoefenaars vragen, met een geldig diploma, een BIG-registratie aan. De kosten voor registratie bedragen € 80. Een inschrijving in het BIG-register betekent het zelfstandig mogen uitvoeren van een beroep in de zorg en een beroepstitel mogen voeren. Vanaf 1 januari 2009 moeten beroepsbeoefenaars die in het register zijn ingeschreven en dit willen blijven, zich na 5 jaar herregistreren. In het BIG-register zijn op dit moment 400.000 mensen ingeschreven.

In het BIG-register is van elke geregistreerde onder meer naam, BSN, geslacht, beroep, specialisme en woonadres opgenomen. Het BIG-register gebruikt in beginsel de woonadressen, omdat die gekoppeld zijn met de GBA en daardoor up-to-date zijn. Het is mogelijk het BIG-register te bevragen met het adres als zoekleutel. Bij veel patiënten en cliënten zal dit woonadres echter niet bekend zijn. Daarom verzoekt het CIBG beroepsbeoefenaars ook het werkadres op te geven. 10% van de beroepsbeoefenaars hebben dit gedaan. Voor patiënten blijkt het ontbreken van een werkadres het grootste probleem te zijn om de zorgverlener in het BIG-register te vinden (Consumentenbond, 2011).

Beschikbaarstelling van de data

De gegevens in het BIG-register worden via een website (www.bigregister.nl) en webservice ter beschikking gesteld. Ongeveer zeven jaar geleden is het CIBG overgegaan op een servicegeoriënteerde architectuur (SOA) voor de ontsluiting van de data op de website. Ook is het CIBG toen begonnen met het vernieuwen van het applicatieportfolio. Het werken binnen een SOA maakt het relatief eenvoudig en goedkoop om ook derden toegang te geven tot de BIG-gegevens.

De gegevens in het BIG-register zijn in 2005 ook via een webservice ontsloten. Het CIBG wilde daarmee (op dat moment) vooral de grote zorginstellingen beter bedienen. Alle instellingen zijn verplicht om de BIG-registratie te controleren bij het in dienst nemen van nieuw medisch personeel. In 2005 was de verwachting dat op korte termijn de HRM-softwareleveranciers gebruik konden gaan maken van de webservice, om de HR-medewerkers van zorginstellingen te ontlasten.

Het BIG-register werd stilzwijgend als webservice beschikbaar gesteld en zonder veel ruchtbaarheid. Binnen een maand waren er al de eerste gebruikers voor de webservice. Een uitgever was een van de eersten. Het CIBG kreeg meteen de vraag naar de beschikbaarheid van de webservice. De uitgever wilde ervan op aan kunnen dat het register in de toekomst beschikbaar zou blijven en dat de webservice de beschikbaarheid van de eigen website niet zou verstoren.

Gebruik en gebruikers

Het CIBG kent de gebruikers niet. Het kent alleen de IP-nummers van bezoekers van de website en ontvangt daarnaast af en toe individuele vragen. Wel kent het CIBG een aantal voorbeelden van hergebruik. Zo geven sommige zorgverzekeraars korting als iemand staat ingeschreven in het BIG-register.

De verhouding gebruik website/webservice is ongeveer fiftyfifty. Zorginstellingen zijn grote gebruikers van de CIBG website, maar zoeken nog regelmatig telefonisch contact. Ze toetsen of een nieuwe medewerker geregistreerd is. Het is opvallend dat grote zorginstellingen weinig gebruik maken van de webservice, aangezien de webservice in eerste instantie voor hen bedoeld was. Het merendeel van de gebruikers die de webservice gebruikt, zijn nieuwe gebruikers. Uitgeverij Springer is een voorbeeld. Springer Verlag geeft klanten die geregistreerd zijn in het BIG korting op de producten. Wanneer een persoon zijn/haar BIG-nummer en huisadres opgeeft voor de factuur, dan controleert de website automatisch of de klant BIG-geregistreerd is en geeft vervolgens korting.

Het aantal hits op de website is 80.000 tot 90.000 per maand. Hetzelfde aantal geldt voor de webservice. Uit een top-10 van IP-adressen over de maanden juli en augustus 2011, blijkt dat er intern bij het CIBG het meeste van gebruik gemaakt wordt. Zorginstellingen zijn een andere belangrijke hergebruiker. Individuen controleren ook in het BIG-register of bijvoorbeeld de arts bevoegd is of niet. Gemiddeld genomen worden maandelijks 1944 telefonische vragen bij de informatielijn gesteld (periode januari - oktober 2011). Via de mail zijn dit er maandelijks 934 in dezelfde periode. De website en webservice zijn wat betreft aantallen het belangrijkste communicatiekanaal.

Strategie

Het CIBG is verplicht bij een vraag van een burger over het BIG-register informatie te verstrekken. De organisatie heeft er bewust voor gekozen dit via het goedkoopste kanaal te doen (kanaalsturing): via website en webservices. De servicegeoriënteerde architectuur waar het CIBG op over is gegaan, maakte dit relatief gemakkelijk mogelijk. Webservices zijn volgens het CIBG de manier waarop data-uitwisseling zou moeten plaatsvinden. Het hoofdargument hiervoor is dat het potentieel veel efficiencywinst en administratieve lastenreductie oplevert bij data-uitwisselingen tussen overheidsorganisaties en tussen overheidsorganisaties en burgers en bedrijven.

Het openen van de data via webservices is verder ongeschreven beleid. De belangrijkste motivatie was de vraag 'waarom niet?'. De kosten voor het CIBG zijn zeer laag. CIBG visie tegenover open data: "Je moet het gewoon doen, met als randvoorwaarde dat je een goede infrastructuur hebt". Het aanbieden van een applicatie past ook binnen deze kanaalsturingsfilosofie. Hier heeft het CIBG al mee geëxperimenteerd. Er is een applicatie beschikbaar waarin via layar - een applicatie die omgevingsinformatie die met de camera van een smartphone wordt ontsloten, combineert met andere informatie - informatie over een zorginstelling gezien kan worden. Het CIBG is hierbij wel van mening dat zij het niet meer hoeft aan te bieden als er voldoende concurrenten zijn - zowel wat betreft applicaties als de website en webservice.

Het CIBG heeft er bewust voor gekozen om de gebruikers niet te kennen. De organisatie ziet het als taak om data op toegankelijke wijze openbaar te maken. De creativiteit om iets met de data te doen is aan de markt. Op dit moment is er geen communicatie met gebruikers, hoewel het CIBG daar meer behoefte aan krijgt, bijvoorbeeld met zorginstellingen die nog weinig gebruikmaken van de webservice.

Kosten en opbrengsten

De kosten voor de webservice bedroegen eenmalig € 15.000. Dit wordt afgeschreven over 6 jaar. Er zijn geen exploitatiekosten.

De opbrengsten zijn kwalitatief. Het CIBG stelt zich op als een dienstverlenende overheidsorganisatie. Een aantal bedrijven kunnen leuke dingen doen met overheidsinformatie. Intern zijn er nog geen opbrengsten, omdat weinig andere overheden met webservices werken. Het maatschappelijk gebruik van het BIG is wel toegenomen.

Als andere overheden ook met webservices gaan werken, dan kan het geld opleveren. Het CIBG zou bijvoorbeeld drie à vier fte (€ 300.000,- op jaarbasis), kunnen besparen wanneer DUO automatisch doorgeeft als iemand in de sector afgestudeerd is en ingeschreven kan worden in het BIG-register. Dit scheelt ook administratieve lasten voor afgestudeerden, omdat zij de gegevens niet hoeven aan te leveren bij de BIG-registratie.

Recente ontwikkelingen

Het CIBG denkt verder in de rol als dienstverlenende publieke organisatie. Een persoon die in het BIG-register staat ingeschreven, kan dit in de toekomst exporteren naar LinkedIn. Via een 'knop' in iemands LinkedIn-pagina kan hij/zij laten zien of hij/zij BIG-geregistreerd is. Wederom kost dit het CIBG niets, terwijl de dienstverlening aan burgers en het belang van het register toenemen. Hetzelfde geldt voor de applicatie die het CIBG heeft ontwikkeld.

Hergebruik door Verpleegkundigen en Verzorgenden Nederland

Verpleegkundigen en Verzorgenden Nederland (V&VN) is een beroepsvereniging van en voor verpleegkundigen en verzorgenden. V&VN dicht veel waarde toe aan het BIG-register. Het registreert immers wie wel en niet bevoegd is. Vaak attenderen zij patiënten op het BIG-register, zodat zij kunnen achterhalen of de verpleegkundige bevoegd is.

V&VN gebruikt het BIG-register voor haar eigen kwaliteitsregister. V&VN vindt het niet voldoende dat bij herregistratie in het BIG-register alleen wordt gekeken naar het aantal uur patiëntenzorg dat iemand geeft. Kwalitatieve aspecten, zoals het volgen van een opleiding, tellen voor V&VN ook mee. Verpleegkundigen en verzorgenden kunnen zich in dit kwaliteitsregister inschrijven. Als iemand zich inschrijft, wordt hij/zij via een webservice gecontroleerd in het BIG-register. V&VN vindt de website door de jaren heen gebruiksvriendelijker geworden. Zo is het bijvoorbeeld gemakkelijker geworden om iemand op te sporen.

De waarde van het BIG-register en het openstellen daarvan, is volgens V&VN niet in euro's uit te drukken. Een registratie in het BIG-register is een kwaliteitsstempel. Het beschermt een bepaalde beroepsgroep tegen het illegaal dragen van een titel. Het is dan ook niet meer dan normaal dat je het openbaar maakt, want een patiënt moet dit kunnen controleren. Die transparantie zorgt voor bescherming van de patiënt en geregistreerde beroepsbeoefenaars. Als het niet openbaar wordt gemaakt, dan is er geen controle op de Wet BIG mogelijk en is het niet bekend of er met bevoegdheden geknoeid wordt. Nu kan in luttele seconden gecheckt worden of iemand bevoegd is. Als dat niet kan, dan kan een persoon schade aandoen, zowel aan patiënt als aan de beroepsgroep. Openbaarheid is cruciaal om kwaliteit naar de buitenwereld te kunnen uitdragen. Transparantie blijkt de belangrijkste waarde van het beschikbaar stellen van data uit het register.

Erfgoed informatie van de RCE

Gesproken met:

- Carla Schulte - hoofd Archief, Bibliotheek en Collecties, RCE
- Kees Hendriks - projectmanager Kennisinfrastructuur modernisering monumentenzorg, RCE
- Dirk Houtgraaf - hoofd sector Kennisuitwisseling, RCE
- Hans Berrens - controller, RCE
- Arjan den Boer - oprichter en eigenaar, ABC Media (hergebruiker)

De Rijksdienst voor het Cultureel Erfgoed

De Rijksdienst voor het Cultureel Erfgoed (RCE) houdt zich bezig met het behoud van het Nederlands cultureel erfgoed. Het gaat hierbij om archeologie, monumenten, cultuurlandschappen en roerend goed. Op grond van de Monumentenwet 1998 is de wettelijke taak van de RCE het aanwijzen van monumenten, het verlenen van vergunningen (monumentenvergunning, opgravingsvergunning) en het verlenen van subsidies voor de instandhouding en restauratie van rijksmonumenten. Om het cultureel erfgoed te behouden en toegankelijk te maken, draagt de RCE zorg voor een kennisinfrastructuur, die niet alleen door de RCE, maar ook door andere erfgoedpartijen gebruikt wordt. De RCE heeft geen publieksopdracht, maar richt zich voornamelijk op (het faciliteren van) een netwerk van ketenpartners die zich gezamenlijk bezighouden met het beheer van het Nederlands cultureel erfgoed.

De rijksdienst valt onder het ministerie van Onderwijs, Cultuur en Wetenschap en wordt daar ook door gefinancierd. Hoewel de RCE een onderdeel is van het ministerie, voert zij toch een relatief zelfstandige bedrijfsvoering. Zo zijn locaties en voorzieningen gescheiden en is ook de informatievoorziening zelfstandig.

Kenmerken van de data

De RCE beschikt over veel verschillende typen data, van boeken en tijdschriften tot foto's en tekeningen, en de RCE beheert ook een aantal registers waaronder het rijksmonumentenregister en het archeologische vondsten- en onderzoeksregister (ARCHIS). Omdat de RCE ook verantwoordelijk is voor vergunningverlening en subsidieverstrekking houdt het ook uitgebreide correspondentie bij over deze vergunningsprocessen.

Alle bezittingen van de RCE zijn gecategoriseerd in de afdeling Archief, Bibliotheek en Collecties. Het archief omvat onder andere de documentatie van de RCE betreffende de 60.000 rijksmonumenten en meer specifiek, de correspondentie rondom subsidieaanvragen en vergunningverleningen, fotoafdrukken en bouwkundige tekeningen. De bibliotheek bevat de eigen uitgaven en circa 126.000 boeken over onderwerpen die van belang zijn voor de

RCE en haar partners. De collectie bestaat uit de beeldbank met circa 500.000 afbeeldingen (voornamelijk foto's) en referentiecollecties die door de RCE beheerd worden. De data die beheerd worden zijn de metadata over collectie en bibliotheek, en ook de data die in het archief besloten liggen. Ook beeld- en geluidsmateriaal in de collectie kan als data gezien worden.

Veel van de data die de RCE bezit is oud, al lang in het bezit van de RCE, en daardoor niet digitaal. In 2010 is een project gestart om de post en het archief - met een totale omvang van 2000 meter - te digitaliseren. Het doel van deze digitalisering was zowel een efficiëntere en effectievere bedrijfsvoering, als een effectievere ontsluiting van data voor zowel interne als externe doelen.

Veel van de data die de RCE beheert, is inmiddels openbaar. Dat geldt bijvoorbeeld voor de beeldbank en het monumentenregister en de metadata van de collecties en bibliotheek. De data zijn niet altijd openbaar geweest, omdat de RCE in het verleden van mening was dat dit vanwege privacy-aspecten ongewenst was. Bij het openbaar maken van de data bestond, ondanks een kwaliteitsslag die van te voren gemaakt was, het risico dat burgers of bedrijven protest zouden aantekenen tegen de openbaarmaking; dat kan bijvoorbeeld gebeuren bij het openbaarmaken van interieurfoto's. In 2009 heeft de RCE, ondanks de bestaande risico's een bewuste keuze voor openbaarmaking gemaakt, Belangrijk resultaat was echter dat door massale en snelle openstelling ook energie ontstond in de organisatie voor verdere digitalisering en openbaring. De RCE heeft bij alle digitaal beschikbare data de volgende regel opgenomen:

"Meent u desondanks dat de digitale beschikbaarstelling van bepaald materiaal inbreuk maakt op enig recht dat u toekomt of uw (privacy)belangen schaadt, dan kunt u dit onderbouwd aan ons laten weten. Wij zullen dan op korte termijn contact met u opnemen en bij een gegronde klacht vanzelfsprekend het materiaal van de website verwijderen" (RCE, 2011).

In sommige gevallen lijken er juridische beperkingen te zijn, doordat derden auteursrechten hebben op stukken die in de archieven zijn opgenomen. Dat is bijvoorbeeld het geval bij rapporten die in het archeologische-vindplaatsenregister zijn opgenomen. Juridisch onderzoek door de RCE wees echter uit dat de openbaarmaking zwaarder woog dan het auteursrecht dat op de stukken lag (RCE Intranet, 25 juli 2011).

De RCE gebruikt op dit moment slechts één licentiemodel. Voor bijna alle materialen geldt dat hergebruik onder de Creative Commons Share Alike 3.0 licentie (CC-BY-SA) valt. Dat betekent dat iedereen de data vrijelijk mag hergebruiken, ook voor commerciële doeleinden, maar dat afgeleide producten ook onder dezelfde licentie moeten worden verstrekt. Dit vanuit het argument; 'De belastingbetaler heeft er al voor betaald', en ook omdat de

marginale kosten voor verspreiding toch al nul zijn. Het blijft overigens wel zo dat voor specifieke handelingen (bijvoorbeeld hoge kwaliteit scans op opdracht) er een kostprijs zal worden gerekend.

Voor inzage van het archief en de registers, alsmede ook het inschrijven van nieuwe objecten in de registers worden op dit moment geen leges gerekend. De RCE vindt dat zolang deze voorzieningen niet onoverkomelijk belast worden, dit als onderdeel van het aan haar toegedeelde budget gezien moet worden.

Beschikbaarstelling van de data

De Rijksdienst voor het Cultureel Erfgoed stelt op vele manieren data ter beschikking aan hergebruikers en ketenpartners. Een aantal plekken waar RCE data aangetroffen kan worden:

- In de Beeldbank (<http://beeldbank.cultureelerfgoed.nl/>) zijn meer dan 550.000 foto's en afbeeldingen opgenomen.
- Op Wikipedia is het monumentenregister ontsloten (inclusief monument nummer en redengevende omschrijving). Ook via het [monumentenregister online](#) zijn deze gegevens beschikbaar.
- Verder ontsluit de RCE gedigitaliseerde museale collecties van aangesloten musea via een zoekomgeving, gebaseerd op open source, in de Digitale Museale Collectie Nederland. Daarbovenop zijn geavanceerde zoektoepassingen ontwikkelt (www.dimcon.nl).
- Via Kennisinfrastructuur Cultuurhistorie ([KICH](#)) wordt een aantal geo-datasets ontsloten.
- Vanaf januari 2011 is de RCE nationale aggregator voor roerend erfgoed binnen het Europeana netwerk.

De RCE biedt (in tegenstelling tot de gebruikelijke filosofie) geen 'ruwe' data aan. De RCE hanteert een informatiestrategie waarin bronnen (de data laag) door middel van een semantische informatielaag (verbindingslaag) ontsloten worden voor applicaties en gebruikers (de e-dienstenlaag).i. Doordat de bronnen niet 'ruw' ontsloten worden, kan de RCE ervoor zorgen dat alleen betekenisvolle entiteiten ontsloten worden, die eenvoudiger voor hergebruik in te zetten zijn. Elk object dat uit de verbindingslaag opgevraagd wordt, is immers al voorzien van identificerende nummers en gebruikt alleen begrippen die al in de thesaurus zijn opgenomen. Ook zijn objecten vanuit verschillende databronnen (maar van hetzelfde type) geüniformeerd.

De objecten die uit de verbindingslaag op te vragen zijn, worden door de RCE "halffabricaten van informatieproducten" genoemd. De halffabricaten van de RCE zijn te vergelijken met informatiebouwstenen zoals die in "Linked Open Data" gebruikt worden.

Zoals ook in de architectuur afbeelding te zien is, is de informatievoorziening van de RCE zoveel mogelijk via services vormgegeven. Deze moderne service oriented architecture (SOA) is een van de resultaten van het project KiMoMo (Kennisinfrastructuur modernisering monumentenzorg), een project dat met subsidiegelden uit de PRIMA-regeling (ter ondersteuning van de uitvoering van de ICT Agenda van het kabinet) is uitgevoerd. Een van de belangrijke uitgangspunten bij het KiMoMo-project was dat de informatievoorziening zowel voor de RCE zelf, alsook voor partners goed bruikbaar moest zijn. Een gevolg van SOA is ook dat het ontsluiten van halffabricaten altijd via webservices gedaan wordt.



Gebruik en gebruikers

De RCE richt zich in haar functioneren vooral op eigenaren van erfgoed, intermediairs en andere erfgoed beherende organisaties en kent zogezegd geen publieksopdracht. Dat wil zeggen dat de meeste hergebruikers geen individuele burgers, maar professioneel betrokkenen en intermediaire organisaties zijn. De belangrijkste hergebruikers die er tot nu toe zijn:

- Wikipedia - hergebruikt data uit het monumentenregister. Van elk rijksmonument is het adres, het monumentnummer en de oorspronkelijke functie opgenomen. Omdat niet van alle monumenten ook foto's beschikbaar waren, is er door Wikipedia Nederland, de RCE en nog een aantal andere partners een wedstrijd uitgeschreven: Wiki Loves Monuments. In zowel 2010 als 2011 zijn er in september door honderden fotografen meer dan 10.000 foto's van monumenten gemaakt en vrijgegeven (in Wikimedia Commons) onder de CC-BY-SA licentie. De RCE beschikt daardoor over veel meer beeldmateriaal van de door haar beheerde monumenten. Ook blijkt dat Wikipedia-gebruikers de omschrijvingen van de monumenten aanvullen met eigen kennis. Zowel de kwaliteit als de kwantiteit van de meldingen hebben de RCE verrast. Inmiddels wordt er gewerkt aan een proces om deze toegevoegde informatie ook weer in de systemen van de RCE op te nemen. Bijkomend voordeel is overigens ook nog dat de RCE geen eigen tool meer hoefde te bouwen voor het publiek ontsluiten van het monumentenregister. Medewerkers zoeken op Wikipedia snel en eenvoudig aan de hand van het monumentnummer of de plaats- en straatnaam het betreffende pand op.
- AB-C Media - zie laatste paragraaf / grijze blok.
- Meerdere heemkringen, lokale archieven en andere ketenpartners gaan gebruik maken van de erfgoedsuite. De erfgoedsuite is een Software as a Service (SaaS)-

oplossing voor het beheren en presenteren van collecties, die gebruik maakt van de thesaurus, semantische standaarden en halffabricaten die binnen de RCE ontwikkeld zijn. Op die manier krijgen de ketenpartners een kosteneffectieve oplossing om collectiebeheer te doen, en kunnen ook kleine collecties (digitale) verbindingen maken met andere partijen. Op haar beurt kan de RCE gebruik maken van en kennis verzamelen over de collecties van de ketenpartijen door deze tool aan te bieden.

- Vereniging Hollandse Molen levert molengegevens, die worden ingezet in de molenwidget, waarmee een groep experts kan werken aan de beschrijvingen van molens en aan de thesauri op dit terrein.

De RCE hoopt op nog meer initiatieven, maar andere grootschalige informatie uitwisselingen zijn nog niet opgezet.

Het aantal bezoeken aan de centrale balie van de RCE voor het kopen van beeldmateriaal is afgenomen van enige honderden tot bijna nul, omdat dit nu helemaal online kan. Wel is er nog steeds regelmatig - en in toenemend aantal - bezoek bij de RCE voor de bibliotheek en de archieven (met onder andere de uitgebreide informatie over de monumenten). Het archief van 2000 tot heden is digitaal beschikbaar tot de plaatsnaam beginnend met de letter 'L' (1/1/2012) en slechts gedeeltelijk online ontsloten. Het digitaliseringstraject van het archief voor de periode 1920 tot 2000 volgt aansluitend.

Strategie

De RCE heeft tot taak om het beheer en behoud van het cultureel erfgoed te ondersteunen en stimuleren. Binnen de budgettaire, juridische en organisatorische kaders waarin de RCE opereert, is dit echter niet te doen voor één organisatie. De RCE heeft vele partners nodig die gezamenlijk bijdragen aan de zorg voor het cultureel erfgoed. De RCE kan daardoor gezien worden als een facilitator van een veel grotere netwerkorganisatie, waarin de onderdelen van het netwerk ieder op hun eigen specifieke gebied verantwoordelijkheid (kunnen) nemen voor het Nederlandse erfgoed.

Vanuit haar wettelijke taken, heeft de RCE onder andere de verantwoordelijkheid voor het ontsluiten van een aantal datasets, waaronder de diverse registers. Vanuit het overkoepelende doel om juist ook andere organisaties in het netwerk te helpen, is er voor gekozen om de ontsluiting van data veel breder op te pakken. Onder andere door de ontsluiting van halffabricaten, de thesauri en meta-datastandaarden kan de RCE een veel grotere impact hebben op de gehele erfgoedsector. De RCE is van mening dat ontsluiting van data via internet een goede en goedkope manier is, om de impact die zij in de samenleving en de erfgoedsector heeft, te vergroten.

De achtergrond van deze strategie ligt in de beleidsbrief modernisering monumentenzorg (2010) en de start van de digitalisering en de beeldbank. Met modernisering monumentenzorg vond een verschuiving van gerichtheid op regelgeving en controle naar

zelfregulering (belanghebbende als drager van monumentenzorg) en kennisontsluiting plaats. Een interne oriëntatie gericht op kennisbehoud maakte plaats voor een externe oriëntatie gericht op kennisdelen en samenwerken. Daarbij was het essentieel dat ook de informatievoorziening op de schop ging.

De RCE is ervan overtuigd dat je als moderne overheidsorganisatie niet moet willen proberen alles onder controle te houden. Haar werkwijze kenmerkt zich dan ook door design by trial. De RCE ziet open data als een continu proces dat winst voor de organisatie en maatschappij oplevert.

Kosten en opbrengsten

De kosten die de RCE in verband brengt met open data, bestaan uit twee categorieën kosten.

- Ten eerste zijn er de projectkosten voor KiMoMo en andere digitalisatieprojecten.
 - Digitaliseren pandsdossiers: € 3,2 miljoen (project van 4/5 jaar);
 - Project digitaal werken € 1,8 miljoen
 - KiMoMo Kennisinfrastuctuur ontwikkelen: € 5 miljoen (project loopt tot eind 2012), gefinancierd voor 50% uit een externe subsidie uit de PRIMA-gelden.
 - Digitaal
- Ten tweede zijn er de kosten voor het beheren en onderhouden van de gerealiseerde voorzieningen.
 - Beeldbank: € 60.000,- per jaar;
 - Hiernaast zijn er hosting, opslag, verwerkingskosten en onderhoudskosten - hiervoor is geen inschatting gemaakt.

Er zijn ook baten voor de RCE. In tegenstelling tot de kosten zijn de baten allemaal structureel.

- De interne baten zijn:
 - Besparing van 7 fte op jaarbasis door efficiëntere organisatie met minder administratieve lasten (als gevolg van de drie projecten). Deze reductie kan vanaf 2013 met de taakstelling ook gerealiseerd worden. De fte-reductie komt neer op € 500.000,- per jaar. Op duurdere onderhoudscontracten kan er € 300.000,- per jaar bespaard worden. Dat is een totaal van € 800.000,- en betekent een terugverdientijd van ongeveer 10 jaar ten opzichte van de projectkosten voor de organisatie (van € 7,5 miljoen). De uiteindelijke baten in fte's zouden op termijn nog wel eens veel hoger kunnen uitvallen, maar daar is nu nog te weinig zicht op.
 - Effectiviteitswinst: snellere, correctere en completere inzage in archieven, online tracking van subsidieaanvragen, minder fouten en een nauwkeurigere procesgang, de mogelijkheid van het koppelen van bestanden en het uitvoeren van analyses (projectplan digitalisering post/archief);

- Structurering kennis en kwalitatief hoogwaardiger archief (door opschoning).
- De externe baten zijn:
 - Gemeenten hebben minder tijd nodig voor het opvragen van monumenten dossiers (ca. 7 minuten per opvraging)
 - Ook vele andere partijen in het erfgoed-veld hebben sneller toegang tot meer complete informatie.

Wanneer we deze baten uitzetten tegen de lasten, blijkt dat de investeringen in ongeveer 10 jaar terugverdiend worden door de RCE. Het is echter onmogelijk om een scheiding aan te brengen tussen de kosten voor het 'open-data'-beleid aan de ene kant en de kosten voor het verbeteren van de interne informatievoorziening aan de andere kant. Vanuit de filosofie van de RCE is dat overigens ook onzin, omdat de interne informatievoorziening vooral geoptimaliseerd wordt om met meer ketenpartners meer informatie te kunnen delen.

Recente en voorziene ontwikkelingen

De RCE ziet open data als onderdeel van de bedrijfs- en informatiestrategie en als een continu proces. In ieder geval streeft de RCE naar een opvolger voor KiMoMo om de infrastructuur, het gereedschap en de toepassingen verder te brengen. De beeldbank wordt opengesteld voor deelname door andere partners in het erfgoedveld en bij betrokkenen en geïnteresseerden in het erfgoedveld, wordt gekeken of er via een crowdsourcing-instrument aan kwaliteitsverbetering kan worden gedaan.

Hergebruik door AB-C Media

AB-C Media is een bedrijf dat websites, apps en ander publicatiemateriaal voor culturele instellingen maakt. AB-C Media heeft een aantal apps gebouwd waarin informatie over monumenten ontsloten wordt, waaronder Erfgoed.mobi (<http://www.ab-c.nl/erfgoed-mobi>). Erfgoed.mobi maakt gebruik van meerdere API's van de RCE. Erfgoed.mobi is een demonstratie app, die vooral de technische kwaliteiten van AB-C Media moet tonen en die in acquisitie gebruikt wordt. De waarde van dit instrument is volgens Arjan den Boer "niet in geld te meten", maar desgevraagd gaf Arjan aan niet te willen stoppen met het gebruiken van deze tool, zelfs niet als hij daarvoor € 100.000,- zou ontvangen.

Luchtkwaliteit- en emissiedata bij het RIVM

Gesproken met:

- Bennie Bloemberg - hoofd i-organisatie
- Wim van der Maas - projectleider emissieregistratie
- Ivo Stumpe - senior inhoudelijk medewerker, Milieudefensie (hergebruiker)
- Daan Swart - hoofd afdeling luchtkwaliteitsmetingen
- Reinout Woittiesz - plaatsvervangend directeur-generaal
- Piet van Zoonen - hoofd centrum monitoring

Het Rijksinstituut voor Volksgezondheid en Milieu (RIVM)

De voornaamste taak van het RIVM, is het leveren van ondersteuning op het gebied van volksgezondheid en milieu door hierover kennis te genereren, te behouden en in te zetten. Het RIVM wil een neutrale kennisbron zijn voor overheden en het publiek. De organisatie zal echter nooit een handelingsperspectief toevoegen aan de informatie die ter beschikking gesteld wordt. Alle taken van het RIVM worden in opdracht van overheden uitgevoerd. In de Wet op het RIVM is een grote mate van vrijheid vastgelegd voor de wijze waarop de opdracht wordt uitgevoerd. De organisatie is een baten-lastendienst van het ministerie van VWS. In deze case zijn twee datasets onderzocht, die zich beide binnen de sector Milieu van het RIVM bevinden: namelijk data over de luchtkwaliteit en over emissies.

Kenmerken van de data

Via circa 50 meetstations, verspreid over heel Nederland, worden met hoge frequentie de concentratie van verschillende stoffen in de lucht (dicht bij de grond) gemeten. Het gaat onder andere om fijnstof, broeikasgassen en ammoniak. De data worden door de meetstations automatisch doorgestuurd naar de database van het RIVM. De data worden vervolgens gepubliceerd als niet-gevalideerde data. Achteraf worden de metingen gevalideerd, onder andere door afwijkende waarden te analyseren. De gevalideerde data worden opnieuw gepubliceerd.

Naast de publicatie van de data, gebruikt het RIVM de data in een aantal publicaties, waarbij de gemeten waarden worden vergeleken met wettelijke normen. Het RIVM stelt een luchtkwaliteitsverwachting op. Het smogalarm is daarvan een onderdeel.

De tweede dataset betreft de emissieregistratie. Dat is een database die de emissies richting bodem, lucht en water bevat, zoals die zijn gemeten of zijn afgeleid uit andere metingen. Circa tien organisaties werken samen om de data aan te leveren, het RIVM beheert de dataset.

Beschikbaarstelling van de data

De luchtkwaliteitsdata worden beschikbaar gesteld op een website (www.lml.rivm.nl), waarbij het mogelijk is tabellen en kaartjes te genereren op basis van een eenvoudige query. Voor professionele gebruikers die daarom verzoeken is daarnaast toegang via XML mogelijk. De beschikbaarstelling van luchtkwaliteitsdata is overigens niet nieuw; vroeger werden al tabellenboeken uitgegeven.

De website voor de emissieregistratie (www.emissieregistratie.nl) maakt het eveneens mogelijk tabellen en kaartjes te genereren. Het is ook mogelijk om een uitgebreide query op de database te maken en de resultaten daarvan te exporteren. Vragen over de data kunnen gesteld worden via een helpdesk, die jaarlijks enkele honderden vragen beantwoordt.

Gebruik en gebruikers

Grofweg zijn drie typen gebruikers te onderscheiden:

- Overheden. Zij gebruiken de milieudata om beleid te maken en/of te evalueren. Een bijzonder proces daarin is de periodieke rapportage aan de Europese Commissie over het voldoen aan Europese normen.
- Professionele gebruikers, met name ingenieursbureaus en milieu- en natuurorganisaties.
- Burgers. Zij beschikken volgens het RIVM vaak over minder kennis om de data te kunnen duiden dan professionele gebruikers.

De data van het RIVM kunnen worden aangeboden door intermediairs, maar vanuit de organisatie zelf is daar weinig zicht op. Hergebruik door het Twitter-account @RIVMalarm is wel opmerkelijk. Het account is geen eigendom van het RIVM, maar het RIVM heeft desalniettemin geen actie ondernomen om gebruik van haar naam tegen te gaan, uit een pragmatische afweging: het gebruik was niet schadelijk.

De website van de emissieregistratie kent ongeveer 200 unieke hits per dag, waarvan vermoed wordt dat het in 80% van de gevallen om professionele gebruikers gaat.

Strategie

De verzameling en publicatie van de data volgt uit wettelijke verplichtingen. De belangrijkste daarvan is het Verdrag van Aarhus aangaande de toegang tot informatie over milieu-aangelegenheden. Verzameling en publicatie zijn onlosmakelijk met elkaar verbonden; het een is zonder het ander niet mogelijk of doelloos.

Het RIVM bewaakt de positie als onafhankelijke kennisleverancier. Zij voorziet de data daarom niet van betekenis, anders dan dat zij goed beschrijft wat wel en niet gemeten is en hoe betrouwbaar de cijfers zijn.

Er is geen overkoepelende open-datastrategie binnen de sector Milieu of het RIVM als geheel. Afwegingen over de beschikbaarstelling van data worden gemaakt op het niveau van een individuele dataset. Momenteel wordt wel gepoogd om de verschillende elektronische kanalen onder een label te brengen (RIVM-website), maar er is nog geen sprake van integratie van de datasets of harmonisatie van de semantiek.

Vanwege de wettelijke verplichting tot publicatie is de vraag niet of de data worden gepubliceerd, maar resteert vooral de vraag op welke wijze dat gebeurt.

Kosten en opbrengsten

De luchtkwaliteitswebsite kost circa € 100.000,- per jaar. De website voor het ontsluiten van de data van de emissieregistratie brengt € 250.000,- aan jaarlijkse kosten met zich mee. Financiële opbrengsten zijn niet direct aanwezig, anders dan de financiering vanuit de ministeries van IenM en EL&I. Met de beschikbaarstelling via de website zijn de kosten voor de papieren kanalen (met name tabellenboeken) wel vervallen of stevig teruggebracht.

Recente en voorziene ontwikkelingen

Publieke organisaties en kennisinstellingen werken steeds meer samen, in netwerkvorm, om de doelen te bereiken. Het RIVM staat open voor die samenwerking, zoals ook blijkt uit de wijze waarop de data voor de emissieregistratie wordt verzameld. Zij zal echter vasthouden aan een onafhankelijke positie en om die reden altijd een bijzondere verantwoordelijkheid voor de kwaliteit en correctheid van de data willen behouden. Het toekennen van het 'label RIVM' aan de data heeft ook een waarde voor gebruikers en zal om die reden niet snel aan een andere organisatie worden overgelaten.

Hergebruik door Milieudefensie

Milieudefensie is een organisatie die zich op verschillende niveaus (lokaal, nationaal, internationaal) inzet voor een beter milieu. Verkeer is een van haar bijzondere aandachtsgebieden. Milieudefensie is tevreden met het aanbod en de onafhankelijkheid van het RIVM op het gebied van luchtkwaliteitsdata.

Milieudefensie is een professionele gebruiker en heeft geen nadere wensen ten aanzien van de ontsluiting van de data. Zij combineert de luchtkwaliteitsdata niet met andere datasets; gezamenlijke ontsluiting van verschillende typen datasets is dan ook niet nodig.

De dataset is vooral van belang om te kunnen zien of er overschrijdingen ten opzichte van normen optreden. Daarmee kan Milieudefensie zich voorbereiden op discussies in de beleids- of media-arena's, of zelf aandacht vragen voor de overschrijding en het belang van maatregelen.

Metingen over luchtkwaliteit worden door organisaties buiten het RIVM ook gebruikt om

voorspellingen te doen, waarbij vaak verschillende beleidsscenario's worden vergeleken. Denk aan de bouw van een nieuwe weg. De rekenmodellen zijn veel minder objectief dan de metingen zelf, omdat de uitkomst van de modellen erg gevoelig is voor de invoerwaarden en gedane aannames. Het is daarom van belang dat de rollen van objectieve leverancier van meetdata en voorspeller niet vermengd worden, om geen discussie te krijgen over de betrouwbaarheid van de meetresultaten.

De wettenbank van het ministerie van BZK

Gesproken met:

- Dick van den Berg - functioneel beheerder, ICTU
- Hans Flier - beleidsmedewerker directie Regeldruk, Dienstverlening en Informatie, ministerie van BZK
- Kees Keuzenkamp - waarnemend directeur Regeldruk, Dienstverlening en Informatie, ministerie van BZK
- Jaco de Vroed - Chief Technology Officer, Legal Intelligence (hergebruiker)
- Hylke Wierda - hoofd CIO-office, ministerie van BZK

Wettenbank

De wettenbank is een dataset met geconsolideerde teksten van wet- en regelgeving die op rijksniveau zijn ontstaan. Het gaat onder andere om wetten, Algemene Maatregelen van Bestuur en ministeriële regelingen. Via www.wetten.nl en een webservice is deze dataset toegankelijk. Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) is vanuit twee taken verantwoordelijk voor de wettenbank. Aan de ene kant borgt zij democratische waarden zoals transparantie van het openbaar bestuur en aan de andere kant treedt zij op als aanjager en facilitator voor de samenwerking binnen de rijksdienst.

Kenmerken van de data

Als resultaat van het wetgevingsproces ontstaan officiële publicaties van de Rijksoverheid. Dergelijke publicaties kondigen meestal algemeen verbindende regelgeving af. Vaak gaat het daarbij om wijzigingen op de tekst, of de geldigheid van eerdere regelgeving.

Uit de losse publicaties wordt niet zonder meer duidelijk wat de geldende regelgeving op een specifiek moment is. Regelingen kunnen immers worden aangevuld of ingetrokken. Om de actuele tekst te kennen, is het nodig de wetgeving te consolideren. Die consolidatie is een redactioneel proces dat de overheid inkoop en waarvan de resultaten openlijk toegankelijk zijn. Voor de consolidatie wordt de wetgeving systematisch verzameld, door een beperkt aantal 'publicatiekanalen' (officiële publicaties zoals het Staatsblad en de Staatscourant) te monitoren op regelgeving.

Sommige regelgeving, bijvoorbeeld verdragen van voor 1952, zijn nog niet opgenomen in de wettenbank met geconsolideerde teksten. Beleidsvoorbereidende notities die niet met de Kamer zijn gedeeld, worden in de regel niet gepubliceerd.

Wetsteksten worden kort na publicatie van de onderliggende officiële stukken aangepast. Vaak verstrijkt er enige tijd tussen de publicatie van een wet en de inwerkingtreding ervan, hierdoor is een wetstekst vaak ruim voor de inwerkingtreding al te raadplegen via wetten.nl.

Beschikbaarstelling van de dataset

Staatssecretaris Kohnstamm heeft de Tweede Kamer in 1996 medegedeeld, dat de Rijksoverheid naar zijn mening een verantwoordelijkheid heeft in het toegankelijk maken van wetten voor burgers (TK20644/30). De Ministerraad heeft later ingestemd met de eerste realisatie van de wettenbank.

Tot 2001 nam de Rijksoverheid geconsolideerde wetsteksten af via de wettenbank van Kluwer (Algemene Databank Wet- en Regelgeving). Kluwer had een de facto monopolie op de geconsolideerde wetsteksten van de (vrijwel) gehele collectie wetgeving. Anderen - burgers en bedrijven - konden ook een cd-rom afnemen van Kluwer, tegen betaling. De teksten waren onder andere toegankelijk in openbare bibliotheken. In 2001 is de wettenbank aanbesteed, waarbij de Staat het eigendom van de geconsolideerde teksten verwierf en een website opende waarop iedereen de teksten kosteloos kon opvragen.

De wettenbestanden worden zowel via een eigen presentatievoorziening (wetten.nl) als via een webservice (in XML) aangeboden. Deze laatste wordt gebruikt door intermediairs en professionele gebruikers. Beide kanalen zijn voor iedereen kosteloos toegankelijk. Aan het gebruik van het wettenbestand worden geen voorwaarden gesteld.

Het ministerie van BZK treedt op als opdrachtgever en financier naar twee partijen. Enerzijds naar SDU, die na een aanbesteding verantwoordelijk is geworden voor de redactie van de wetsteksten en anderzijds ICTU (programma e-Overheid voor Burgers), dat het functionele beheer en de technische doorontwikkeling verzorgt.

Gebruik en gebruikers

De wettenbank kent drie doelgroepen:

1. Burgers, die geacht worden de wet te kennen.
2. Ambtenaren, die wetsteksten gebruiken in het proces van beleidsvoorbereiding, het opstellen van wetgeving en in uitvoerende processen.
3. Professionele juristen bij bedrijven, in de advocatuur of de magistratuur.

Burgers gebruiken teksten via wetten.nl en via sites van anderen, zoals lexius.nl. Ambtenaren en bedrijven raadplegen de wettenbank rechtstreeks, doen dat via derde partijen, of integreren het met de eigen systemen (zogenaamde 'content integration', waarbij de wetten zijn geïntegreerd of verrijkt met andere bronnen). Er zijn geen cijfers bekend over de omvang van dergelijk hergebruik.

De communicatie van wetten is onder andere van belang zodat burgers en rechtspersonen weten aan welke regels zij zich moeten houden. Daarnaast wordt er veel beroep gedaan op wetsteksten in de voorbereiding van juridische procedures. Verder zijn de wetten van belang

voor uitvoeringsorganisaties, in het inrichten van de processen en het motiveren van besluiten. Tot slot worden de teksten gebruikt in het wetgevingsproces zelf. Via wetten.nl worden jaarlijks circa 5,5 miljoen sessies opgestart.

Strategie

De wettenbank is ontwikkeld op basis van de politieke wil (in 1996) tot het beschikbaar stellen van wetsteksten aan burgers (transparantie). De strategie om dit doel te realiseren is vooral van financiële aard: de voordelen die samenhangen met het centraal realiseren van de toegang tot een open wettenbank voor rijksambtenaren waren voldoende om het doel van transparantie te realiseren binnen de bestaande rijksmiddelen. De kosten van een centrale open voorziening liggen lager dan die van decentrale abonnementsmodellen waarin particuliere uitgevers zijn betrokken. Het initiatief bleek later goed te passen in nieuwe beleidsdoelen zoals de compacte overheid (vergroten interne efficiency) en economische stimulering door middel van open data. Het accent heeft in de loop van de tijd op verschillende motivaties gelegen, maar tezamen hebben deze gezorgd voor continuïteit in het bestaan en de doorontwikkeling van de wettenbank.

Verder is de wettenbank de eerste stap geweest in een reeks van vervolginiciatieven om verzamelingen overheidsinformatie actief openbaar en herbruikbaar te maken. Deze beheerste, stapsgewijze strategie heeft er volgens BZK toe bijgedragen dat geen van deze projecten terecht is gekomen in de hoek van de mislukte overheids-ICT-projecten.

Kosten en opbrengsten

De kosten voor de redactie, de techniek en het beheer van de wettenbank bedragen samen € 2,2 miljoen per jaar. De wettenbank wordt gefinancierd via de begroting van het ministerie van BZK.

Met het openstellen van de wetsteksten is het de facto oligopolie van Kluwer en SDU op geconsolideerde wetsteksten doorbroken. Uitgevers kunnen, net als andere partijen, informatiediensten baseren op de teksten, maar de teksten zelf worden niet langer exclusief door hen aangeboden.

Parallel met de ontwikkeling van de wettenbank is in het Juriconnect-forum gewerkt aan de standaardisatie van de identificatie van regelingen en regelingselementen.

De kosten van de redactie en het beheer van de wettenbank zijn lager dan wat de Rijksoverheid vroeger betaalde aan licenties voor de Algemene Databank Wet- en Regelgeving van Kluwer (ca. 3,5 miljoen gulden per jaar). Rekening houdend met prijsindexatie liggen de kosten van het ontwikkelen en beheren van de wettenbank niet ver uit de buurt van de prijs die werd betaald voor de licenties van Kluwer.

Doordat verschillende soorten regelgeving en regelgeving van verschillende departementen, centraal beschikbaar zijn, worden daarnaast veel redactionele werkzaamheden uitgespaard. Voorheen kwam het voor dat departementen de eigen 'wetgevingsinformatievoorziening' hadden. Op verschillende plaatsen (bij uitgevers, bij departementen, andere overheden en wellicht ook in de private sector), werden dezelfde wetsteksten geconsolideerd en gestructureerd. Hier zijn echter geen gegevens over bekend.

Recente en voorziene ontwikkelingen

Vanuit de wet- en regelgeving worden steeds meer verbanden aangelegd naar andere informatie. Zo is het mogelijk om de wetsgeschiedenis op te vragen, waarin verwijzingen zijn opgenomen naar Kamerstukken en de oorspronkelijke tekst in de officiële publicatie.

De inhoud van de wettenbank wordt ook verbreed. Momenteel wordt er samen met het ministerie van Buitenlandse Zaken gewerkt aan het opnemen van oudere verdragsteksten. Daarnaast wordt getracht om de relatie tussen Europese regelgeving en de implementatie in Nederlandse wet- en regelgeving zichtbaar te maken.

De voorzieningen en methoden rondom de wettenbank (en wetten.nl) hebben model gestaan voor de Centrale Voorziening Decentrale Regelgeving, die sinds kort de regelgeving van decentrale overheden ontsluit. Die data zijn geen eigendom van de Rijksoverheid, maar BZK treedt wel op als 'shared service centre' voor het verspreiden ervan.

Hergebruik door Legal Intelligence

Legal Intelligence biedt een commercieel content-integratieproduct voor professionele juridische gebruikers. Zij ontsluiten meerdere bronnen op een gerelateerde manier en hebben daarvoor een eigen zoekdienst ontwikkeld. De wetsteksten zijn daar onderdeel van, maar het kan ook gaan om bijvoorbeeld juridische tijdschriften. Legal Intelligence is tevreden over de toegang die zij heeft tot de wettenbank, maar geeft wel aan dat het enkele jaren heeft geduurd voordat de teksten op een goede manier geautomatiseerd te verwerken waren. De organisatie werkt met een 'crawler' die de inhoud van alle webpagina's ophaalt. Die crawler is gebouwd op het moment dat er nog geen webservice van de wettenbank beschikbaar was. De standaardisatie van de identificatie van regelingselementen was een belangrijk hulpmiddel in het kunnen hergebruiken van de gegevens.

Scholeninformatie van DUO

Gesproken met:

- Bram Voermans - Adviseur strategisch informatiebeleid van het ministerie van Onderwijs, Cultuur en Wetenschap
- David Bronsgeest - Analist/ontwikkelaar afdeling Informatieproducten van de Dienst Uitvoering Onderwijs
- Jan Douwe Kooistra - oprichter 10.000scholen.nl (hergebruiker)

De Dienst Uitvoering Onderwijs

Sinds 1 januari 2010 vormen de IB-Groep en CFI één organisatie: Dienst Uitvoering Onderwijs (DUO). DUO financiert en informeert onderwijsdeelnemers en onderwijsinstellingen. Ook organiseert DUO verschillende examens. De dienst maakt onderdeel uit van het ministerie van Onderwijs, Cultuur en Wetenschap (OCW). DUO is een uitvoeringsorganisatie van dit ministerie. Gekeken vanuit het perspectief van opdrachtgever en opdrachtnemer, is DUO opdrachtnemer en het ministerie opdrachtgever. Deze rollen zijn in de praktijk goed waarneembaar. Het ministerie stuurt DUO aan en DUO voert uit.

Kenmerken van de data

DUO beheert gegevens over het Nederlandse onderwijs in een 'datawarehouse'. Binnen DUO worden vier gegevensdomeinen onderscheiden:

- Onderwijsdeelname en -resultaten
- Instellingen en opleidingen
- Onderwijspersoneel
- Financiële gegevens

De informatie in een gegevensdomein, kan het primair onderwijs (PO), het voortgezet onderwijs (VO), beroepsonderwijs en volwasseneducatie (BVE) en het hoger onderwijs (HO) betreffen.

Voor beleid en bekostiging van scholen mogen alleen geanonimiseerde gegevens gebruikt worden. Basisgegevens mogen hiervoor niet worden gebruikt. Het gaat dan onder andere om de basisgegevens uit het Basisregister Onderwijs (BRON), dat alle relevante persoonsgegevens, gekoppeld aan een persoonsgebonden nummer, van alle deelnemers bevat. Daarom worden er vanuit BRON geanonimiseerde bestanden aangemaakt waarin alleen nog relevante persoonsgegevens over blijven: geslacht, geboortjaar (leeftijd), postcodegebied (postcode zonder lettercode), etniciteit en herkomst ouders. Deze geanonimiseerde bestanden worden 1-cijferbestanden genoemd en vormen de basis voor leveringsverzoeken die bij DUO binnenkomen. De bestanden worden 1-cijferbestanden genoemd, omdat DUO en andere overheidsorganisaties die deze gegevens voor hun wettelijke taakuitvoering nodig hebben, zoals de Onderwijsinspectie en gemeenten,

gebruikmaken van dezelfde bestanden en daarmee één definitie (cijfer) hanteren (Ministerie van Onderwijs, Cultuur en Wetenschappen, z.d.).

Vanuit de Wet Openbaarheid Bestuur gaat het hier om openbare informatie die in principe integraal openbaar moeten zijn. OCW maakt echter een onderscheid tussen actieve en passieve openheid. Geanonimiseerde en geaggregeerde, op het niveau van onderwijsinstellingen, gegevens worden actief voor hergebruik te beschikking gesteld. Andere datasets, veelal maatwerkleveringen van gegevens, worden alleen reactief, naar aanleiding van een specifieke vraag met een specifiek doel en veelal met een beroep op de WOB voor hergebruik ter beschikking gesteld.

Omdat de 1-cijferbestanden maar één keer per jaar gewijzigd worden, wordt de informatie die als open data beschikbaar wordt gesteld ook maar één keer per jaar geüpdatet. De adresbestanden - die niet op 1-cijferbestanden gebaseerd zijn - worden maandelijks geüpdatet.

Voor elke onderwijssector is er andere informatie als open data beschikbaar. Dat wil zeggen dat een gegevensset die voor bijvoorbeeld het PO beschikbaar is, niet ook altijd voor het VO en/of het BVE beschikbaar is. Sommige gegevenssets kunnen sectoroverstijgende informatie bevatten, zoals de intersectorale stroominformatie (het aantal leerlingen dat een onderwijssector in- of uitstroomt of het aantal leerlingen dat van de ene onderwijssector doorstroomt naar een andere onderwijssector). Deze informatie wordt 'berekend' door de dienst informatieproducten van DUO.

Naast een gesloten document formaat (.xls) wordt de data waar mogelijk ook in een open document formaat aangeboden (.csv). Stroombestanden zijn alleen als .csv beschikbaar, omdat Excel het aantal records niet aan kan. De bbo-ROM bestanden (basisgegevens bekostigd onderwijs) die aan gemeenten geleverd worden, zijn opgebouwd volgens een vaste lay-out. De gemeenten kunnen middels eigen software de bbo-ROM bestanden inlezen. De standaard voor de bbo-ROM bestanden wordt door gemeenten en DUO gezamenlijk beheerd, de overige bestanden zijn niet volgens een standaard ingedeeld. Wel is het zo dat DUO probeert om de indeling zo min mogelijk te wijzigen.

Beschikbaarstelling van de data

De databestanden worden beschikbaar gesteld via data.duo.nl. De site is nog niet zo lang in de lucht en is (bij schrijven) nog in bèta. Het moet de opvolger worden van 'onderwijsincijfers' (www.onderwijsincijfers.nl), 'Feiten en Cijfers' (cijfers.minocw.nl) en een aantal andere vergelijkbare sites.

De 'oude' sites kennen allemaal enkele gebreken. De gebreken vertonen een vicieuze cirkel. Het is moeilijk om de openbare data te vinden (zowel de websites zelf, als binnen de data),

waardoor er weinig hergebruik is, wat ertoe leidt dat weinig data beschikbaar komt, waardoor de data die beschikbaar komt ook niet te hergebruiken is, wat weer tot een slechtere vindbaarheid leidt (Ministerie van Onderwijs, Cultuur en Wetenschappen, 2011). Om deze reden is data.duo.nl als nieuw portaal ontwikkeld. Gelijktijdig met de ontwikkeling van dit portaal liep het project datawarehouse dat ervoor moest zorgen dat de aanlevering van gegevens structureel en geautomatiseerd gebeurt. De website is het ontsluitingsmechanisme.

Waar op de oude site van Feiten en Cijfers nog een enkele webservice werd aangeboden, wordt deze bewust niet op de nieuwe website aangeboden. Het zou volgens DUO de site nodeloos ingewikkeld maken zonder een duidelijke meerwaarde. Op dit moment is er namelijk maar een partij die van de webservice gebruik maakt. OCW en DUO gaan er vanuit dat iedereen, dus ook deze partij, .xls en/of .csv bestanden kan lezen.

Gebruik en gebruikers

Een van de belangrijkste gebruikers van de data - naast beleidsmedewerkers van OCW zelf - zijn de sectorraden van de verschillende onderwijssectoren, zoals de PO-raad (sectorraad primair onderwijs) en de VO-raad (de sectorraad voortgezet onderwijs). Het project Vensters voor Verantwoording (VvV) is een belangrijk project van de VO-raad. Met Vensters voor Verantwoording bieden scholen door middel van indicatoren inzicht in hun resultaten. In Venster voor Verantwoording wordt informatie over het voortgezet onderwijs van DUO, informatie van de Onderwijsinspectie en informatie van de scholen zelf verzameld in één systeem en beschikbaar gesteld via www.schoolvo.nl en www.schoolkompas.nl. Vensters voor Verantwoording wordt op dit moment ook voor het primair onderwijs ontwikkeld. Dit is een lang traject, omdat de voorbereiding en het op orde brengen van gegevens veel tijd vraagt. Denk hierbij aan het harmoniseren van gegevensdefinities, het omnummeren van gegevens en het opzetten van een telling (in welk bestand telt een leerling wel en in welk bestand telt een leerling niet?). Vensters voor Verantwoording krijgt vanwege haar bijzondere positie als vertegenwoordiger van de VO-scholen een eigen levering van DUO. In feite gaat het hier om een teruglevering aan de VO-scholen van hun eigen data. Ook andere overheidsinstellingen met een wettelijke taak in het onderwijs zoals de Inspectie voor het Onderwijs, het Centraal Bureau voor de Statistiek en diverse ministeries krijgen een eigen levering.

Naast de sectorraden gebruiken gemeenten, onderzoeksbureaus, adviesbureaus en burgers de informatie. Bij gemeenten gaat het vooral om de G4, die onder meer de bbo-ROM bestanden (waar zij zelf initiatiefnemer van is) gebruikt. De data worden verder voor toegepast en wetenschappelijk onderzoek gebruikt. Adviesbureaus, zowel grote als kleine, werken vaak in opdracht van schoolbesturen met de opdracht om de data te analyseren, te interpreteren en schoolbesturen te adviseren. Er is niet bekend hoe groot de markt is voor adviesdiensten die gebruik maken van DUO-data. Tot slot maken ook burgers gebruik van

de informatie. Opvallend is dat weinig ouders gebruikmaken van de informatie. Althans, zij lijken de ruwe data niet te downloaden.

Naast deze hergebruikers, zijn er meerdere bedrijven die de DUO-data via een website of via een ander medium aanbieden. In sommige gevallen gaat het daarbij om non-profit-instellingen, in andere gevallen zijn dit commerciële bedrijven. Een aantal interessante bedrijven:

- Rinus de Loos is adviseur en maakt benchmarks voor het onderwijs en gemeenten, <http://deloos.net>.
- 10.000scholen.nl is een relatief nieuwe site die gebruikt kan worden voor het maken van een schoolkeuze.
- www.scholenkaart.nl gebruikt DUO-data.
- Dagblad Trouw ontsluit de gegevens die in haar jaarlijkse scholentest zijn gebruikt via haar site schoolprestaties.trouw.nl.
- Ook Elsevier maakt jaarlijks een overzicht van de beste en slechtst presterende scholen in Nederland.

DUO heeft regelmatig contact met de (her)gebruikers van de data. Gebruikers geven zowel gevraagd als ongevraagd feedback. Dit contact is van oudsher. Toen de data nog niet via internet beschikbaar was, belden of mailden de gebruikers en dat contact is gebeven. Het contact vindt ad hoc plaats. Voor het beheer van de bbo-ROM standaard is er wel een overleg dat op een gebruikersraad lijkt, bestaande uit de G4 gemeenten. Zowel het ministerie als DUO zien op dit moment niet voldoende toegevoegde waarde in een gebruikersraad voor de andere data. De reden hiervoor is dat het inrichten van een gebruikersraad veel tijd kost, terwijl gebruikers DUO en het ministerie op dit moment ook goed weten te vinden. De nieuwe website biedt gebruikers wel een feedbackfunctionaliteit. Dit biedt een extra contactmogelijkheid.

Hoe vaak de gebruikers gebruikmaken van de informatie is moeilijk te zeggen. Op de oude website werden alleen page views geregistreerd. Op de nieuwe website kan op rapportniveau worden gekeken. Hier is echter nog geen informatie over beschikbaar.

Strategie

Het ministerie en DUO hebben de volgende stelregel voor het openbaar maken van data: 'alles is openbaar, tenzij'. Het 'tenzij' betreft weigeringsgronden genoemd in de WOB zoals in het kader van privacy. De informatie is dan ook niet tot personen herleidbaar. Ook gegevens over etniciteit worden niet (meer) openbaar gemaakt. Zoals gezegd, maakt het ministerie een onderscheid tussen actieve en passieve openbaarheid. Actief worden de data via de website beschikbaar gesteld. Specifieke wensen omtrent datasets worden alleen op aanvraag en in opdracht van het ministerie (passief) beschikbaar gesteld.

Een directe aanleiding voor het open maken van onderwijsdata is een verzoek van dagblad Trouw in 1999. Trouw diende een WOB-verzoek in voor gegevens van de Onderwijsinspectie. De Onderwijsinspectie wilde deze gegevens niet leveren, maar de rechter oordeelde uiteindelijk dat Trouw de gegevens mocht hebben. Toen was de gedachte op het ministerie 'als we deze data aan Trouw geven, dan moeten ook anderen over deze data kunnen beschikken'.

Interne overwegingen vormden de tweede driver om tot actieve openbaarmaking van datasets over te gaan. Vanaf 2000 konden standaardinformatieproducten via mail en telefoon opgevraagd worden. Vanaf 2003 waren deze via OLAPs op internet verkrijgbaar. In 2009 kwam er druk op de interne organisatie te staan. DUO zat op de top qua uren en had een capaciteitsprobleem. Er was geen ruimte voor innovatie meer. Toen is besloten niet langer informatie, maar data te leveren via Feiten en Cijfers. Door het leveren van data, kunnen gebruikers zelf informatie maken en hoeft DUO minder tijd te besteden aan telefoontjes, mailtjes en aparte informatieverzoeken. Samen met het ministerie heeft DUO toen beleid gemaakt. Wie kan welke producten krijgen? Het antwoord op deze vraag is dat ministeries alles kunnen krijgen en gemeenten, provincies en burgers alleen standaardproducten. Maatwerkvragen door gemeenten, provincies en burgers moeten via het ministerie aangevraagd worden.

Voor de data hoeft niet betaald te worden. Vanuit DUO wordt wel verkend of het mogelijk is om een vergoeding op basis van marginale kosten voor de totstandkoming van maatwerkproducten te vragen. Vanuit het perspectief van één overheid wordt het door het ministerie echter niet aannemelijk geacht dat dit ook daadwerkelijk gaat gebeuren.

Kosten en opbrengsten voor DUO

De kosten en opbrengsten van de open-datastrategie bij DUO hebben meerdere facetten. Ze zijn zowel kwantitatief als kwalitatief en hebben betrekking op zowel projectmatige als structurele activiteiten.

DUO heeft in een projectplan voor het project 'Openbare Onderwijs Informatie' (OOI) beschreven welke baten zij verwachtte te realiseren door data pro-actief via de website aan te bieden (Ministerie van Onderwijs, Cultuur en Wetenschappen, 2011). De belangrijkste kwalitatieve baten waren:

- Betere vindbaarheid van de website. Het beschikbaar stellen van één centrale plek (website) waar openbare geaggregeerde onderwijsdata en informatie is te vinden (in plaats van de huidige versnipperde informatieverstrekking via diverse sites en ontsluitingsmechanismen);
- Betere vindbaarheid binnen de site. Het aanbieden van zoekfunctionaliteit, zodat afnemers binnen de website openbare onderwijsdata en informatie eenvoudig kunnen vinden;

- Transparantie door het verstrekken van heldere gegevensdefinities en toelichtingen (metadata) bij de geleverde producten;
- Heldere relatie met andere sites. Afnemers worden doorverwezen naar andere sites/data/informatie als de data of informatie binnen DUO niet voor handen is en andere partijen wel in de specifieke behoefte kunnen voorzien;
- Het aanbieden van de mogelijkheid aan afnemers om commentaar te leveren op hetgeen geleverd wordt. Met deze feedback kan de dienstverlening worden doorontwikkeld en toegesneden op de daadwerkelijke behoefte uit de markt;
- Verbeterd imago van de Rijksoverheid: Een werkelijke “open en transparante” communicatie over aanpak en resultaat hoort bij het onderwerp “openbare informatie”. In het verleden is het onderwerp echter gemonopoliseerd. Een andere houding zal daarom niet onopgemerkt blijven, en zal leiden tot een beter imago van de Rijksoverheid.
- Voorkomen van schade bij eindgebruikers dankzij interactie met betrekking tot leverkalender en publicatievorm. De schade is onbekend.
- Tot slot, worden de voorwaarden gecreëerd voor ontwikkeling van nieuwe informatiediensten door marktpartijen. Dit is een besparing ten opzichte van het model dat de Rijksoverheid (DUO) alle diensten ontwikkelt.

Daarnaast waren er ook een aantal baten die te kwantificeren zijn:

- Uitfaseren van andere websites en ontsluitingsmechanismen levert besparing in beheer van 0,5 fte op.
- Snellere time-to-market voor het publiceren van nieuwe of gewijzigde datasets en informatieproducten. Zonder verbetering van de website zouden wijzigingen op de oude websites duurder uitvallen en langer duren. Uitgaande van 2 majeure releases per jaar worden de meerkosten geschat op 1 fte.
- Verbeteringen in interne informatielogistiek. Hierbij gaat het voornamelijk om het voorkomen van handmatige werkzaamheden in de informatieketen. Deze zijn geraamd op 6 fte.
- Totaal baten: 7,5 fte.

Het project OOI hangt nauw samen met een al eerder gestart project. In dit project “Datawarehouse Informatieproducten” had men de informatielogistiek al onder handen genomen. OOI was zonder die verbeterde logistiek van het datawarehouse nooit mogelijk geweest. Omdat de baten van het project datawarehouse sterk leken op de hierboven laatst genoemde baten van 6 fte, beredeneert DUO dat die 6 fte voor de helft toegerekend zou moeten worden aan het project datawarehouse. Voor de businesscase moet men echter de volledige 6 fte tellen, omdat de 3 aan het project datawarehouse toegerekende fte’s niet gerealiseerd kunnen worden zonder OOI.

De kosten van het project OOI zijn uiteindelijk rond de € 500.000,- geweest. Wanneer we nu rekening houden met een besparing van 7,5 fte betekent dit een terugverdiendtijd van iets meer dan 1 jaar (uitgaande van 60K/fte aan lasten). Gaan we uit van 4,5 fte besparing (zonder de 3 fte die toegerekend worden aan het project datawarehouse) dan is de terugverdiendtijd van het project iets meer dan 2 jaar.

De baten die in het projectplan opgenomen zijn, zijn voor een groot gedeelte ook gerealiseerd. Zo worden er inmiddels bijvoorbeeld bijna geen uren meer geschreven op leveringen naar burgers. Ook kan DUO veel selectiever zijn in het leveren van specifieke datasets, en worden een aantal van de websites die vroeger onderwijsinformatie verstrekten uitgefaseerd en wordt het beheer afgebouwd.

Andere gerealiseerde baat is dat de 'time to market' is verbeterd. Voorheen duurde het drie maanden voordat nieuwe informatieproducten geproduceerd konden worden, nu kan dat binnen een dag. Tot slot is de gegevenskwaliteit verbeterd. Immers, als je transparant bent, wordt je gedwongen je gegevens op orde te brengen. De belangrijkste winst die hiermee geboekt wordt, is dat er geen discussie meer over de cijfers is, maar dat de discussie over de inhoud gevoerd wordt.

Recente en voorziene ontwikkelingen

Voor OCW en DUO staat een aantal ontwikkelingen op stapel. Deze ontwikkelingen betreffen vooral de borging en verdere professionalisering. Zo wil OCW de data geschikt maken om als Linked Open Data te gebruiken. Ook is het ministerie bezig om de relatie met afnemers te structureren. Zo worden met Vensters voor Verantwoording service level agreements (SLA) opgesteld over de levering en het niveau van de data. Op die manier zien afnemers hun eisen geborgd en heeft OCW op papier vastgelegd wat zij maximaal levert.

Hergebruik door 10.000scholen.nl

Ouders die een (basis- of middelbare) school voor hun kinderen zoeken, kunnen terecht op de site 10.000scholen.nl. Zij kunnen op deze site alle scholen in Nederland vinden en met elkaar vergelijken. Het gaat dan om aspecten als het aantal en de opbouw van leerlingen en docenten, maar ook om kwaliteitsoordelen van de Onderwijsinspectie en de locatie van de school.

Achtergrond

10.000scholen is een particulier initiatief van Jan Douwe Kooistra en Mateo Mol. Toen zij destijds bij de Volkskrant werkten, waren zij op zoek naar een simpele database met schoolgegevens van Nederland. Deze data was niet gemakkelijk te vinden en/of niet betrouwbaar. Ook was het combineren van Excel-sheets niet altijd gemakkelijk. Binnen de Volkskrant was het niet mogelijk om dit op te pakken, dus hebben Jan Douwe en Mateo het op eigen houtje opgepakt. Hun voornaamste doel en drijfveer is dat schoolinformatie voor

ouders begrijpelijk, toegankelijk en gratis (dus niet via de markt) beschikbaar moet zijn.

Gebruik

Enkele maanden geleden is 10.000scholen met een bèta-versie live gegaan en het lijkt nu al een succes te zijn. Zonder veel communicatie - en de site via google vindbaar te maken - zijn er nu al regelmatig meer dan 1000 bezoekers per dag. Dit betekent eigenlijk dat er heel veel Nederlanders op zoek zijn naar informatie over scholen.

Gegevens

Voor 10.000scholen worden gegevens van DUO en de Onderwijsinspectie gebruikt. Op dit moment wordt gekeken of ook gegevens uit de Basisregistratie Adressen en Gebouwen (BAG) gehaald kunnen worden. Op die manier kan bijvoorbeeld de grootte van de school (oppervlakte) en hoe oud een school is (bouwjaar) worden aangegeven.

Kosten en opbrengsten

Op dit moment is er veel (vrije) tijd geïnvesteerd in het toegankelijk en vergelijkbaar maken van de website en het bouwen en onderhouden van de website. De investering is gedaan vanuit een intellectuele uitdaging.

Er zijn nog geen opbrengsten met de website gemaakt. Wel wordt er naar creatieve mogelijkheden gekeken om opbrengst te creëren. Zo wordt er een API¹ gebouwd. Voor deze API zou betaald moeten worden. De initiatiefnemers denken dat de API waardevol is, omdat via 10.000scholen.nl bestanden gekoppeld zijn, kwalitatief goede data geselecteerd zijn en data via de API realtime van de site wordt gehaald en niet via Excel-sheets moet worden gedownload. Een andere optie is dat bijvoorbeeld gemeenten in hun website een overzicht bieden van scholen binnen de gemeente. Deze informatie kan via 10.000scholen automatisch worden geüpdatet. Gemeenten zouden via de API ook een eigen scholenzoeker kunnen maken. Verder biedt het scholen een mogelijkheid zich te profileren door informatie aan de data toe te voegen, bijvoorbeeld promotiefilmpjes.

Ruimte voor verbetering

Op persoonlijke basis is er goed contact met DUO. Vragen worden altijd snel en goed beantwoord. Toch worden er een aantal verbetermogelijkheden gezien. De belangrijkste is dat er geen afspraken - in de vorm van een service level agreement - met DUO over de continuïteit van de data zijn. Dit komt erop neer dat 10.000scholen een probleem heeft, als DUO data niet meer publiceert. Dat komt nu ook al voor, omdat DUO statusupdates niet direct communiceert. Als er alleen al een header in een Excel-sheet wordt aangepast, moet

¹ Application Programming Interface: koppelvlak tussen twee softwaresystemen. Zie ook <http://en.wikipedia.org/wiki/API>.

10.000scholen haar code aanpassen. De gedachte van de initiatiefnemers van 10.000scholen is dat als OCW de data publiceert met de bedoeling dat mensen er wat mee doen, ze een minimale SLA moet bieden. Op dit moment is het voor hen een black box. Het is afwachten wat er gebeurt. Een gebruikersoverleg zou voor hen ook welkom zijn. Zij zouden er in ieder geval zeker naartoe gaan.

Verder zijn de BRIN-codes (identificatienummers voor onderwijsinstellingen en locaties) van de Onderwijsinspectie en DUO niet dezelfde. Het zou handig zijn als dit gestroomlijnd werd. Ook zou het aan te bevelen zijn om richtlijnen of standaarden voor het aanbieden van de data op te stellen. Op dit moment is er geen helderheid over hoe data beschikbaar gesteld zou moeten worden en worden vaak Excel-sheets gemaald.

Diverse datasets van de gemeente Rotterdam

Gesproken met:

- Roel van der Bolt - projectleider Rotterdamse Innovatie Agenda, gemeente Rotterdam
- Jaap Dekker - programmamanager, gemeente Rotterdam
- Léon Gommans - verantwoordelijk voor ROD bij Hogeschool Rotterdam
- Paul Hendriks - adviseur Stadsontwikkeling, dienst Stadsontwikkeling, gemeente Rotterdam
- Hans Nijman - CIO, gemeente Rotterdam
- Peter-Jan Speerstra - adviseur informatiebeleid, dienst stadsontwikkeling.

Kenmerken van de data

De gemeente Rotterdam stelt verschillende datasets ter beschikking voor hergebruik. Deze datasets zijn ontsloten via de Rotterdam Open Data Store (RODS). Oorspronkelijk bevatte de RODS een aantal datasets die door zeven verschillende gemeentelijke diensten zijn vrijgegeven in het kader van het pilotproject 'Rotterdam Open Data', een samenwerkingsverband van de Hogeschool Rotterdam (HR), de gemeente en enkele bedrijven. Inmiddels is de pilot omgezet in een stabielere vorm en is het de bedoeling dat er meer gemeentelijke data in de RODS ontsloten gaat worden.

Op dit moment zijn de volgende data vrijgegeven door de gemeente Rotterdam (gemeente Rotterdam, 2010):

- Bibliotheek
 - Bibliografische gegevens
 - Leden (geanonimiseerd)
 - Leenhistorie (transacties in het bibliotheekstelsel - koppeling tussen collectie en leden)
 - Online catalogus
- Dienst Stadsontwikkeling
 - Actuele status van 6 parkeergarages (vol/leeg)
 - Actuele status van 5 bruggen (gestremd/vrij voor verkeer)
 - Fietsstallingen op diverse locaties
 - NDW-data binnen gemeente Rotterdam
 - Bezettingcijfers over P+R-garages
- Gemeentearchief
 - API's om gebruik te maken van:
 - digitale stamboom
 - adresboeken
 - archieven (incl. kranten en notariële akten)

- beeld en geluid
 - Gemeentewerken / Buitenruimte
 - 3D-bestanden van de gemeente per wijk (op gemeentewebsite)
 - RegioGis / Openspatial geo-gegevens inclusief diverse kaartlagen. RegioGis en Openspatial zijn informatiesystemen waarin geo-informatie is opgeslagen over allerlei objecten in de openbare ruimte zoals wegen, bruggen, riolering, luchtkwaliteit, verkeer, waterhoogten etc. Een aantal van deze typen objecten worden door gemeentewerken vrijgegeven.
- GGD
 - Gezondheidsenquêteresultaten
 - Documenten binnen GGD-portal Rijnmond
- Sport en Recreatie
 - Shapefiles bebouwing, complexen, sportvelden
 - Adressen + kenmerken van sport- en recreatielocaties
 - Reserveringen op recreatie/sportvoorzieningen

In de komende maanden zullen er meer datasets van de dienst Stadsontwikkeling vrijgegeven gaan worden. Welke dat zijn, is nog niet bekend.

Voor de meeste datasets is er een webservice beschikbaar via de RODS, die data in XML- of JSON²-formaat kan leveren. De RODS is overigens een technische voorziening die niet alleen in de link naar de API of de downloadlink naar het XLS- of CSV-bestand voorziet, ze verzorgt ook een transformatie naar de meest gebruikte standaarden (CSV, JSON, XML). Daardoor kunnen hergebruikers over het algemeen het voor hen bruikbaarste format kiezen en is het oorspronkelijke formaat van de data minder relevant.

De data die ontsloten zijn via RODS en de site van de gemeente zelf, zijn niet voorzien van een licentie.

De datasets van de gemeente zijn niet aangemeld bij het nationale dataportaal data.overheid.nl. De gemeente Rotterdam ziet daarin niet genoeg meerwaarde, omdat daar alleen data-uitwisseling is en er geen contacten gelegd worden tussen hergebruiker en databeheerder (de gemeente).

Beschikbaarstelling van de data

Het pilotproject Rotterdam Open Data was geen initiatief van de gemeente zelf. Vooral de HR en enkele bedrijven hebben samengewerkt om de gemeentelijke diensten over te halen data beschikbaar te stellen.

² JSON is een data-uitwisselformaat dat veel gebruikt wordt voor het uitwisselen van open data. Het is zowel goed leesbaar voor computers als voor mensen.

De data die de diensten ter beschikking gesteld hebben in de pilot, zijn ook nu nog alleen beschikbaar voor de studenten van de HR die aan het project meededen. Het is wel de bedoeling dat de data die al in RODS zit ook voor anderen bruikbaar wordt. Alle geïnterviewden geven aan dat dit eigenlijk al had moeten gebeuren.

In september 2011 heeft het gemeentebestuur besloten om alle data (die geen persoonsinformatie is) en die door de dienst Stadsontwikkeling beheerd wordt openbaar te maken, zonder daar verdere voorwaarden aan te verbinden (gemeente Rotterdam, 2011a). Eind november 2011 waren er echter nog geen nieuwe datasets beschikbaar in de RODS. Er is daardoor vooral een principebesluit genomen dat het voor ambtenaren makkelijker moet maken om de data te ontsluiten. De ambtelijke top van de gemeente is op het gebied van open data eerst vooral met communicatie bezig geweest zonder dat aan concrete acties te verbinden.

Gebruik en gebruikers

Een van de bijzondere aspecten van de manier waarop het Rotterdam Open Data (ROD)-platform tot stand is gekomen, is de initiatie van het project ROD vanuit de HR en een aantal bedrijven. Hierdoor was het voor de gemeente vanaf het begin duidelijk wie de hergebruikers van de door haar vrijgegeven data zouden zijn, namelijk studenten en Rotterdamse bedrijven. In het eerste jaar van het project hebben 39 groepen studenten (onder begeleiding van bedrijven) applicaties gemaakt op basis van de gemeentelijke data. De zeven meest aansprekende projecten waren volgens de diensten zelf:

- Enquête 2.0 - Een online enquêtetool waarbij de antwoorden van een individu gekoppeld worden aan antwoorden van andere respondenten en infographics het eigen gedrag in vergelijking met anderen visualiseren.
- Rotterdam Onbepakt - Een applicatie voor rolstoelgebruikers waarin rolstoelvriendelijke routes, voorzieningen e.d. te vinden zijn en die het mogelijk maakt om eigen observaties met andere rolstoelers te delen.
- Studeerplekken - Een app toont waar in de bibliotheek nog vrije studieplekken zijn, om zo beter gebruik van de voorzieningen mogelijk te maken.
- Greenway - een routeplanner voor fietsers die je langs de mooiste route laat fietsen. Houdt rekening met verkeer en type bebouwing.
- Yourigin - een iPad-app waarmee archiefdata wordt ontsloten rondom stambomen. Maak je eigen stamboom en importeer automatisch stamboom informatie van het Rotterdams gemeentearchief.
- Start Hier - is een marktplaats die op basis van leegstandscijfers startende ondernemers en (beheerders van) leegstaande panden koppelt.
- Spotfinder - Door middel van Augmented Reality en plattegronden kunnen 'freerunners' de stad op een andere manier verkennen en nieuwe freerunspots vinden. De complexiteit van de objecten, alsmede aantekeningen van andere freerunners worden via de applicatie ontsloten.

- RODS - De Rotterdam Open Data Store is weliswaar een app voor andere hergebruikers, en niet voor eindgebruikers, maar is toch een bijzonder resultaat van het eerste jaar ROD.

De gemeente koppelt hergebruik van de data vooral aan de eigen economische agenda. Zij ziet de hergebruikers dus vooral in het kader van zowel bestaande bedrijven die data slimmer hergebruiken, alsook nieuwe bedrijven die ontstaan vanuit de kennisinstellingen (zoals de HR).

Een belangrijke gebruiker van de open data is op dit moment de HR zelf. Zij gebruiken de data feitelijk als lesmateriaal. De HR heeft in het kader van ROD inmiddels een aantal projecten met Nederlandse en Europese subsidies binnengehaald, waarbij er voor bijna € 700.000,- aan subsidiegeld binnengehaald is. Dit bedrag wordt door de HR en diverse Rotterdamse bedrijven gebruikt om verder onderzoek te doen. Ook heeft de HR voor € 200.000,- a € 300.000,- aan publicitaire waarde gegenereerd met publicaties over (de diverse onderdelen van) ROD.

Strategie

Rotterdam Open Data is een onderdeel van de Rotterdamse Innovatie Agenda (RIA) (gemeente Rotterdam, 2011b). De agenda beschrijft de motivatie om te kiezen voor open data (gemeente Rotterdam, 2011c):

“De Gemeente Rotterdam heeft eind 2010 de ambitie kenbaar gemaakt om alle niet-persoonsgebonden, overheidsdata openbaar te gaan maken. Het belang voor innovatie ligt er in dat dit een stimulans voor ondernemers kan zijn om hier nieuwe toepassingen en diensten op te ontwikkelen. Bovendien zijn er nog niet veel voorbeelden van vergelijkbare projecten bekend en zou een succes Rotterdam voorop doen lopen op dit gebied. Om dit proces te faciliteren zijn overheden, bedrijfsleven en onderwijs een open netwerk gestart - Rotterdam Open Data - van waaruit tal van initiatieven worden ontplooid, gericht op het toegankelijk en inzichtelijk maken van de beschikbaar komende informatie van, over en voor de stad Rotterdam. Rotterdam Open Data geeft hiermee invulling aan onderzoek en kennisopbouw over een groot strategisch vraagstuk van de stad, nl. de betekenis van informatie in de stad.”

Op de lange termijn wil de gemeente vooral zichzelf transformeren naar een organisatie die zich meer richt op de wensen van burgers en ondernemers:

“Eén van de initiatieven [Van de RIA - red] is het SIA-RAAK-project ‘Professionals Supported - Rotterdam Open Data’. Met dit RAAK-project willen de Hogeschool Rotterdam, Gemeente Rotterdam en partners de ontwikkeling van de

informatiedienstverlening en -beheer bij de gemeentelijke diensten zodanig mogelijk maken dat deze is afgestemd op de beoogde innovatie die het open netwerk Rotterdam Open Data voorstaat. Hiertoe is een consortium geformeerd tussen de Hogeschool Rotterdam en vier uitvoerende gemeentelijke diensten; Bibliotheek Rotterdam, Gemeentewerken, Stadsontwikkeling en het Gemeentearchief. In de opstartfase van het SIA-RAAK programma zullen naar verwachting alle andere gemeentelijke diensten zich ook aansluiten (ruimte voor groei).

Met het SIA-RAAK project beoogt het consortium in de periode van september 2011 t/m augustus 2013 praktijkgericht onderzoek te doen naar de ontsluiting van overheidsdata bij 4 uitvoerende diensten van de Gemeente Rotterdam, teneinde een nieuwe informatie-infrastructuur en interface te creëren waarmee een open data Space ontstaat waarin Gemeentelijke medewerkers proactief kunnen inspelen op de maatschappelijke behoefte aan en de mogelijkheden van dataontsluiting voor samenleving en overheid.”

De gemeente Rotterdam heeft het ter beschikking stellen van open data sterk verbonden met innovatie en economische groei aan de ene kant, en een kleinere overheid - Rotterdam heeft op dit moment een stevige bezuinigingsuitdaging - aan de andere kant. De gemeente zoekt op dit moment naarstig naar manieren om (het leefklimaat van) de stad te verbeteren, waarbij de overheid ook meer taken aan de burgers en bedrijven kan overlaten. Innovatieve manieren van samenwerken, zoals bijvoorbeeld crowdsourcing, en ook het samenwerken rondom open data, hebben daarbij bijzondere aandacht van het CIO-office en de gemeente als geheel.

Voor de gemeente is het verbinden van bedrijfsleven, onderwijs en overheid en het daardoor verbeteren van de economische positie en leefbaarheid van de stad dan ook de kern van het initiatief ROD. Zonder zo'n verbinding tussen een maatschappelijk thema en het project, zou dit waarschijnlijk ook niet van de grond komen.

Op dit moment hebben 18 bedrijven een handtekening gezet onder intentieverklaringen of voeren daadwerkelijk onderdelen van het SIA-RAAK- cq. het ROD2.0-project uit. Daarmee kan gezegd worden dat het Rotterdamse open-dataplatform het in termen van impact in, en samenwerking met het bedrijfsleven een van de meest aansprekende voorbeelden is van alle Nederlandse cases.

Kosten en opbrengsten

De gemeente Rotterdam heeft op dit moment slechts beperkt budget beschikbaar om innovatieve samenwerking met de Hogeschool Rotterdam en de andere partners te financieren. Toch heeft zij op een aantal punten het initiatief ondersteund. Zo zijn er een aantal evenementen gesponsord en heeft ROD een eigen ruimte (Stadslab) gekregen in een

(leegstaand) gemeentelijk gebouw. De inschatting van het ROD platform is dat de gemeente in totaal ter waarde van € 130.000,- aan kosten heeft gehad voor haar activiteiten. Het leeuwendeel daarvan zit in het beschikbaar stellen van uren voor het SIA-RAAK-project.

In termen van open-dataproducten die nu al op de markt gebracht worden is de oogst van het eerste collegejaar programmeren nog niet heel groot. Er zijn bijvoorbeeld nog geen apps die al door studenten in appstores beschikbaar worden gesteld. De projecten zijn veelal niet goed genoeg om daadwerkelijk in de markt gezet te worden, maar zijn meer een soort proof-of-concept. Aan de andere kant gaat er nu wel ingezet worden op het verder uitbouwen van de app Rotterdam Onbeperkt, onder andere in samenwerking met Stichting MEE. Ook zijn de voorwaarden voor toekomstige ontwikkeling van open-dataproducten bij bedrijven die participeren in ROD veel beter geworden in de afgelopen jaren. Niet alleen zijn er meer potentiële werknemers die ervaring hebben met open data, ook de ondernemingen die participeren in ROD zelf hebben in het afgelopen en de komende jaren alle mogelijkheden om te leren ondernemen met open data.

De financiële opbrengsten van het project vallen tot nog toe vooral buiten de gemeente zelf.

- De HR heeft, samen met de gemeentelijke diensten Bibliotheek, Gemeentewerken, het Archief en stadsontwikkeling een SIA-RAAK-subsidie (subsidies om onderzoek/onderwijs aan hoge scholen te verbinden met lokale thema's en bedrijfsleven) ter waarde van € 430.000,- binnengehaald. De gemeente heeft voor € 100.000,- aan uren (dus sunk costs) ter beschikking gesteld van dit project.
- Voor het ROD2.0-project (het vervolg) hebben drie bedrijven gezamenlijk capaciteit (ureninzet) ter waarde van € 260.000,- en de HR capaciteit € 150.000,- geïnvesteerd om zo gezamenlijk een subsidie van € 135.000,- van de Rotterdam Media Commission binnen te halen.

In deze projecten doen HR, gemeente en bedrijven gezamenlijk ervaring op met het ondernemen op basis van open data, het samenwerken rondom open data en het onderwijs geven op basis van open data. De HR heeft in het kader van de beide projecten minimaal 3 fte aan nieuwe mensen aan kunnen stellen voor onderwijs en onderzoek. Voor de gemeente is het leren samenwerken met andere partijen in de stad een essentiële waarde die in de komende jaren steeds vaker toegepast zal moeten worden.

De Hogeschool Rotterdam telt op basis van publicaties en media uitingen ook hoeveel publicitaire waarde haar projecten genereren. Dit gaat op basis van het budget dat nodig zou zijn om voor dezelfde hoeveelheid aandacht advertentieruimte in te kopen. De HR schat dat de door haar gegenereerde media aandacht op dit moment zo'n € 200.000,- tot € 300.000,- waard is.

Het expliciteren van de totale kosten en baten van de open-datastrategie blijkt lastig. Verreweg het grootste gedeelte van de kosten (en de baten) zijn onderdeel van grotere budgetten, die toch al vast lagen. *“Met Open Data hebben studenten in het kader van het projectonderwijs een bijdrage aan Rotterdam geleverd. Was er geen open data geweest dan waren er wel projecten geweest. Echter de kwaliteit en de aantrekkelijkheid van het onderwijs is vele malen hoger en de resultaten relevanter!”*, aldus Léon Gommans. Het voordeel van open data is volgens alle geïnterviewden nauwelijks kwantificeerbaar. Iedereen is het ook met elkaar eens dat het belangrijkste resultaat van de open-dataprojecten tot nog toe voornamelijk het leertraject is geweest.

Recente en voorziene ontwikkelingen

Binnen tenminste één project van de dienst Stadsontwikkeling worden de open-dataprincipes al gebruikt. Voor het creëren van een applicatie die als zelftoets voor vergunningverlening werkt, worden diverse databronnen van de eigen dienst, maar ook van anderen ontsloten in een applicatie. Al deze bronnen komen te zijner tijd samen met de app ook onafhankelijk van de app beschikbaar voor hergebruik door anderen. Het beschikbaar stellen van open data, volgt in dit geval dus op het ontwikkelen van functionaliteit in de website van Rotterdam.

Geo-informatie van de provincie Noord-Brabant

Gesproken met:

- Erik Dietvorst (senior adviseur vakgebieden, provincie Noord-Brabant)
- Johan van Aragon (teamleider Geo, provincie Zuid-Holland)
- Marjan Bevelander (senior beleidsadviseur geo-informatie, IPO)
- Kees Schotten (Planbureau voor de Leefomgeving) (hergebruiker)

Het Provinciaal Georegister

Het Provinciaal Georegister (PGR) is een website (www.provinciaalgeoregister.nl) van de gezamenlijke provincies, verenigd in het Interprovinciaal Overleg (IPO). Het Provinciaal Georegister bevat beschrijvingen van en toegang tot datasets van alle Nederlandse provincies.

Het PGR is onder andere opgezet om te voldoen aan de eisen die werden gesteld door de INSPIRE-richtlijn. INSPIRE verplicht alle lidstaten om geo-informatie over 34 thema's openbaar beschikbaar te maken. Omdat een gedeelte van die informatie in Nederland door de provincies beheerd wordt, ontsluiten de provincies die informatie zelf.

Kenmerken van de data

De provincie Noord-Brabant stelt meer dan 50 geografisch gecodeerde datasets (kaarten) ter beschikking aan hergebruikers via het PGR. Het Brabantse gedeelte van het PGR bevat informatie over ruimtelijk beleid, zoals cultuur, waterhuishouding, verkeer, bestuurlijke grenzen etc.

De provincie Noord-Brabant levert via het PGR alleen data waarvan zij zelf de bronhouder is. Er worden dus geen datasets die van anderen gekocht zijn of datasets waarvan anderen (bv. Rijk, gemeenten of waterschappen) bronhouder zijn ontsloten door de provincie. Een aantal gegevenssets worden niet ontsloten, omdat zij privacygevoelige informatie bevatten. Ook worden gegevenssets die vertrouwelijke informatie bevatten niet ontsloten. Gegevens over toekomstige bestemmingsplannen die door grondspeculanten kunnen worden gebruikt om ten koste van de provincie winsten te behalen zullen bijvoorbeeld niet ontsloten worden.

Van alle datasets die in het PGR zijn opgenomen zijn uitgebreide metadateringen beschikbaar. De hiervoor gebruikte standaard is "ISO 19115 voor geografie 1.2". De INSPIRE-verplichtingen kunnen hiermee nageleefd worden.

De datasets verschillen sterk van elkaar in actualiteit. Sommige datasets bevatten archiefdata die nooit meer veranderen, andere worden dagelijks (via koppelingen) bijgewerkt.

Alle datasets worden aangeboden via webservices. Daarnaast is het mogelijk om een groot aantal datasets handmatig te downloaden. Het format waarin de data wordt aangeboden is in vrijwel alle gevallen zowel WMS- als WFS-bestand. Een WMS-bestand is een door de server gegenereerd plaatje van een kaart, terwijl een WFS-bestand alleen de data bevat die in het WMS-plaatje getoond wordt.

Beschikbaarstelling van de data

De datasets die in het PGR zijn opgenomen vallen in drie klassen uiteen:

- Klasse 1: Datasets die verplicht beschikbaar moeten worden gesteld. De belangrijkste vigerende wetgeving is op dit vlak de INSPIRE-richtlijn (en de Nederlandse vertaling daarvan). De provincie Noord-Brabant voldoet op dit vlak aan haar verplichtingen.
- Klasse 2: Datasets waarbij de provincies gezamenlijk hebben afgesproken dat elke provincie deze beschikbaar stelt. Noord-Brabant is over het algemeen een voorloper en heeft datasets in deze categorie vaak al eerder opgenomen. Alle datasets in deze klasse worden volgens vaste structuren opgebouwd. Over de beschikbaarstelling van de datasets worden afspraken gemaakt met afnemers in de gebruikersraad; de uitkomst van het overleg daar is te vinden in de publicatiekalender³.
- Klasse 3: Datasets die door een provincie op eigen initiatief in het PGR geplaatst kunnen worden. Noord-Brabant maakt gebruik van deze mogelijkheid om datasets die vaak bij de provincie opgevraagd worden pro-actief aan te bieden. Wanneer een dataset meer dan drie keer per jaar wordt opgevraagd, wordt deze opgenomen in het PGR. Noord-Brabant heeft dan minder kosten aan beheer van de opgenomen dataset dan wanneer zij deze telkens genereert bij een opvraag.

Alle in het PGR ontsloten data van de provincie Noord-Brabant zijn voorzien van het 'public domain mark', waarmee de provincie aangeeft dat er geen copyrights (of andere rechten) op de data rusten en dat eenieder de data voor elk doel mag hergebruiken.

Gebruik en gebruikers

Er zijn veel gebruikers van het Provinciaal Georegister. Ten eerste zijn de provincies onderling vaak afhankelijk van elkaars data. Daarnaast zijn er veel ketenpartijen- van ingenieursbureaus tot het Planbureau voor de Leefomgeving - die naast de overige bestuurslagen (gemeenten, waterschappen en Rijk) gebruikmaken van de door de provincie verzamelde gegevens. Het PGR maakt deze uitwisselingen van informatie zowel organisatorisch, financieel en technisch eenvoudiger.

De data wordt door deze gebruikers vooral gebruikt voor onderzoek (op diverse gebieden) en voor het plannen en uitvoeren van ruimtelijke ontwikkelingen in diverse sectoren. Een

³ <http://pgrwiki.wikispaces.com/Publicatieafspraken>

voorbeeld van hergebruik ligt bij het Planbureau voor de Leefomgeving, dat veel gebruikt maakt van de PGR-data.

Om de vraagarticulatie te organiseren is er een nationaal georganiseerde gebruikersraad die de belangen van de afnemers van het PGR vertegenwoordigd. Naast vertegenwoordigers van het bedrijfsleven en enkele onderzoeksinstituten neemt hierin ook het collectief HackDeOverheid zitting.

Een bijzondere gebruiksrelatie bestaat hiernaast met het Nationaal Georegister dat sinds kort integraal de data uit het PGR overneemt en publiceert. Dit wederom in het kader van de INSPIRE-richtlijn dat de lidstaten verplicht om een centraal portaal te hebben om alle INSPIRE-informatie te ontsluiten.

Strategie

De strategie om deel te nemen in het PGR, maar ook de gebruiksstrategie van het PGR door Noord-Brabant zijn beide sterk gerelateerd aan het beperken van kosten van het beschikbaar stellen van informatie.

De provincie Noord-Brabant ziet dat zij zelf een verplichting heeft om data ter beschikking aan burgers te stellen. Zowel vanuit de morele gronden (de burger heeft er tenslotte voor betaald) als vanuit wettelijke verplichtingen. De afdeling informatiebeheer heeft al sinds jaar en dag de taak om namens de provincie informatie aan burgers en ketenpartners te verstrekken. Door datasets die enkele keren per jaar opgevraagd worden, eenmalig middels het PGR ter beschikking te stellen, kan deze afdeling efficiënter werken.

Door het PGR in gezamenlijkheid met de andere provincies te ontwikkelen en te beheren (onder de vlag van het IPO) worden ook kosten bespaard. De kosten voor het gezamenlijk beheren en ontwikkelen van het PGR zijn veel lager dan wanneer elke provincie dat afzonderlijk zou doen. De provincie Noord-Brabant heeft bijvoorbeeld haar eigen download-omgeving voor geodata op internet, inclusief de kosten voor de licenties, kunnen beëindigen door deelname aan het PGR.

De provincies zijn gezamenlijk verantwoordelijk voor het PGR en leveren alle een (personele) bijdrage aan de organisatie. Dit is in IPO-verband de gebruikelijke methode om dit soort projecten te financieren. Op die manier hoeven er geen facturen en ingewikkelde verrekeningen plaats te vinden, maar kan er wel gekeken worden of alle provincies over het geheel aan IPO-projecten evenwichtig bijdragen.

Het ligt, vanuit de samenwerkingstraditie van de provincies, overigens niet voor de hand dat er rond het PGR op de middellange termijn een soort 'shared service centre GEO' opgericht

gaat worden. Daardoor zou teveel eigenheid van de provincies verdwijnen en dat wordt binnen de huidige politiek-organisatorische context niet als wenselijk gezien.

Kosten en opbrengsten

De kosten en opbrengsten van het Provinciaal Georegister kunnen op twee niveaus beoordeeld worden. Ten eerste op het niveau van de gezamenlijke provincies en ten tweede op het niveau van de individuele provincie.

Er is door de gezamenlijk provincies in totaal circa € 450.000,- uitgegeven (out-of-pocket) in de afgelopen 2 jaar aan het project Provinciaal Georegister waar het portaal PGR een onderdeel van is. Daarnaast hebben de provincies gezamenlijk nog zo'n 100 dagen (of 800 uur) per jaar besteed aan dit project. In het project Provinciaal Georegister (onderdeel van het programma ProGideon) zijn meer werkzaamheden uitgevoerd dan alleen de werkzaamheden voor het Portaal. Zo is er bijvoorbeeld ook aandacht geweest voor voor de provinciale referentie-architectuur PETRA, standaarden afstemmen op de INSPIRE-richtlijn, het inrichten van redactieraad en gebruikersoverleg, het afstemmen van vraag en aanbod en het inrichten van de beheerorganisatie. De kosten voor de bouw van het portaal zelf bedroegen ca. € 60.000,-.

Het PGR heeft (in combinatie met een deel standaardisatie dat ook het resultaat is geweest van Progideon) per provincie een jaarlijkse baat van naar schatting 100 tot 500 uur opgeleverd. De besparing ontstaat doordat veel minder informatieaanvragen binnenkomen of doordat eenvoudiger kan worden doorverwezen naar het PGR. Wanneer we de kostprijs van een uur op € 50,- stellen, betekent dat een bruto-urenbesparing ter waarde van € 60.000,- tot € 300.000,- per jaar voor de gezamenlijke provincies. De geïnvesteerde tijd moet daar vanaf worden getrokken. Het is overigens waarschijnlijk dat de provincies ook in de toekomst gezamenlijk nog zo'n 100 dagen per jaar kwijt zijn aan beheer en onderhoud van het PGR. De tijdsinvestering ligt daarmee grofweg tussen de 10 en 70 uur per jaar per provincie.

Overigens liggen de baten niet alleen in het voldoen aan INSPIRE en tijdbesparing, maar ook in kwaliteitsverbetering en een vergrote bruikbaarheid. Daarvoor zijn rond het PGR de redactieraad en de werkgroep standaarden actief.

Samenvattend kan gesteld worden dat het ontwikkelen van een stevige gezamenlijke infrastructuur voor het pro-actief beschikbaar stellen van data voor de provincies van grote waarde is. Het is (afhankelijk van de besparingen, op kortere of langere termijn) financieel aantrekkelijk, en verhoogt de mogelijkheden om ook andere processen efficiënter en effectiever in te regelen.

De provincie Noord-Brabant maakt meer gebruik van de mogelijkheden van het PGR dan het gemiddelde, en zij zal dus relatief hogere besparing kennen dan de andere provincies.

Recente en voorziene ontwikkelingen

In de komende maanden en jaren zal het PGR worden uitgebreid met meer datasets van klasse 2, en naar verwachting van klasse 3. Er zal (nog) meer gestuurd worden op het realiseren van baten bij hergebruikers en provinciale organisaties door beter te prioriteren welke datasets in het PGR opgenomen worden. In klasse 1 zal een uitbreiding plaatsvinden met bestanden uit de INSPIRE-annexen II en vooral III.

Ook de aansluiting van het PGR op het nationaal georegister (NGR), een website die qua inhoud zeker niet voorloopt op het PGR, wordt in de komende jaren een belangrijker thema. Ook de mogelijkheden voor aansluiting bij data.overheid.nl, het landelijke portaal met open data, zullen onderzocht en gerealiseerd worden. Verder wordt de aansluiting op portalen zoals RO-online, Risicokaart, het Zwemwaterregister en het Landelijk Grondwaterregister overwogen.

Hergebruik door Planbureau voor de Leefomgeving

Het Planbureau voor de Leefomgeving (PBL) evalueert het beleid. Daarvoor is zij sterk afhankelijk van databronnen van andere organisaties. Het planbureau gebruikt de gegevens om tot inzicht in de implicaties van beleid te komen. Naast beleidsanalyses worden gegevens soms ook op een andere manier gepresenteerd en gecombineerd. Een van de bekendste producten op dit terrein is het Compendium voor de Leefomgeving.

De Rijksoverheid is de belangrijkste gebruiker van producten van het PBL. De data achter de resultaten van het PBL zijn in beginsel openbaar, tenzij daarin op een herleidbare wijze data van derden wordt gebruikt die niet verder mag worden verspreid.

Uit het Provinciaal Georegister gebruikt het PBL data die betrekking hebben op milieu, natuur en ruimte. Daarbij gaat het enerzijds om administratieve en beleidsdata, zoals de begrenzing van provinciale beleidsgebieden (bijvoorbeeld stilte- en grondwaterbeschermingsgebieden). Anderzijds gaat het om meetdata, bijvoorbeeld uit het meetnet grondwater.

Het PBL heeft mede aanleiding gegeven tot de totstandkoming van het PGR, door bij het IPO te pleiten voor één plek waar provinciale gegevens worden ontsloten.

De ontsluiting van provinciale data via het PGR heeft het PBL verschillende voordelen gebracht. Allereerst zijn de data actueler; vroeger kon er meer dan een jaar overheen gaan voordat de data van alle provincies was verzameld en geschikt gemaakt voor analyse of publicatie. Een tweede effect is dat er een flinke kostenbesparing is gerealiseerd op de tijd die nodig was om de provinciale gegevens te verzamelen; het gaat omgerekentdom zo'n

€ 50.000,- per jaar. Ten derde hebben de provincies veel definities geharmoniseerd, wat de onderlinge vergelijkbaarheid ten goede komt. Ten slotte zijn ook de gebruiksvoorwaarden van de data voor een groot gedeelte geüniformeerd. De vier effecten tezamen maken dat het PBL zich meer op haar kerntaak kan richten: het analyseren in plaats van het verzamelen van data. Het PBL is van mening dat de belangrijkste maatschappelijke waarde van het PGR schuilt in toegenomen transparantie en het toegenomen bereik van de data.

Er is regelmatig contact tussen het PBL als hergebruiker en de provincies als bronhouder. Dat loopt meestal via de gebruikersraad van het PGR, waarin het PBL zitting heeft. De raad spreekt onder andere over harmonisatie van definities, de prioritering van datasets en de technische wijze van beschikbaar stellen.

Nu de ontsluiting via het PGR goed geregeld is, zijn de verschillen in definities het grootste knelpunt bij het landsdekkend gebruik van de data. Zo is er een verschil tussen 'stiltegebieden' en 'stille gebieden', en kennen ook niet alle provincies beide typen gebieden. Verdere harmonisatie van (beleids)definities helpt om de data gemakkelijk te kunnen interpreteren en vergelijken.

Het PBL verwacht dat het belang van het PGR toeneemt, nu een verdere decentralisatie van taken naar de provincies op komst is. Voor beleidsanalyses is het van belang om over langlopende tijdreeksen te kunnen beschikken. Bij de decentralisatie ziet het PBL de continuïteit van de gegevensvoorziening daarom als een belangrijk aandachtspunt waar het PGR een rol kan spelen.

Bijlage II. Interviewguide

Inleiding

- Achtergrond onderzoek
- Doel onderzoek
- Interview is niet vertrouwelijk
- Concept casebeschrijving wordt ter validatie voorgelegd
- Inhoud interview

Organisatie en omgeving

1. Wat zijn de kernactiviteiten van de organisatie?
2. Wie zijn de belangrijkste klanten / doelgroepen van de organisatie?
3. Wie zijn de belangrijkste (samenwerkings)partners / opdrachtgevers van de organisatie?
4. Waar leeft de organisatie van en wat zijn de belangrijkste kostenposten?
5. Welke wet- en regelgeving is op de organisatie van toepassing?
6. Wat zijn op dit moment de uitdagingen waar de organisatie voor staat / krijgt de organisatie veel aandacht van politiek of media?

Open-datastrategie

1. Wat is de informatiestrategie?
 - a. Welke rol heeft informatie in de bedrijfsstrategie?
 - b. Heb je meerdere datasets voor dezelfde data?
 - c. Is er een centraal aanspreekpunt?
 - d. Is er een centraal overzicht van alle datasets?
 - e. Heeft elke dataset een eigenaar?
 - f. Zijn er organisatiebrede standaarden voor het naar buiten brengen van informatie? / Hoe zit het met data.overheid.nl?
2. Over welke informatie en data beschikt de organisatie?
3. Welke data zijn open data en welke niet? En waarom wel of niet?
 - a. Welk type data betreft het? (documenten, datasets, beeldmateriaal)
 - b. Wordt de data systematisch verzameld?
 - c. Wordt de data digitaal of analoog bewaard?
 - d. Zijn er juridische beperkingen?
 - e. Wat is de kwaliteit van de data?
 - f. Op welk detailniveau is de data?
 - g. Wat is de actualiteit van de data?
4. Wat is de achtergrond van de open data strategie?
 - a. Sinds wanneer?

- b. Waarom? Waarom past deze open data strategie bij de bedrijfsstrategie?
 - c. Hoe heeft het zich in de loop der tijd ontwikkeld?
- 5. Wat wil men met open data bereiken? Welk doel en welke doelgroep?
- 6. Zijn er concurrerende organisaties, zo ja, welke?
- 7. Wat is de strategie?
 - a. Wordt data direct of via intermediairs geleverd?
 - i. Indien via intermediairs, welke?
 - b. Wat is de prijsstrategie? Gratis, marginaal, integraal, anders namelijk...
- 8. Waar staat de organisatie op dit moment aangaande open data (strategie)? In welk stadium bevindt de organisatie zich?

Uitvoering open-datastrategie: kosten

- 1. Wat is ervoor nodig om de strategie uit te voeren?
 - a. Is er inzicht in de kosten? Wat zijn de kostenposten en wat is de omvang daarvan?
 - i. Hoeveel mensen zijn er mee bezig?
 - ii. Distributiekanaal
 - 1. Website
 - 2. Ontsluitingsinstrument
 - 3. PR
 - 4. Functioneel beheer/helpdesk
 - iii. Afhandeling/verwerking financiële transacties
 - iv. Extra verzameling
 - v. Datasets voor publicatie geschikt maken
 - vi. Privacy checks / verzekering voor privacy fouten
 - vii. Verrijking
- 2. Tegen welke problemen wordt er in de uitvoering aangelopen?

Resultaten strategie: opbrengsten

- 1. Welke data worden afgenomen en welke niet?
- 2. Door wie, waarvoor en op welke wijze?
- 3. Hoeveel wordt er afgenomen?
- 4. Zijn er spin-offs ontstaan? Nieuwe bedrijvigheid etc.
- 5. Wat zijn de opbrengsten (in euro's)? Denk aan:
 - a. Directe opbrengsten:
 - i. Abonnementen
 - ii. Transacties
 - iii. Afkoop uit algemene middelen
 - b. Vermeden kosten
 - i. Transparantievereisten / verantwoording, specifiek: WOB
 - ii. Presentatievoorzieningen

- iii. Voorlichting
- iv. Tools voor eigen primaire proces
- v. Kwaliteitsverbetering
- c. Misgelopen inkomsten
 - i. Inkomsten uit eigen diensten (value-added services)
- d. Waar slaan de opbrengsten neer?
- e. Wat levert dit de organisatie op?
- 6. Zijn de doelen behaald? Waarom wel of niet? Zijn er andere doelen bijgekomen of belangrijker geworden?
- 7. Zijn er aanpassingen in de toekomst voorzien?

Conclusies

1. Als u reflecteert op kosten en opbrengsten van open data in uw organisatie, hoe slaat dan voor u de balans uit?
2. Wat zijn voor u de belangrijkste 'variabelen' aangaande open data?

Bijlage III. Begeleiding van het onderzoek

De begeleidingscommissie van het onderzoek bestond uit:

- Mirjam Kalverda - ministerie van Binnenlandse Zaken en Koninkrijksrelaties
- Imke Arts-Vrijling - ministerie van Binnenlandse Zaken en Koninkrijksrelaties
- Mildo van Staden - ministerie van Binnenlandse Zaken en Koninkrijksrelaties
- Jan den Elzen - ministerie van Financiën
- Richard Blad - ministerie van Economische Zaken, Landbouw en Innovatie

De volgende experts waren - naast de begeleidingscommissie - aanwezig bij een of beide expertbijeenkomsten:

- Robbin te Velde - Dialogic
- Marcel Thaens - HEC/ROI, bijzonder hoogleraar aan de Erasmus Universiteit Rotterdam
- Arre Zuurmond - Zenc
- Léon Gommans - Rotterdam Open Data
- Peter Conradie - Hogeschool Rotterdam
- Paul Suijkerbuijk - De Werkmaatschappij, ministerie van Binnenlandse Zaken en Koninkrijksrelaties
- Tijs van den Broek - TNO

Bijlage IV. Elementsgewijze analyse van de cases

In deze bijlage zijn de cases in een tabel 'gescoord' op de verschillende elementen van de open-datastrategie, kenmerken van de organisatie en omgeving, kenmerken van de data en de effecten. Onderaan elke tabel staat een toelichting.

Open-datastrategie

Open-datastrategie
Beoogde effecten Doelgroep Relatie met hergebruikers Prijs Investeringsbereidheid Gebruiksbeperkingen (Semantische) vorm van gegevensuitwisseling Granulariteit Ruwheid Aanbieden van webservice Aanbieden van presentatievoorziening

Beoogde effecten

Organisaties streven veelal meerdere doelen na met open data. Hieronder staan de hoofddoelen per organisatie weergegeven.

CIBG	Financiële besparing
RCE	Versterking van de legitimiteit Financiële besparing Vergroting van de externe oriëntatie van de organisatie
RIVM	Vergroting transparantie
BZK/wettenbank	Vergroting transparantie Financiële besparing
Rotterdam	Economische waardecreatie

Noord-Brabant/PGR	Vergroting transparantie Financiële besparing
OCW/DUO schoolinformatie	Vergroting transparantie Financiële besparing

Het CIBG heeft met open data bewust gekozen voor het goedkoopste communicatiekanaal. Internet is goedkoper dan telefoon en baliecontact. De RCE kent niet één hoofddoel dat zij met open data wil bereiken. Het voornaamste doel is misschien wel dat zij haar organisatie meer extern georiënteerd wil inrichten, maar ook financiële overwegingen spelen een rol. BZK, de provincie Noord-Brabant en OCW/DUO hebben twee hoofddoelen met open data: het vergroten van de transparantie van de overheidsorganisatie en het maken van een efficiëntieslag en het boeken van een financiële besparing. Rotterdam is de enige case die voornamelijk maatschappelijke effecten beoogt. Met open data wil Rotterdam innovatie en bedrijvigheid in de gemeente stimuleren.

In de loop van de tijd zijn er wel verschuivingen in de beoogde doelen zichtbaar. Vooral bij het ministerie van BZK lijkt het relatieve belang van de financiële voordelen aan belang gewonnen te hebben ten opzichte van het transparantiebelang. De verantwoordelijkheid die het ministerie heeft rond de verbetering van de bedrijfsvoering van de Rijksoverheid valt sterk samen met de wettenbank. Ook bij de provincie Noord-Brabant is een verschuiving te zien. De provincie startte met het PGR en openbaarmaking van gegevens in het kader van onder andere INSPIRE. Inmiddels is het PGR echter ook een tool om de kosten van informatiebevraging te verlagen. De discussie over wat er verder nog bereikt kan worden met het PGR gaat inmiddels al veel verder dan alleen de verplichting, en focust voornamelijk op het uitbuiten van de voordelen van open data voor de burger en vooral ook voor de provincies zelf.

Doelgroep

CIBG	Eindgebruikers
RCE	Eindgebruikers + intermediairs
RIVM	Eindgebruikers
BZK/wettenbank	Eindgebruikers
Rotterdam	Eindgebruikers + intermediairs
Noord-Brabant/PGR	Eindgebruikers
OCW/schoolinformatie	Eindgebruikers

Vrijwel alle organisaties richten zich op het beschikbaar stellen van data voor eindgebruikers: burgers en professionele eindgebruikers. Om het hergebruik voor professionele hergebruikers te vergemakkelijken wordt er veelal een webservice aangeboden. Natuurlijk bereiken deze organisaties ook intermediairs met de data, maar daar wordt niet bewust voor gekozen. De RCE en gemeente Rotterdam vormen hier een uitzondering op. De RCE richt zich veelal op andere organisaties in de erfgoedsector en wil dit netwerk faciliteren. Zo stelt zij data beschikbaar aan Wikipedia en aan erfgoed.mobi. Rotterdam levert (voorlopig) alleen aan partijen die aangesloten zijn bij het Rotterdam Open Data initiatief. Tegelijkertijd leveren deze organisaties echter ook informatie aan andere hergebruikers. Het is dus niet zo dat er exclusief aan intermediairs geleverd wordt.

Relatie met hergebruikers

De relatie met hergebruikers is in sommige gevallen sterk en in andere gevallen non-existent. Enkele cases kennen een gebruikersraad.

CIBG	Niet
RCE	Wel
RIVM	Niet
BZK/wettenbank	Wel + gebruikersraad
Rotterdam	Wel
Noord-Brabant/PGR	Wel + gebruikersraad
OCW/DUO schoolinformatie	Wel

Het CIBG heeft geen enkele bemoeienis met de afnemers die informatie uit de webservice afnemen. Doordat er vragen gesteld worden, wordt de naam van de hergebruiker bekend, maar dat is de uitzondering, niet de regel.

OCW kent in principe de hergebruikers niet, maar krijgt veel meer gevraagd en ongevraagd commentaar van deze groep, waardoor er toch een beeld bestaat van en er een contactmogelijkheid is met de hergebruikers.

Het PGR en de wettenbank hebben een gebruikersraad waarin vooral professionele afnemers van de informatie zitten. Dat zijn in het geval van het PGR vooral milieu-instanties en in het geval van de wettenbank vooral wetgevingsjuristen van andere ministeries.

De RCE en de gemeente Rotterdam kennen de hergebruikers goed. De RCE vanwege de centrale rol in het netwerk van erfgoedbewaarders en het idee dat de organisatie alle spelers in dat veld moet kennen en ondersteunen. Rotterdam heeft vooral een goede relatie met de hergebruikers, doordat gebruikt gemaakt kan worden van de organisatie die al in het extern georganiseerde Rotterdam Open Data initiatief aanwezig is.

Prijs

Behalve de RCE, bieden alle organisaties de data gratis aan. De RCE vraagt een kleine vergoeding, maar deze is ver beneden het niveau van de marginale verstrekingskosten. Alle organisaties hechten zeer aan de (bijnae) gratis beschikbaarheid van de data en zouden alleen met tegenzin de prijs van de data verhogen. Het CIBG, het RIVM, BZK en provincie Noord-Brabant kunnen zich overigens wel voorstellen dat, wanneer excessief gebruik van de data tot hoge kosten voor henzelf leidt, ze op termijn overwegen om alsnog kosten voor het gebruik van de data in rekening te brengen. Zolang de huidige gebruikspatronen echter blijven of slechts in beperkte mate groeien, gebeurt dit niet.

Investeringsbereidheid

CIBG	Beperkt
RCE	Groot
RIVM	Beperkt
BZK/wettenbank	Gemiddeld
Rotterdam	Gemiddeld
Noord-Brabant/PGR	Gemiddeld
OCW/DUO schoolinformatie	Gemiddeld

De investeringsbereidheid van RIVM en CIBG is zeer beperkt, daar zij nauwelijks geld uitgeven aan open data, wanneer dit niet door een opdrachtgever gedekt is. Dit is geen diskwalificatie, want dat is de manier waarop beide instituten normaliter gefinancierd worden.

Alleen bij de RCE is de investeringsbereidheid echt groot, omdat de investering hier € 7,2 miljoen bedroeg, en de terugverdienperiode rond de 5 tot 6 jaar bedraagt. Er kan wel een kanttekening gemaakt worden, omdat een gedeelte van de investering is opgebracht via een PRIMA-subsidie.

De investeringsbereidheid bij de andere cases is niet heel groot, maar ook niet heel klein. Zij investeren tussen de € 40.000,- en € 2,2 miljoen, maar doen die investeringen met korte terugverdientijden, tot maximaal 2,5 jaar.

Gebruiksbeperkingen

CIBG	Geen
RCE	Met bronvermelding
RIVM	Met bronvermelding
BZK/wettenbank	Geen
Rotterdam	Alleen deelnemers in project Rotterdam Open Data (voorlopig)
Noord-Brabant/PGR	Geen
OCW/DUO schoolinformatie	Geen

De RCE en het RIVM vinden het belangrijk om de eigen rol in het tot stand komen van de data te benadrukken. Deels om maatschappelijke relevantie van de eigen instelling voor het voetlicht te brengen, deels om de kwaliteit van de gegevens te duiden. Voor dit laatste is het belangrijk dat hergebruikers de data op de juiste manier interpreteren en representeren. De organisaties geven aan dat het soms moeilijk is om een goede afweging te vinden tussen het voorkomen van verkeerde interpretatie en het mogelijk maken van gebruik van ruwe data.

(Semantische) vorm van gegevensuitwisseling

CIBG	Semantische structuur
RCE	Semantische structuur
RIVM	Semantische structuur
BZK/wettenbank	Juriconnect-standaard
Rotterdam	Divers
Noord-Brabant/PGR	Diverse geostandaarden, o.a. door formele verplichtingen INSPIRE
OCW/DUO schoolinformatie	Semantische structuur

De semantische structuren verschillen in de mate waarin ze zijn uitgewerkt. Waar het RIVM, OCW en het CIBG precies aangeven welke informatie geleverd wordt en welke betekenis deze heeft, is er bij de RCE een onderliggende thesaurus en objectmodel beschikbaar.

De standaarden die in het PGR gebruikt worden zijn voorlopig nog onderwerp van discussie tussen de diverse provincies. De provincie Noord-Brabant probeert wel zoveel mogelijk aan te sluiten bij (inter)nationale open standaarden.

Granulariteit

CIBG	Zeer granulair
RCE	Zeer granulair
RIVM	Matig granulair
BZK/wettenbank	Zeer granulair
Rotterdam	Divers (verschilt sterk per dataset)
Noord-Brabant/PGR	Divers (verschilt sterk per dataset)
OCW/DUO schoolinformatie	Matig granulair (veelal gegroepeerd per school of BRIN-nummer)

In de data van RCE zijn ondermeer individuele monumenten en adressen te onderscheiden. In de BIG-data zijn individuele zorgverleners te herkennen. De wetsteksten in de wettenbank zijn op zinsniveau individueel op te vragen. Dit zijn dus zeer granulaire datasets. OCW heeft in sommige sets (adresbestanden) ook deze granulariteit, maar veel cijfers over bijvoorbeeld prestaties, leeftijd en geslacht van zowel docenten als leerlingen zijn op school gegroepeerd en daarmee minder granulair. Gebruikers kunnen bij het RIVM toegang krijgen tot het kleinste element in de dataset. Wel is het mogelijk dat er (tussentijdse) meetwaarden bestaan die geen onderdeel uitmaken van de dataset die de basis vormt voor de beschikbaarstelling. Dit kan bijvoorbeeld voorkomen bij het berekenen van emissies op basis van verschillende meetwaarden. Niet de metingen zelf, maar de berekende emissie maakt dan onderdeel uit van de emissiedata.

Ruwheid

CIBG	Ruwe data
RCE	Bewerkte data
RIVM	Bewerkte data en ruwe data

BZK/wettenbank	Bewerkte data en ruwe data
Rotterdam	Ruwe data
Noord-Brabant/PGR	Bewerkte data
OCW/DUO schoolinformatie	Bewerkte data

De RCE zorgt ervoor dat objecten altijd ontsloten worden via de informatie-architectuur die ook intern gebruikt wordt. Dit betekent dat de informatielaag in die architectuur (zie casebeschrijving voor uitleg) altijd een bewerking doet op de data die beschikbaar gesteld wordt. De RCE spreekt zelf van het beschikbaar stellen van 'informatie-halfabrikaten'. OCW is ook zeer stellig in het feit dat ze geen ruwe data ter beschikking stelt. De bestanden worden zorgvuldig door de afdeling informatieproducten opgebouwd om de vragen die door burgers (in diverse rollen) worden opgevraagd, zo goed mogelijk te kunnen beantwoorden.

Het RIVM en de wettenbank bieden gebruikers zowel de mogelijkheid om ruwe data als om bewerkte data te gebruiken. In het geval van het RIVM zit dat in het verschil tussen de ongevalideerde en de gevalideerde datasets. Bij de wettenbank kan de gebruiker bijvoorbeeld zowel de officiële bekendmaking (de wijzigingswet) als de geconsolideerde versie van de gewijzigde wet inzien.

Aanbieden van webservice

CIBG	Ja
RCE	Ja
RIVM	Luchtkwaliteit: op aanvraag Emissieregistratie: Nee
BZK/wettenbank	Ja
Rotterdam	Ja
Noord-Brabant/PGR	Ja
OCW/DUO schoolinformatie	Nee

Bijna alle organisaties bieden een webservice aan. RIVM en OCW/DUO zijn hier de uitzondering op. Het RIVM biedt voor de emissiedata geen webservice en vermoedt dat hier ook geen behoefte aan is, omdat al uitgebreide querymogelijkheden op de database bestaan en de data een beperkte ververvingsfrequentie heeft. Voor de luchtkwaliteitsgegevens biedt zij wel een webservice aan, zij het dat de details hiervan alleen

op aanvraag worden verstrekt. OCW/DUO heeft in haar nieuwe website er bewust voor gekozen om geen webservices aan te bieden. Eerder bood zij dit wel aan, maar omdat er maar door één gebruiker gebruik van werd gemaakt en hij de data ook op andere manieren kan lezen, is het aanbieden van webservices geen rendabele investering.

Aanbieden van presentatievoorziening

CIBG	Website + app
RCE	Website + app
RIVM	Website
BZK/wettenbank	Website + app
Rotterdam	Geen
Noord-Brabant/PGR	Website
OCW/DUO schoolinformatie	Website

Rotterdam is de enige organisatie zonder eigen presentatievoorziening. Via het RODS wordt data aan hergebruikers geleverd. Het CIBG, BZK en de RCE bieden bovenop de website ook apps aan. Vanuit het CIBG is dit gewoon ontwikkeld, Een medewerker van het CIBG heeft er twee dagen aan besteed en de app was een feit. De RCE heeft in haar strategie bewust gekozen om apps (zoals de Beeldbank) aan te bieden. Wel stelt zij de data ook ter beschikking aan gebruikers van Wikipedia en andere apps (bijvoorbeeld erfgoed.mobi).

CIBG en de RCE kunnen zich voorstellen dat de apps die zij hebben op termijn verdwijnen, wanneer er voldoende 'marktwerking' is tussen diverse apps die informatie ontsluiten. De overheid ontwikkelt, volgens hen, dan pas ontsluitingstools wanneer de markt hiervoor niet voldoende zorg draagt. BZK kan zich echter niet voorstellen dat de wettenbank ooit alleen maar via apps van derden ontsloten wordt, ook niet als dit er veel zijn en deze voldoende kwaliteit bieden.

Kenmerken organisatie en omgeving

Kenmerken organisatie en omgeving
Trigger Informatie: hoofdtaak of bijzaak? Discretionaire ruimte

Trigger

Voor alle organisaties geldt dat er een trigger was om met open data aan de slag te gaan. Bij de ene organisatie is deze intern gemotiveerd en bij de andere extern.

CIBG	Intern
RCE	Intern
RIVM	Extern
BZK/wettenbank	Extern
Rotterdam	Intern
Noord-Brabant/PGR	Extern
OCW/DUO schoolinformatie	Intern

De strategie van het CIBG is intern gemotiveerd. Omdat de kosten slechts € 2.500,- per jaar waren en er een vermoeden was dat het beschikbaar maken van een webservice mogelijk een verlichting van de (dure) callcenteractiviteiten kan opleveren, is door het CIBG besloten om het BIG-register als webservice aan te bieden.

De RCE wilde zichzelf opnieuw uitvinden, omdat deze vond dat er een geheel nieuwe toolset nodig was om het Nederlandse erfgoedbeheer naar de 21e eeuw te brengen. De werkwijze die voorheen gebruikelijk was, gaf te weinig resultaten. Interne overwegingen speelden ook hier een hoofdrol in de open-datastrategie.

Het RIVM heeft in bijna alle gevallen (open) data gepubliceerd, omdat opdrachtgevers expliciet wilden dat verzamelde data ter beschikking werd gesteld aan derden. Opdrachtgevers wilden dat de data eenvoudig hergebruikt kan worden, omdat daardoor het maatschappelijk belang van de data (en dus van de opdracht) kan groeien.

BZK startte met de wettenbank, omdat het van mening was dat burgers op internet alle wetgeving moesten kunnen vinden en omdat het ministerie vond dat de niet geconsolideerde wijzigingswetten niet voldoende toegankelijk waren. Naast dit 'betere overheid' argument waren echter ook de financiële voordelen voor de rijks- en lagere overheden groot.

Rotterdam is voornamelijk gestart met open data, vanuit de vraag van Hogeschool Rotterdam (HR) en een aantal ondernemingen. De gemeente ging in op het verzoek, omdat het de economie in de havenstad wil stimuleren en verwacht dat open data hieraan bijdraagt.

De provincie Noord-Brabant is, als reactie op de INSPIRE richtlijn, gestart met het pro-actief openbaar maken van gegevens die ze al langer op verzoek verstrekten. Later bleek dat de tools die voor de INSPIRE verstrekkingen ontwikkeld waren ook geschikt waren voor andere data. Inmiddels wordt door data pro-actief online te plaatsen de werklast van de afdeling informatieverstrekking verlaagd.

Bij OCW/DUO bestond een vergelijkbare situatie. De afdeling informatieproducten van DUO was overbelast met aanvragen voor (vaak) gecompliceerde queries op de door DUO beheerde informatie. Door pro-actief informatie openbaar te maken, hoeven de ambtenaren minder individuele queries uit te voeren, omdat deze analyses door de burger zelf kunnen worden uitgevoerd op basis van de bestanden die al online te vinden zijn.

Informatie: hoofdtaak of bijzaak?

CIBG	Hoofdtaak
RCE	Hoofdtaak
RIVM	Hoofdtaak
BZK/wettenbank	Hoofdtaak
Rotterdam	Bijzaak
Noord-Brabant/PGR	Bijzaak
OCW/DUO schoolinformatie	Bijzaak

Voor de provincie Noord-Brabant geldt dat zowel de INSPIRE-data als de andere gegevens die de provincie ontsluit verzameld zijn in het kader van andere taken. Hierdoor is provincie Noord-Brabant geclassificeerd als een organisatie met een andere hoofdtaak dan informatieverzameling.

De RCE beheert voornamelijk informatie, bijvoorbeeld diverse registers en archieven. Hiernaast heeft de RCE een aantal taken die gekoppeld zijn aan die informatie, zoals diverse vergunningsverleningsprocessen. Dit zijn voornamelijk informatiegebaseerde processen. De RCE is daarom geclassificeerd als een organisatie die informatieverzameling en beheer als hoofdtak heeft.

Discretionaire ruimte

CIBG	Klein
RCE	Groot
RIVM	Klein
BZK/wettenbank	Groot
Rotterdam	Groot
Noord-Brabant/PGR	Groot
OCW/DUO schoolinformatie	Groot

RIVM en CIBG zijn instituten, die een aantal beslissingen zelfstandig kunnen nemen. Beide worden echter sterk projectgedreven aangestuurd door de opdrachtgevers en kunnen zelf nauwelijks kiezen hoe bepaalde taken worden ingevuld. De opdrachtgever voor het CIBG is het ministerie van VWS. Voor het RIVM geldt dat een scala aan opdrachtgevers (vaak ministeries) bepalen hoe en welke taken uitgevoerd worden. Het RIVM heeft overigens wel een onafhankelijke positie in de onderzoeksmethodologie, maar veel minder in bijvoorbeeld de presentatie van de resultaten.

Hoewel de RCE een agentschap is van OCW, staat deze organisatie relatief op afstand van het ministerie en heeft het meer autonomie om te bepalen op welke manier de taken worden uitgevoerd. DUO is een baten-lastendienst van OCW en heeft ook relatief veel vrijheid om uitvoeringsbeslissingen te nemen, maar onderhoudt wel intensief contact met OCW als beleidsdepartement.

Kenmerken van de data

Kenmerken data
Privacygevoeligheid Datakwaliteit Actualiteit Kunnen ontlenen van rechten

Privacygevoeligheid

CIBG	Persoonsgegevens
RCE	Afhankelijk van dataset. Interieurfoto's, persoonsgegevens in monumentendossiers wel, archieffoto's bijvoorbeeld vaak niet
RIVM	Geen persoonlijke gegevens
BZK/wettenbank	Geen persoonlijke gegevens
Rotterdam	Geen persoonlijke gegevens
Noord-Brabant/PGR	Geen persoonlijke gegevens
OCW/DUO schoolinformatie	Enige privacy-issues

Het CIBG geeft burgers inzicht in het BIG-register waarin bij uitstek privacygevoelige informatie wordt ontsloten. De RCE heeft mogelijk ook privacygevoelige informatie in de systemen zitten, maar probeert zo min mogelijk van deze data, naar buiten te brengen. Als een persoon bezwaar maakt tegen publicatie van een foto van privébezit (bijvoorbeeld een interieur in een rijksmonument) dan wordt deze direct van de website van de RCE verwijderd.

De informatie die door OCW wordt ontsloten is bewerkt, waardoor docenten en studenten niet persoonlijk te herleiden zijn. Zo zijn er bijvoorbeeld wel gegevens beschikbaar van de gemiddelde leeftijd van de docenten op een school, maar geen gegevens over individuele docenten.

Datakwaliteit

CIBG	Hoog
RCE	Middelmatig

RIVM	Hoog (gevalideerde data)
BZK/wettenbank	Hoog
Rotterdam	Middelmatig
Noord-Brabant/PGR	Middelmatig
OCW/DUO schoolinformatie	Hoog

Het RIVM geeft direct actuele data vrij. In het meetproces kunnen echter fouten optreden. Hiervoor worden er statistische analyses gedaan die na enkele maanden leiden tot gecorrigeerde gegevens. Deze gevalideerde data is daarmee van hoge kwaliteit, terwijl de niet-gevalideerde data van hoge actualiteit is.

OCW gebruikt de jaarlijks vastgestelde datasets die voor de bekostiging van scholen gebruikt wordt, als basis voor de schoolinformatie die vrijgegeven wordt. Er zijn verschillende mechanismen georganiseerd om fouten te corrigeren. Daarmee zijn ook deze datasets van hoog niveau.

De wettenbank koopt redactiediensten bij een grote juridische uitgever om de consolidatie van teksten uit te voeren. In dit redactionele proces wordt zorggedragen voor een hoge kwaliteit van de (geconsolideerde) data.

Rotterdam en provincie Noord-Brabant stellen veel meer data 'as is' ter beschikking. Het gaat hier om veel verschillende data, met diverse achtergronden. Hierdoor is deze data als middelmatig geclassificeerd.

De RCE publiceert bewust data die niet altijd is gevalideerd en gecontroleerd. De organisatie reageert wel op klachten en verzoeken tot verbetering. Voor de informatie die de RCE via Wikipedia heeft ontsloten wordt nu een apart informatieverbeteringstraject ingezet. De kwaliteit van de door de RCE geleverde informatie wordt geclassificeerd als middelmatig.

Actualiteit

Hoe actueel zijn de gegevens die beschikbaar worden gesteld?

CIBG	Zeer actueel
RCE	Matig actueel
RIVM	Zeer actueel

BZK/wettenbank	Matig actueel
Rotterdam	Divers
Noord-Brabant/PGR	Zeer actueel
OCW/DUO schoolinformatie	Ten hoogste 1 jaar na vaststelling 'foto'

Bij het CIBG, het PGR en het RIVM worden zeer actuele cijfers geleverd, doordat de databevraging direct inpikt op de eigen data. Zodra de organisatie over de gegevens beschikt, zijn deze ook beschikbaar voor de hergebruiker. Voor de RCE geldt dit in beperktere mate ook. Een ander gedeelte van de RCE-gegevens kan met vertraging worden vrijgegeven.

De wettenbank kan ook worden geclassificeerd als matig actueel. Aan de ene kant zijn wetten bij inwerkingtreding altijd al in de wettenbank vindbaar, maar andersom duurt het soms een tijd voordat een wet die in de Kamer is aangenomen vindbaar is in de wettenbank.

In Rotterdam zijn sommige datasets zeer actueel (status verkeer, stoplichten), terwijl andere datasets zeer statisch zijn (groenvoorzieningen) en niet vaak ververs worden.

OCW/DUO maakt voor een groot deel van de geleverde informatie gebruik van informatie die in zogenoemde foto's is vastgelegd. Dit zijn statische versies van de databases die een heel jaar voor statistiek en beleidsonderzoek worden gebruikt. Hierdoor loopt de informatiestroom ver achter bij de werkelijkheid, maar is het antwoord op een statistische vraag over de onderwijssector wel dezelfde bij twee opeenvolgende identieke bevragingen.

Kunnen ontlene van rechten

CIBG	Mogen rechten aan ontleend worden
RCE	Mogen rechten aan ontleend worden (registers)
RIVM	Aan gevalideerde data mogen rechten ontleend worden
BZK/wettenbank	Geen
Rotterdam	Geen
Noord-Brabant/PGR	Aan sommige data mogen rechten ontleend worden op basis van wetgeving
OCW/DUO schoolinformatie	Vermelding van DUO als auteur, alsmede versie en toelichting op begrippen.

Niet alle informatie die de RCE beheerd kan door de burger gebruikt worden om zijn/haar rechten af te dwingen. Dit geldt wel voor de informatie die in de wettelijk voorgeschreven registers zit, bijvoorbeeld informatie over het monumenten- of het archeologisch vindplaatsenregister. De data die als open data worden gepubliceerd, zijn daarbij niet anders dan de data die op andere wijzen kunnen worden verkregen.

Op wetten.nl staat (op elke pagina) een disclaimer⁴ waarin staat dat burgers zich alleen mogen verlaten op de originele wetsteksten, niet op de geconsolideerde wetsteksten die door wetten.nl worden aangeboden.

Effecten

Effecten
Voor organisatie Voor overheid als geheel Voor maatschappij

Voor eigen organisatie

CIBG	Versterking legitimiteit
RCE	Financiële besparing Verhoging datakwaliteit Versterking legitimiteit Versterking externe oriëntatie
RIVM	Versterking legitimiteit
BZK/wettenbank	Versterking legitimiteit Financiële besparing

⁴ “De informatie in dit onderdeel vormt geen bekendmaking in de zin van de Grondwet. Alleen publicatie in het Tractatenblad, het Staatsblad, de Staatscourant en andere vanwege de overheid verkrijgbaar gestelde publicatiebladen heeft een officieel karakter.”

Rotterdam	Versterking externe oriëntatie
Noord-Brabant/PGR	Financiële besparing Versterking legitimiteit
OCW/DUO schoolinformatie	Financiële besparing

Het aantal bevestigingen van het BIG-register bij CIBG dat niet per telefoon verloopt, is na de invoering van de webservice verdubbeld (van 80.000 bevestigingen via de website naar 80.000 bevestigingen via website en 80.000 via webservice). Een financiële besparing is nog niet gerealiseerd, omdat er nog geen echte verschuiving is opgetreden in de kanaalmix. Het aantal telefoontjes daalt niet en het aantal websitebezoekers neemt ook niet af.

De RCE ondervindt met de reorganisatie de meeste effecten van de nieuwe open data-activiteiten. De organisatie ziet zowel besparingen in de operatie (in fte) als versterking van de legitimiteit door groeiend gebruik van de digitale kanalen. De RCE krijgt ook feedback op de data en gebruikt deze om de datakwaliteit te verhogen. Tenslotte is de oriëntatie van de dienst steeds meer naar buiten gericht. De oorzaak van de open-datastrategie was wellicht de oriëntatie van (een deel van) het management, maar de verandering in cultuur door de organisatie heen is wel degelijk een gevolg van de invoering van deze nieuwe open dataprincipes.

RIVM wint aan autoriteit en dus aan legitimiteit bij meer gebruik van haar data.

BZK heeft als opdrachtgever van de wettenbank een bescheiden besparing op inkoop van juridische materialen gerealiseerd. Hetzelfde effect is groter voor de hele Rijksoverheid. Daarnaast wint het ministerie vooral aan legitimiteit door de positie als inkoopcentrum voor de hele Rijksoverheid te versterken en doordat meer burgers, sneller en eenvoudiger toegang hebben tot de Nederlandse wetten.

OCW (DUO) ziet net als de provincie Noord-Brabant de kosten van informatieverstrekkingen dalen. Beiden zien een toename van het gebruik van de data. Noord-Brabant krijgt (samen met de andere provincies) meer legitimiteit omdat de positie van de gezamenlijke provincies als informatieleverancier wint, door de standaardisatieslagen die rond het PGR georganiseerd zijn.

Voor de overheid als geheel

De open data-activiteiten van twee cases hebben ook een impact op de overheid als geheel. Dit geldt voor de wettenbank (financiële besparing en standaardisatievoordelen) en voor het PGR (financiële besparing).

De wettenbank heeft voor veel organisaties, zowel binnen als buiten de Rijksoverheid, de kosten van het raadplegen van wetten verlaagd. Niet alleen doordat er geen of minder inkoopkosten voor wettenbundels noodzakelijk waren, ook de snelheid van zoeken en de toegankelijkheid van de data is sterk vergroot. Daarbij is het door de beschikbaarheid van de BWB-standaard mogelijk geworden om de wetgevingsteksten te integreren in diverse overheidsprocessen. Zo wil de belastingdienst automatische links naar de vigerende regelgeving in de (online) documenten opnemen om de beslissingen beter aan de burger te kunnen uitleggen.

Doordat in het PGR een standaardisatieslag is gemaakt over de data van de provincies, kunnen afnemers van deze informatie eenvoudiger informatie op nationaal niveau verzamelen. Dit scheelt tijd voor bijvoorbeeld waterschappen en instituten als het Planbureau voor de Leefomgeving bij het verzamelen van data over heel Nederland, die op provincieniveau wordt beheerd.

Andere organisaties kunnen hier ook een voordeel mee behalen. Het CIBG zou met de webservice in potentie een besparingseffect bij de zorginstellingen kunnen hebben. Dit is echter nog niet het geval, omdat deze instellingen nog nauwelijks gebruik maken van de webservice en nog vaak bellen.

Voor de maatschappij

De effecten die acties van individuele instellingen hebben op de maatschappij in zijn geheel, zijn nauwelijks meetbaar. Toch zijn in kwalitatieve zin effecten te zien die impact hebben op de gehele maatschappij.

CIBG	Betere dienstverlening aan burgers en bedrijven Economische waardecreatie
RCE	Betere dienstverlening aan burgers en bedrijven Economische waardecreatie
RIVM	Vergroting transparantie Betere dienstverlening aan burgers en bedrijven Economische waardecreatie
BZK/wettenbank	Betere dienstverlening aan burgers en bedrijven Vergroting transparantie Economische waardecreatie
Rotterdam	Economische waardecreatie

Noord-Brabant/PGR	Betere dienstverlening aan burgers en bedrijven Vergroting transparantie Economische waardecreatie
OCW/DUO schoolinformatie	Betere dienstverlening aan burgers en bedrijven Vergroting transparantie Economische waardecreatie

Bij het CIBG maken commerciële partijen als Springer Verlag reeds gebruik van de BIG informatie. Springer heeft meerwaarde kunnen realiseren in de huidige productportfolio en processen. Daarnaast heeft het beschikbaar stellen van het BIG-register als open data geleid tot een betere dienstverlening aan burgers geleid, onder andere door de zorglayar, die hierop gebaseerd is.

Bij de RCE is het niveau van dienstverlening aan burgers duidelijk verbeterd. Mensen hoeven niet meer aan de balie langs te komen voor foto's of documenten, kunnen meer gegevens online inzien en gratis (vooral foto's) krijgen. Daarnaast hebben bedrijven als AB-C media producten gebouwd met commerciële waarde (Erfgoed.mobi, waarde circa € 100.000,-).

De data die door het RIVM beschikbaar worden gesteld leiden ertoe dat burgers meer inzicht hebben in de milieukwaliteit van de eigen leefomgeving (cf. Aarhus). Daarnaast gebruiken veel adviesbureaus de data in de analyses bij bijvoorbeeld planologische opdrachten. De data zijn voor hen nu goedkoper te gebruiken, omdat de digitale vorm veel gebruiksvriendelijker is en direct in eigen systemen in te laden is. Noord-Brabant en OCW zijn hiermee vergelijkbaar.

In Rotterdam heeft vooral de hogeschool veel profijt gehad van het open-dataproject. Naast de diverse onderzoekssubsidies die zijn binnengehaald vormden de open data ook een gratis leermiddel.