

Maatschappelijke kosten-batenanalyse open data

dr. ir. F. Welle Donker, dr. ir. B. van Loenen, prof. dr. W.K. Korthals Altes

Dit onderzoek is uitgevoerd in opdracht van:
Het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Auteurs:

dr. ir. F. Welle Donker, dr. ir. B. van Loenen, prof. dr. W.K. Korthals Altes

20 maart 2017

Kenniscentrum Open Data
Faculteit Bouwkunde, Afdeling OTB – Onderzoek voor de gebouwde omgeving
E-mail: opendata-bk@tudelft.nl
<http://www.opendata.bk.tudelft.nl/>



Dit werk is gelicenseerd onder een Creative Commons Naamsvermelding 4.0 Nederland. Bezoek <https://creativecommons.org/licenses/by/4.0/deed.nl> om een kopie te zien van de licentie of stuur een brief naar Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Inhoudsopgave

Samenvatting	i
1 Inleiding	1
1.1 Opendatabeleid	1
1.2 Verschillende vormen van open data	3
1.3 Linked data.....	4
1.4 Korte uitleg maatschappelijke kosten-batenanalyses	5
1.5 Onderzoeksvraag	7
1.6 Onderzoeksmethodologie en leeswijzer.....	8
2 Literatuuronderzoek kosten-batenstudies	10
2.1 Algemene bevindingen	10
2.2 Baten van open data.....	11
2.3 Kosten van open data	15
2.3.1 Kosten verhoudingsgewijs marginaal	17
2.3.2 Niet alle kosten meegenomen/toe te schrijven aan open data?	19
2.4 Baten groter dan de kosten	20
2.5 Reflectie op onderzoeksmethodologie bestudeerde literatuur	20
2.5.1 Gebruik van verschillende methodieken.....	20
2.5.2 Aannames en extrapolaties	21
2.5.3 Verschillende kengetallen of ontbreken van kwantificering.....	21
2.5.4 Geen vergelijkbare uitkomsten	22
2.6 Conclusies	22
2.7 Methodologische lessen uit het literatuuronderzoek	22
3 Basisregistraties Adressen en Gebouwen	25
3.1 Probleemanalyse BAG-data	25
3.2 Huidige situatie: 0-alternatief	26
3.3 Alternatieven voor BAG semi-open data	27
3.4 Kwalitatieve omschrijving van investeringskosten en effecten.....	28
3.5 Monetarisering	32
3.6 Conclusies	33
4 Actuele verkeersinformatie NDW	35
4.1 NDW-samenwerkingsverband en doelen	35
4.2 Huidige situatie NDW-dataverstrekking	36
4.3 Projectalternatieven voor NDW open data	37
4.4 Kwalitatieve omschrijving van investeringskosten en effecten.....	38
4.5 Monetarisering	42
4.6 Conclusies	44

5	Landelijke voorziening WOZ-waarden	45
5.1	Inleiding WOZ-waarden	45
5.2	Huidige situatie ontsluiting WOZ-waarden.....	45
5.3	Projectalternatieven voor de WOZ-waarden.....	47
5.4	Kwalitatieve omschrijving van investeringskosten en effecten.....	48
5.5	Monetarisering kosten en baten	50
5.6	Conclusies	52
6	Kindermishandeling in NL anno 2005	53
6.1	Probleemanalyse.....	53
6.2	Huidige situatie	53
6.3	Projectalternatieven	54
6.4	Kwalitatieve omschrijving van investeringskosten en effecten.....	56
6.5	Monetarisering	58
6.6	Conclusies	60
7	Polisbus: Banen en lonen van werknemers in Nederland	61
7.1	Inleiding.....	61
7.2	Huidige situatie: Polisbus alleen voor gemachtigde gebruikers	61
7.3	Projectalternatieven Polisbusdata	62
7.4	Kwalitatieve omschrijving van investeringskosten en effecten.....	64
7.5	Monetarisering	67
7.6	Conclusies	70
8	Conclusies	71
	Dankwoord	76
	Lijst van geïnterviewden	77
	Geraadpleegde literatuur	78
	Onderzochte kosten-batenstudies	82
	Lijst van afkortingen	86
	Bijlage 1 High value-datasets (lijst BZK)	87
	Bijlage 2 Analyse van alle onderzochte studies	88
	Bijlage 3 Casestudy-effecten van open data op microniveau	95
	B.1. Effecten basisregistratie topografie open data	95
	B.2. Effecten Actueel Hoogtebestand Nederland (AHN)	97
	B.3. Effecten Adresdata Denemarken.....	98
	B.4. Effecten Australian Bureau of Statistics.....	99
	B.5. Effecten Companies House	99
	B.6. Kosten-baten van open data: BZK-onderzoeken	100
	Bijlage 4 ITS Richtlijn 2010/40/EU	104

Bijlage 5 WOZ internationale vergeleken	107
B.1. Open data voor referentiegrondwaarde in Duitsland	107
B.2. Vastgoedwaarde als open data in Noordrijn-Westfalen	111
B.2.1 Overzicht maatschappelijke kosten en baten Noordrijn-Westfalen	112
B.2.2 Conclusie.....	115
Bijlage 6 Internationale vergelijking inkomen/belastingdata	116
B.1. Noorwegen	116
B.2. Verenigd Koninkrijk.....	118

Samenvatting

Dit onderzoek geeft inzicht in de (potentiële) maatschappelijke kosten en baten van open data. Achtergrond vormt de ambitie die de Minister van Binnenlandse Zaken en Koninkrijksrelaties in de Nationale Open Data Agenda (NODA) heeft geformuleerd om zoveel mogelijk overheidsgegevens die zich daarvoor lenen, als open data beschikbaar te stellen. Daarbij geeft het kabinet prioriteit aan 'high value'-datasets: datasets die met voorrang beschikbaar worden gesteld, omdat zij een hoge sociaal-economische meerwaarde hebben of in hoge mate bijdragen aan een transparante en beter controleerbare overheid. In dit onderzoek wordt de verhouding onderzocht tussen de kosten van het beschikbaar stellen van overheidsdata als open data en de maatschappelijke baten daarvan. Het onderzoek bestaat uit twee delen: (1) een internationaal vergelijkend literatuuronderzoek naar de kosten en maatschappelijke baten van het openen van overheidsdata, en (2) een maatschappelijke kosten-batenanalyse van vijf overheidsdatasets. Op grond van beide delen van het onderzoek kan in algemene zin worden geconcludeerd dat de maatschappelijke baten hoger zijn dan de maatschappelijke kosten. Uit de casestudies blijkt dat er wel verschillen zijn tussen de datasets. In veel gevallen kunnen kosten voor open data worden meegenomen in de algemene informatiebeheerkosten en leiden nauwelijks tot meerkosten. In bepaalde gevallen zijn er echter kostbare bewerkingen nodig om te voorkomen dat persoonsgegevens als open data beschikbaar worden gesteld. Voor deze datasets leidt dit tot een minder gunstige verhouding tussen maatschappelijke kosten en baten.

Alle onderzochte studies in de internationale literatuurstudie geven een positief beeld van de verhouding tussen kosten en baten van open data. De kwalitatieve effecten zijn divers en overtuigend. De kwantitatieve studies laten een verhouding tussen kosten en baten zien van 1:1,6 tot 1:70. Dat wil zeggen dat de baten van het openstellen van overheidsdata ruim 1,5 tot 70 keer zo groot zijn als de kosten daarvan. Uit de internationale literatuurstudie blijkt echter ook dat het lastig is om de baten methodologisch goed te kwantificeren en uit te drukken in geld. De meeste studies waarin de baten gemonetariseerd zijn, hebben een zwakke onderbouwing. De resultaten zijn hierdoor lastig op hun merites te beoordelen. Daarentegen zijn veel niet-monetaire effecten zoals meer raadplegingen en downloads van de datasets door het bedrijfsleven, het gebruik van de data in andere domeinen, kwaliteitsverbeteringen van de datasets door terugmeldingen, en efficiëntieslagen binnen de overheid ex-post, 'hard', vastgesteld. De literatuurstudie laat verder zien dat de kosten van het verstrekken van beschikbare data als open data marginaal zijn ten opzichte van de totale kosten van informatievoorziening van een organisatie. Deze literatuurstudie laat ook zien dat de (meer)kosten van open data moeilijk te onderscheiden zijn van de kosten die een dataverstrekker altijd al moet maken om zijn kerntaken uit te voeren.

Naast de literatuurstudie zijn vijf datasets met een potentieel grote toegevoegde waarde onderzocht – zogenaamde high value-datasets. De onderzochte cases zijn de Basisregistraties Adressen en Gebouwen (BAG), de Nationale Databank Wegverkeersgegevens (NDW), de landelijke voorziening WOZ-waarden (LV WOZ), de dataset Kindermishandeling in Nederland anno 2005 en Polisbus (met data over banen en lonen van werknemers in Nederland). De BAG en NDW zijn reeds als open data beschikbaar. Voor de andere datasets zou dat op termijn het geval kunnen worden. In de cases zijn voor drie varianten de maatschappelijke kosten en baten onderzocht: (1) van het aanbieden van de data als open data, (2) het niet als open data aanbieden, en (3) het aanbieden van de data als 'linked open data'. Door data te linken kunnen de gegevens digitaal worden gekoppeld. Dit maakt het makkelijker om nieuwe toepassingen te ontwikkelen met open data uit diverse bronnen.

Uit de cases blijkt dat de kosten van het als open data beschikbaar stellen van overheidsdata laag zijn, mits de data met een geringe bewerkingslag gepubliceerd kunnen worden als open data. Met

name voor data die nu reeds door de overheid tegen betaling worden aangeboden zijn de kosten voor de data als open data beschikbaar stellen marginaal. Deze overheidsdata blijven geproduceerd worden, ook als hier geen inkomsten uit verkoop meer tegenover staan. Het kosteloos toegankelijk maken van de data trekt een nieuwe groep gebruikers aan. Tegenover het wegvallen van de inkomsten staat dezelfde maatschappelijke baat bij de afnemers die eerst betaalden. Ook verdwijnen administratieve lasten rond betalingen. De directe maatschappelijke baten zijn hierdoor groter dan de kosten. Uit de cases blijkt dat ook het hergebruik door de overheid zelf door het verstrekken van de data als open data toeneemt. Dit draagt bij aan de efficiëntie van de overheid.

Daar waar de data niet geschikt zijn om zonder bewerking te publiceren, zijn de verhoudingen binnen de cases anders. Indien er bijvoorbeeld sprake is van persoonsgegevens in een dataset moeten de data worden geanonimiseerd om als open data beschikbaar gesteld te kunnen worden. Het anonimiseren van data brengt extra kosten met zich mee en resulteert in een dataset met een lagere waarde voor hergebruikers. De algemene positieve kosten-baten balans van open data zal hierdoor kleiner worden of mogelijk omslaan naarmate de datasets meer persoonsgegevens bevatten.

De Polisbusdata bevatten bijvoorbeeld gegevens over lonen van individuele personen. Deze persoonsgegevens kunnen niet als open data beschikbaar worden gesteld. Het is sterk afhankelijk van de dataverzameling zelf – en ook de technische karakteristieken van deze dataverzameling – welke kosten er gemoeid zijn met de bewerking van de data om deze niet meer herleidbaar tot personen te maken. Een dergelijke aggregatieslag kan ook de waarde van de data voor hergebruik verlagen. Toch verwachten wij dat de open beschikbaarheid van met name high value-datasets op een iets hoger aggregatieniveau dan de brondata nog steeds een positieve maatschappelijke kosten-batenverhouding zal hebben.

De kosten van het alternatief om de data 'linked' te maken, zijn – op grond van ervaringen in enkele van de cases – geraamd. De kosten van het 'linked' maken betreffen slechts een fractie van de totale kosten van het ontwikkelen en beheren van deze data. De effecten zijn lastig te monetariseren omdat linked (open) data nog in een erg pril stadium is. De verwachting is dat veel van de verwachte positieve effecten van open data veel gemakkelijker bereikt kunnen worden bij het aanbieden van dezelfde data als 'linked data', omdat hierdoor veel betere aansluitingen tussen databestanden binnen de data-infrastructuur mogelijk zijn. Indien de kosten van het 'linked' maken overeenkomstig de geraamde bedragen blijven, verwachten we dat de maatschappelijke baten van high-value datasets hoger zijn dan de maatschappelijke kosten.

1 Inleiding

1.1 Opendatabeleid

De overheid genereert informatie om publieke taken te kunnen uitvoeren. Volgens de Wet openbaarheid van bestuur is deze informatie in principe (er zijn uitzonderingsgronden zoals persoonsgegevens, staatsveiligheid, de eenheid van de Kroon en vertrouwelijke bedrijfsgegevens) openbaar toegankelijk. Openbare overheidsinformatie levert een bijdrage aan de transparantie van de overheid, en aan het informeren en betrekken van de burger in democratische processen. Deze informatie kan passief (op verzoek), actief (uit eigen beweging) of actief als open data toegankelijk worden gemaakt. Bij het actief als open data beschikbaar stellen van informatie is de data niet alleen actief gepubliceerd of op een website of dataportal geplaatst, maar ook geschikt gemaakt voor hergebruik. Dit betekent dat de informatie kan worden verwerkt en gecombineerd zodat er nieuwe producten en diensten kunnen worden ontwikkeld.

Open data vormen een van de pijlers van onze informatie-economie. Open data worden geassocieerd met vele baten, zoals een transparante en efficiënte overheid, het aanpakken van maatschappelijke problemen, en het creëren van economische meerwaarde. Door de verschuiving van het verstrekken op aanvraag (passieve verstrekking) naar proactief verstrekken, kunnen burgers en bedrijven van overheidsinformatie zelf interpreteren en (her)gebruiken voor eigen toepassingen. De Algemene Rekenkamer heeft in het Trendrapport Open Data van 2014 open overheidsdata gedefinieerd als data die:

- Uit publieke middelen bekostigd en gegenereerd zijn bij of voor de uitvoering van een publieke taak, en
- openbaar zijn, en
- vrij zijn van auteursrechten of andere rechten van derden, en
- computer-leesbaar zijn en bij voorkeur aan open standaarden voldoen, en
- voor hergebruik beschikbaar zijn zonder beperkingen zoals kosten of verplichte registratie.

In 2015 heeft de Minister een Nationale Open Data Agenda (NODA) opgesteld om duidelijkheid te bieden over het proces voor tempo, prioriteitstelling en kwaliteit waarmee overheidsdata beschikbaar komen.¹ In de NODA is de opendatadefinitie van de Algemene Rekenkamer overgenomen. In de Leidraad Open Data Gebruik is deze definitie vertaald met acht karakteristieken waaraan open data moeten voldoen: openbaar, gratis, vrij van rechten, zonder registratie toegankelijk, computer verwerkbaar, voorzien van metadata, beschikbaar 'as-is' (zo volledig en onbewerkt mogelijk en qua kwaliteit en actualiteit zoveel mogelijk gelijk aan binnen de overheid gebruikte vorm), en vindbaar.²

Het kabinet heeft in de NODA de ambitie gesteld om zoveel mogelijk overheidsgegevens die zich daarvoor lenen, als open data beschikbaar te stellen. Daarbij geeft het kabinet prioriteit aan high value-datasets: datasets met een potentieel grote toegevoegde waarde.³ Als onderdeel van de NODA

¹ Kamerstukken (2015). Nationale open data agenda 2016 (NODA). Den Haag, Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. kst II 2015/16, 32 802, nr. 20, p. 2.

² <https://data.overheid.nl/over-open-data>.

³ De NODA gebruikt de definitie voor high-value datasets die door Independent Reporting Mechanism (2015). Aligning supply and demand for better governance, open data in the Open Government Partnership is geformuleerd.

is een inventarisatie uitgevoerd waarbij door de departementen zelf is vastgesteld welke datasets als 'high value' aangemerkt kunnen worden.⁴ Bij de inventarisatie van 2016 hebben de departementen in totaal 27 datasets als 'high-value'-dataset aangemerkt.⁵ Daarvan zijn er reeds negentien beschikbaar via data.overheid.nl en staan er vier gepland om met voorrang openbaar te worden gemaakt. Daarnaast zijn er vier datasets waarvan nog wordt onderzocht of ze als open data beschikbaar kunnen komen.

Van open overheidsdata wordt vooral verwacht dat ze marktpartijen zullen stimuleren tot innovatie met hogere omzetten en meer banen als resultaat. Deze potentiële economische impact van open databeleid is het onderwerp van vele studies in binnen- en buitenland geweest. Deze studies voorspellen grote baten: zo zou de Europese informatie-economie met €40 miljard per jaar groeien.⁶ Er is zelfs een potentiële waarde van open publieke sector informatie (PSI) in de EU van USD 900 miljard genoemd.⁷ Maar ook eerder onderzoek uit 2000 beloofde al een groei van €68 miljard per jaar voor herbruikbare overheidsinformatie.⁸ Deze groei zou met name in het midden- en kleinbedrijf plaatsvinden.

Data zullen pas waarde voor een hergebruiker hebben als aan een aantal voorwaarden wordt voldaan (Figuur 1.1). Het aanbieden van data als open data helpt om aan voorwaarden te voldoen. De eerste voorwaarde is dat de gebruiker weet dat een dataset bestaat (kenbaar) en waar deze kan worden verkregen (vindbaar). Als dit niet het geval is, zal de dataset niet kunnen worden gebruikt. Ten tweede moet een dataset bereikbaar zijn: de gebruiker moet de dataset voor zijn doel kunnen gebruiken (is de dataset beschikbaar). Voor de *juridische bereikbaarheid* gaat het om de toegang van gegevens die op basis van wet- en regelgeving verkregen kan worden of beperkt kan worden. Voorbeelden hiervan zijn de toegang via de Wet openbaarheid van bestuur (Wob) en de licentievoorwaarden en de beperkingen van de Wet bescherming persoonsgegevens (privacy). Bij de *financiële bereikbaarheid* is het relevant dat de gebruiker zich kan veroorloven de gevraagde prijs voor een dataset te betalen. Hierbij spelen ook de extra transactiekosten bij het betalen voor data een rol.⁹ Als een dataset bereikbaar blijkt, is, ten derde, de bruikbaarheid van de data van belang. Bruikbaarheid betreft de technische eigenschappen van de dataset en de hanteerbaarheid van de dataset (gezien de middelen waarover de gebruiker beschikt en de doelen die de gebruiker voor ogen heeft). Daarbij is de bestendige beschikbaarheid van de dataset essentieel (geen gebroken links als de dataset wordt verplaatst). Met open data worden met name de bereikbaarheid en bruikbaarheid van de data vergroot.

⁴ Kamerstukken (2015). Nationale open data agenda 2016 (NODA), p.2.

⁵ Zie Bijlage 1 voor de volledige lijst.

⁶ Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Paris, Information Economics: 44. <http://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments>.

⁷ Manyika, J., M. Chui, P. Groves, D. Farrell, S. Van Kuiken & E. Almasi Doshi (2013). Open Data: Unlocking innovation and performance with liquid information, McKinsey Global Institute: 116. http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information, p.6.

⁸ Pira International Ltd, University of East Anglia & KnowledgeView Ltd (2000). Commercial exploitation of Europe's public sector information - Final report, European Commission Directorate General for the Information Society: 132. ftp://ftp.cordis.lu/pub/econtent/docs/commercial_final_report.pdf.

⁹ Zie van Loenen, B. & F. Welle Donker (2014). De stand in opendataland. Delft, Kenniscentrum open data.

Figuur 1.1 Randvoorwaarden voor hergebruik van data (Backx, 2003)



1.2 Verschillende vormen van open data

Tim Berners-Lee heeft het zogenaamde TBL-sterrenmodel, ontwikkeld voor het beschikbaar stellen van open data op verschillende niveaus.¹⁰

★	Beschikbaar op het web (eender welk formaat) maar met een open licentie (Open Data)
★★	Beschikbaar als machine-leesbare gestructureerde data (zoals in Excel in plaats van een plaatje of een tabel)
★★★	Als (★★) met non-proprietary (eigendomsrechtenvrije) formaat (bijvoorbeeld CSV in plaats van een eigendomsmatig formaat zoals Excel)
★★★★	Als (★★★) met het gebruik van de open standaarden Resource Description Framework (RDF) en SPARQL om dingen te identificeren zodat anderen makkelijk naar de data-objecten kunnen verwijzen (gebruik hiervoor URI's)
★★★★★	Als (★★★★) plus: Link je data aan data van anderen om meer context te geven aan de data

Eerste ster-niveau

Het sterrenmodel van Berners-Lee beschouwt het publiceren van data en documenten in pdf of als illustratie (tiff/jpg) als het laagste niveau (1 ster) omdat de data wel voor mensen in te zien en te begrijpen zijn, maar machines de data niet kunnen analyseren of combineren met andere data. Vanuit een kostenooptpunt kan het een goedkope optie zijn om bijvoorbeeld rapporten die al in een pdf-formaat zijn opgemaakt voor intern gebruik, ook voor extern gebruik in die vorm te publiceren.

Tweede ster-niveau

Bij de tweede ster worden de data in een gestructureerd bestandsformaat beschikbaar gesteld, bijvoorbeeld als een Excel-spreadsheet of Word-document. Gebruikers kunnen de data/documenten

¹⁰ Zie <http://5stardata.info/> voor voorbeelden van de sterreniveaus.

nu door computers bewerken/combineren/analyseren/converteren/transformeren, zij het met propriëtaire (eigendomsmatige) software. Het beschikbaar stellen op dit niveau brengt kosten met zich mee omdat de data aan bepaalde kwaliteitseisen moeten voldoen.

Derde ster-niveau

Bij de derde ster worden de data/documenten in een open formaat beschikbaar gesteld, die door (open source) software verwerkt kunnen worden. Gebruikers zijn niet meer afhankelijk van (dure) propriëtaire (eigendomsmatige) software. De conversiekosten zullen afhankelijk zijn van het type data: voor een document of (simpele) spreadsheet zijn die kosten marginaal; voor andere bestanden kunnen de conversiekosten hoger zijn.

Vierde ster-niveau

Het vierde niveau bouwt verder op het vorige niveau, met als aanvulling dat er op dit niveau ook een zogenaamde Uniform Resource Identifier (URI) wordt toegekend. Een URI is een internetprotocol dat eenmalige demonstratie van een bron, informatie of gegevens geeft. Het identificeren van data aan de hand van URI's maakt het mogelijk om gegevens met elkaar te linken. Dat kan op verschillende locatie gebeuren, bijvoorbeeld op internet of in een lokale database. Het voordeel van een URI is dat wanneer de dataset op een andere plaats wordt opgeslagen, de dataset van buitenaf toegankelijk blijft, in tegenstelling tot een URL die een dode link oplevert wanneer dit gebeurt.

Vijfde ster-niveau

Data kunnen ook worden gelinkt door het gebruik van zogenaamde Resource Definition Framework (RDF) schema's (RDF-vocabulaire om de classificering en de eigenschappen van de classificaties te beschrijven) en web ontology language. Hoewel data daarmee meer betekenis krijgen omdat ze gelinkt kunnen worden aan andere data en de context duidelijk is, kost het ook een grotere inspanning om de data op dat niveau te krijgen.

In de NODA wordt gestuurd op het beschikbaar stellen van open data op driesterrenniveau.

1.3 Linked data

Het vierde en vijfde niveau van Tim Berners-Lee beschrijft zogenaamde 'linked data'. Door data te linken kunnen de gegevens digitaal (machine-verwerkbaar) met elkaar in verband worden gebracht en wordt er samenhang aangebracht in de informatie. Ter illustratie: stel, een gebruiker zoekt via een zoekmachine naar informatie over Den Haag. Deze zal dan geen informatie vinden over 's-Gravenhage of The Hague, ook al zijn het namen voor dezelfde stad. Dat komt omdat zoekmachines naar sleutelwoorden zoeken, en niet naar de inhoud kijken. Met 'linked data' worden van woorden of dingen unieke concepten gemaakt. Hoe meer beschrijvingen worden toegevoegd aan een concept, hoe meer betekenis dit concept krijgt. En daarmee wordt de informatie waardevoller door dat de informatie beter vindbaar wordt, en voor meer toepassingen kan worden ingezet. Zo wordt 'linked data' bijvoorbeeld ook gebruikt voor het opsporen van illegale praktijken.¹¹ De kern van 'linked data' bestaat uit semantiek en standaarden. Met semantiek wordt de betekenis van gegevens vastgelegd. Door standaardisatie van het vastleggen van betekenissen en relaties kunnen data deze digitaal worden uitgewisseld.¹²

¹¹ Zie bijvoorbeeld <https://youtu.be/Qhk9ciHlzzo> als voorbeeld van het opsporen van illegale hennepsteelt.

¹² Platform Linked Data Nederland (2014). Linked data in beeld. Amersfoort: 24, p.3. http://www.pilod.nl/w/images/3/32/Linked_Data_in_beeld_2014.pdf.

Om data geschikt te maken als 'linked data', moeten er een aantal stappen worden doorlopen. De dataset moet eerst opgeschoond worden (bijvoorbeeld dubbele gegevens verwijderen). Daarna moeten er een aantal technische testen worden uitgevoerd en moet de dataset gemodelleerd worden en moet er URI's worden toegekend. Vervolgens moet de dataset worden gepubliceerd. Dat kan op een simpele manier als een plat bestand met alleen een URI, maar ook op een complexere manier via speciale toegangskanalen, waarmee data bevroegd en/of gedownload kunnen worden.¹³ Omdat er een grote diversiteit is onder ontwikkelaars (bijvoorbeeld qua technische achtergrond en ervaring), is het van belang dat (linked) open data in verschillende (open) formaten wordt aangeboden.

Het opendataportaal van de Europese Unie stelt de catalogus van metadata beschikbaar als linked data zodat ontwikkelaars makkelijker informatie uit verschillende bronnen kunnen gebruiken. Via het EU opendataportaal zijn nu ook een aantal van de datasets zelf (bijvoorbeeld de data van de European Environment Agency) beschikbaar als linked data.¹⁴

1.4 Korte uitleg maatschappelijke kosten-batenanalyses

Een maatschappelijke kosten-batenanalyse (MKBA) geeft inzicht in de positieve en negatieve effecten van beleidsopties op de welvaart van de maatschappij. Het draait bij een MKBA om het afwegen van verschillende alternatieven op de vraag hoe deze verschillen op het gebied van de totale maatschappelijke welvaart. MKBA's kunnen worden gebruikt om plannen te verbeteren, om prioriteiten te kunnen stellen, of dat het wel verstandig is om een bepaald beleid in te voeren dan wel te wijzigen. Om tot een maatschappelijke kosten-batenanalyse te komen, moeten een aantal stappen worden doorlopen.

1. Allereerst moet het probleem worden geanalyseerd dat een beleidsmaatregel tracht op te lossen.
2. Hierna wordt het nulalternatief – in dit geval de huidige situatie – beschreven en afgezet tegen de beleidsalternatieven.
3. Vervolgens worden de effecten van het nulalternatief en de beleidsalternatieven ingeschat. Bij deze verschillenanalyse worden zowel de directe effecten als de indirecte (afgeleide) effecten en de effecten op de welvaart ingeschat.
4. De geïdentificeerde effecten worden gemonetariseerd, dat wil zeggen, de effecten worden in geld uitgedrukt. Het resultaat is dat voor verschillende jaren in de toekomst de kosten en baten in geld zijn uitgedrukt.
5. Vervolgens worden de kosten en baten van de verschillende jaren in de toekomst met een discontovoet naar één basisjaar teruggerekend. In dit rapport hebben wij een discontovoet van 5,5% gehanteerd.¹⁵
6. Ten slotte worden de uitkomsten van de verschillende beleidsalternatieven gepresenteerd in een tabel.

Een MKBA heeft een aantal beperkingen: MKBA zijn onvolledig omdat onberekenbare welvaartseffecten niet meegenomen kunnen worden.¹⁶ Ten tweede zijn de inschattingen van

¹³ TNO, LOD Stappenplan, <http://www.pilod.nl/wiki/BoekTNO/stappenplan>

¹⁴ Zie <https://data.europa.eu/euodp/nl/linked-data> (geraadpleegd 31/01/2017).

¹⁵ Gebaseerd op advies van de Werkgroep Discontovoet (2015). Rapport Werkgroep Discontovoet 2015, p.4. Dit advies is per 1 april 2016 overgenomen door Ministerie van Financiën, zie <https://www.rijksoverheid.nl/ministeries/ministerie-van-financien/documenten/kamerstukken/2015/11/13/kabinetsreactie-eindrapport-werkgroep-discontovoet>.

welvaartseffecten altijd onzeker omdat in een MKBA de effecten in de toekomst worden ingeschat. Bovendien kunnen niet alle kosten ten behoeve van het uitvoeren van bepaald beleid of een beleidswijziging niet altijd worden gescheiden van kosten die gemaakt moeten worden voor het uitvoeren van primaire taken. Ten slotte zijn er effecten die moeilijk te kwantificeren en/of te monetariseren zijn, en daardoor onevenwichtig worden meegenomen in een MKBA. Dergelijke effecten hebben een zwakkere positie in verhouding tot effecten die makkelijk zijn te monetariseren.¹⁷

Hieronder volgt een korte uitleg van de termen die in de volgende hoofdstukken worden gehanteerd.¹⁸

Consumentensurplus	Het maximumbedrag dat iemand (de consument) bereid is te betalen voor een goed of een dienst, vermindert met het werkelijk te betalen bedrag.
Direct effect	Effecten op de markt waarop een beleidsmaatregel ingrijpt, zijn directe effecten.
Disconteren	Kosten en baten van een project vallen zelden precies gelijk in de tijd. Om de kosten en de baten goed te kunnen vergelijken worden de verwachte kosten en baten in een MKBA teruggerekend naar het moment dat een project start (het zogenaamde basisjaar).
Ex ante effecten	De effecten voordat een beleidsmaatregel wordt ingevoerd.
Ex post effecten	De effecten nadat een beleidsmaatregel is ingevoerd, en of welvaartseffecten daadwerkelijk zijn gerealiseerd als gevolg van die beleidsmaatregel.
Externe effecten	Externe effecten zijn niet-geprijsde effecten van een activiteit op derden waar degene die de activiteit onderneemt geen of onvoldoende rekening mee houdt. Die derden willen het effect niet en ze worden er niet voor gecompenseerd (dit is een negatief extern effect) of ze willen het effect wel maar ze hoeven er niet voor te betalen (dit is een positief extern effect).
Indirecte effecten	Het onderscheid tussen directe effecten en indirecte effecten wordt gemaakt op basis van causaliteit: directe effecten zijn een rechtstreeks gevolg van het project en treden op op de markten waarop een project of beleidsmaatregel ingrijpt. Indirecte effecten zijn daar een afgeleide van en werken door in andere markten.
Netto contante waarde	Het saldo van ingeschatte baten en kosten van een project. Ingeschatte kosten en baten worden in de MKBA voor verschillende jaren gedisconteerd naar het moment waarop het project start. De netto contante waarde berekent men in een MKBA door de contante waarde van de verwachte kosten van een investering af te trekken van de contante waarde van de verwachte opbrengsten.
Nulalternatief	De meest waarschijnlijk ontwikkeling zonder nieuw beleid of beleidswijziging. In een MKBA bepaalt men de beleidseffecten door de

¹⁶ Zie onder meer Eijgenraam, C.J.J., C.C. Koopmans, P.J.G. Tang & A.C.P. Verster (2000). Evaluatie van Infrastructuurprojecten; Leidraad voor kosten-batenanalyse. Deel I: Hoofdrapport. Den Haag, Centraal Planbureau: 84.

¹⁷ Mouter, N. (2012). Voordelen en nadelen van de Maatschappelijke Kosten- en Baten analyse nader uitgewerkt, p.10

¹⁸ Zie <http://www.mkba-informatie.nl/mkba-basics/abc-van-de-mkba/>

ontwikkeling in de wereld zonder nieuw beleid (nulalternatief) te vergelijken met de ontwikkelingen in de wereld met nieuw beleid / beleidswijziging (beleidsalternatief).

PM (pro memorie) post	Wanneer effecten niet met voldoende betrouwbaarheid zijn in te schatten worden ze uitgedrukt als 'pro memorie'. Met 'pm' wordt aan de lezer aangegeven dat deze effecten niet konden worden ingeschat, maar dat dit wel mogelijke effecten van de beleidsmaatregel zouden kunnen zijn en dat deze zouden kunnen worden meegenomen bij de afweging van de beleidsmaatregel.
Rule of Half	Het invoeren van een project, bijv. het aanleggen van een nieuwe weg of het verbeteren van bestaande infrastructuur kan leiden tot een toename van het verkeer over de weg. Aan dit nieuwe verkeer zijn welvaarts-effecten verbonden. De rule of half is een rekenmethode om de totale baten van deze nieuwkomers in te schatten. De rule of half houdt in dat een 'nieuwkomer' gemiddeld genomen half zoveel baat heeft bij een verbetering van de infrastructuur in vergelijking met iemand die in het nulalternatief al gebruik maakt van de infrastructuur.
Welvaartseffecten	Welvaartseffecten zijn zowel financiële als niet-financiële effecten van een project of een beleidsmaatregel op de welvaart van een land of een regio. Wanneer een infrastructuurproject geluidsoverlast voor omwonenden veroorzaakt en dit hun kwaliteit van leven negatief beïnvloed is dit een negatief welvaartseffect van het infrastructuurproject.

1.5 Onderzoeksvraag

Het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties wil inzicht krijgen in de vraag hoe de kosten van het openstellen van overheidsdata – en in het bijzonder van de high value-datasets geïdentificeerd door de departementen – zich verhouden tot de maatschappelijke baten hiervan, een maatschappelijke kosten-batenanalyse (MKBA) van open data dus.

Voor dit onderzoek worden de volgende beleidsvarianten onderzocht:

- Variant 0: Referentiealternatief (0-alternatief). Huidige situatie. Per casus is deze uitgangssituatie verschillend.
- Variant 1: Beleidsvariant - 1 ((Min 1 alternatief): Indien in de casus de data reeds als open data beschikbaar waren, zijn de maatschappelijke kosten en baten onderzocht van het uitsluitend aanbieden van deze data tegen betaling. De dataset is dus alleen als niet-open data beschikbaar.
- Variant 2: Beleidsvariant + 1 (Plus 1 alternatief): Indien de data niet als open data beschikbaar waren, de data in open formaat als open data actief publiceren (TBL 3 sterren).
- Variant 3: Beleidsvariant + 2 (Plus 2 alternatief): Alle data actief publiceren als 'linked open data' (TBL 4 sterren).

Deze rapportage levert een bijdrage aan het onderzoek naar maatschappelijke effecten van open data (en in het bijzonder de kosten en baten) vanuit de volgende twee onderzoekslijnen:

1. Een internationaal vergelijkend literatuuronderzoek.
De deelvraag die hier wordt gesteld is de volgende:

- Welke, waar mogelijk kwantitatieve, conclusies zijn op basis van (inter)nationaal vergelijkend literatuuronderzoek te trekken over de kosten en maatschappelijke baten van het openen van overheidsdata in het algemeen en van high value-overheidsdata in het bijzonder?
2. Een kosten-batenanalyse voor specifieke high value-datasets.
- De deelvraag die hier wordt gesteld is de volgende:
- Hoe verhouden de kosten van het ontsluiten van specifieke high value-datasets zich tot de baten hiervan in verschillende varianten die onderling van elkaar verschillen in ambitieniveaus en snelheden van openstelling?

De eerste onderzoeksvraag wordt in hoofdstuk 2 behandeld. De tweede onderzoeksvraag wordt in hoofdstuk 3 t/m 7 geadresseerd. In deze hoofdstukken worden vijf casestudy's beschreven van datasets die tot high value-datasets kunnen worden beschouwd, namelijk:

1. De Basisregistratie Adressen en Gebouwen (BAG) – Ministerie van Infrastructuur en Milieu (IenM) (Hoofdstuk 3)
2. Actuele verkeersinformatie (NDW) – Ministerie van IenM (Hoofdstuk 4)
3. Landelijke voorziening WOZ-waarden – Ministerie van Financiën (Hoofdstuk 5)
4. Kindermishandeling Dataset – Ministerie van Veiligheid en Justitie (VenJ) (Hoofdstuk 6)
5. Banen en lonen werknemers Nederland Polisbus – CBS (Hoofdstuk 7).

De BAG is reeds beschikbaar als open data volgens beleidsvariant 2 (TBL 3 sterren), en er wordt nu gewerkt aan het beschikbaar stellen volgens beleidsvariant 3 (TBL 4 tot 5 sterren). NDW stelt een aantal van de gegevens beschikbaar volgens beleidsvariant 2 en een aantal gegevens beschikbaar in ruil voor een wederdienst. De WOZ-waarden zijn sinds 3 oktober 2016 openbaar, maar niet beschikbaar als open data voor hergebruik. De Kindermishandeling-dataset is beschikbaar volgens beleidsvariant 2, maar voldoet niet aan alle opendataprincipes uit de NODA omdat de dataset in een propriëtair (eigendomsmatig) formaat beschikbaar is en gebruikers zich eerst moeten registreren. De Polisbus-dataset van CBS is geen openbare dataset omdat de dataset privacygevoelige gegevens bevat. De Polisbus-dataset is nu alleen onder strikte voorwaarden te gebruiken, maar zou theoretisch ook in een aangepaste, anonieme, versie beschikbaar gesteld kunnen worden als open data.

1.6 Onderzoeksmethodologie en leeswijzer

Voor hoofdstuk 2 van dit rapport hebben wij een literatuurstudie uitgevoerd van internationale rapporten en case studies met betrekking tot de baten en, waar mogelijk, de kosten van het laagdrempelig(er) beschikbaar stellen van overheidsinformatie. Hierbij moet worden opgemerkt dat de meeste studies alleen potentiële baten in beschouwing nemen, en geen (maatschappelijke) kosten-baten analyses behelzen. Verder moet worden opgemerkt dat slechts een deel van deze studies de baten en kosten van het beschikbaar stellen van open data beschrijven. Een groot deel van de literatuur is geschreven voordat het concept van 'open data' bestond. Deze rapporten beschrijven veelal de potentiële voordelen van het openbaar maken van en/of het laagdrempelig beschikbaar stellen van overheidsinformatie op macroniveau. Bijlage 2 geeft een uitgebreide analyse van de literatuurstudie uitgevoerd in Hoofdstuk 2, en in Bijlage 3 beschrijven wij de effecten van een aantal case studies op microniveau. Deze case studies betreffen het vrijgeven van specifieke datasets in Nederland, Denemarken, Verenigd Koninkrijk en Australië.

Voor de case studies in hoofdstukken 3 t/m 7 hebben wij gebruik gemaakt desk research, analyse van webstatistieken, en van interviews met data aanbieders, en waar mogelijk, van enquêtes uitgezet

onder gebruikersgroepen, zoals voor de BAG (Hoofdstuk 3). Verder hebben wij voor de BAG gebruik gemaakt van analyses van IP-adressen die gebruik maakten van de BAG.

Voor NDW (Hoofdstuk 4) hebben wij, met dank aan Ecorys, ook gebruik kunnen maken van enquêteresultaten uitgezet onder Data+Diensten-gebruikers. Bijlage 4 beschrijft het Europese juridisch kader van toepassing op NDW.

Voor de WOZ-waarden (Hoofdstuk 5) hebben we een internationale vergelijking gemaakt met open data voor referentiegrondwaarden en vastgoedwaarde in Duitsland. Deze vergelijking wordt in Bijlage 5 beschreven.

Voor de 'Kindermishandeling anno 2005' dataset (Hoofdstuk 6) hebben we gebruik gemaakt van webstatistieken beschikbaar gesteld door DANS

Voor Polisbus-dataset van CBS (Hoofdstuk 7) hebben we gebruik gemaakt van interviews van gebruikers van de dataset. Verder hebben we een vergelijkend onderzoek uitgevoerd naar de beschikbaarheid van inkomsten- en belastingdata als open(bare) data in Noorwegen en in Verenigd Koninkrijk. Deze case studies worden in Bijlage 6 beschreven.

Onze interim bevindingen zijn voorgelegd aan de geïnterviewde partijen, de begeleidingscommissie, en aan de Stuurgroep Open Data. In hoofdstuk 8 geven wij onze conclusies.

2 Literatuuronderzoek kosten-batenstudies

In dit hoofdstuk wordt ingegaan op de vraag: “Welke, waar mogelijk kwantitatieve, conclusies zijn op basis van (inter)nationaal vergelijkend literatuuronderzoek te trekken over de kosten en maatschappelijke baten van het openen van overheidsdata in het algemeen en van high value-overheidsdata in het bijzonder?”

Om een antwoord te geven op de onderzoeksvraag zijn er 54 internationale studies en 11 Nederlandstalige studies naar de baten en de kosten van open overheidsdata en, waar mogelijk, high value-data¹⁹ onderzocht. Veel van de onderzochte studies zijn *ex-antestudies* die potentiële effecten in beeld brengen. Deze ex-ante, macro-economische, studies beschrijven de te *verwachten* effecten van open data. Op datasetniveau zijn er een aantal *ex-poststudies* die terugkijken op wat open data daadwerkelijk voor effect hebben gehad. In dit hoofdstuk worden de belangrijkste bevindingen uit de literatuurstudie van beide typen onderzoek samengevat.

2.1 Algemene bevindingen

De onderzochte studies laten een positieve kosten-batenratio zien. De verhouding tussen de kosten en baten in de onderzochte studies varieert van 1:1,57²⁰ tot 1:70.²¹ De baten bestaan onder andere uit een toename het gebruik, een vergroting van het aantal toepassingsgebieden, een toename van de ontwikkeling van nieuwe producten, efficiëntiewinst voor zowel gebruiker als aanbieder, betere datakwaliteit, een grotere tevredenheid van de gebruiker en een beter imago voor de opendata-aanbieder.

Voor zover de kosten van open data zijn meegenomen in de onderzoeken zijn deze in verhouding tot de totale begroting van een organisatie die de open data beschikbaar stelt marginaal.

De onderzoeken beschrijven deels kwalitatieve en deels kwantitatieve effecten. De kwantitatieve studies hebben vaak een zwakke onderbouwing van de baten. Met name bij de ex-antestudies worden discutabele aannamen en extrapolaties gedaan. Zo zouden bijvoorbeeld de effecten van open data in een specifiek domein in Nieuw-Zeeland representatief zijn voor alle open data in de gehele Europese Unie. Ook bij sommige ex-poststudies wordt op onnavolgbare wijze getracht de effecten in geld uit te drukken. De verhouding tussen het aantal ambtenaren en aantal mensen uit het bedrijfsleven dat een online-enquête heeft ingevuld over open data zou bijvoorbeeld representatief zijn voor de werkelijke verhouding van opendatagebruikers bij de overheid en het bedrijfsleven. Verder wordt er bij het bepalen van de baten van open data in sommige gevallen voor burgers en ambtenaren een dagtarief van €1000 gerekend voor gebruik van de dataset. De resultaten van de kwantitatieve studies moeten dan ook met de nodige scepsis worden bekeken.

¹⁹ High-value-datasets zijn datasets met een potentieel grote toegevoegde waarde. De indicatoren, om te bepalen wat een high-valuedataset is, zijn: vraag van de maatschappij naar een dataset, transparantie, betreft een wettelijke plicht of taak, kostenbesparing, de verwachte impact op een specifieke doelgroep (Nationale Open Data Agenda 2016 (NODA), p.2).

²⁰ OSTP [Ocean Science and Technology Partnership] (2011). Lessons learned from OOS in Canada: Preliminary Assessment of OOS Value: 19. http://www.qc.dfo-mpo.gc.ca/publications/science/documents/Preliminary%20OOS%20value%20assessment_e.pdf

²¹ DECA [Danish Enterprise and Construction Authority] (2010). The value of Danish address data: Social benefits from the 2002 agreement on procuring address data etc. free of charge: 8. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

2.2 Baten van open data

Er is veel onderzoek gedaan naar de directe, indirecte en welvaartseffecten (impact) van open data. Directe effecten zijn een rechtstreeks gevolg van een project en zijn waar te nemen in de markten waarop een project of beleidsmaatregel ingrijpt. Indirecte effecten zijn daar een afgeleide van en treden op andere markten op. Voor open data treden de directe effecten op bij de data-aanbieders en de gebruikers van open data, en indirecte effecten ontstaan als een gevolg van de directe effecten die doorwerken op andere markten. Welvaartseffecten zijn zowel financiële als niet-financiële effecten van een beleidsmaatregel op de welvaart van een land of een regio.²²

In de onderzochte studies worden verschillende baten gemeten. Genoemde baten zijn opgesomd in de tabellen 2.1 en 2.2. De meeste studies benoemen als directe effecten: efficiëntieslagen en lagere transactiekosten bij zowel de overheid als bij de gebruikers, betere dienstverlening, hogere datakwaliteit en economische waardecreatie.

De baten ten gevolge van de indirecte effecten en welvaartseffecten worden vaak als maatschappelijk baten aangeduid zonder deze direct te benoemen. Waar de baten wel benoemd zijn, worden deze uitgedrukt in verbetering van werkcondities, minder luchtverontreiniging of een positief effect op de handelsbalans.

²² <http://www.mkba-informatie.nl/mkba-basics/abc-van-de-mkba/welvaartseffecten/>

Tabel 2.1: Baten ten gevolge van directe effecten genoemd in onderzochte studies

Directe effecten/Baat	Bron
Extra baten bedrijfsleven (bijvoorbeeld nieuwe informatieproducten, complementaire producten, consultancy, etc.)	Dekkers <i>et al.</i> 2006; Pollock 2009, 2011, ACIL Tasman 2008; Coote 2010; Koski 2011; Houghton 2011; Vickery 2011; Kronenburg <i>et al.</i> 2012; Carpenter & Watts 2013; Oxera 2013; Bregt <i>et al.</i> 2013, 2014, 2016
Meer investeringen	ACIL Tasman 2008
Efficiëntere overheid (o.a. minder dubbelingen, minder dataverzoeken, gebruik van dezelfde data binnen hele overheid, lagere transactiekosten aanbieder)	Pollock 2009, 2011; DECA 2010; Danish Government 2012; Kronenburg <i>et al.</i> 2012; Jetzek 2013; RDW 2013; Ubaldi 2013; Algemene Rekenkamer 2014; Bregt <i>et al.</i> 2014; de Vries 2014; Lind 2014; Welle Donker & van Loenen 2015
Efficiëntie bestaande gebruikers (o.a. lagere transactiekosten gebruikers)	Pollock 2011; Kronenburg <i>et al.</i> 2012; Carpenter & Watts 2013; Bregt <i>et al.</i> 2013, 2014, 2016; Welle Donker & van Loenen 2015; Baudel <i>et al.</i> 2016; Ploos van Amstel 2016
Lagere transactiekosten aanbieders	Deloitte 2013; de Vries <i>et al.</i> 2011; Omidyar Network 2014; Baudel <i>et al.</i> 2016; Bregt <i>et al.</i> 2013, 2014; de Vries 2014; Welle Donker & van Loenen. 2015
Extra waarde voor PSI-bronhouders en waardetoevoegende bedrijven	OFT 2006; Dekkers <i>et al.</i> 2006; Pettifer 2009; Houghton 2011; Koski 2011; McKinsey Report 2013; DotEcon 2015
Verbeterde consumentendiensten	Genovese <i>et al.</i> 2010; Coote 2010; McKinsey Report 2103; Omidyar Network 2014; Welle Donker <i>et al.</i> 2015; Ploos van Amstel 2016
Verbeterde datakwaliteit	Genovese <i>et al.</i> 2010; DECA 2010; Kronenburg <i>et al.</i> 2012; McKinsey 2013; Bregt <i>et al.</i> 2013, 2014; Lind 2014; RDW 2013
Eenvoudiger toegang tot data	Genovese <i>et al.</i> 2010; DECA 2010; Kronenburg <i>et al.</i> 2012
Verbeteringen in governance	Omidyar Network 2014
Verbetering van werkcondities	Omidyar Network 2014
Welvaartsbaten	Houghton 2011; Newbery <i>et al.</i> 2008; Pollock 2008, 2011
Verhoging arbeidsproductiviteit	Coote 2010
Toename gebruik	DECA 2010; de Vries <i>et al.</i> 2011; Danish Government 2012; Bregt <i>et al.</i> 2014, 2016; Welle Donker & Van Loenen 2015
Verbreiding toepassingsgebied	Bregt <i>et al.</i> 2014, 2016; Welle Donker & van Loenen 2015
Rijker gebruik	Bregt <i>et al.</i> 2014
Directe financiële baten voor primaire gebruikers	DECA 2010, Bregt <i>et al.</i> 2016
Verandering aard vragen aan de helpdesk	Algemene Rekenkamer 2014; Welle Donker & van Loenen 2015; Bregt <i>et al.</i> 2016
Effect op bestaande betaalde dienstverlening nihil	RDW 2013; Welle Donker & van Loenen 2015
Meer onderzoeks-/onderwijsprojecten	Kronenberg <i>et al.</i> 2012

Tabel 2.2: Baten ten gevolge van indirecte effecten en welvaartseffecten genoemd in onderzochte studies

Indirecte effecten & welvaartseffecten Baat	Bron
Verbetering reputatie aanbieder	Kronenberg <i>et al.</i> 2012; Welle Donker & Van Loenen 2015; Bregt <i>et al.</i> 2016
Uitgaven per huishouden lager	ACIL Tasman 2008
Positief effect op de handelsbalans	ACIL Tasman 2008; Carpenter & Watts 2013
'Maatschappelijke' baten	Deloitte 2013; SCGOE 2014; Pham 2011; Booz Allen Hamilton 2012; Oxera 2013; Baudel <i>et al.</i> 2016; Ploos van Amstel 2016
Betere beslissingen (burgers)	Genovese <i>et al.</i> 2010; McKinsey Report 2013; Ubaldi 2013; Oxera 2013
Vermindering van corruptie/fraude	Eaves 2010; Jetzek 2013; Omidyar Network 2014
Verbetering van werkcondities	Omidyar Network 2014
Welvaartsbaten	Houghton 2011; Newbery <i>et al.</i> 2008; Pollock 2008, 2011
Vermeden kosten/schade	Pham 2011; Booz Allen Hamilton 2012; Oxera 2013; Ploos van Amstel 2016
Meer competitieve markt	Weiss & Pluijmers 2002; OFT 2006; OECD 2006; Pettifer 2009; Pollock 2008, 2011
Grotere gebruikerscommunity	Grus <i>et al.</i> 2015

Kwantitatieve effecten

De macro-economische studies voorspellen grote baten voor open data. Deze baten variëren van een groei van de Europese informatie-economie van €27 miljard²³, tot €40 miljard²⁴ per jaar tot €68 miljard per jaar.²⁵ Er is zelfs een potentiële waarde van publieke sector informatie (PSI) in de EU van USD 900 miljard genoemd.²⁶ Deze groei zou met name in het midden- en kleinbedrijf plaatsvinden.

Voor specifieke datasets zijn de kosten-batenstudies eveneens positief. Het onderzoek naar de effecten van open adresdata in Denemarken bepaalde dat de directe en indirecte baten van gratis adresdata €63 miljoen bedroegen in de periode 2005-2009.²⁷ In Nederland is het effect van de openbasisregistratie topografie geschat op €11,5 tot 14,5 miljoen (in 2013), ten minste €9 miljoen (in 2014) €13,6M (in 2015).²⁸ Voor het open Actueel Hoogtebestand (AHN) wordt door meer dan de helft van de zakelijke gebruikers aangegeven dat het openstellen een positief effect heeft op de omzet van de organisatie. Het opendata-effect wordt geschat op een bedrag van €5,5 miljoen aan directe financiële investering door gebruikers van het AHN. Dit zou met een 30%-belastingtarief tot minimaal €1,65 miljoen per jaar aan belastinginkomsten leiden.²⁹

Meer gebruik en meer uiteenlopende gebruikers

Een direct effect van open data is het toenemende gebruik van de data. Alle studies op datasetniveau laten dit effect zien: er is sprake van een significante toename van het aantal views en/of het aantal downloads. De toename in het gebruik is niet alleen waargenomen bij zakelijke gebruikers, maar ook bij overheden en burgers.

Daarnaast worden de open data ook door nieuwe groepen gebruikers gebruikt in nieuwe toepassingsdomeinen. In het geval van het AHN voerden in het verleden waterbeheersing, archeologie en wetenschappelijk onderzoek de boventoon. Nu wordt de data ook gebruikt in domeinen als Bouw en infrastructuur, Milieu Bodem en Natuur en Ruimtelijke Ordening.³⁰

Ten slotte meldt één onderzoek dat het gebruik ook rijker is geworden: “De basisregistratie topografie wordt steeds vaker gebruikt om deze voor analyses en te koppelen met andere gegevens, en wordt steeds minder vaak als onbewerkt bestand doorgeleverd (van 30% in 2013 naar 10% in 2015)”.³¹

²³ Dekkers, M., F. Polman, R. te Velde and M. de Vries (2006). Measuring European Public Sector Information Resources (MEPSIR) Study. Final report of study on exploitation of public sector information - benchmarking of EU framework conditions, HELM Group of Companies of Moira, Northern Ireland & ZENC, the Netherlands: 94. http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1198.

²⁴ Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Paris, Information Economics: 44. <http://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments>.

²⁵ Pira International Ltd, University of East Anglia & KnowledgeView Ltd (2000). Commercial exploitation of Europe's public sector information - Final report, European Commission Directorate General for the Information Society: 132. ftp://ftp.cordis.lu/pub/econtent/docs/commercial_final_report.pdf.

²⁶ Manyika, J., M. Chui, P. Groves, D. Farrell, S. Van Kuiken & E. Almasi Doshi (2013). Open Data: Unlocking innovation and performance with liquid information, McKinsey Global Institute: 116. http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information, p.6.

²⁷ DECA [Danish Enterprise and Construction Authority] (2010). The value of Danish address data: Social benefits from the 2002 agreement on procuring address data etc. free of charge: 8. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

²⁸ Zie Bregt *et al.* 2014, Grus *et al.* 2015.

²⁹ Bregt *et al.* 2016, p.27.

³⁰ Bregt *et al.* 2016; maar ook Welle Donker & Van Loenen 2015; Bregt *et al.* 2014, 2015.

³¹ Bregt *et al.* 2015.

Verhoogde overheidsefficiëntie en -effectiviteit

De werkelijke effecten op de overheidseffectiviteit en -efficiëntie zijn minder bekend en nauwelijks onderzocht. Veel studies op datasetniveau constateren dat het aantal verzoeken (op basis van de Wet openbaarheid van bestuur) is afgenomen nadat een dataset als open data werd beschikbaar gesteld. Dit is het geval bij de DUO, het AHN, de basisregistratie topografie, bij de data van de Australian Bureau of Statistics, bij de RDW en bij de Companies House in de VK. Bij de DUO heeft open data geleid tot 60% minder verzoeken.³² Bij sommige organisaties heeft (de voorbereiding op) het opendatabeleid geleid tot een vermindering van het aantal dubbele of vergelijkbare datasets. Dit is bijvoorbeeld het geval bij het Ministerie van IenM in het kader van hun opendata-inventarisatie. Het effect is echter niet gekwantificeerd.

Bij de Britse Companies House heeft de verbeterde dienstverlening een positief effect gehad op de efficiëntie, voor zowel Companies House als voor de bedrijven. De kosten per bedrijf waren al gedaald van £24.01 in 2005/2006 naar £15.88 in 2012/2013, en Companies House verwacht dat dit als gevolg van (linked) open data nog verder zal dalen.

Kwaliteitsverbetering van de data

Door de data open te stellen is het voor eenieder mogelijk om de kwaliteit te controleren en eventuele fouten te melden. Dit zorgt voor een verbetering van de kwaliteit van de data. Daarbij komt dat data-aanbieders voordat ze tot open data overgaan veelal de datakwaliteit een impuls geven door bijvoorbeeld metadata aan de data toe te voegen.

Het beschikbaar maken van kadastrale data via TIM-online (web viewing service) leidde bijvoorbeeld tot meer terugmeldingen en een hogere updatefrequentie in Noordrijn-Westfalen.³³ In Denemarken heeft het beschikbaar stellen van adresdata ook geleid tot meer terugmeldingen van fouten, zeker nadat de adresdata werden toegevoegd aan de kaarten van OpenStreetMap.³⁴ Ook bij de RDW werd een kwaliteitsverbetering gesignaleerd.³⁵ Overig onderzoek in Nederland is nog beperkt tot anekdotisch bewijs.³⁶ Wat er ontbreekt, is overkoepelend kwantitatief onderzoek om de effecten van open data voor andere organisaties te staven.

Meldingen helpdesk veranderen van aard

In een aantal onderzoeksrapporten wordt geconstateerd dat de aard van de meldingen bij de helpdesk van een organisatie is veranderd na beschikbaarstelling van data als open data. In plaats van basale vragen over bestellingen (prijs en voorwaarden) van de datasets worden er nu meer inhoudelijke vragen over de datasets gesteld.³⁷

Effect op betaaldiensten van opendata-aanbieder

In twee onderzoeken wordt expliciet ingegaan op het effect van open data op de betaaldiensten van de organisatie: de RDW en Australian Bureau of Statistics (ABS). In het geval van de RDW heeft open data niet of nauwelijks effect gehad op het aantal klanten van de betaalde (en actuelere) datadienst. In het geval van ABS is er zelfs sprake van een lichte toename van de maatwerkopdrachten, mede vanwege de specifieke kennis en bevoegdheden van ABS.

³² Zie Algemene Rekenkamer 2014.

³³ van Loenen, B., J. Zevenbergen, G. Giff and J. Crompvoets (2007). Open toegankelijkheidsbeleid voor geo-informatie vergeleken: het gras leek groener dan het was, Technische Universiteit Delft: 116..

³⁴ Lind, M. (2014). Addresses and Address Data. Socio-economic benefits of Open Address Data experiences in Denmark State of the Map France. Paris, OpenStreetMap France (OSM-FR) http://www.slideshare.net/slideshow/embed_code/33158858.

³⁵ RDW (2013). Eindevaluatierapport - PoC Open Data Voertuigen: 24.

³⁶ zie o.a. Bregt *et al.* 2013, 2014; de Vries 2014; Grus *et al.* 2015; Welle Donker *et al.* 2015.

³⁷ Welle Donker & van Loenen 2015; Bregt *et al.* 2016.

We merken voor beide cases wel op dat de inkomsten uit data relatief laag waren in verhouding tot de voornaamste bron van inkomsten en dat inkomsten uit informatieverstrekking onafhankelijk zijn van de hoofdbron van inkomsten.³⁸

Imago opendata-aanbieder verbeterd

Het blijkt dat open data een positief effect heeft op het imago van een organisatie. Gebruikers van het AHN hebben dit expliciet aangegeven, terwijl ABS dit afleidt aan de hand van de constante stroom positieve berichten in de media.³⁹

De Hogeschool Rotterdam gaf aan voor €200.000 à €300.000 aan publicitaire waarde gegenereerd te hebben met publicaties over (de diverse onderdelen van) Rotterdam Open Data platform.⁴⁰

Impact van verhoogde transparantie

De effecten op de andere beleidsdoeleinden zijn nauwelijks onderzocht. Een verhoogde transparantie van overheidshandelen wordt door de Nationale Ombudsman als een groot goed beschouwd.⁴¹ Echter, niet altijd hoeven de effecten van een transparante overheid positief te zijn. Zo toonde Stephan Grimmelikhuijsen in 2012 aan dat een transparantere overheid niet altijd zal leiden tot meer vertrouwen in de overheid en op korte termijn zou het vertrouwen zelfs kunnen dalen door meer transparantie.⁴²

Buitenlandse studies laten zien dat de effecten van verhoogde transparantie nog nauwelijks waarneembaar zijn, bijvoorbeeld omdat de benodigde infrastructuur nog niet toereikend is.⁴³ Er zijn aanwijzingen dat vooral in ontwikkelingslanden, verhoogde transparantie van overheidsfinanciën een positief effect kan hebben, bijvoorbeeld tegen corruptie.⁴⁴

Overige effecten: Meer onderzoeks-/onderwijsprojecten

De Hogeschool Rotterdam gebruikt open data als lesmateriaal, en heeft bijna €700.000 aan nationale en Europese subsidies binnengehaald voor projecten die betrekking hebben op de open data.⁴⁵

2.3 Kosten van open data

Om data beschikbaar te stellen voor hergebruik, zal een data-aanbieder kosten moeten maken. Deze kosten vallen in drie categorieën: aanpassingskosten, infrastructurele kosten en structurele beheerkosten. Voordat data gepubliceerd kunnen worden, moeten er bewerkingslagen gemaakt worden zoals de data anonimiseren/aggregeren, metadata aanvullen en dataformaten aanpassen. Er moet verder een infrastructuur zijn om de data te ontsluiten. Als er gebruik wordt gemaakt van de

³⁸ zie Welle Donker & van Loenen 2015, p.50.

³⁹ Bregt *et al.* 2016; Australian Bureau of Statistics 2014.

⁴⁰ Kronenburg *et al.* 2012, p.80.

⁴¹ zie https://www.nationaleombudsman.nl/uploads/bijlage/naar_een_open_nederlandse_overheidfinaal_2.pdf

⁴² Grimmelikhuijsen, S.G. (2012). Transparency and trust. An experimental study of online disclosure and trust in government. Ph.D. Dissertation, Utrecht University. <http://dspace.library.uu.nl/handle/1874/218113>

⁴³ Davies, T. (2013). Open Data Barometer. 2013 Global Report Web Foundation and Open Data Institute: 45. <http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf>

⁴⁴ zie bijvoorbeeld Heusser, F.I. (2012). Understanding Open Government Data and addressing its impact (Draft Version) en Khalil, S., W. Saffar and S. Trabelsi (2015). "Disclosure Standards, Auditing Infrastructure, and Bribery Mitigation." *Journal of Business Ethics* **132**(2): 379-399 DOI: 10.1007/s10551-014-2321-6.

⁴⁵ Kronenburg *et al.* 2012, p.80.

bestaande infrastructuur, zullen er toch extra kosten zijn zoals de aanschaf van servers om aan de extra capaciteit te voldoen en om het opendatakanaal fysiek af te schermen van de beveiligde data. Om gebruik van open data verder te bevorderen, zal er mogelijk ook geïnvesteerd moeten worden in het bouwen van tools en apps, en zal de beschikbaarheid van open data gepromoot moeten worden. Ook zijn er kosten voor databeheer: data moeten worden bijgehouden en er zal een faciliteit moeten worden ingericht om met vragen van gebruikers om te gaan.

Tabel 2.3: Algemeen overzicht kosten van open data

Kosten om data geschikt te maken voor ontsluiting (aanpassingskosten)
Kwaliteitsverbeteringen data
Aanbrengen van structuur
Digitalisering van data
Kosten voor een opendatakanaal (infrastructurele kosten)
Bouwen van een website (indien deze er nog niet is)
Beheren van een website
Kosten voor de webservice
Capaciteit voor downloaden en uploaden
Servers
Bouwen van apps, indien de overheidsorganisaties die zelf willen exploiteren
Beheren van apps
Kosten voor operationele opendata-activiteiten (beheerskosten)
Databeheer
Beheren van klantrelaties
Verwerken van financiële transacties
Publiciteitskosten
Kosten voor helpdesk/serviceorganisatie

Naast de kostenposten in tabel 2.3 worden in de door ons onderzochte studies een aantal afwijkende posten genoemd zoals de gevolgen van open data voor de privacy van individuen, de extra kosten voor het trainen van medewerkers en kosten met betrekking tot inkomstendering nu de data niet meer verkocht worden (zie tabel 2.4).⁴⁶

⁴⁶ Bij de onderzochte studies op datasetniveau valt op dat de gedeelde inkomsten vaak klein waren in verhouding tot de totale begroting van de data-aanbieder (zie RDW 2013; Welle Donker & van Loenen 2015).

Tabel 2.4: Overzicht kosten open data genoemd in de onderzoeken

Kosten	Bron
Aanpassingskosten	
Kosten om de kwaliteit van de data te verbeteren	DECA 2010; Deloitte 2012; Denmark Government 2012; Kronenburg <i>et al.</i> 2012; Indecon 2014
Kosten met betrekking tot openbreken van huidige samenwerkingsverbanden voor dataverstrekking	DECA 2010; Oslo Economics 2011
Ontwikkelen van nieuwe vaardigheden van medewerkers aanbieder open data/training	Ubaldi 2013; Omidyar Network 2014; de Vries 2014; PwC & Uscreates 2015
Infrastructurele kosten	
Upgrade netwerkinfrastructuur	DECA 2010; Kronenburg <i>et al.</i> 2012; Ubaldi 2013; de Vries 2014
Het online zetten van de open data	DECA 2010; Kronenburg <i>et al.</i> 2012; Ubaldi 2013
Beheerskosten	
Kosten met betrekking tot het stimuleren van open data (o.a. organiseren open data challenges)	Kronenburg <i>et al.</i> 2012; PwC & Uscreates 2015; Krishnamurthy & Awazu 2016
Kosten gerelateerd aan misbruik	McKinsey Report 2013
Overige kosten	
Privacy	McKinsey Report 2013
Inkomstenderving door open data	Oslo Economics 2011; Houghton 2011; Deloitte 2013; Carpenter & Watts 2013; Bregt <i>et al.</i> 2013, 2014; Omidyar Network 2014; de Vries 2014
Verlies van baten door geen open data aan te bieden (geen toegang tot open data)	PIRA <i>et al.</i> 2000; OFT 2006; ACIL Tasman 2009; Pham 2011
Extra kosten uit algemene middelen betaald	OFT 2006; Pollock 2008, 2011; DECA 2010; Vickery 2011; de Vries <i>et al.</i> 2011; de Vries 2014

2.3.1 Kosten verhoudingsgewijs marginaal

Voor zover de kosten van open data zijn meegenomen in de onderzoeken, zijn deze verhoudingsgewijs marginaal.

De Algemene Rekenkamer heeft berekend dat de kosten voor het aanbieden van open data door de Rijksoverheid over het algemeen slechts een fractie van de begroting bedragen: gemiddeld 0,01%.⁴⁷ Ook bleek uit hetzelfde onderzoek dat het aantal mensen dat specifiek voor open data wordt ingezet beperkt blijft tot 0,14% van het totale aantal fte's van een organisatie.⁴⁸

De Vries (2014) schat per data-aanbieder deze kosten in op gemiddeld circa €50.000 veranderkosten (eenmalig) en daarna circa €15.000-45.000 per jaar (infrastructuur en beheer).⁴⁹ Voor sommige organisaties zijn de veranderkosten veel hoger, bijvoorbeeld de Rijksdienst voor het Cultureel Erfgoed (RCE) raamde de totale veranderkosten op €7,5 miljoen⁵⁰. RIVM gaf aan dat er geen extra kosten zijn gemaakt om luchtkwaliteit- en emissiedata beschikbaar te stellen, maar dat open data een verschuiving van kosten binnen de eigen organisatie inhield. RIVM, RCE en DUO gaven aan dat de meeste veranderkosten gemaakt zijn voor het geschikt maken van open data (datakwaliteit verhogen, anonimiseren/aggregeren, etc.).⁵¹

⁴⁷ Hoewel procentueel sprake is van een marginale kostenpost staat 0,01% van de begroting van het Ministerie van IenM wel gelijk aan ruim €800.000 (0,01% van €8.175.961.000).

⁴⁸ Algemene Rekenkamer, 2014, Trendrapport open data, p.13.

⁴⁹ De cases bestonden uit open data aangeboden door het KNMI, het Kadaster, de RCE, en de gemeenten Enschede en Rotterdam.

⁵⁰ De RCE beschouwde alle kosten die gemaakt moeten worden voor het digitaliseren van informatie en het verbeteren van de informatievoorziening, in totaal €7,5M, in zijn geheel als investeringen voor open data (Kronenburg *et al.* 2012, p.31).

⁵¹ Kronenburg *et al.* 2012.

Onderzoek door RDW in 2013 komt op €92.000 voor de voorbereiding en €100.000-120.000 per jaar aan kosten voor de infrastructuur en beheer van de RDW open data.⁵² De beheerkosten van de Deense adresdata zijn circa €200,000 per jaar.⁵³ Ook Welle Donker en van Loenen komen in 2015 met vergelijkbare kosten van open data voor de Australian Bureau of Statistics (ABS) en de Britse Companies House (CH), waarbij opgemerkt kan worden dat deze organisaties die kosten beschouwden als integraal onderdeel de reguliere ICT-uitgaven. Liander heeft een extra server aangeschaft om kleinverbruik-energiegegevens beschikbaar te stellen als open data in 2013. De aanpassingskosten van het opendataproject waren 3,3 fte, incl. 1,1 fte voor wetenschappelijk onderzoek naar de randvoorwaarden van open data.⁵⁴ De beheerkosten van open data zijn marginaal omdat deze nu als een integraal onderdeel worden beschouwd van de reguliere beheerkosten.⁵⁵

De Vries (2014) concludeert uit de vijf casestudy's dat er grote verschillen zijn tussen het type data-aanbieder: Aan de ene kant van het spectrum staan de organisaties van wie het een kerntaak is om data te verstrekken (KNMI, Het Kadaster) en aan de andere kant staan de andere organisaties zoals de gemeenten en RCE. In de tabellen 2.5 en 2.6 wordt een samenvattend overzicht van de baten en kosten van de BZK-onderzoeken gegeven.⁵⁶

⁵² Welle Donker & Van Loenen 2015.

⁵³ DECA 2010.

⁵⁴ Welle Donker, F., B. Van Loenen & A. Bregt (2014). Open Data and Beyond II: Measuring the impact and modelling the data. Delft, Kenniscentrum Open Data, TU Delft: 40. http://www.bk.tudelft.nl/fileadmin/Faculteit/BK/Over_de_faculteit/Afdelingen/OTB/Kenniscentrum_Open_Data/NGI_monitoring_rapport_Liander_TUD_WUR.pdf

⁵⁵ Welle Donker, F. & B. van Loenen (2015). Duurzame businessmodellen voor open data. Delft, Kenniscentrum Open Data: 60. <http://repository.tudelft.nl/view/ir/uuid%3Ac5b18289-81d9-4692-ab2e-62b3a96c02d2/>

⁵⁶ Zie de Vries, M. (2014). Wah kos'dah dan? Onderzoek naar de incrementele kosten van aan Open Data doen: 60. <http://open-overheid.nl/open-overheid/open-data-wah-kos-dah-dan/>. Kronenburg *et al.* hebben in 2012 in opdracht van het Ministerie van BZK een *ex-ante* onderzoek uitgevoerd naar de benodigde kosten en potentiële baten van het beschikbaar stellen van overheidsdata als open data. De casestudy's betroffen het beschikbaar stellen van bestanden van de Rijksdienst voor het Cultureel Erfgoed (RCE), luchtkwaliteit en -emissiegegevens van de Rijksinstituut voor Volksgezondheid en Milieu (RIVM), het BIG-Register van het Centraal Informatiepunt Beroepen Gezondheidszorg (CIBG), de Wettenbank van het Ministerie van BZK, gemeentelijke data van de gemeente Rotterdam, provinciale geografische (meta)data via het Provinciaal GeoRegister (PGR) en onderwijsdata van de Dienst Uitvoerend Onderwijs (DUO). De Vries heeft in 2014 een *ex-post* onderzoek uitgevoerd naar de kosten van overheidsdata die als open data beschikbaar waren gesteld. De casestudy's betroffen data van het KNMI, het Kadaster, de RCE en de gemeenten Enschede en Rotterdam.

Tabel 2.5: Samenvattend overzicht van de bevindingen in de 5 cases in euro's (de Vries 2014, p.25)

Case	Begin uitvoering	Data	Kosten in euro's			Dekking	Baten
			Infra-structuur	Verander-kosten	Beheer-kosten		
<i>Rotterdam</i>	2011	Zeer divers, 137 sets en API's	20.000/jr.	50.000 1x	45.000/jr.	Extern (2011), daarna budget CIO	Onbekend
<i>Enschede</i>	2012	Zeer divers, 23 sets	10.000 1x	8.000 1x	Minimaal	Budget afdeling IM	Onbekend
<i>KNMI</i>	2013/14	real-time, klimaat, seismologisch	4.000/jr.	50.000 1x	10.000/jr.	Vooralsnog eigen middelen	80.000/jr. vanaf 2015/17
<i>Kadaster</i>	2012	Basisregistratie Topografie	0, lift mee in PDOK	75.000 1x	25.000/jr.	250.000 van Ministerie van IenM, rest eigen middelen, incl. inkomstenderving €30.000/jr.	115.000/jr. vanaf 2014
<i>RCE</i>	2010	Cultureel erfgoed, zoals monumenten	0	0	0	Ministerie van OCW, subsidies, eigen middelen	Onbekend

Tabel 2.6: Samenvattend overzicht van de bevindingen in de 7 cases in euro's (Kronenburg et al. 2012)

Case	Begin invoering	Data	Kosten in euro's		Baten in euro's
			Infrastructureel	Veranderkosten	
RCE	2011	Cultureel erfgoed	60.000/jaar voor de Beeldbank	7,5 miljoen	Intern: 7 fte/jaar (500.000/jr) minder + onderhoudscontracten (300.000/jr) extern: gemeenten besparen ca. 7 minuten per opvraging monumentendossiers
BZK	2001	Wettenbank		300.000	Lagere transactiekosten
CIBG	2012	BIG Register	15.000	Moesten toch al gemaakt worden	Intern: 80.000 meer bevestigingen per maand zonder extra personeel
RIVM	-*	Luchtkwaliteit en -emissiedata	100.000	250.000	Extern: waarschijnlijk lagere transactiekosten bijv. Milieudefensie
Rotterdam	2012	Divers		130.000, waarvan 10.000 faciliteren HR	Extern: HR: bijna 700.000 aan subsidies + 200.000-300.000 aan publicitaire waarde
DUO	2011	Onderwijsdata		500.000	4.5 fte besparen in 2 jaar
Provincies	2007	Opzetten PGR	60.000	450.000 + 800 mensuren in 2 jaar	Intern: Transactiekostenbesparing provincies 100-500 mensuren extern: bijv. PBL 50.000/jaar

* Invoerdatum niet vermeld in rapport.

2.3.2 Niet alle kosten meegenomen/toe te schrijven aan open data?

De studies geven slechts beperkt inzicht in de kosten van open data. De meeste onderzoeken richten zich zelfs alleen op de baten en gaan niet in op de kosten.⁵⁷ Een enkel onderzoek richt zich slechts op de kosten van een enkel project of een platform.⁵⁸

⁵⁷ Zie Castelein et al. 2010; Genovese et al. 2010; Vickery 2011; McKinsey Report 2013; Oxera 2013; Deloitte 2014; SCGOE 2014.

⁵⁸ Bijvoorbeeld PwC & Uscreates 2014 en Krishnamurthy & Awazu 2016.

Waar kosten worden genoemd, wordt vaak gesteld dat de data-aanbieders toch al die kosten moeten maken om de data op kwaliteitsniveau te krijgen voor de eigen bedrijfsvoering⁵⁹ of dat de kosten voor open data niet te splitsen zijn van de reguliere bedrijfsvoering.⁶⁰

Als voorbeeld kunnen we de investering noemen die het Deense Ministerie van Volkshuisvesting, Stedelijke en Plattelandszaken (MBBL) heeft gepleegd om de IT-infrastructuur te verbeteren. Tussen 2010 en 2014 werden €5 miljoen (eenmalig) geïnvesteerd om de IT-infrastructuur voor een nieuw register en nieuwe distributiedienst in te richten, met 85 extra fte bij de gemeenten en het MBBL.⁶¹

2.4 Baten groter dan de kosten

In de onderzochte studies worden de baten (zoals beschreven in paragraaf 2.2) vele malen hoger geschat dan de kosten (zoals beschreven in paragraaf 2.3). De verhouding tussen de kosten en baten varieert van 1:1,57⁶² tot 1:70.⁶³

In Bijlage 2 staat het totale overzicht aan bevindingen van de studies, inclusief een kwantitatieve uitkomst.

2.5 Reflectie op onderzoeksmethodologie bestudeerde literatuur

2.5.1 Gebruik van verschillende methodieken

In de studies wordt zo'n variëteit aan onderzoeksmethodes en -onderwerpen gebruikt dat we kunnen concluderen dat ze niet of nauwelijks met elkaar te vergelijken zijn.⁶⁴ Vaak ontbreken de brongegevens zodat de uitkomsten niet geverifieerd kunnen worden. Dit betreft met name studies die een grote impact voorspellen van open data⁶⁵, maar waarin kennelijk wat het delen van de eigen onderzoeksdata betreft een andere keuze gemaakt is. In veel studies wordt de methodologie niet onderbouwd.

Slechts in een deel van de studies is een maatschappelijke kosten-batenanalyse (MKBA) gehanteerd; wat gegeven het abstractieniveau van veel studies ook een terechte keuze is. Al een halve eeuw is bekend dat MKBA's niet voor alle doeleinden geschikt zijn.⁶⁶ Met name voor relatief vage en alomvattende alternatieven (als de maatschappelijke effecten van de utopische situatie dat alle overheidsdata ter wereld open zouden zijn) is deze methode niet geëigend. De in Nederland ontwikkelde Algemene leidraad voor maatschappelijke kosten-batenanalyse is ontwikkeld om ervoor te zorgen dat deze methode op een verantwoorde wijze wordt gebruikt. Een studie naar het effect van open data op het gebied van het gehele overheidsapparaat valt buiten de scope van deze leidraad. Cruciaal bij een MKBA is immers dat de "probleemanalyse aansluit bij de voorgenomen beleidsmaatregelen".⁶⁷ Op een zeer algemeen niveau kunnen geen beleidsmaatregelen worden genomen. Daarvoor is een zekere mate van concreetheid noodzakelijk.

⁵⁹ Zie Danish Government 2012; en Indecon 2014.

⁶⁰ Zie Kronenburg *et al.* 2012; de Vries 2014; Welle Donker *et al.* 2015; Welle Donker & van Loenen 2015.

⁶¹ Zie Lind 2014.

⁶² OSTP 2011.

⁶³ DECA 2010.

⁶⁴ Zie ook Häggquist & Söderholm 2015.

⁶⁵ O.a. PIRA *et al.* rapport uit 2000, Vickery rapport uit 2011 en McKinsey Report uit 2013.

⁶⁶ Bijvoorbeeld Prest & Turvey 1965; Wildavsky 1966.

⁶⁷ Romijn & Renes 2013, p.12.

2.5.2 Aannames en extrapolaties

De uitkomsten van een aantal onderzoeken zijn veelal gebaseerd op geaggregeerde waarden van een bepaalde marktsector zoals omzet van gerelateerde producten en diensten en/of de productiekosten, en deze studies zijn vaak sterk afhankelijk van aannames en extrapolaties.⁶⁸ Soms worden de resultaten van een sector in een land geëxtrapoleerd naar het BNP van dat land.⁶⁹ Andere onderzoeken bouwen voort op deze studies en extrapoleren verder naar regionaal niveau. Zo zou volgens het Vickery rapport uit 2011 de Nederlandse geosector in combinatie met de Australische geosector representatief zijn voor de gehele EU27-markt voor hergebruik van alle overheidsinformatie.

Hergebruik van onderzoeksresultaten

Een voorbeeld van de wijze waarop wordt voortgebouwd op eerder onderzoek vormt de rapportage van Vickery (2011) die volledig gebruik maakt van eerder uitgevoerd onderzoek en de resultaten van nationaal onderzoek in Nederland en Australië (Castelein *et al.* 2010 en ACIL Tasman 2008) extrapoleert naar het niveau van de EU. Vickery komt tot een omvang van de EU27-markt op basis van onderzoek van Castelein *et al.* (2010) naar de werkgelegenheid in de geosector in Nederland. Deze werd op basis van €100.000 per fte per jaar in totaal (dus inclusief onderzoekinstellingen, de publieke sector en de private sector) geschat op €1,4 miljard. Dit stond gelijk aan 0,23% van het Nederlandse BNP. Vervolgens is, volgens de methodologie van Vickery, de omvang van de EU27 PSI-geosector gelijk gesteld aan 0,23% van het Europese bruto product (€27 miljard). ACIL Tasman (2008) toonde aan dat in de deelstaat Victoria (Australië) geodata ongeveer 50% van de totale PSI-markt uitmaakte. Hiervan gebruikmakend komt Vickery uit op een omvang van de EU27 PSI-markt van €54 miljard. We kunnen kanttekeningen maken bij de mate van representativiteit voor de onderzochte regio (bijv. Australische studies om de waarde van overheidsdata te bepalen in de EU). Door (mogelijk optimistische) extrapolaties krijgen relatief kleine onzekerheden een steeds grotere betekenis.

2.5.3 Verschillende kengetallen of ontbreken van kwantificering

In de cases zijn niet-overeenkomende kengetallen gehanteerd. Zo hebben alle hergebruikers soms hetzelfde tarief van €1.000 per dag⁷⁰ of zijn de personele kosten van een ambtenaar werkzaam op een geografische informatieafdeling €100,000 per jaar,⁷¹ worden die personele kosten in de publieke sector onderdeel gemaakt van de markt,⁷² en wordt gebruik gemaakt van single casestudy's/narratief en/of anekdotisch bewijs dat vervolgens representatief wordt verondersteld voor een hele markt.⁷³

Een andere tekortkoming van veel van deze studies is dat de maatschappelijke baten niet of nauwelijks zijn onderzocht. Als de maatschappelijke baten worden genoemd, is dat op basis van anekdotische aanwijzingen die niet gekwantificeerd zijn of zeer grof zijn geschat. Zo benoemt Ploos van Amstel verminderde CO₂ en betere luchtkwaliteit als een significante maatschappelijke baat maar kwantificeert deze baten niet.⁷⁴

⁶⁸ Zie onder meer Omidyar Network 2014; Vickery 2011; Dekkers *et al.* 2006.

⁶⁹ Bijvoorbeeld in Castelein *et al.* 2010; ACIL Tasman 2008 en 2009.

⁷⁰ Bregt *et al.* 2014; Grus *et al.* 2015; Bregt *et al.* 2016.

⁷¹ Castelein *et al.* 2010.

⁷² Vickery 2011 op basis van Castelein *et al.* 2010.

⁷³ Bijvoorbeeld ACIL Tasman 2008; Vickery 2011.

⁷⁴ Ploos van Amstel, W. (2016) "Gebruik actuele verkeersdata levert mogelijk 1 miljard op." [De Laatste Meter](#),

2.5.4 Geen vergelijkbare uitkomsten

Het vergelijken van de studies is ook lastig doordat de baten in uiteenlopende eenheden worden uitgedrukt. Zo worden in een aantal onderzoeken de baten in percentages van het BNP uitgedrukt.⁷⁵ In sommige onderzoeken zijn de baten uitgedrukt in absolute bedragen⁷⁶ en worden er verschillende valuta gebruikt die onderhevig zijn aan onderlinge valutaschommelingen. Andere eenheden die gebruikt zijn, zijn toegenomen productiviteit,⁷⁷ consumentensurplus,⁷⁸ en anekdotisch kwalitatief bewijs.⁷⁹

Het is opmerkelijk dat de baten van open data ruim worden geschat hoewel deze in de praktijk moeilijk te meten zijn. Het is bijvoorbeeld lastig om onderscheid te maken tussen de baten die toe te rekenen zijn aan alleen open data en baten die het resultaat zijn van andere bedrijfsprocessen. Zo zijn bijvoorbeeld verbeterde metadata, aanvullende documentatie en andere verbeteringen aan de kwaliteit van de dataset niet altijd geheel aan open data toe te schrijven, maar kunnen deze ook het gevolg zijn van goed informatiemanagement in algemene zin.

Overigens konden in geen van de onderzochte cases de incrementele kosten exact worden bepaald aangezien in geen van de cases *ex-ante* een duidelijk budgettering heeft plaatsgevonden, er niet bijgehouden is hoe de kosten *ex-post* precies verliepen. Bovendien is het moeilijk voor de organisaties om de kosten van het uitvoeren van publieke taken te scheiden van de kosten van open data.

2.6 Conclusies

Alle onderzochte studies geven een positief beeld van de kosten-batenverhouding van open data. Hoewel er veel kwalitatieve effecten worden beschreven, blijkt het lastig om de baten goed te kwantificeren. De meeste kwantitatieve studies hebben dan ook een zwakke onderbouwing, waardoor de resultaten lastig op hun merites kunnen worden beoordeeld.

De studies laten zien dat de kosten zeer wisselend zijn, afhankelijk van type data en type aanbieder. Voor data-aanbieders van wie het de kerntaak is om dynamische en hoge volumes aan data te verstrekken, zijn de kosten marginaal en er is vaak al een infrastructuur en beheersysteem aanwezig. Baten zijn vaak wel aantoonbaar op de korte termijn, zoals meer raadplegingen en downloads van de datasets door het bedrijfsleven, kwaliteitslagen op de datasets door terugmeldingen, en efficiëntieslagen binnen de overheid. De studies laten ook zien dat de (meer)kosten van open data moeilijk te onderscheiden zijn van de kosten die een dataverstrekker in ieder geval moet maken om zijn kerntaken uit te voeren. Ten slotte zijn de genoemde kosten en (potentiële) baten in de onderzochte studies vaak gebaseerd op discutabele aannames en extrapolaties.

2.7 Methodologische lessen uit het literatuuronderzoek

We concluderen dat er op dit moment geen objectieve, betrouwbare en bruikbare methodologie wordt gehanteerd die het mogelijk maakt de effecten van open data te meten en te vergelijken. De nadruk ligt op macro-economische schattingen met zwak gefundeerde aannames. Verder worden relevant beoogde baten van verhoogde transparantie/burgerparticipatie en overheidseffectiviteit en

⁷⁵ Onder andere ACIL Tasman 2008, 2009; Omidyar Network 2014; Pham 2011.

⁷⁶ Zie o.a. DECA 2010; Vickery 2011; PwC 2014; Pollock 2011; Oslo Economics 2011.

⁷⁷ Coote 2010.

⁷⁸ Oslo Economics 2011.

⁷⁹ onder andere Nilson 2010; Shakespeare 2013; Newbery *et al.* 2013; Jetzek 2013; Ubaldi 2013; World Bank 2014; SCGOE 2014.

-efficiëntie niet of nauwelijks meegenomen/onderzocht of gemeten. Door deze lacunes blijft het benodigde inzicht in het werkelijke effect van open data onduidelijk, gefragmenteerd, onvolledig, en mogelijk zelfs ongeloofwaardig.

Het is op grond van dit literatuuronderzoek daarom niet mogelijk deze onderzoeksresultaten direct te vertalen naar de maatschappelijke kosten en baten van het als open data beschikbaar stellen van alle high value-datasets van de rijksoverheid. De les van het literatuuronderzoek is juist dat specifieke maatschappelijke kosten en baten van het beschikbaar stellen van een specifieke dataset niet volgt uit deze studies, omdat deze veelal slechts maatschappelijke voordelen duiden van een veel bredere overgang naar open data. Het is gegeven deze context verstandig om specifiek te kijken naar de wijze waarop de dataset beschikbaar wordt gesteld en een concreet overzicht op te stellen van maatschappelijke kosten en baten van alternatieven.

Voor een dergelijke maatschappelijke kosten-batenanalyse kiezen we voor de volgende aanpak:

- Volg de standaard MKBA-methodologie.
Veel van de eerdere studies volgen geen heldere structuur van maatschappelijke kosten en baten. Door aan te sluiten bij de reeds bestaande ervaringen met MKBA kan een kwaliteitsslag worden gemaakt. De methode maakt een helder onderscheid naar directe, indirecte en overige effecten en zorgt ervoor dat op een gestructureerde wijze maatschappelijke kosten en baten in beeld worden gebracht. Door bij elke studie dezelfde methodiek te gebruiken kan ook vooruitgang worden geboekt bij de toepassing ervan, waardoor toekomstige studies kunnen voortbouwen op eerdere studies. Hierbij moet – gezien de ervaringen uit het voorgaande – wel steeds kritisch worden gekeken naar eerder gemaakte veronderstellingen. Bij de open data onderzoeken bleek weinig selectief met dubieuze aannames in volgende rapportages te zijn omgegaan. Vanuit deze methode behoren onder meer de volgende principes
 - Beschrijf zo gedetailleerd mogelijk de definities van termen en concepten (met name de kosten en baten).
 - Definieer zo helder mogelijk de alternatieven met open data (definieer hierbij ook specifiek hoe dit zijn beslag kan krijgen bij de dataverzameling die wordt bestudeerd) en zonder open data.
Op deze wijze kan voorkomen worden dat baten die samenhangen met reguliere ontwikkelingen worden toegerekend aan open data. Hierbij is het ook van belang om niet alleen de meeste open vorm van open data als alternatief mee te nemen. Deze is bij sommige datasets niet de meest realistische vorm. Ook een alternatief waarbij een gegeven dataset een zekere mate van openheid verkrijgt, kan worden meegenomen in de vergelijking. Dit vraagt om een goede omschrijving van het alternatief.
 - Beschrijf zo duidelijk mogelijk de onderzoeksmethodologie en de keuzes in de aanpak.
- Bepalen van de maatschappelijke baten is lastig: leg uit wat eronder verstaan wordt en wat wordt gemeten (en wat niet).
Maak hierbij (op grond van de afbakening van diverse markten) een onderscheid in directe effecten, indirecte effecten en overige effecten. Het verdient aanbeveling om bij directe effecten zoveel mogelijk ‘harde’ gegevens te vinden. Bij indirecte effecten en overige effecten kan het noodzakelijk zijn om een kwalitatieve beschrijving van mogelijke effecten te maken. Geef deze effecten ook apart weer.
- Voor wat betreft de kosten: neem in de eenmalige en structurele kosten ook de kosten gerelateerd aan het veranderen van het beleid richting open data mee.
Dit sluit aan bij wat hierboven over het formuleren van alternatieven is gesteld. Het gaat om het verschil in kosten (en baten) tussen de alternatieven. Kosten (en baten) die in ieder geval

gemaakt moeten worden (wat de keuze ook zal zijn) zijn niet relevant voor deze analyse. Ook wat betreft maatschappelijke kosten dient onderscheid te worden gemaakt tussen directe, indirecte en overige effecten.

- Extrapoler de onderzoeksresultaten niet/bezint eer gij extrapoleert!
Elke dataverzameling is uniek. Maatschappelijke kosten en baten van het open maken van deze verzameling zullen sterk uiteenlopen, mede omdat ook er grote verschillen zijn wat betreft het 0-alternatief in relatie tot openheid en de mogelijkheden om bijvoorbeeld vanuit privacyoverwegingen tot volledige openheid te komen. De belangrijkste bijdrage die de MKBA hier levert is het komen tot een methode waarmee keuzes die gemaakt kunnen worden bij verschillende dataverzamelingen, kunnen worden geanalyseerd. Op die manier ontstaat niet een beeld van de meerwaarde van 'alles open data maken', maar ontstaat inzicht in wat de maatschappelijke kosten en baten zijn van verschillende manieren om een specifieke dataset open te maken. De studies worden op die manier gekoppeld aan concrete beleidsalternatieven.
- Maak een helder onderscheid tussen maatschappelijke kosten en baten en kosten en baten voor specifieke actoren.
Het kan soms verstandig zijn om naast een maatschappelijke kosten-batenanalyse een actorenanalyse te maken waarin wordt weergegeven welke actoren deze kosten en baten dragen. Op het moment dat dezelfde dataset bijvoorbeeld door twee partijen wordt samengesteld en verkocht en een van die partijen verstrekt deze data als open data, dan verliest de andere partij inkomsten. Maatschappelijke kosten en baten van meerdere partijen onderling afwegen vergt een andere insteek dan kosten en baten van één partij afwegen. Bij de besluitvorming over alternatieven is het verstandig inzicht te hebben in de kosten en baten bij alle betrokken actoren. Hierbij is de situatie anders dan bij de meeste infrastructuurprojecten. Bij infrastructuur liggen publieke en private tolwegen zelden naast elkaar en is het zeker niet zo dat de publieke weg extra verkeer kan accommoderen als de tol afgeschaft. Bij dataverzameling kan deze situatie wel voorkomen. In het verleden werd bijvoorbeeld deze discussie gevoerd over het beschikbaar stellen van de wegenlegger van Rijkswaterstaat en de belangen van actoren die vergelijkbare gegevens op de markt brachten (bijvoorbeeld voor navigatieapparatuur) op grond van eigen dataverzameling. Voor een dergelijke actorenanalyse is met zoals voor de inschatting van indirecte effecten en overige effecten het noodzakelijk om de bredere context van de dataverzameling te bezien.

3 Basisregistraties Adressen en Gebouwen

3.1 Probleemanalyse BAG-data

De Basisregistraties Adressen en Gebouwen (BAG) bevatten gegevens van vrijwel alle adressen en gebouwen in Nederland, zoals bouwjaar, oppervlakte, gebruiksdoel (bijvoorbeeld woonfunctie) en locatie van het gebouw op de kaart. De BAG is onderdeel van het stelsel van basisregistraties. Dat houdt in dat organisaties met een publieke taak verplicht zijn de authentieke gegevens uit de BAG te gebruiken; zo gebruikt bijvoorbeeld de Belastingdienst BAG-gegevens voor het toekennen van huurtoeslag.

De gemeenten zijn bronhouders van de BAG en zijn verantwoordelijk voor de kwaliteit van de gegevens. De BAG-gegevens van alle gemeenten komen bij elkaar in de Landelijke Voorziening BAG (LV BAG). Het Kadaster voert het functioneel en technisch beheer van de LV BAG uit en stelt de gegevens via diverse kanalen (zoals de BAG Viewer en PDOK) beschikbaar aan afnemers. Een aantal afnemers levert deze data weer door.

Figuur 3.1: Uitsnede uit de Basisregistratie Adressen en Gebouwen (BAG)

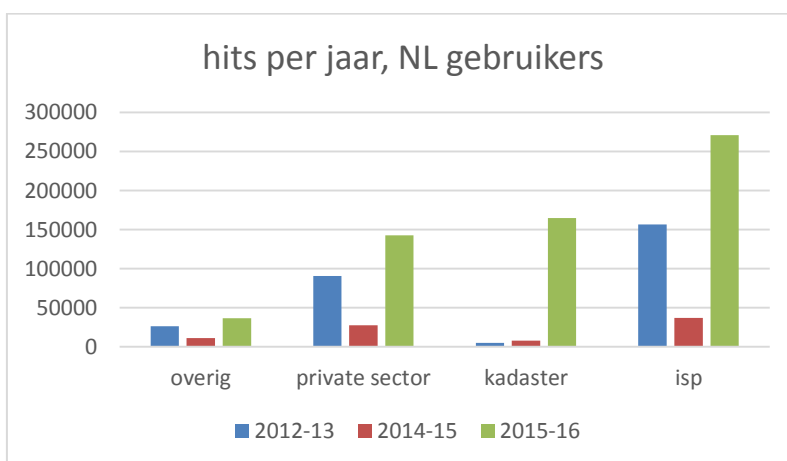
The screenshot displays the BAG Viewer interface. At the top, the Kadaster logo and the title 'Basisregistraties Adressen en Gebouwen (BAG)' are visible. Below the title, there are navigation links: 'Opnieuw beginnen', 'Terugmelding', 'Gebruik viewer', 'Wat is BAG', 'PDF', and 'Help'. The main area is divided into three sections: a search panel on the left, a map in the center, and a data panel on the right. The search panel shows a search for 'enhof den haag' with one result: 'Binnenhof 's-Gravenhage'. The map shows a street layout with a red building highlighted. The data panel on the right provides details for the selected building, including its ID, bouwjaar (1906), status (Pand in gebruik), gebruiksdoel (bijeenkomstfunctie), oppervlakte (3236 m2), and status (Verblijfsobject in gebruik). It also lists the numeraanduiding (ID: 051820000039208E, Postcode: 2513AG, Huisnummer: 20, Huisletter: toev.) and the openbare ruimte (ID: 051830000020059J, Naam: Buitenhof, Status: Naamgeving uitgegeven). The woonplaats is listed as 's-Gravenhage'.

De eerste vraag is nu wat de toegevoegde waarde is van de BAG als open data, ten opzichte van de situatie voordat de BAG landsdekkend en zonder beperkingen beschikbaar was als open data. De tweede vraag is hoe de kosten van het beschikbaar stellen van de BAG als linked open data zich verhouden tot de maatschappelijke baten hiervan.

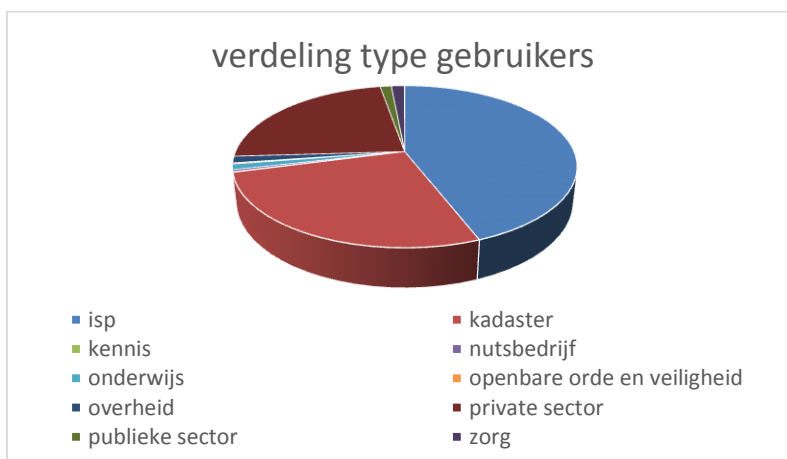
3.2 Huidige situatie: 0-alternatief

Sinds maart 2011 is de BAG landsdekkend beschikbaar. Sinds juni 2012 wordt de BAG beschikbaar gesteld aan iedereen, zonder voorwaarden aan hergebruik.⁸⁰ Voor de overheid zijn BAG-gegevens gratis beschikbaar, voor niet-overheden worden verstrekingskosten in rekening gebracht. Het eenmalig downloaden van de gehele BAG kost €168. Wel kan iedereen BAG-gegevens online gratis raadplegen via de BAG Viewer en via PDOK.⁸¹ Daarmee voldoet de BAG aan 7 van de 8 uitgangspunten van de NODA, en daarom beschouwen we de BAG als semi-open data. Figuren 3.2 en 3.3 laten de toename van gebruik sinds 2012 en het type gebruiker zien, waarbij isp (internet serviceproviders, zoals Ziggo) een indicatie is van het aantal burgers. Figuur 3.4 laat zien dat de meeste hits voornamelijk uit Nederland komen.

Figuur 3.2: Hits per jaar per type gebruikers



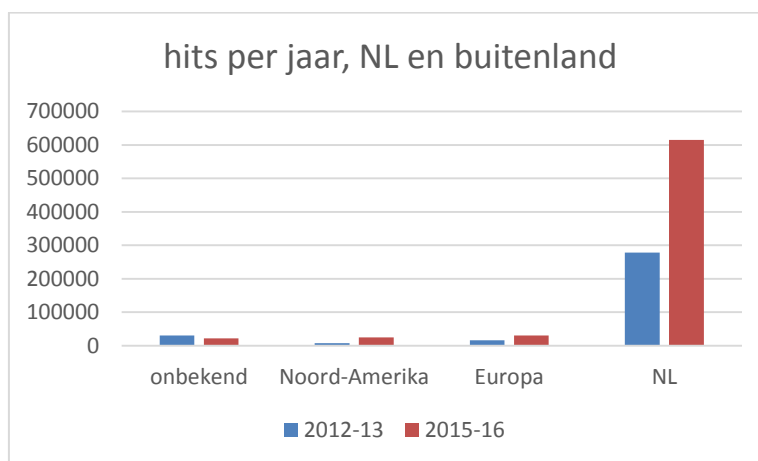
Figuur 3.3: Verdeling van type gebruikers



⁸⁰ De BAG is beschikbaar onder een Creative Commons ZERO verklaring.

⁸¹ Zie <https://bagviewer.kadaster.nl/>

Figuur 3.4: Aantal hits per jaar per regio



3.3 Alternatieven voor BAG semi-open data

De waarde van de BAG als open data zal op twee manieren worden bepaald voor de vier alternatieven:

- 0-alternatief: huidige situatie, publieke BAG-data zijn open voor overheden en tegen verstrekingskosten voor ieder ander (semi-open data);
- Min 1 alternatief: de BAG is voor niemand als open data beschikbaar, maar alleen tegen betaling beschikbaar in een machine-leesbare vorm;
- Plus 1 alternatief: de BAG is open data en gratis voor iedereen (3 sterren op de TBL-index);
- Plus 2 alternatief: de BAG is voor iedereen beschikbaar als 'linked open data' (4 tot 5 sterren op de TBL-index).

Het Min 1 alternatief is niet heel realistisch – er zijn geen signalen dat de huidige situatie zal worden teruggedraaid – maar biedt wel een mogelijkheid om kosten en baten van open data te beredeneren. Veel onderzoek is gedaan vanuit de context dat de data niet open beschikbaar zijn en de vraag is dan wat maatschappelijke kosten en baten zijn van het als open data beschikbaar stellen. Deze analyse biedt de spiegel: de data zijn open beschikbaar; wat als dit niet zo zou zijn?

Het Plus 1 alternatief houdt in dat de huidige financiële drempel voor hergebruik buiten de overheid zal worden verwijderd.

Het Plus 2 alternatief biedt de mogelijkheid om door middel van het linken van data dat iedereen dezelfde betekenis aan de informatie ontleent. Het Kadaster heeft aangekondigd dat begin 2017 een proefversie van de BAG beschikbaar zal komen als linked data, en de dagelijkse geactualiseerde linked data vanaf medio 2017.⁸²

⁸² Zie <https://www.kadaster.nl/-/vooraankondiging-bag-als-linked-data-en-adressen-api>

3.4 Kwalitatieve omschrijving van investeringskosten en effecten

In algemene zin onderscheiden wij effecten met betrekking tot de kosten van de alternatieven en effecten met betrekking tot de baten ten opzichte van het 0-alternatief. Belangrijke verschillen tussen de alternatieven zitten in de volgende aspecten.

Investeringskosten

Kosten om data tot linked data te maken: Het gaat hierbij om de kosten die gepaard gaan met het linken van de BAG aan andere data. Deze kosten zijn afhankelijk van de keuze van linken op 4-sterrenniveau of 5-sterrenniveau. Bij deze kosten moeten ook de beheerkosten worden meegenomen: bij elke update (en dat is bij de BAG frequent) moeten de links gecontroleerd worden. Er zullen ook meer vragen bij de data-aanbieder binnenkomen.

Directe effecten

Serverkosten: Meer gebruik (door open data) kan een hogere servercapaciteit vereisen.

Administratieve lasten: Data beschikbaar stellen tegen een vergoeding brengt extra transactiekosten voor de aanbieder van de data met zich mee (administratiekosten voor de onderhandeling van een contract, verwerken van betalingen, contractmanagement, handhaving etc.) Bij de Plus 1 en Plus 2 alternatieven vallen deze transactiekosten voor de aanbieder weg. Daar staat het verlies van de marginale inkomsten uit dataverstrekking tegenover.

Met open data gaan de transactiekosten van gebruikers omlaag. De data zijn gratis (geen transactiekosten verbonden aan betalingen) en gebruiksvoorwaarden zijn niet beperkend.

Dubbel verzamelen: open beschikbaarheid van data kan leiden tot besparingen daar waar data dubbel worden verzameld of waar meerdere licenties voor dezelfde dataset binnen één organisatie worden aangeschaft. Indien open data beschikbaar zijn, zullen kosten voor het beschikbaar maken en aanbieden van alternatieve datasets afnemen.⁸³

Consumentensurplus: Gratis beschikbaar stellen van data kan leiden tot een consumentensurplus. Dit geldt met name waar het extra gebruik is. Immers het consumentensurplus van bestaande data kan worden weggestreept tegen verlies van de producenten van de data die deze nu gratis aanbieden.

Indirecte effecten

Toegenomen transparantie: Dankzij open data wordt de transparantie van de overheid richting de burger verhoogd. In de huidige context neemt de waarde van transparantie toe.

Verbeterde datakwaliteit: Door open data wordt de kwaliteit van de data verhoogd omdat er meer terugmeldingen van fouten, e.d. komen.

Hergebruik van basisregistratie data: ('Linked') open data versterken de bekendheid van authentieke data uit de basisregistratie. Als authentieke data hergebruikt worden voor andere toepassingen, bijvoorbeeld de BAG als onderlaag in de WOZ-Waardeloket, worden die toepassingen betrouwbaarder.

Nieuwe producten en diensten overheid: Het beschikbaar stellen van open data kan leiden tot meer verzoeken voor maatwerkproducten, adviesdiensten, etc.⁸⁴

⁸³ Zie bijvoorbeeld het opzetten van een parallelle infrastructuur van private weerstations in Duitsland nadat onderhandelingen tussen een private partij en de Duitse Nationale Meteorologische Dienst op niets uitliepen (Fornefeld, M. 2009. The value to Industry of PSI: the business sector perspective The Socioeconomic Effects of Public Sector Information on Digital Networks: Towards a Better Understanding of Different Access and Reuse Policies. Workshop Summary, P. Uhler, Ed., pp. 10-16. Washington: The National Academies Press).

⁸⁴ Zie onder meer Welle Donker, F. & B. van Loenen (2015). Duurzame businessmodellen voor open data. Delft, Kenniscentrum Open Data: 60. <http://repository.tudelft.nl/view/ir/uuid%3Ac5b18289-81d9-4692-ab2e-62b3a96c02d2/>

Nieuwe producten en diensten bedrijfsleven: Open data kunnen een bijdrage leveren aan het realiseren van deze ambitie. De gedachte is dat, omdat de data open beschikbaar zijn, kansen worden gecreëerd voor het bedrijfsleven. Naast de vorm (open data) is het natuurlijk ook van belang dat er innovatieve inhoud beschikbaar is die mogelijkheden biedt tot het ontwikkelen van nieuwe producten en diensten.

Overige effecten

Vestigingsklimaat: Hergebruik van basisregistratiegegevens voor intelligente ('smart') toepassingen kan het vestigingsklimaat van Nederland inzichtelijker en aantrekkelijker maken, en daarmee economische groei genereren.

Fraudebestrijding: Betere informatie over adressen en gebouwen kan ertoe leiden dat fraude makkelijker te detecteren is, bijvoorbeeld door koppelingen met het Handelsregister. En daarmee kunnen kosten worden vermeden. Het als open data beschikbaar stellen is vooral van belang voor fraudedetectie door partijen buiten de overheid (er zit geen verschil in de alternatieven tussen beschikbaarheid binnen de overheid). Het 'linked' maken van de data kan zorgen voor betere detectie door alle partijen.

Privacy: Het linken van allerlei data aan gebouwen en adressen (en deze zonder restricties ter beschikking stellen) kan leiden tot op natuurlijke personen herleidbare gegevens. Dit kan negatieve effecten hebben op de privacy van individuen.

Tabel 3.1: Kwalitatief overzicht van verschillen effecten (ordinaal)

BAG	0	-1 (betaalversie voor allen)	+1 (open voor iedereen)	+2 (Linked Open Data)
Investeringskosten				
-Kosten om data 'linked' te maken	0	0	0	-
Directe effecten				
-Serverkosten	0	0	-	--
-Administratieve lasten	0	-	+	+
-Transactiekosten gebruikers	0	-	+	+
-Dubbel verzamelen		-	+	++
-Consumentensurplus	0	-	+	++
Indirecte effecten				
-Toegenomen transparantie: beter geïnformeerde burger/meer participatie	0	-	+	++
-Toegenomen transparantie: betere datakwaliteit	0	-	+	++
-Toegenomen betrouwbaarheid van afgeleide diensten	0	-	+	++
-Nieuwe producten/diensten aanbieder	0	+	+	+
-Nieuwe producten/diensten bedrijfsleven		-	+	+
Overige effecten				
-Vestigingsklimaat		-	+	+
-Fraudebestrijding/vermeden kosten		-	+	+
-Privacy			-	--
Totaal		-	+	+pm

+ = meer baten

- = meer lasten

pm = pro memorie (effecten kunnen niet worden ingeschat, maar zijn wel waarschijnlijk)

Tabel 3.1 geeft een overzicht van de verschillen tussen de alternatieven. Dit is per effect ordinaal vastgesteld. Het 0-alternatief is het ijkpunt van de verschillen. De serverkosten in alternatief -1 zijn bijvoorbeeld lager dan in het 0-alternatief. In alternatief 1 zijn deze hoger en in alternatief 2 zijn deze hoger dan in alternatief 1. Dit zegt niets over de omvang van deze kosten, noch over de kwantitatieve verschillen tussen de alternatieven.

Per alternatief gaan wij hieronder nader in op de te verwachte effecten.

-1 alternatief: geen open data meer, alleen tegen betaling beschikbaar

In het Min 1 alternatief wordt de stap naar open data teruggedraaid maar de BAG blijft wel bestaan.

- *Investeringskosten.* Geen.
- *Directe effecten:* Door het opwerpen van een 'betaalmuur' zullen de data minder gebruikt gaan worden. Dit leidt tot besparingen op de server die deze data moet verwerken. Het leidt tot stijging van administratieve lasten aan zowel de aanbiderszijde als de gebruikerszijde. Deze kosten bestaan onder andere uit handelingen om betalingen te verwerken en andere activiteiten gerelateerd aan de verkoop van de data. Er is wel een directe koppeling tussen aanbieder en gebruikers en daarmee kunnen de behoeften van gebruikers makkelijker geadresseerd worden. Het consumentensurplus zal lager zijn.
- *Indirecte effecten:* De transparantie naar de burger is zeer beperkt en daarmee het draagvlak voor beslissingen die door de overheid worden genomen. De datakwaliteit zal minder zijn doordat er minder gelegenheid is om de data in te zien en fouten te kunnen terugmelden. Er zullen minder nieuwe producten en diensten door het bedrijfsleven worden gecreëerd. Dat schept ruimte voor de data-aanbieder om dat zelf te doen op basis van eigen data.
- *Overige effecten:* Deze zullen zonder open data niet plaats kunnen vinden.

+1 alternatief: alle data gratis beschikbaar voor iedereen

- *Investeringskosten.* Marginaal. De bestaande betaaldienst moet worden opgeheven en de (personele) middelen van die betaaldienst zullen waarschijnlijk in andere diensten worden ingezet.
- *Directe effecten.* De marginale verstrekingskosten voor niet-overheden vallen weg, maar de gerelateerde administratielasten vallen ook weg. Door de laatste financiële drempel weg te halen, zal het dataverkeer op de server toenemen. Bovendien zal door het extra gebruik van de data het consumentensurplus stijgen.
- *Indirecte effecten.* Het bedrijfsleven zal de data meer gaan gebruiken om de eigen bedrijfsvoering efficiënter te maken en om nieuwe producten en diensten te ontwikkelen.
- *Overige effecten.* De verdere integratie van BAG-data leidt tot verbetering van het vestigingsklimaat. Koppelingen met andere bestanden kunnen leiden tot fraudebestrijding; koppeling met andere bestanden kan negatieve invloed hebben op de privacy van individuen.

+2 alternatief: alle open data 'linked' maken

- *Investeringskosten* Het voorbereiden en de beheren van het 'linked' maken van de data brengt extra kosten met zich mee.
- *Directe effecten.* Voor het eventueel linken tot 5-sterrenniveau houdt in dat er veel meer opslagcapaciteit nodig zal zijn. Deels zal deze capaciteit buiten de BAG gevonden kunnen worden, maar dit leidt wel tot extra kosten. De datakwaliteit zal (beperkt) hoger worden door alle extra kwaliteitscontroles die uitgevoerd moeten worden om data geschikt te maken als linked data. Doordat de data beter vindbaar zullen zijn, zullen deze vaker worden opgevraagd. Het toegenomen gebruik leidt tot hogere serverkosten. Wat betreft administratieve lasten om betalingen te verwerken is er geen verschil met het +1 alternatief: open blijft open. Er zullen wel

meer vragen komen, waardoor de kosten van een eventuele helpdesk omhoog gaan. Bovendien zal door het extra gebruik van de data het consumentensurplus stijgen.

- *Indirecte effecten.* Door het 'linked' maken kunnen nog betere relaties gelegd worden met andere bronnen en zullen gebruikers beter in staat zijn om hun taken uit te voeren ook buiten het domein van de basisregistraties. Er is een positief effect op de betrouwbaarheid van de data omdat de datakwaliteit hoger zal zijn met de extra kwaliteitscontroles. Door het 'linked' maken van de data ontstaan verdere perspectieven voor de ontwikkeling van nieuwe producten en bedrijven.
- *Overige effecten.* Door verdere verbeteringen en het maken van koppelingen zullen de overige effecten versterkt worden ten opzichte van alternatief plus 1.

Nieuwe producten en diensten: De afgelopen vijf jaar zijn het gebruik en de gevarieerdheid van het gebruik van de BAG sterk toegenomen. Op basis van de BAG zijn er diverse nieuwe toepassingen ontwikkeld. Zo zijn er toepassingen op het gebied van analyse, beleid en trends, van veiligheid en zorg, op het gebied van adresverrijking en -validatie, van route, navigatie en kaarten, van technische koppelingen en toepassingen met betrekking tot duurzaamheid en milieu.⁸⁵

Door de BAG volledig open te stellen voor iedereen zullen de drempels voor hergebruikers minimaal zijn. Er zullen naar verwachting meer nieuwe producten en diensten worden ontwikkeld. Een deel van de diensten zal bestaan uit het al dan niet verrijkt doorleveren van de open data. Bedrijven als ESRI converteren de BAG-data zelf naar hun eigen bestandsformaten en leveren de aangepaste BAG direct aan hun klanten. Bedrijven zoals Geon en Geomij hebben BAG converteersoftware ontwikkeld die meerdere conversiemogelijkheden biedt. Uit onze enquête onder BAG-gebruikers blijkt dat zonder de open BAG-data het ontwikkelen van nieuwe toepassingen niet mogelijk zou zijn geweest. De BAG wordt gezien als een enabler voor betere dienstverlening, zowel binnen als buiten de overheid.

Gebruiker DAT.Mobility

De BAG gebruiken wij onder andere bij de EU-geluidskartering 2012 waar we op een kaart de geluidsbelasting tonen, op iedere plek en op ieder pand. Deze geluidsbelastingskaarten maken we voor een flink aantal provincies en steden, zoals de provincies Noord-Holland en Limburg en de steden Enschede, Hengelo, Zwolle, Heerlen en Nijmegen.

Daarnaast worden gegevens over gebouwen gebruikt om voor het nationale programma NSL Monitoring Luchtkwaliteit de honderdduizenden toetslocaties in Nederland te bepalen. Zo komen de toetslocaties op de plekken waar de meeste relevante informatie wordt verzameld. Ten slotte maakt de BAG het mogelijk dat we veel eenvoudiger [regionale verkeersmilieukaarten](#) kunnen produceren.

Bron: <http://www.goudappel.nl/actueel/2011/08/30/basisregistratie-adressen-en-gebouwen-vandaag-geopend/>

⁸⁵ Zie voor een overzicht: <http://www.kadaster.nl/bag-toepassingen>.

Het Asbestdossier

Alle huizen gebouwd voor 1994 worden geacht asbest te bevatten, tenzij kan worden aangetoond dat dit niet zo is. Het Ministerie van IenM houdt toezicht op sanering en verwijdering van asbest. Daartoe zijn gegevens nodig over bouwjaren en sloopvergunningen van panden. Dit is makkelijk terug te vinden in de landelijke voorziening van de BAG.

Bron:

<http://www.kadaster.nl/documents/20838/87954/BAG+toepassing+Asbestdossier/5f4cfa19-2cd6-4a59-9711-fd9bb6ca8cac>

3.5 Monetarisering

In deze paragraaf gaan we in op de monetarisering van de effecten van de verschillende alternatieven.

Investeringskosten

- *Kosten om data 'linked' te maken:* De ervaringen in de pilot 'linked data' geven inzicht in de extra kosten van het 'linked' maken van de BAG. Binnen dit alternatief wordt uitgegaan van een realistische kostenpost van €120.000 per jaar, oftewel €2,2 miljoen in contante waarde.

Directe effecten

- *Serverkosten (pm, de serverkosten van de BAG zijn niet te isoleren van de totale serverkosten).*
- *Administratieve lasten.* Bij de afname van administratieve lasten gaan we uit van ongeveer 200 contracten die met private partijen zijn afgesloten (voor BAG Extract en BAG Compact).⁸⁶ Tevens is het nu noodzakelijk om specifieke (gratis) toegang te faciliteren voor overheidsafnemers. Verwacht wordt dat dit een besparing oplevert van €70.000 per jaar (ofwel een contante waarde van €1,3 miljoen).
- *Transactiekosten gebruikers.* Verondersteld mag worden dat gebruikers dezelfde kosten hebben als de aanbieder: €70.000 per jaar (ofwel een contante waarde van €1,3 miljoen).
- *Consumentensurplus.* In de pilot 'Open data en BAG' stelde het Kadaster gebruikers een maand lang in de gelegenheid om de BAG gratis te downloaden. In ruil moesten de gebruikers gebruikersinformatie geven en input leveren op gebruikerswensen. Van deze mogelijkheid hebben 163 gebruikers gebruik gemaakt.⁸⁷ Op grond van de 'rule of half' zou dit een consumentensurplus van ca €14.000 (=163/2*€168), ofwel een contante waarde van €250.000 vertegenwoordigen. Dit is geen heel precieze schatting, er zijn op dit moment geen exacte gegevens beschikbaar. Bij 'linked data' zal deze hoger zijn.

Indirecte effecten

Toegenomen transparantie: de transparantie hebben wij niet kunnen kwantificeren.

Verbeterde datakwaliteit: Een hogere datakwaliteit heeft een doorwerking op alle toepassingen gebaseerd op BAG-data, maar hebben wij ook niet kunnen kwantificeren.

Hergebruik van authentieke data: Hetzelfde geldt voor bekendheid en betrouwbaarheid van data.

⁸⁶ Zie Kwartaalrapportage BAG 2015 1^e -4^e kwartaal.

⁸⁷ Zie Kwartaalrapportage BAG 2015 1^e -4^e kwartaal. "Bij de pilot zijn ook veel leveringen gedaan aan private partijen en particuliere mailadressen".

Nieuwe producten en diensten overheid: Open data kan leiden tot meer verzoeken voor maatwerkproducten, adviesdiensten, etc. Dergelijke extra baten zijn echter moeilijk volledig toe te schrijven aan open data.

Nieuwe producten en diensten bedrijfsleven: Hoewel er voldoende bewijs lijkt te zijn dat open data een positief effect hebben op de bestaande bedrijfsvoering en de ontwikkeling van nieuwe producten en diensten, zijn bedrijven terughoudend in het kwantificeren van de baten.

Overige effecten

Vestigingsklimaat: Hiervoor kunnen alleen kwalitatieve ramingen worden gemaakt.

Fraudebestrijding: Ook hiervoor kunnen alleen kwalitatieve ramingen gebaseerd op voorbeelden worden gemaakt.

Privacy: Het 'linked' maken van de data kan extra kosten met zich meebrengen om te voldoen aan de privacywetgeving en deze te handhaven. Deze extra kosten zijn lastig in te schatten omdat er nu nog weinig ervaring is met data 'linked' maken. Deze kosten zullen dalen na het doorlopen van dit leerproces.

Tabel 3.2: Gemonetariseerd overzicht van verschillen effecten, in miljoenen euro's

BAG	0	-1 (betaalversie voor allen)	+1 (open voor iedereen)	+2 (Linked Open Data)
Kosten				
-Kosten om data 'linked' te maken	0			-2,2
Directe effecten				
-Serverkosten	0			
-Administratieve lasten			1,3	1,3
-Lagere transactiekosten gebruikers	0		1,3	1,3
-Dubbel verzamelen	0	-pm	+pm	+pm
-Consumentensurplus*	0		0,25	0,25 + pm
Indirecte effecten				
-Toegenomen transparantie: betere datakwaliteit	0			
-Nieuwe producten en diensten	0			
Overige effecten				
-Verbetering reputatie aanbieder	0			
-Vermindering corruptie/fraude	0			
Privacy				
Totaal			2,85 + pm	0,65 + pm

* Zie indirecte kosten.

3.6 Conclusies

Het beschikbaar stellen van de BAG heeft positieve effecten gehad: het heeft een efficiëntere en effectievere overheid opgeleverd. Er worden minder data dubbel ingewonnen en er wordt meer gebruik gemaakt van authentieke gegevens. Zo gebruikt de Landelijke Voorziening WOZ de BAG voor het recent geopende WOZ-Waardeloket. De maatschappelijke baten zijn ook groter dan voor 2012: bedrijven pakken de BAG steeds meer op als een bron voor het verbeteren van hun dienstverlening en voor het creëren van nieuwe producten en diensten. Hergebruik van authentieke gegevens

versterkt het imago van de overheid als aanbieder van transparante en betrouwbare data. De pilot met het verwijderen van de laatste financiële drempel toonde aan dat dat een positief effect heeft op bedrijvigheid, vooral bij kleine bedrijfjes en in de vorm van startups. Het handhaven van die financiële drempel lijkt ook geen meerwaarde te bieden ten opzichte van geheel open data. Open data aanbieden is slechts een kleine stap binnen het gehele programma van het Kadaster om te komen tot betere gegevens over ruimtelijke informatie. Het effect van open data is daardoor lastig af te zonderen van de totale innovatie. Zowel qua kosten als opbrengsten is het effect van open data bescheiden ten opzichte van het totale overzicht.

De stap naar bijvoorbeeld mogelijk 'linked data' brengt extra kosten met zich mee. Maar gezien de potentiële baten voor de overheid (betere uitvoering van een complexe Omgevingswet), zullen de uiteindelijke baten hoger zijn dan de kosten. Als gevolg zullen de maatschappelijke baten ook hoger zijn: beter geïnformeerde burgers, betere vindbaarheid van de data en betere mogelijkheden om koppelingen te maken voor nieuwe toepassingen.

- Gerealiseerde of geschatte reistijd
- Voertuigcategorie (afgeleid van de lengte van het passerend voertuig)
- Vanaf eind 2016 Floating Car Data (FCD) (rijrichting, -afstand en snelheden ingewonnen via informatie uitgezonden vanuit het voertuig, bijv. mobiele telefoons of navigatiesystemen).

NDW verzamelt verder de volgende gegevens over de status van de weg:

- Wegwerkzaamheden en tijdelijke verkeersmaatregelen bij evenementen
- Filemeldingen en meldingen van ongevallen en incidenten
- Veiligheidgerelateerde meldingen
- Status (open of dicht) van bruggen
- Status (open of dicht) van spitsstroken, plusstroken en rijstroken
- Vanaf eind 2016
 - Beeldstanden van signaalgevers boven de snelweg
 - Beeldstanden van Dynamisch Route-informatiepanelen (DRIP-panelen)
- Gepland
 - Status gegevens van verkeerslichten.

NDW biedt sinds 2013 actuele verkeersinformatie en sinds 2015 ook historische gegevens aan als open data. Dit geldt alleen voor data die door NDW en haar partners zelf wordt ingewonnen. De data die via marktpartijen wordt gepubliceerd is dit beleid niet van toepassing. Cruciaal aan het Europese kader (zoals de ITS-richtlijn 2010/40/EU)¹ is dat open data geen doel is maar een middel om te komen tot een verbeterde informatie-infrastructuur op het gebied van verkeer en vervoer. Intelligente vervoerssystemen vragen data en daarvoor dienen deze data te worden uitgewisseld en geharmoniseerd. De ITS-richtlijn schrijft voor dat er nationale toegangspunten komen voor gestandaardiseerde uitwisseling van data. Daarbij wordt onderscheid gemaakt in verplichtingen voor aanbieders van publieke en private data en verplichtingen die relateren aan veiligheidgerelateerde data en overige data. Dit leidt tot verschillende condities als het gaat om toepassen van leveringsvoorwaarden. DATEX 2 is de standaard voorgeschreven door de ITS-richtlijn.

4.2 Huidige situatie NDW-dataverstrekking

Van de actuele verkeersgegevens is een historische databank opgebouwd, die gebruikt kan worden voor verkeerskundige analyses. Vooralsnog worden alle NDW-gegevens (behalve de historie van statusgegevens) als open data beschikbaar gesteld. Dit geldt niet voor private data, zoals Floating Car Data, indien er auteursrecht rust op deze data. Bij gebruik van NDW real-time Open Data Service geeft NDW geen garanties op databeschikbaarheid (er is wel een leveringsplicht!), actualiteit en continuïteit van de datafeed, maar in de praktijk loopt deze real-time levering gelijk met de datalevering uit de productie-omgeving voor verkeersmanagement (VM) en verkeersinformatie (VI), die beschikbaar is voor partners en afnemers met een Data+Diensten-overeenkomst. Dit maakt het mogelijk dat afnemers relevante en adequate real-time data voor hun VI services kunnen gebruiken. Het gebruik van deze gegevens is echter voor eigen risico. NDW kan de service beëindigen en zal dat, behoudens onvoorziene omstandigheden, ruim van tevoren aankondigen. Voor afnemers van open data zijn geen servicedesk of andere vormen van dienstverlening beschikbaar. Deze zijn alleen beschikbaar indien de afnemer een Data+Diensten-overeenkomst met NDW sluit. Overigens kunnen afnemers van real-time open data zelf ook historische bestanden aanmaken.

Zoals hierboven al uit blijkt, biedt NDW ook datalevering uit de zogenaamde productie omgeving. Data via deze levering is alleen verkrijgbaar wanneer de afnemer een Data+Diensten-overeenkomst met NDW sluit. Hierbij wordt niet langer gewerkt met (betaalde) licentieovereenkomsten, maar met

uitruil van data en/of diensten met gesloten beurzen.⁸⁹ Een van die diensten kan het uitwisselen van gegevens zijn die door een van de private partijen zijn verzameld, zoals Floating Car Data.

Alternatieven NDW open data

Omwille van eenvoudigheid, toegankelijkheid en kwaliteit zijn bepaalde zaken gelijk in alle typen datalevering. Het structureren van de data volgens de standaard DATEX 2⁹⁰ en het via een gemeenschappelijk portal beschikbaar stellen van de data zijn daar een voorbeeld van. Deze kosten zullen in alle alternatieven moeten worden gemaakt. Ook wordt ervan uitgegaan dat in alle gevallen metadata moeten worden verstrekt.

4.3 Projectalternatieven voor NDW open data

Gegeven de beleidscontext zullen de volgende alternatieven met elkaar worden vergeleken:

- 0-alternatief: publieke NDW-data zijn open data. Dit alternatief scoort 3 tot 4 sterren (er zijn verschillen tussen de data die worden aangeboden) op de TBL (Tim Berners-Lee)-index.
- Min 1 alternatief: publieke NDW-data zijn niet open, maar zijn tegen betaling beschikbaar in een machine-leesbare vorm. Dit alternatief kan worden bereikt als NDW de data die nu open zijn gaat beprijzen.
- Plus 1 alternatief: 'linked data' volgens 4 tot 5 sterren op TBL-index.
- Plus 2 alternatief: 'linked data', waarbij ook private data open is en gelinkt.

In het 0-alternatief zijn de publieke NDW-data open data. Het is het huidige beleid en het is het waarschijnlijk dat de verkeersdata als open data beschikbaar blijven. Dat naast de huidige open data eventueel andere data (zoals de private Floating Car Data) niet als open data beschikbaar komen, heeft geen betrekking op het huidige data-aanbod en past binnen dit alternatief.

Hierbij is het Min 1 alternatief eveneens niet realistisch – er zijn geen signalen dat open data worden teruggeschroefd (anders dat extra data die thans niet als open data beschikbaar zijn onder andere condities ter beschikking kunnen worden gesteld) – maar biedt wel een mogelijkheid om kosten en baten van open data te beredeneren. Veel onderzoek is gedaan vanuit de context dat de data niet open zijn en de vraag is dan wat maatschappelijke kosten en baten zijn van het open aanbieden van deze data. Deze analyse biedt de spiegel: de data zijn open; wat als dit niet zo zou zijn? In dit alternatief is er wel sprake van het aanbieden van de data in DATEX 2 vorm via een centraal portal conform de huidige werkwijze in Nederland.

In het Plus 1 alternatief is de technische stap meegenomen naar het aanbieden van 'linked data'. Dit vraagt om een extra investering die het mogelijk maakt externe betekenis toe te voegen aan de data. In dit alternatief is gekozen voor de (relatief) simpele versie van 'linked data'.⁹¹ Bij deze versie is geen extra semantiek nodig. Het is mogelijk 'linked data' met extra handwerk verder door te ontwikkelen,

⁸⁹ Momenteel voert Ecorys een evaluatie uit van Data+Diensten bij NDW; wij danken Ecorys voor het delen van hun ervaringen.

⁹⁰ Zie ook 'Overzicht resultaten op het gebied van Smart Mobility in Nederland' Bijlage 2015D47518, bij de brief van de Minister van Infrastructuur en Milieu aan de Tweede Kamer, Kamerstuk 31305, nr. 214, 3 december 2015.

⁹¹ In deze versie worden de data omgezet naar RDF (Resource Description Framework) en met URI's (Uniform Resource Identifiers) volgens standaarden van het World Wide Web Consortium (W3C) beschikbaar gesteld. Zie hiervoor ook Bizer, C., R. Cyganiak & T. Heath, How to Publish Linked Data on the Web (<http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>, bezocht: 29/11/2016).

waarbij bij elke datahouder extra stappen dienen te worden genomen. Hiervoor is niet gekozen in dit alternatief, omdat dit minder aansluit bij wat een relevant alternatief kan zijn in deze context.

Het realiseren van het Plus 2 alternatief is de ambitie van de Rijksoverheid.⁹² De cruciale stap die in dit beleid op dit gebied wordt geformuleerd is de maximale openheid en beschikbaarheid van data (publiek en privaat). Dus naast de openheid van data van de publieke partij via NDW betreft het hier ook de beschikbaarheid van private data (bijvoorbeeld van de Data+Diensten-afnemers). Voor het Plus 2 alternatief geldt dat de gedachte dat alle data gratis als open data beschikbaar zou zijn, niet realistisch is. Dit alternatief zal daarom worden geoperationaliseerd als de stap die voor 2020 realistisch kan worden bereikt zonder gedwongen onteigening van data (waar wetwijziging voor nodig is) en is gebaseerd op het beschikbaar komen van de data door samenwerking en niet door afkoop van de eigendomsrechten op data. Deze stap sluit aan bij het Data+Diensten-concept van NDW.

4.4 Kwalitatieve omschrijving van investeringskosten en effecten

Belangrijke verschillen tussen de alternatieven zitten in de volgende aspecten:

Investeringskosten

Kosten om data 'linked' te maken: Het gaat hierbij om kosten die verder gaan dan het structureren van de data volgens DATEX 2.

Directe effecten

Serverkosten: Meer gebruik (door open data) kan ervoor zorgen dat een hogere capaciteit aan computerservers nodig is. Bij verkeersgegevens gaat het om real-time data en niet om historische bestanden; dit vraagt om systemen die deze vraag aankunnen.

Administratieve lasten: Minder financiële transacties leiden tot besparingen op administratieve lasten bij zowel aanbieder als gebruikers. Het verwerken van betalingen en de handhaving van betalingen en contracten kost geld. Daar waar het om real-time data gaat moeten deze gegevens worden afgeschermd voor niet betalende potentiële gebruikers en dienen de data wel toegankelijk te worden gemaakt voor betalende gebruikers. Er dient een administratie te worden gevoerd, waarin koppelingen tussen betalingstoestand en toegang wordt geregeld.

Dubbel verzamelen: Open beschikbaarheid van data kan leiden tot besparingen daar waar data dubbel worden verzameld. Indien open data beschikbaar zijn, zullen kosten voor het beschikbaar maken en aanbieden van alternatieve verkeersmanagementinformatie (wegvallen van aanschaf van wegwantsystemen, beheer, onderhoud en vervanging) afnemen. Door het volgens de ITS-richtlijn op een centraal punt beschikbaar stellen van de data wordt op dit gebied al een belangrijke winst geboekt. De data zijn immers reeds kenbaar, vindbaar en beschikbaar. Toch zijn er verschillen tussen de alternatieven wat betreft betaalbaarheid en bruikbaarheid ('linked data' zijn beter hanteerbaar).

Consumentensurplus: Het gratis beschikbaar stellen van data kan leiden tot een consumentensurplus. Dit geldt met name waar het extra gebruik betreft. Het consumentensurplus van bestaande data kan immers worden weggestreept tegen het verlies van de producenten die de data nu gratis aanbieden.

Indirecte effecten

Minder voertuigverliesuren: Open data kunnen leiden tot minder voertuigverliesuren, omdat informatie beschikbaar komt over de actuele verkeerssituatie en daar sneller op geanticipeerd kan

⁹² Beter geïnformeerd op weg: Routekaart 2013-2023: Hoofddocument (2013), p.55.

worden (bijvoorbeeld het beter verdelen van voertuigen over het wegennet, het beter gebruiken van de beschikbare capaciteit en daarmee het voorkomen van oplopende reistijden en kosten).

Omrijkosten: Wanneer op grond van open data een ander alternatief wordt gekozen, kan dit leiden tot omrijkosten (er wordt een alternatief gekozen dat minder gunstig is dan het optimale alternatief dat bijvoorbeeld zonder congestie zou zijn gekozen).

Betrouwbaarheid: Minder voertuigverliesuren en tijdige informatie over de verkeerssituatie leiden tot een hogere betrouwbaarheid.

Brandstofgebruik: Een en ander kan invloed hebben op brandstofgebruik.

Nieuwe producten en bedrijven: Nederland als voorloper op het gebied van ITS kan zorgen voor startups en nieuwe bedrijven op het gebied van ITS. Dat kan leiden tot economische ontwikkeling en export van kennis naar landen die minder ver zijn op het gebied van ITS. De Minister van Infrastructuur en Milieu heeft dit als volgt omschreven in een brief aan de Kamer over smart mobility:

“Het creëren van het juiste ondernemersklimaat en het vooroplopen in internationale initiatieven vergroot de kans dat Nederlandse producten en diensten internationaal vermarkt worden. Daarnaast wordt de kans vergroot dat Nederland een rol speelt in het bepalen van de standaarden. Daarmee kan smart mobility nationaal en internationaal net zo waardevol voor Nederland worden als watermanagement.”⁹³

Open data kunnen een bijdrage leveren aan het realiseren van deze ambitie. De gedachte is dat, omdat de data open beschikbaar zijn, voorwaarden worden gecreëerd voor het bedrijfsleven. Naast de vorm (open data) is het natuurlijk ook van belang dat er innovatieve inhoud beschikbaar is die mogelijkheden biedt tot het ontwikkelen van nieuwe producten en diensten.

Overige effecten

Vestigingsklimaat: Verbetering van vervoerssysteem door intelligent gebruik van open data kan leiden tot een verbetering van het vestigingsklimaat van Nederland en daarmee economische groei genereren.

Aansluiting arbeidsmarkt en woningmarkt: Betere bereikbaarheid leidt ertoe dat meer banen bereikbaar zijn vanuit bepaalde woongebieden en andersom dat meer werknemers wonen binnen een bereikbare afstand van een baan. Dit kan positief uitwerken op de arbeidsmarkt. Dit kan effect hebben op de grondwaarde.

⁹³ Brief van de Minister van Infrastructuur en Milieu aan de Tweede Kamer, Kamerstuk 31305, nr. 214, 3 december 2015, blz. 2.

Tabel 4.1: Kwalitatief overzicht van verschillen effecten (ordinaal)

NDW	0	-1	1	2
Investeringskosten				
-Kosten om data 'linked' te maken			-	--
Directe effecten				
-Serverkosten		+	-	--
-Administratieve lasten		-		+
-Dubbel verzamelen		-	+	++
-Consumentensurplus		-	+	++
Indirecte effecten				
-Minder voertuigverliesuren		-	+	++
-Omrijkosten		+	-	--
-Betrouwbaarheid		-	+	++
-Brandstofgebruik		pm	pm	pm
-Nieuwe producten en bedrijven		-	+	++
Overige effecten				
-Vestigingsklimaat		-	+	++
-Aansluiting arbeidsmarkt en woningmarkt		-	+	++
Totaal		pm	pm	pm

+ = meer baten

- = meer lasten

pm = pro memorie

Tabel 4.1 geeft een overzicht van de verschillen tussen de alternatieven. Dit is per effect ordinaal gebeurd. De serverkosten in alternatief -1 zijn bijvoorbeeld lager dan in het 0-alternatief. In alternatief 1 zijn deze hoger en in alternatief 2 zijn deze hoger dan in alternatief 1. Dit zegt niets over de omvang van deze kosten, noch over de verschillen tussen deze stappen. Toch kunnen op grond van tabel 4.1 (in vergelijking met het 0-alternatief) de volgende conclusies worden getroffen.

Min 1 alternatief: geen open data meer

- *Investeringskosten.* Geen.
- *Directe effecten.* Door het opwerpen van een 'betaalmuur' zullen de data minder gebruikt gaan worden. Dit leidt tot besparingen op de serverkosten. Het leidt tot stijging van administratieve lasten om betalingen te verwerken. Aan de data worden geen extra kosten gespendeerd om deze 'linked' te maken. Afhankelijk van de prijs en andere verstrekingsvoorwaarden kunnen partijen in de verleiding komen kosten te maken om eigen meetdata te verzamelen en tegen betaling te verstrekken. Door prijsverhoging verdampt een deel van het consumentensurplus. Omdat er ook minder gebruik van de data wordt gemaakt wordt dit niet volledig gecompenseerd door het gestegen producentensurplus.
- *Indirecte effecten.* Door het gebruik van NDW-data – real-time informatie over de toestand op de weg – kunnen gebruikers hun gedrag aanpassen aan deze toestand, waardoor er minder of kortere files ontstaan en er minder voertuigverliesuren optreden. Wanneer de data niet meer open beschikbaar worden gesteld neemt dit gebruik af en leidt dit tot kosten: extra voertuigverliesuren, omdat door de prijsverhoging (ook wanneer de gegevens via alternatieve bronnen worden verkregen) er vraaguitval zal optreden. Weggebruikers zullen minder alternatieven kiezen op grond van de informatie en de omrijkosten zullen daarom lager zijn (verwacht mag overigens worden dat dit effect kleiner is dan het effect van de voertuigverliesuren; hieronder zullen deze effecten worden gemonetariseerd). Minder informatie en meer voertuigverliesuren gaan ten koste van de betrouwbaarheid van de weg. Het

effect op brandstofgebruik is onduidelijk. Er zullen door de verminderde toegankelijkheid van de data minder bedrijven komen met nieuwe producten op grond van deze data.

- *Overige effecten.* Extra vertraging op de weg schaadt het vestigingsklimaat in Nederland. Door verhoging van de reistijden zijn vanuit woongebieden minder banen bereikbaar met een acceptabele reistijd. Bij de keuze van werk kan dit leiden tot additionele fricties op de arbeidsmarkt; er zijn immers transactiekosten gemoeid met het verkopen van een woning bij het aanvaarden van een nieuwe baan (het effect dat reistijd heeft op woningprijzen; het aantal banen binnen 45 minuten reistijd heeft een voorspellende waarde op de prijs van een woning; zit al verdisconteerd in de waardering van voertuigverliesuren: dit slaat neer in de woningprijs).

Plus 1 alternatief: NDW open data 'linked' maken

- *Investeringskosten.* Er moeten kosten worden gemaakt om de data 'linked' te maken.
- *Directe effecten.* Door het 'linked' maken van deze data wordt de relevantie ervan vergroot en zullen deze vaker worden opgevraagd. Dit toenemend gebruik leidt tot hogere serverkosten. Wat betreft administratieve lasten om betalingen te verwerken is er geen verschil met de huidige situatie: open blijft open. Doordat 'linked data' meer betekenis heeft, kan het meer alternatieve databestanden vervangen en zijn er kostenbesparingen op het gebied van dubbel verzamelen te behalen. Door het 'linked' maken van de data neemt de waarde ervan toe. Bovendien zal door het extra gebruik van de data het consumentensurplus stijgen.
- *Indirecte effecten.* Door het 'linked' maken kunnen nog beter relaties gelegd worden met andere relevante bronnen en zullen daardoor gebruikers beter in staat zijn om vertragingen te vermijden. Hierdoor zullen wel de omrijkosten toenemen. Er is verder een positief effect op de betrouwbaarheid. De brandstofkosten zijn onduidelijk. Door het 'linked' maken van de data ontstaan verdere perspectieven voor de ontwikkeling van nieuwe producten en bedrijven.
- *Overige effecten.* De verdere verbetering op het gebied van bereikbaarheid leidt tot verbetering van het vestigingsklimaat. Ook verbetert de aansluiting van de woningmarkt op de arbeidsmarkt. Zonder transactiekosten verbonden met een verhuizing kunnen vanuit een woning meer (nieuwe) banen worden bereikt.

Plus 2 alternatief: alle data als 'linked open data' (incl. private data)

- *Investeringskosten.* Er zullen extra kosten moeten worden gemaakt om de data 'linked' te maken. Dit betreft Immers niet alleen publieke data, maar ook private data.
- *Directe effecten.* De data omvatten meer gegevens en winnen aan relevantie. Extra gebruik zal ertoe leiden dat meer serverkosten moeten worden gemaakt. De administratieve lasten van het bedrijfsleven zullen worden verlaagd: bedrijven hoeven geen betaaladministratie meer te voeren bij gratis levering van de data. Verdere integratie van data verlaagt verder de kosten om dubbele bestanden bij te houden. Het consumentensurplus zal verder stijgen. Door extra gebruik overstijgt dit het verlies aan producentensurplus.
- *Indirecte effecten.* Door de nog hogere kwaliteit van de publieke en private 'linked data' kunnen voertuigverliesuren verder worden beperkt. Deels zal dit leiden tot extra omrijkosten. Wel levert dit een grotere betrouwbaarheid op van het vervoerssysteem. Het is onduidelijk wat de effecten op brandstofgebruik zijn. Wat betreft nieuwe producten en diensten is enerzijds sprake van meer en relevante data die een basis vormen voor deze producten en diensten. Deze data zijn laagdrempelig (open) beschikbaar, waardoor deze toegankelijk zijn voor het midden- en kleinbedrijf. Het inbrengen van private data in dit alternatief kan leiden tot verlies aan exclusiviteit van informatieproducten, waardoor een prikkel kan verdwijnen om waarde toe te voegen. In dit alternatief is echter sprake van een vrijwillige inbreng van deze data op grond van de gedachte dat meerwaarde – ook voor de private bedrijven – kan ontstaan. Deze condities dragen eraan bij dat de uitkomst positief is.

- *Overige effecten.* Verdere verbetering van betrouwbaarheid en vermindering van voertuigverliesuren leiden tot een verbetering van het vestigingsklimaat. Dit heeft (in het verlengde van het Plus 1 alternatief) waarschijnlijk ook een positief effect op de aansluiting van woningmarkt en arbeidsmarkt.

4.5 Monetarisering

Op grond van het bovenstaande en op grond van interviews met betrokken deskundigen kunnen de kosten en effecten van de alternatieven als volgt worden gemonetariseerd.

Investeringskosten

- *Kosten om data 'linked' te maken.* Geschat wordt dat voor variant Plus 1 de voorbereidingskosten ca. €25.000 en voor Plus 2 €100.000 zullen zijn. De beheerkosten voor Plus 1 zullen ca. €50.000 per jaar zijn en voor variant Plus 2 €100.000 per jaar. Contant gemaakt zijn hiermee de kosten €0,9 miljoen voor plus 1 en €1,9 miljoen voor plus 2.

Directe effecten

- *Serverkosten.* Op grond van het thans parallel lopend onderzoek van Ecorys kunnen de extra kosten van open data ten opzichte van verstrekking van de data tegen beperkende voorwaarden (niet gratis voor eenieder beschikbaar) worden geschat op circa €10.000 per jaar (netto contante waarde €0,18 miljoen). Dit is slechts een fractie van de kosten van het verstrekken van de verkeersdata. Op dit gebied is er echter geen verschil tussen de alternatieven. Verondersteld mag worden dat voor de stap van alternatief 0 naar Plus 1 en/of Plus 2 vergelijkbare kosten (€10.000 per jaar) gemaakt dienen te worden.
- *Administratieve lasten.* Deze lasten zijn afhankelijk van het aantal gebruikers en zijn mogelijk €50.000 per jaar (contante waarde €0,9 miljoen).
- *Dubbel verzamelen.* De voortgang die met NDW gemaakt is, maakt dat de residuele verschillen hier klein zijn tussen de alternatieven.
- *Consumentensurplus.* Het consumentensurplus voor het extra beschikbaar hebben van verkeersdata interfereert met de indirecte effecten die kunnen worden bereikt door met deze informatie in de hand andere keuzen te maken, waardoor reistijdverlies kan worden beperkt (door een onder ideale omstandigheden suboptimale oplossing (omrijkosten) te kiezen), waardoor betrouwbaarheid kan worden vergroot. Dit betekent dat het consumentensurplus van verkeersdeelnemers overeenkomt met deze effecten. Daarnaast zijn er andere gebruikers die deze data benutten om producten en diensten aan te bieden. Deze producten en diensten staan ook opgevoerd bij de indirecte effecten. Invulling van dit consumentensurplus kan zo achterwege blijven, omdat de effecten elders worden meegenomen. Dit betekent dat het verschil in consumentensurplus voor de gebruiker van deze data die hier vervolgens niets mee doet (de eindgebruiker die bevrediging vindt in het consumeren van real-time-informatie over de weg, zonder dat hij hier iets mee doet) niet wordt meegenomen in deze analyse. Verwacht wordt dat dit deel bescheiden is ten opzichte van de overige effecten.

Indirecte effecten

- *Vermindering voertuigverliesuren.* Verwacht wordt dat het actieprogramma *Beter geïnformeerd op weg* per 2020 zorgt voor een vermindering van 3 miljoen voertuigverliesuren per jaar.⁹⁴ Een zesde deel hiervan (er zijn in dit actieprogramma zes acties voorzien die tot dit resultaat moeten

⁹⁴ Mobiliteitsbeleid; Brief regering; Beter geïnformeerd op weg: Koersbepaling Reisinformatie & Verkeersmanagement, 12 februari 2013, kamerstuk 31 305 nr. 202.

leiden): een vermindering van 0,5 miljoen voertuigverliesuren kan specifiek worden toegerekend aan het verschil tussen de alternatieven 0 en Plus 2 wat betreft open data. Dit kan worden gewaardeerd op €8 miljoen per jaar vanaf 2020; ofwel €114 miljoen in contante waarde per 2016 (op grond 5,5%).

- *Omrijkosten*. Verondersteld wordt dat de helft van de besparingen van voertuigverliesuren vanwege suboptimale omrijkosten teniet worden gedaan.⁹⁵
- *Betrouwbaarheid*. Conform de algemene praktijk van MKBA's wordt deze geschat op 25% van de vermindering van voertuigverliesuren.⁹⁶
- *Brandstofkosten*: Verwacht wordt dat de effecten ongeveer in evenwicht zijn.
- *Het effect op nieuwe producten en bedrijven* is niet gemonetariseerd.
- *Het verminderen van kosten van alternatieve systemen*: door het niet meer dubbel bijhouden van vergelijkbare gegevens. Dit is niet gemonetariseerd.

Overige effecten

- *Vestigingsklimaat*. Bij ons zijn alleen kwalitatieve ramingen bekend van de relatie tussen voertuigverliesuren en het vestigingsklimaat.
- *Aansluiting tussen arbeidsmarkt en woningmarkt*. Op dit gebied is enige literatuur beschikbaar die aangeeft dat er een effect is,⁹⁷ maar het is onvoldoende uitgewerkt om rechtstreeks te kunnen gebruiken.

Tabel 4.2: Gemonetariseerd overzicht van verschillen effecten in miljoenen euro's

NDW	0	-1	1	2
Investeringskosten				
-Kosten om data 'linked' te maken			-0,9	-1,9
Directe effecten				
-Serverkosten		-0,18		0,18
-Administratieve lasten		-0,9		0,9
-Dubbel verzamelen		pm	pm	pm
-Consumentensurplus		*	*	*
Indirecte effecten				
-Minder voertuigverliesuren		-	57	114
-Omrijkosten		+	-29	-57
-Betrouwbaarheid		-	14	29
-Brandstofgebruik		pm	pm	pm
-Nieuwe producten en bedrijven		-	+	++
Overige effecten				

⁹⁵ Dit is een toepassing van de 'rule of half': gemiddeld zal de keuze halverwege uitvallen. In het 'Eindrapport praktijkproef Amsterdam' (Amsterdam Mobiel, december 2015) bleek overigens dat in de praktijk verkeerskundige effecten tegen kunnen vallen. De gemeten reistijd van on-trip app-gebruikers was aanmerkelijk hoger dan zowel de verwachte reisduur als de gemiddelde reisduur van overige alternatieve routes. Hiervoor zijn verschillende verklaringen. Een deel van de problemen lag aan de datakwaliteit; bepaalde essentiële data, zoals over het onderliggend wegennet, de doorstroming bij de afritten, de stabiliteit van de gegevens over het parkeren (deze data lagen er vaak uit) en de betrouwbaarheid van de data lieten nog te wensen over.

⁹⁶ Hilbers, H., J. van Meerkerk, A. Verrips, W. Weijtschede- van der Straaten, P. Zwaneveld (2015) Maatschappelijke Kosten en Baten Prijsbeleid Personenauto's, CPB/PBL Achtergronddocument, Den Haag.

⁹⁷ Zie bijvoorbeeld Van Ommeren J., Rietveld P. (2007) Compensation for commuting in imperfect urban markets, Papers in Regional Science, 86 (2), pp.241-259.

-Vestigingsklimaat		-pm	+pm	++pm
-Aansluiting arbeidsmarkt en woningmarkt		-pm	+pm	++pm
Totaal		-1,08+pm	41,1+pm	85,18+pm

*Consumentensurplus is meegenomen onder indirecte effecten.

4.6 Conclusies

Open data zijn slechts een kleine stap binnen het gehele programma van NDW om te komen tot betere gegevens over wegen en verkeer. Het effect van open data is daardoor lastig af te zonderen van de totale innovatie. Zowel qua kosten als opbrengsten is het effect van open data bescheiden ten opzichte van het totale takenpakket van NDW. Veel bestaande professionele partijen kiezen ervoor om naast open data afspraken vast te leggen in een Data+Diensten-overeenkomst. Onderzoek hiernaar is geen onderdeel van deze MKBA.⁹⁸ De maatschappelijke kosten en baten van open data binnen deze verdergaande integratie zijn overigens wel positief en het kan daardoor als een sluitstuk worden beschouwd van data-innovaties.

⁹⁸ Maar wel van een onderzoek dat Ecorys thans uitvoert. Ons eerder onderzoek naar businessmodellen voor open data heeft uitgewezen dat er behoefte kan zijn aan betaalde datadiensten naast gratis levering van open data, zie Welle Donker, F. & B. van Loenen (2015). Duurzame businessmodellen voor open data. Delft, Kenniscentrum Open Data: <http://repository.tudelft.nl/view/ir/uuid%3Ac5b18289-81d9-4692-ab2e-62b3a96c02d2>.

5 Landelijke voorziening WOZ-waarden

5.1 Inleiding WOZ-waarden

De Basisregistratie Waardering Onroerende Zaken (WOZ) bevat gegevens van alle onroerende zaken in Nederland. Voor belastingheffingen kunnen gemeenten, waterschappen en het Rijk uitgaan van eenzelfde waarde van een onroerende zaak: de WOZ-waarde. De WOZ-waarde wordt gebruikt voor de belastingheffing door het Rijk (inkomstenbelasting, vennootschapsbelasting, erf- en schenkingsbelasting), de waterschappen (waterschapsysteemheffing) en de gemeenten (onroerendezaakbelastingen). Daarnaast wordt de WOZ-waarde gebruikt door CBS voor statistische doeleinden, en door het Openbaar Ministerie (OM) en geautoriseerde partijen (notarissen, banken en verzekeraars) ter bestrijding van vastgoedfraude.

Gemeenten bepalen jaarlijks de WOZ-waarde van 8,7 miljoen panden aan de hand van modellen waarin de gemiddelde waarden van historische transacties van vergelijkbare woningen in een wijk worden gewogen.

Gemeenten mogen zich aansluiten bij de Landelijke Voorziening WOZ (LV WOZ) nadat zij met goed gevolg een aansluitingstoets hebben afgelegd.⁹⁹ Het Ministerie van Financiën is de opdrachtgever van de LV WOZ, en de Waarderingskamer houdt toezicht en toetst de gemeenten. De LV WOZ is primair bedoeld om WOZ-gegevens te delen met andere overheden en organisaties die geautoriseerd zijn om WOZ-gegevens te gebruiken. Met het oog op een zorgvuldig gebruik van de gegevens door de afnemers stelt de LV WOZ eisen aan de consistentie en kwaliteit van de aangeleverde gegevens. Voordat een gemeente kan aansluiten op de LV WOZ moet er voor een aansluittoets gedaan worden.¹⁰⁰ De data bevatten een groot aantal (persoons)gegevens van eigenaren en gebruikers van onroerende zaken, en daarom zijn er ook beveiligingsmaatregelen nodig voor de LV WOZ. Op 31 januari 2017 waren 192 van 388 gemeenten aangesloten op de LV WOZ.

Ten behoeve van deze casus is gekeken naar de effecten van het per 2016 aanbieden van informatie over referentiegrondwaarden in de Duitse deelstaat Noordrijn-Westfalen (Bijlage 5). De uitkomst van maatschappelijke kosten en baten is daar positief. De administratieve lasten zijn verlaagd en de gegevens worden op veel grotere schaal gebruikt dan voorheen. Dit leidt niet alleen tot een consumentensurplus, maar hierdoor wordt ook de transparantie van de vastgoedmarkt vergroot.

5.2 Huidige situatie ontsluiting WOZ-waarden

In principe kan een inwoner via het WOZ-waardeloket WOZ-waarden inzien. Als de betreffende gemeente nog niet is aangesloten op de LV WOZ, dan kan een inwoner kan bij de gemeente kosteloos de WOZ-waarde van de eigen woning opvragen, hetzij via het eigen dataportaal van de gemeente of via een gemeenschappelijk dataportaal van bijvoorbeeld Regionale Belasting Groep.¹⁰¹

⁹⁹ <https://www.waarderingskamer.nl/basisregistratie-woz-lv-woz/aansluiten-gemeenten-lvwoz/>

¹⁰⁰ <https://www.rijksoverheid.nl/documenten/kamerstukken/2016/11/07/beantwoording-kamervragen-over-bericht-woz-inzage-gemeenten>.

¹⁰¹ Een inwoner mag maximaal drie WOZ-waarden van vergelijkbare woningen uit de directe omgeving, maximaal drie WOZ-waarden van woningen uit een andere categorie (ander type) en maximaal drie WOZ-waarden van woningen uit een andere locatie (binnen de gemeente) inzien. Bij een WOZ-waarde rapport maakt de gemeente zelf een selectie van woningen die vergelijkbaar worden geacht.

Voor alle inwoners geldt dat als men het niet eens is met de vastgestelde waarde er een bezwaar kan worden ingediend. Als de WOZ-waarde te hoog wordt vastgesteld zullen de verwante belastingaanslagen te hoog uitvallen. Maar als de WOZ-waarde te laag wordt vastgesteld, heeft dat een negatief effect op de verkoopwaarde van de woning. Onder het oude regime kon een inwoner alleen de WOZ-waarde van de eigen woning inzien en niet van vergelijkbare woningen, waardoor de taxatiewaarde niet als transparant werd beschouwd.

In 2016 is in circa 1% van alle beschikkingen een bezwaar ingediend, een daling ten opzichte van 2013 toen in 3% van alle beschikkingen bezwaren werden ingediend. In ruwweg een kwart van de gevallen maken de bezwaarmakers gebruik van een gespecialiseerd bureau. Rond het indienen van bezwaarschriften is een hele industrie ontstaan van zogenaamde 'no cure no pay'-bureaus. Bij een gegrond bezwaar ontvangen zij de daarvoor geldende wettelijke vergoedingen: €246 voor het indienen van een bezwaarschrift, €246 voor het bijwonen van een hoorzitting, €50 (+ BTW) per uur (meestal 2 uur) voor het maken van een taxatierapport. Een gegrond bezwaar levert al snel 2x €246 + €121 (taxatierapport) = €613. In 2016 werden in totaal 87.000 bezwaren ingediend waarvan ca. 53% werd gehonoreerd. Van de 87.000 bezwaren werden er in 2016 ca. 19.500 bezwaren ingediend via 'no cure no pay'-bureaus.¹⁰² Gemeenten betalen jaarlijks enkele miljoenen euro's aan 'no cure no pay'-bureaus. De gemeenten hebben in 2014 ruim €7 miljoen uitbetaald aan proceskosten,¹⁰³ en in 2015 werd dit bedrag geschat op €4,5 tot 11 miljoen.¹⁰⁴

Om meer transparantie te bieden in de onderlinge verschillen tussen woningen, zijn op 1 oktober 2016 de WOZ-waarden openbaar gemaakt. Dit wil zeggen dat iedereen de WOZ-waarden kan raadplegen voor persoonlijk gebruik. Het betekent dus niet dat de WOZ-waarden open data zijn. Voor de gemeenten die aangesloten zijn bij de LV WOZ kunnen de WOZ-waarden worden ontsloten via het WOZ-waardeloket. Om te voorkomen dat het WOZ-waardeloket wordt overbelast door massale afname, is een drempelmechanisme ingebouwd. Dit betekent dat er een maximum is gesteld aan het aantal woningen dat binnen een bepaalde tijdsperiode opgevraagd mag worden. De getoonde objectkenmerken zijn afkomstig uit de Basisregistraties Adressen en Gebouwen (BAG) en worden aangeroepen via de BAG Viewer. Voor het taxeren zijn veel meer kenmerken van de woning gebruikt dan alleen bouwjaar en oppervlakte. De objectkenmerken die gebruikt zijn bij de taxatie zijn echter niet openbaar en daarom niet te zien via dit loket.

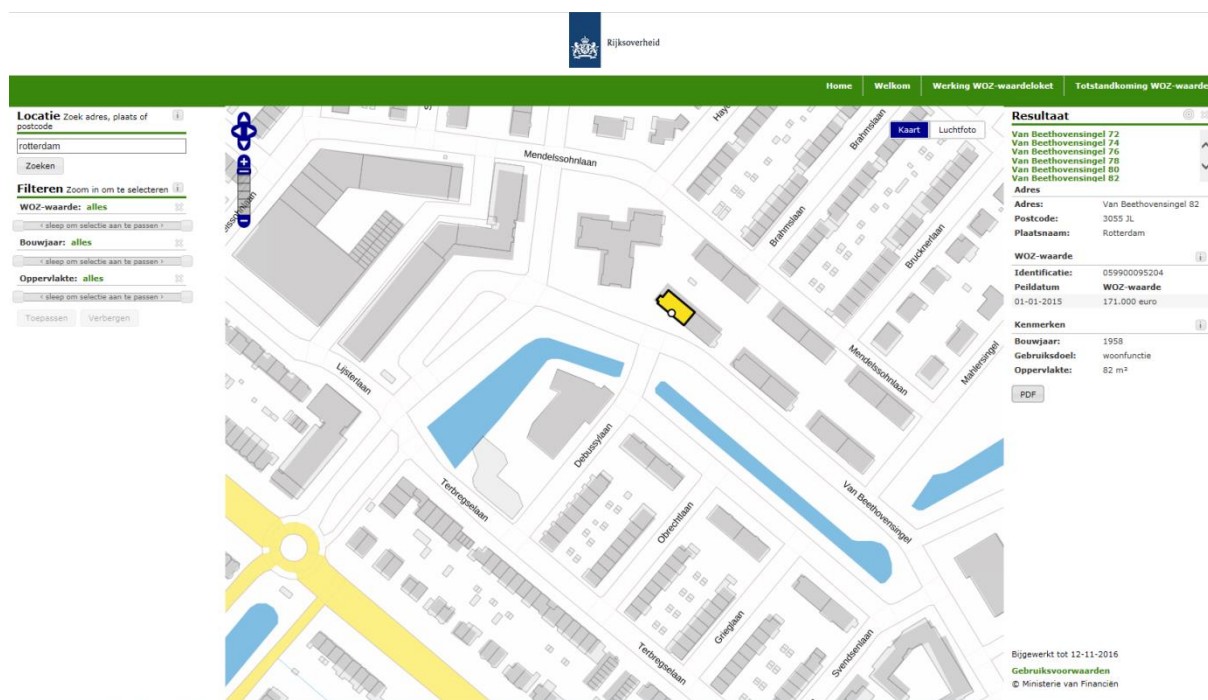
De gemeenten die nog niet zijn aangesloten bij de LV WOZ, kunnen de WOZ-waarden via de eigen website aanbieden of gebruik maken van alternatieve platforms.

¹⁰² Dit getal is een schatting: niet alle gemeenten konden uitsplitsing maken, zie <https://www.waarderingskamer.nl/woz-in-cijfers/feiten-over-de-woz/>

¹⁰³ <http://radar.avrotros.nl/uitzendingen/gemist/14-03-2016/gratis-advies-woz-bezwaar/>

¹⁰⁴ Zie <http://nos.nl/artikel/2026390-gemeenten-miljoenen-kwijt-aan-no-cure-no-paybureaus.html>

Figuur 5.1: Screenshot WOZwaardeloket.nl



De gemeenten zijn huiverig voor een vrij toegankelijk WOZ-Waardeloket: niet alle gemeenten hebben hun processen ingericht om de verwachte toename aan vragen snel af te handelen. Een andere angst bij de gemeenten is dat het zal leiden tot meer bezwaarschriften aangezwengeld door 'no cure no pay'-bedrijfjes, gespecialiseerd in bezwaarschriften.¹⁰⁵

5.3 Projectalternatieven voor de WOZ-waarden

We zullen de volgende beleidsopties voor LV WOZ-gegevens in beschouwing nemen:

- 0-alternatief: WOZ-waarden van woningen zijn openbaar en centraal beschikbaar via LV WOZ.
- Loket Min 1 alternatief: Voor bewoners zijn WOZ-waarden alleen op verzoek beschikbaar als pdf-document en alleen voor de eigen woning en door de gemeente geselecteerde vergelijkbare woningen (passieve openbaarheid). Voor een aantal geselecteerde partijen (notarissen, e.d.) zijn WOZ-waarden beschikbaar onder strikte voorwaarden en tegen betaling, de situatie voor 1 oktober 2016.
- Plus 1: alle WOZ-waarden zijn in een aangepaste versie beschikbaar als open data voor een ieder (volgens TBL 3-sterrenniveau).
- Plus 2: WOZ-waarden zijn in een aangepaste versie beschikbaar als 'linked open data' (volgens ten minste TBL 4 sterrenniveau).

Hierbij biedt het Min 1 alternatief de mogelijkheid om de effecten van de andere alternatieven in kaart te brengen.

¹⁰⁵ <https://www.rijksoverheid.nl/documenten/kamerstukken/2016/11/07/beantwoording-kamervragen-over-bericht-woz-inzage-gemeenten>.

Het 0-alternatief is de situatie die waarschijnlijk in de nabije toekomst zal worden bereikt. De aansluiting van gemeenten wordt nu nog door technische en organisatorische factoren afgeremd, maar er zijn stappen ondernomen om die belemmeringen zo veel mogelijk op te heffen. Het actief openbaar maken van WOZ-waarden van niet-woningen is momenteel in onderzoek (maar maakt geen onderdeel uit van deze analyse).

Voor het realiseren van het Plus 1 alternatief zijn er nog grote stappen te nemen, maar het is wel de uiteindelijke ambitie van het Ministerie van Financiën. Omdat de LV WOZ-data persoons- en bedrijfsgegevens bevatten, zijn de WOZ-waarden niet beschikbaar als open data. Voordat de data als open data vrijgegeven kunnen worden, zullen er bewerkingslagen nodig zijn om de gevoelige aard van de gegevens weg te nemen. De koppeling tussen de WOZ-waarde en personen dient op een of andere manier te worden doorbroken.

Voor het realiseren van het Plus 2 alternatief zijn er nog grotere stappen te nemen dan voor het Plus 1 alternatief. Dit alternatief zal in de nabije toekomst niet realistisch zijn.

5.4 Kwalitatieve omschrijving van investeringskosten en effecten

In deze sectie beschrijven we de verschillen in investeringskosten en effecten van de beleidsalternatieven.

Investeringskosten

Bij het opzetten van het WOZ-Waardeloket is er bewust voor gekozen om het WOZ-Waardeloket via het Kadaster (PDOK) te hosten. Er zijn extra beveiligingsmaatregelen genomen om te voorkomen dat het WOZ-Loket misbruikt wordt door middel van crawlen¹⁰⁶ of scrapen,¹⁰⁷ of door mensen door te frequent aanklikken. Als er vanaf een bepaald IP-adres te veel verzoeken in een bepaalde tijd komen, wordt dat IP-adres geblokkeerd. Het Kadaster heeft extra kosten moeten maken om de kwaliteit van de luchtfoto's die worden gebruikt, te verbeteren.

Kosten open data: Voor een overstap naar het Plus 1 alternatief zullen de kosten hoger zijn omdat de gegevens bewerkt zullen moeten worden (om de link tussen waarde en persoon te doorbreken) voordat deze als open data vrijgegeven kunnen worden.

'Linked open data': Alhoewel de open data niet aan personen gelinkt kunnen worden, kunnen er wel kosten worden gemaakt om de data 'linked' te maken.

Directe effecten openbare WOZ-waarden

Sinds het WOZ-Waardeloket live is gegaan op 1 oktober, zijn er in de eerste weken veel page views geweest. Een consequentie daarvan is dat het aantal bevragingen via de BAG Viewer daardoor ook veel hoger is dan voor 1 oktober, en een stijging van circa 250% kent. Die kosten van de extra belasting op de BAG-server wordt door het Kadaster gedragen.

Een tweede effect is dat op basis van de ervaringen gedurende de eerste dagen enkele veiligheidsmaatregelen zijn aangescherpt, zoals het beperken van het aantal bevragingen per tijdseenheid. Er zijn ook al een aantal IP-adressen geblokkeerd, waaronder een IP-adres uit Nigeria.

Er is voorsnog geen bewijs van een toestroom van bezwaren ingediend door 'no cure no pay'-bedrijven.

¹⁰⁶ Crawl is een computertechniek om met bots het web op een methodische en geautomatiseerde manier door te bladeren en een lokale kopie van de gevonden pagina's te maken om deze later te kunnen verwerken en indexeren.

¹⁰⁷ Scrapen is een computertechniek waarbij software of scripts wordt gebruikt om informatie van webpagina's te extraheren en al dan niet te analyseren.

Het ligt in de lijn der verwachting dat er rond februari 2017 weer veel page views zullen komen met de nieuwe aanslagronde voor de gemeentelijke belastingen in 2017 als bewoners de WOZ-waarde van hun woning willen vergelijken met andere woningen. Er wordt ook verwacht dat afnemers van de LV WOZ zelf meer gebruik zullen gaan maken van het WOZ-Waardeloket.

De Waarderingskamer verwacht ook van het WOZ-Waardeloket dat er meer terugmeldingen in de vorm van wijzigingsverzoeken zullen komen, waardoor de kwaliteit van de data zal verbeteren.

Serverkosten: De LV WOZ maakt gebruik van servercapaciteit van het Kadaster. Die servercapaciteit zal moeten toenemen naarmate meer gemeenten aansluiten op de LV WOZ, ook omdat alle historische gegevens beheerd moeten worden. Omdat het WOZ-Waardeloket gebruik maakt van bestaande diensten (BAG Viewer) zullen de infrastructurele meerkosten van het 0-alternatief t.o.v. het Min 1 alternatief relatief laag zijn. Voor het Plus 1 en Plus 2 alternatief zullen de meerkosten hoger zijn omdat de open data vaak op een aparte server worden beheerd. De meerkosten zullen ook afhankelijk zijn van de keuzes die gemaakt zullen worden voor de ontsluiting van de data.

Administratieve lasten: Hoewel er waarschijnlijk meer wijzigingsverzoeken zullen komen met openbare WOZ-waarden, hoeft dit niet noodzakelijk te leiden tot meer bezwaarschriften. Er zijn in een aantal gemeenten pilots gehouden waarin bewoners een voormelding van de WOZ-beschikking ontvingen met de bedoeling dat bewoners een wijzigingsverzoek konden indienen in plaats van een formeel bezwaarschrift. De voormeldingen hebben in die gemeenten weliswaar geleid tot meer wijzigingsverzoeken maar juist tot minder bezwaarschriften, soms wel tot een halvering van het aantal bezwaarschriften. De voormeldingen hebben niet geleid tot minder bezwaarschriften via 'no cure no pay'-bedrijven. Het afhandelen van een wijzigingsverzoek vergt minder tijd dan het afhandelen van een bezwaar, en er zijn geen proceskosten. De bespaarde kosten door verminderde bezwaren en bestandscontrole waren gemiddeld €30.000 per gemeente.¹⁰⁸ De pilot in Tilburg toonde ook aan dat er nauwelijks een effect was op de OZB-inkomsten.¹⁰⁹ De transactiekosten voor de burgers zullen omlaag gaan omdat het nu makkelijk wordt om zelf een vergelijking van WOZ-waarden uit te voeren via één centraal loket.

Kwaliteitsslag: De pilots met WOZ-voormeldingen toonden aan dat openbaarheid van WOZ-waarden heeft geleid tot een hogere kwaliteit van de data, zoals betere registratie van objectkenmerken).¹¹⁰ Voor de Plus 1 en de Plus 2 alternatieven is het de vraag of de kwaliteitsslagen nog verder verbeterd kunnen worden omdat er waarschijnlijk beperkingen zullen gaan gelden voor de koppeling van WOZ-waarden aan concrete (tot personen te herleiden) WOZ-objecten.

Consumentensurplus: Gratis openbaar maken van WOZ-waarden zal leiden tot een consumentensurplus, omdat de gratis data toch een waarde voor de burgers vertegenwoordigen. Bij het Plus 1 en Plus 2 alternatief zal het consumentensurplus nog hoger zijn omdat een opendataversie zou kunnen leiden tot meer producten en diensten.

Indirecte effecten

De indirecte effecten van openbare WOZ-waarden lijken voorsnog beperkt.

De pilots van de WOZ-voormeldingen hebben aangetoond dat door de hogere transparantie, de acceptatie van de WOZ bij inwoners omhoog gaat.¹¹¹ Het ligt in de lijn der verwachting dat openbare WOZ-waarden hetzelfde effect zal hebben.

¹⁰⁸ Onderzoekcentrum Drechtsteden (2014). Doorbraak WOZ 2014. Ervaringen en effecten van een nieuwe ronde: 71. http://www.onderzoekcentrumdrechtsteden.nl/dds/up/ZixzvkdJG_def_rapportage_doorbraak_WOZ_2014.pdf.

¹⁰⁹ Hypercube Business Innovation (2013). WOZ voormelding. Resultaten fase 2 Tilburg: 32. https://vng.nl/files/vng/woz_voormelding_fase_2_tilburg_-_resultaten_v1.1.pdf, p.3.

¹¹⁰ Onderzoekcentrum Drechtsteden (2014).

¹¹¹ Zie o.m. Onderzoekcentrum Drechtsteden (2014), Hypercube Business Innovation (2013), en <https://www.wozinformatie.nl/public/nieuwsarchief/item/?itemID=7708&bc=nieuws> voor de pilot in Borne.

De overheid zal zelf ook meer gebruik gaan maken van de Basisregistratie WOZ via het WOZ-Waardeloket, en daarmee zal de overheid efficiënter worden.

In de beperkte tijd dat het WOZ-Waardeloket live is gegaan, lijkt het er niet op dat dat al heeft geleid tot meer nieuwe producten of bedrijven, maar is het nog te vroeg om conclusies te trekken. Bedrijven hebben tijd nodig om nieuwe toepassingen te ontwikkelen en de hele opzet van het systeem is gericht op beperkt gebruik. Het aantal toegestane raadplegingen beperkt de mogelijkheden voor professioneel hergebruik aanzienlijk. Het ligt in de lijn der verwachting dat in de Plus 1 en Plus 2 alternatieven het bedrijfsleven de open WOZ-waarden meer zullen oppakken, vooral rond dienstverlening op de vastgoedmarkt en de markt van vastgoedfinancieringen. Bij het Plus 2 alternatief kunnen op dit gebied meer links worden gelegd. Het in de open data varianten doorbreken van de relatie tussen WOZ-waarde en persoon, betekent wel dat bepaalde vormen van waarde voor bedrijven - het direct benaderen van consumenten op grond van directe kennis op grond van hun specifieke eigenschappen - niet mogelijk wordt gemaakt.

Externe effecten

De transparante openbare WOZ-waarden kunnen leiden tot beter geïnformeerde keuzes op de woningmarkt met betrekking tot verwachte belastingheffingen.

Tabel 5.1: Meerkosten WOZ Loket t.o.v. Beleidsvariant 0 (geen actief gepubliceerde data)

WOZ-waarden	0	Min 1	Plus 1	Plus 2
Investeringskosten				
-Kosten om landelijke voorziening in te richten		-		
Directe effecten				
-Serverkosten		-	0	+
-Administratieve lasten		+	0	0
-Kwaliteitsslag		-	+	+
-Consumentensurplus		-	+	++
Indirecte effecten				
-Vertrouwen en draagvlak		-	+	+
-Nieuwe producten en bedrijven		-	+	++
Overige effecten				
-Beter geïnformeerde woningkeuze		-	+	+
Totaal		-	+	+

Tabel 5.1 geeft een overzicht van de verschillen tussen de alternatieven. Dit is per effect ordinaal gebeurd. Het 0-alternatief is het ijkpunt van de verschillen. De serverkosten in alternatief Min 1 zijn bijvoorbeeld lager dan in het 0-alternatief. In alternatief Plus 1 zullen deze kosten ongeveer hetzelfde zijn omdat die kosten ook al gemaakt moeten worden voor het uitvoeren van de primaire taken van de LV WOZ. In alternatief Plus 2 zijn deze kosten hoger dan in alternatief 1. Dit zegt niets over de omvang van deze kosten, noch over de verschillen tussen deze stappen. Toch kunnen op grond van de tabel de volgende conclusies worden getroffen (zie paragraaf 5.5).

5.5 Monetarisering kosten en baten

Hoewel het openbaar maken van de WOZ-waarden per 1 oktober 2016 een unieke gelegenheid had geboden om de kosten ex ante te bepalen, blijkt in de praktijk dat de kosten moeilijk uit te splitsen zijn van andere kosten. Gebaseerd op interviews hebben wij een inschatting gemaakt van de kosten.

Investeringskosten

Kosten om landelijke voorziening in te richten: De voorbereidingskosten om het WOZ-Waardeloket op te zetten bedroegen circa €200.000 en de kosten van het applicatieonderhoud zullen waarschijnlijk rond de €150.000-200.000 per jaar zijn. Gekapitaliseerd tegen 5,5% is dat in totaal ca €3,4 miljoen.

Kosten om data open te maken: De data hebben een bewerkingslag nodig waarin de link tussen data en individuele WOZ-objecten wordt verbroken, om te voorkomen dat persoonsgegevens open data worden. Dit vraagt een investering in het nadenken en implementeren van dit systeem. Tevens vraagt dit jaarlijks kosten om binnenkomende data in dit format om te zetten. De Waarderingskamer schat de kosten op eenmalig ca. €50.000 en de beheerkosten ca €5.000 per jaar. Totaal gaat het dan om een contante waarde van €140.000.

Extra kosten om data 'linked' te maken: Indien dit in één slag kan worden gedaan met de voorgaande stap, zijn de extra kosten hiervan beperkt. Geschat wordt een contante waarde van €20.000.

Directe effecten

Serverkosten: De serverkosten zullen alleen voor het Plus 1 en Plus 2 alternatief hoger zijn omdat er aparte servers nodig zullen zijn. Voor de andere alternatieven zullen de serverkosten gelijk zijn.

Administratieve lasten: Verwacht wordt dat in het 0-alternatief ten opzichte van het Min 1 alternatief de kosten van bezwaren afnemen. Als de bezwaren met 5% dalen, dan kunnen de gemeenten €2,66 miljoen (5% van 87.000 (aantal bezwaren ingediend in 2016) x €613 (vergoeding voor een gegrond verklaard bezwaarschrift)) jaarlijks besparen op procedurekostenvergoedingen voor derden. Gekapitaliseerd tegen 5,5% betekent dat €48,5 miljoen. Vermindering van het aantal bezwaren leidt ook tot verlaagde kosten binnen de gemeenten zelf. Verwacht mag worden dat deze geringere administratieve lasten van dezelfde orde van grootte zullen zijn als de hogere administratieve lasten om de grotere aantallen wijzigingsverzoeken af te handelen. Voor het Plus 2 alternatief zullen de voorbereidings- en beheerkosten hoger zijn. Deze kosten zullen afhankelijk zijn van de keuzes die er nog gemaakt moeten worden, zoals welke partij de werkzaamheden gaat uitvoeren (gemeenten, LV WOZ, uitbesteden aan het Kadaster).

Kwaliteitsslag: Hoewel hogere datakwaliteit zal leiden tot een efficiëntere overheid, is dit effect niet te kwantificeren.

Consumentensurplus: Door de zeer recente invoering van het loket is het toegenomen gebruik nog onbekend. Bovendien zijn gebruikscijfers in de eerste weken na openstelling niet altijd representatief voor het vervolg: is er sprake van een piek of van een langzame start?

Indirecte effecten

De indirecte effecten kunnen niet gemonetariseerd worden omdat de WOZ-waarden via het WOZ-Waardeloket nog te kort openbaar zijn om de effecten waar te nemen.

Tabel 5.2: Gemonetariseerd overzicht van verschillende effecten (in miljoenen euro's)

WOZ-waarden	0	Min 1	Plus 1	Plus 2
Investeringskosten				
-Kosten om LV in te richten en te beheren		3,4		
-Kosten om dat te bewerken tot open data			-0,14	-0,14
-Extra kosten om data 'linked' te maken				-0,02
Directe effecten				
-Serverkosten		0	0	+
-Administratieve lasten		-48,5		
-Kwaliteitsslag				
-Consumentensurplus		*	*	*
Indirecte effecten				
-Vertrouwen en draagvlak				
-Nieuwe producten en bedrijven		-pm	+pm	++pm
Overige effecten				
-Beter geïnformeerde woningkeuze		-pm	+pm	+pm
Totaal		-45,1	-0,14 + pm	-0,16 +pm

* Zie indirecte kosten.

5.6 Conclusies

Het WOZ-Waardeloket is nog te jong om harde conclusies te trekken over de maatschappelijke baten van openbare WOZ-waarden. De beheer- en infrastructurele kosten van de LV WOZ moeten in ieder geval gemaakt worden voor het uitvoeren van de primaire processen waaronder het openbaar maken van WOZ-waarden. De pilots van de WOZ-voormeldingen hebben aangetoond dat meer transparantie van WOZ-waarden leidt tot minder bezwaren, en daarmee tot lagere administratieve lasten voor de gemeenten. Bovendien hebben de pilots aangegeven dat meer transparantie heeft geleid tot meer draagvlak voor de belastingheffingen. Of een openbare WOZ dezelfde effecten zal hebben, zal in 2017 duidelijk worden wanneer de volgende beschikkingen zullen worden verstuurd. De WOZ-waarden zijn nu alleen openbaar en niet beschikbaar als open data, omdat de WOZ-waarden persoonsgegevens bevatten. Het anonimiseren van de WOZ-dataset brengt extra kosten met zich mee maar als een geanonimiseerde dataset beschikbaar komt als open data, zouden de maatschappelijke baten kunnen stijgen mits het bedrijfsleven deze anonieme data kunnen hergebruiken. Maar de waarde voor het bedrijfsleven zal afhankelijk zijn van de keuzes die gemaakt zullen worden voor de mate van anonimiseren en/of aggregeren, en de kanalen die voor de open data beschikbaar worden gemaakt.

6 Kindermishandeling in NL anno 2005

6.1 Probleemanalyse

Het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) is een zelfstandig onderdeel van het Ministerie van Veiligheid en Justitie (VenJ) en verricht wetenschappelijk (beleidsgericht) onderzoek. In kader van die onderzoeken worden datasets gecreëerd. De ‘Kindermishandeling in Nederland anno 2005 (projectnr. 1172)’ dataset is in 2006 samengesteld op basis van een vragenlijst onder scholieren van het voortgezet onderwijs. Tussen december 2005 en april 2006 hebben 1.845 leerlingen meegedaan aan het onderzoek. Het onderzoek is uitgevoerd in het gehele land, op in totaal 14 willekeurig geselecteerde scholen voor voortgezet onderwijs. In het Kindermishandeling-onderzoek zijn stappen ondernomen om de anonimiteit van de deelnemers te waarborgen, zowel van de deelnemers als van de docenten en de medeleerlingen van de scholen. Er zitten in de ‘Kindermishandeling in Nederland anno 2005’ dataset geen persoonsgegevens.

WODC heeft in 2007 besloten een aantal datasets beschikbaar te stellen voor hergebruik, waarvan het ‘Kindermishandeling in Nederland anno 2005’ er één was.

Hoewel de ‘Kindermishandeling in Nederland anno 2005’ dataset beschikbaar is voor hergebruik, lijkt hergebruik vooralsnog beperkt.

6.2 Huidige situatie

Technisch gesproken is de ‘Kindermishandeling in Nederland anno 2005’ geen open data volgens de principes vastgesteld in de NODA. De dataset is gepubliceerd in een eigendomsmatig software (SPSS) format (.sav), en gebruikers moeten zich registreren voordat de data kunnen worden gedownload. Daarmee voldoet de dataset aan 2 sterren op de TBL-index (beschikbaar met open licentie als machine-leesbare gestructureerde data die alleen met eigendomsmatige software kan worden bewerkt).

De dataset is beschikbaar bij DANS (Data Archiving and Networked Services) via EASY (Electronic Archiving System),¹¹² en via NARCIS (National Academic Research and Collaborations Information System).¹¹³ Om de dataset te kunnen downloaden moet een gebruiker zich eerst registreren bij DANS. DANS doet dat om een beeld te hebben van type gebruikers en aantal downloads. Behalve de dataset zelf, worden er ook metadata en aanvullende documentatie beschikbaar gesteld. De webstatistieken laten zien dat de dataset sinds publicatie in juli 2013, 66 keer is gedownload door ca. 11 unieke gebruikers.¹¹⁴

¹¹² EASY is een online archiveringssysteem voor het deponeren en downloaden van onderzoeksdata. Naast vele nieuwe datasets en collecties zijn ook de collecties van het voormalige Nederlands Historisch Data-archief (NHDA) en het Steinmetz-archief in EASY ondergebracht. Een onderdeel van EASY is het e-depot Nederlandse archeologie (EDNA). (<https://dans.knaw.nl/nl/over/diensten/data-archiveren-en-hergebruiken/easy>).

¹¹³ NARCIS is dé nationale portal voor wie informatie zoekt over wetenschappers en hun werk. Naast wetenschappers maken ook studenten, journalisten en medewerkers binnen onderwijs, overheid en het bedrijfsleven gebruik van NARCIS. NARCIS biedt toegang tot wetenschappelijke informatie waaronder (open access) publicaties afkomstig uit de repositories van alle Nederlandse universiteiten, KNAW, NWO en diverse wetenschappelijke instellingen, datasets van een aantal data-archieven, alsmede beschrijvingen van onderzoeksprojecten, onderzoekers en onderzoeksinstituten (<http://www.narcis.nl/about/Language/nl>).

¹¹⁴ Geregistreerde gebruikers mogen gebruik maken van de mogelijkheid om als “anoniem” te worden weergegeven. Daardoor is het niet mogelijk om alle unieke gebruikers te identificeren.

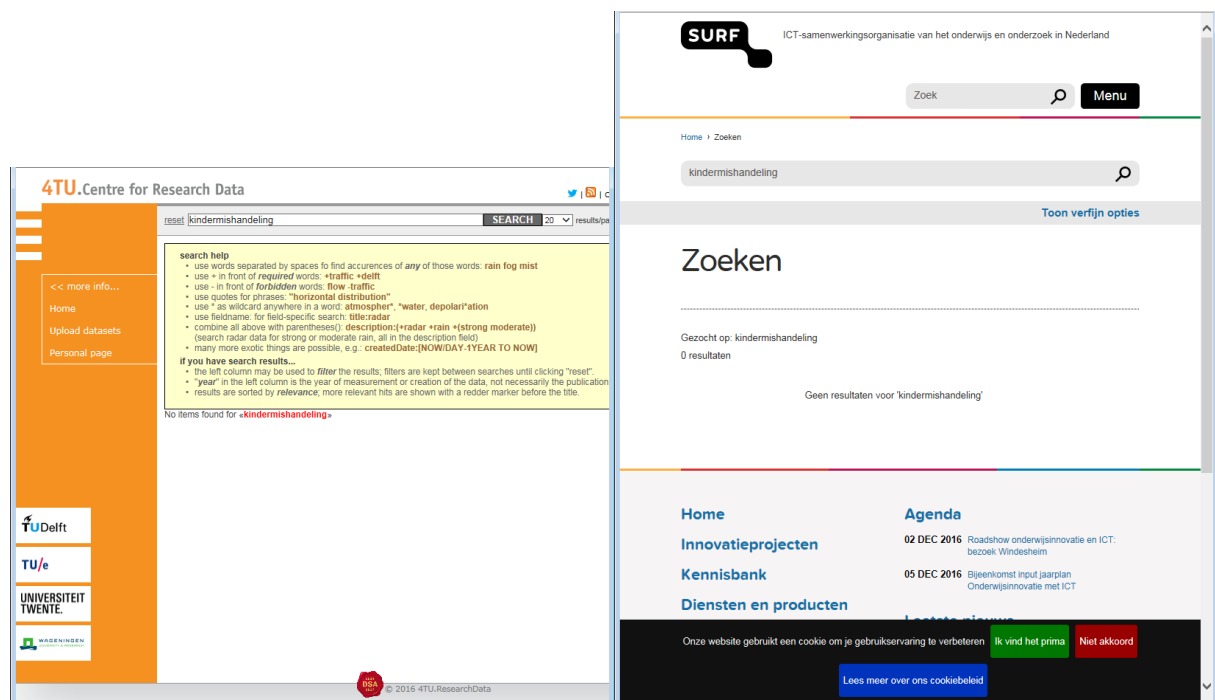
Hoewel de oorspronkelijke ‘Kinder mishandeling in Nederland anno 2005’ dataset geen persoonsgegevens bevat, heeft er toch een bewerkingslag plaatsgevonden voordat de dataset gepubliceerd werd. Die bewerkingslag betrof een controle van de data om te voorkomen dat de gegevens toch herleidbaar zou kunnen zijn tot natuurlijke personen, om de kwaliteit van de data te controleren en om de metadata aan te vullen.

6.3 Projectalternatieven

Er zijn wel maatschappelijke baten: met wetenschappelijk onderbouwd beleid kunnen de problemen omtrent kindermishandeling effectiever worden aangepakt.

De ‘Kinder mishandeling in Nederland anno 2005’ dataset is slecht vindbaar voor buitenstaanders. Een korte scan met een zoekmachine (met het trefwoord ‘kindermishandeling data’) levert geen link op naar DANS, alleen naar CBS-bestanden¹¹⁵ of data.overheid.nl.¹¹⁶ En een zoekopdracht bij SURF en bij de TU Delft (die doorgelinkt zou moeten zijn naar DANS), levert ook geen resultaten. Zoeken via NARCIS leverde wel een positief resultaat op. Maar aangezien Narcis via DANS loopt, is dat niet verwonderlijk.

Figuur 6.1: Zoekresultaat "kindermishandeling" via 4TU Centre for Research Data en via SURF



Om de vindbaarheid, en daarmee de herbruikbaarheid van de dataset te verbeteren, vergelijken we de volgende alternatieven:

- 0-alternatief: publieke WODC onderzoeksdata zijn open en worden via DANS aangeboden. Dit alternatief scoort 2 sterren op de TBL (Tim Berners-Lee)-index (huidige situatie).

¹¹⁵ Bijv. <https://www.cbs.nl/nl-nl/maatwerk/2016/44/kindermishandeling-en-huishelijk-geweld-1e-halfjaar-2016>

¹¹⁶ Bijv. <https://data.overheid.nl/data/dataset/jeugdzorg-kindermishandeling-jongeren-0-tot-18-jaar-regios-2002-2012/resource/3b6d2069-8cf7-4ae7-9a47-61450b362b09>

- Min 1 alternatief: publieke WODC-onderzoeksdata zijn niet open, maar wel na een verzoek beschikbaar. (passieve openbaarheid)
- Plus 1 alternatief: publieke WODC-onderzoeksdata zijn open volgens 3 sterren op de TBL-index, en worden via meerdere opendataplatforms aangeboden.
- Plus 2 alternatief: publieke WODC-onderzoeksdata zijn 'linked data' volgens 4 sterren op TBL-index, en worden via meerdere opendataplatforms aangeboden.

Het 0-alternatief is het databestand 'Kindermishandeling in NL anno 2005', zoals dat thans wordt aangeboden. Voor de specifieke dataset is dit het meest waarschijnlijke alternatief, omdat de datasets die nu als open data beschikbaar zijn, dat zullen blijven.

Hierbij is het Min 1 alternatief niet realistisch voor deze specifieke dataset. Voor andere datasets op het gebied van kindermishandeling is dit echter wel denkbaar. Kindermishandeling is een zeer privacygevoelig onderwerp en het betreft minderjarigen in een kwetsbare situatie. Indien de data-aanbieder niet zeker is dat deze privacy gewaarborgd is, zal nieuwe data niet open worden aangeboden. De huidige dataset biedt geen naam-adres-woonplaats/provincie informatie en de overige informatie (over geslacht, leeftijd, gezinssamenstelling, herkomst, opleidingsniveau e.d.) is onvoldoende specifiek om deze te kunnen koppelen aan individuele personen. Daar waar deze informatie nog meer specifiek wordt, kan de data niet zonder bewerking worden gepubliceerd.

Het realiseren van het Plus 1 alternatief is de ambitie van de het open databeleid van Rijksoverheid voor high value-datasets. Voor de huidige dataset 'Kindermishandeling in NL anno 2005' betekent dit het converteren van de data naar een open formaat. En de dataset moet – naast publiceren via DANS – ook via andere opendataportalen beschikbaar worden gemaakt, bijvoorbeeld via data.overheid.nl. Deze stap zal een extra inspanning vergen omdat de data geconverteerd moeten worden naar bijvoorbeeld .csv en daarna gecontroleerd moeten worden. Het publiceren van de dataset via andere platforms zal ook een inspanning vergen, net als het afhandelen van vragen van gebruikers. Overigens zou het aanbevelingswaardig zijn om naast een openformaatversie de SPSS-versie te handhaven omdat er ook gebruikers zijn die juist wel prijs stellen op eigendomsmatige formaten.¹¹⁷ Voor andere datasets in hetzelfde domein vergt een dergelijke stap een zorgvuldige opschoning van de microdata om ervoor te zorgen dat de gegevens niet naar specifieke personen te herleiden zijn.

Het Plus 2 alternatief is als technische stap meegenomen naar het aanbieden van 'linked data'. Dit vraagt om een extra investering die het mogelijk maakt externe betekenis toe te voegen aan de data. In dit alternatief is gekozen voor de (relatief) simpele versie van 'linked data'.¹¹⁸ Bij deze versie is geen extra semantiek nodig. Het is mogelijk 'linked data' met extra handwerk verder door te ontwikkelen naar vijfsterrenniveau. In dit alternatief is hier niet voor gekozen, omdat dit minder aansluit bij wat een relevant alternatief kan zijn in deze context. Bij het 'linked' maken van de data is het cruciaal dat dit geen aangrijpingspunten biedt om door de gegevens met andere data te linken toch persoonsgegevens te openbaren. Dit vraagt om zorgvuldigheid.

¹¹⁷ Zie van Loenen, B. & F. Welle Donker (2014). De stand in opendataland. Delft, OTB - Research for the built environment. (<http://repository.tudelft.nl/view/ir/uuid%3A7b425db6-2f9a-4be6-b5f0-3492055268d8/>).

¹¹⁸ In deze versie worden de data omgezet wordt naar RDF (Resource Description Framework) en met URI's (Uniform Resource Identifiers), volgens standaarden van het World Wide Web Consortium (W3C) beschikbaar gesteld. Zie hiervoor ook Bizer, C., R. Cyganiak & T. Heath. (2007). How to Publish Linked Data on the Web (<http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>, bezocht: 29/11/2016).

6.4 Kwalitatieve omschrijving van investeringskosten en effecten

Belangrijke verschillen tussen de alternatieven zitten in de volgende aspecten:

Investeringskosten

Kosten om data open te maken: Het converteren van het dataformaat en de additionele kwaliteitscontrole zullen een eenmalige inspanning vergen omdat de dataset een statisch bestand is. Deze kosten zijn marginaal omdat de werkzaamheden in een relatief korte tijd kunnen worden uitgevoerd.

Kosten om data 'linked' te maken: Het gaat hierbij om kosten die verder gaan dan het structureren van de data conform het huidig opendatabeleid van de overheid omdat er ook moet worden nagedacht over het toepassen van RDF-standaarden. Als dit een onderdeel is van het proces om de data geschikt te maken als open data, zullen de extra kosten van het Plus 2 alternatief in verhouding tot het Plus 1 alternatief marginaal zijn.

Directe effecten

Serverkosten: Meer gebruik (door open data) kan ervoor zorgen dat een hogere capaciteit van computerservers nodig is vergeleken met geen data beschikbaar stellen volgens het Min 1 alternatief. Echter, deze kosten zijn marginaal omdat de Kindermishandeling-dataset een relatief klein en statisch bestand is vergeleken met de vele andere (open) data die al door DANS en andere opendataportalen worden beheerd.

Baten voor onderzoekers: Open onderzoeksdata kunnen voor meer onderzoek worden hergebruikt en onderzoeksgelden kunnen effectiever worden besteed omdat er minder tijd gemoeid is met het vinden van de data (lagere transactiekosten).

Consumentensurplus: Gratis beschikbaar stellen van data kan leiden tot een consumentensurplus. Dit geldt met name bij extra gebruik. Gezien de aard van de data (wetenschappelijk onderzoek) zal het consumentensurplus pas op de langere termijn toenemen, als de indirecte effecten vruchten gaan afwerpen.

Indirecte effecten

Bijdrage aan wetenschappelijke kennis: Beter vindbare open onderzoeksdata kunnen effectiever worden hergebruikt.

Betrouwbaarheid: In algemene zin geldt bij open data dat betrouwbaarheid groter kan zijn omdat fouten er door terugmeldingen uit kunnen worden gehaald. Bij geanonimiseerde open data is de kans daarop kleiner.

Nieuwe producten en bedrijven: Beter vindbare onderzoeksresultaten bieden het bedrijfsleven kansen om toepassingen en/of diensten te ontwikkelen op basis van die resultaten.

Overige effecten

Ontwikkelen van beleid: Wetenschappelijk onderzoek kan aantonen wat de verwachte lange termijn ontwikkelingen kunnen zijn op het gebied van kindermishandeling en daar effectief beleid voor ontwikkelen.

Bijdragen aan maatschappelijke oplossingen: Wetenschappelijk onderbouwd beleid, zal beter in staat zijn maatschappelijke problemen aan te pakken, en tijdig bij te sturen.

Tabel 6.1: Kwalitatief overzicht van verschillen effecten (ordinaal)

Kindermishandeling in NL anno 2005	0 (open data 2-sterren via DANS)	-1 (alleen op verzoek beschikbaar)	+1 (open data 3-sterren, via meerdere platforms)	+2 ('linked data' 4-sterren, via meerdere platforms)
Investeringskosten				
-Kosten om data open te maken			-	
-Kosten om data 'linked' te maken				-
Directe effecten				
-Serverkosten		0	0	0
-Baten voor onderzoekers		-	+	++
-Consumentensurplus		-	+	++
Indirecte effecten				
-Wetenschappelijke kennis		-	+	++
-Betrouwbaarheid		-	+	++
-Nieuwe producten en bedrijven		-	+	++
Overige effecten				
-Effectief beleidsontwikkeling		-	+	++
-Maatschappelijke oplossingen		-	+	+
Totaal		-	+	+

+ = meer baten

- = meer kosten

Tabel 6.1 geeft een overzicht van de verschillen tussen de alternatieven. Dit is per effect ordinaal gebeurd. De serverkosten in het Min 1 alternatief zijn bijvoorbeeld lager dan in het 0-alternatief. In de Plus 1 en Plus 2 alternatieven zijn de serverkosten ongeveer gelijk aan die van het 0-alternatief. Dit zegt niets over de omvang van deze kosten, noch over de verschillen tussen deze stappen. Toch kunnen op grond van de tabel (in vergelijking met het 0-alternatief) de volgende conclusies worden getroffen.

Min 1 alternatief: alleen op verzoek beschikbaar

- *Investeringskosten open data.* Dit alternatief vraagt geen investeringskosten.
- *Directe effecten.* Door de data alleen op verzoek beschikbaar te stellen aan onderzoekers zal de dataset beperkt worden gebruikt. Dit leidt slechts tot marginale besparingen op de serverkosten voor de verwerking van de data en de administratieve lasten. Als onderzoeksdata slecht beschikbaar zijn, ontstaat de kans dat gegevens dubbel worden ingewonnen. Kindermishandeling is een onderwerp van zeer groot maatschappelijk belang. Ook niet-onderzoekers, zoals journalisten, berichten hierover en het is van maatschappelijk belang dat deze kunnen beschikken over de juiste gegevens. Doordat anders dan bij open data gegevens niet zorgvuldig gecheckt worden op privacygevoeligheid en gegevens mogelijk te herleiden blijven op personen (omdat de data alleen beperkte kring onderzoekers wordt verstrekt) worden persoonsgegevens ter beschikking gesteld. Dit vraagt om het maken van kosten om zorgvuldig gebruik van deze microdata te waarborgen. Door het niet beschikbaar zijn van de data verdampt een deel van het consumentensurplus, en zal dit niet volledig worden gecompenseerd door het gestegen producentensurplus.
- *Indirecte effecten.* Doordat er minder en/of kwalitatief minder goede dat beschikbaar zijn in het maatschappelijk debat, zal dit een impact hebben op de bijdrage aan de maatschappelijke

discussie. De wetenschappelijke output die op grond van deze dat wordt gerealiseerd zal minder transparant zijn, en niet controleerbaar voor buitenstaanders die geen toegang hebben tot deze data. De spillovereffecten naar het bedrijfsleven zullen lager zijn, waardoor er minder innovatieve toepassingen zullen worden gecreëerd.

- *Overige effecten.* Het beleid van de overheid zal een minder gedegen wetenschappelijke onderbouwing hebben omdat er dan nog maar één onderzoeksbron (WODC) zal zijn, er geen vergelijkingen kunnen worden gemaakt. Daardoor zullen de maatschappelijk baten afnemen.

Plus 1 alternatief: 3-sterren opendataversie via meerdere platforms

- *Investeringskosten.* Er zullen eenmalige kosten moeten worden gemaakt om de dataset aan te passen voor een 3-sterrenversie. Voor het huidige bestand Kindermishandeling in NL anno 2005 zijn deze kosten marginaal. Voor andere bestanden vraagt dit om een zorgvuldige analyse (en zo nodig aanpassing) om er zeker van te zijn dat gegevens niet kunnen worden herleid tot personen.
- *Directe effecten.* Door het beschikbaar stellen van een opendataversie van de dataset via meerdere platforms, zullen de data beter vindbaar zijn, en zullen de data meer worden hergebruikt. Het intensiever gebruik heeft een marginaal effect op de serverkosten omdat de dataset relatief klein is ten opzichte van alle andere datasets die al gehost worden. De administratieve lasten zullen marginaal zijn. De transactiekosten van de onderzoekers zullen lager zijn omdat de data beter vindbaar zullen zijn. Onderzoeksfondsen kunnen efficiënter worden gebruikt. De effecten op het consumentensurplus zullen waarschijnlijk ook marginaal zijn omdat de data voornamelijk voor onderzoek worden gebruikt, en de spillovereffecten voor het bedrijfsleven marginaal zullen zijn. Omdat de kosten 2005 betreffen (relatief lang geleden) zullen deze effecten beperkt zijn. Veel effecten zullen pas optreden als meerdere relevante databestanden als open data beschikbaar zijn.
- *Indirecte effecten.* Er kan een positief effect uitgaan op de maatschappelijke discussie over kindermishandeling indien data hierover open beschikbaar zijn. Gegeven de periode van deze data (2005) zal dit effect vooral optreden als ook recentere data beschikbaar zijn, zodat vergelijkingen kunnen worden getrokken.
- *Overige effecten.* Overheidsbeleid voor de lange termijn zal een betere wetenschappelijke onderbouwing hebben omdat er meer en multidisciplinaire onderzoeksoutput zal komen.

Plus 2 alternatief: 4-sterren open data 'linked data' via meerdere platforms

- *Investeringskosten.* Er zullen extra kosten moeten worden gemaakt om de data 'linked' te maken, maar die zullen marginaal zijn (in vergelijking tot het plus 1 alternatief) als WODC ervoor kiest om dit tegelijk te doen met het converteren van de data naar een open formaat. Dit gelijktijdig doen is verstandig om er zeker van te zijn dat door het linken niet mogelijk wordt om de gegevens te kunnen herleiden tot personen.
- *Directe effecten.* De directe effecten van het Plus 2 alternatief zijn vergelijkbaar met het Plus 1 alternatief. Door het linken van de data wint de dataset aan relevantie. Het extra gebruik zal marginaal zijn. Het consumentensurplus zal waarschijnlijk iets stijgen door extra gebruik.
- *Indirecte effecten.* Deze zijn vergelijkbaar met het Plus 1 alternatief, de mogelijkheden worden op dit gebied versterkt.
- *Overige effecten.* Deze zijn vergelijkbaar met het Plus 1 alternatief. Het linken van data zal mogelijk resulteren in verdere verbetering van betrouwbaarheid van de onderbouwing van beleid.

6.5 Monetarisering

Bij het monetariseren van de kosten en baten van de 'Kindermishandeling in NL anno 2005' dataset moet worden opgemerkt dat de dataset voornamelijk is gebruikt door een klein aantal onderzoekers

(in 2016 slechts 2). Wetenschap heeft een waarde, zowel direct meetbaar in de vorm van een bijdrage aan (de groei van) het bruto binnenlands product (bbp), als breder economisch en maatschappelijk. De waarde van wetenschap betreft echter veel aspecten die moeilijk te kwantificeren zijn, vooral op de korte termijn. Niet alle waarde van wetenschap doet zich voor als een deel van het (toekomstig) bbp en slechts een deel ervan zal in de productiefunctie terug te vinden zijn. Het meten van de totale economische en maatschappelijke waarde van wetenschap is veel moeilijker.¹¹⁹ De kosten van de alternatieven kunnen nog wel bepaald worden, maar de baten zijn niet te monetariseren.

Investeringskosten

Kosten om data open te maken. Gebaseerd op de kosten die in 2013 zijn gemaakt om de data geschikt te maken voor publicatie via DANS, zullen die kosten eenmalig ca. €500 zijn (4 uur werk + 1 uur controle door een derde).

Kosten om data 'linked' te maken. Deze kosten zijn afhankelijk van in welke fase de data gelinkt zullen worden. Als dit in het conversieproces zou gebeuren, zouden deze kosten marginaal zijn. Als dit in een latere fase zou gebeuren, dan zouden de kosten ca. €400 zijn.

Directe effecten

- *Serverkosten.* Afhankelijk van de keuzes die gemaakt worden zullen de extra serverkosten voor het Plus 1 en Plus 2 alternatief marginaal zijn ten opzichte van het 0-alternatief.
- *Baten voor onderzoekers.* Deze kosten zijn afhankelijk van de aantallen onderzoekers, het soort onderzoek en type onderzoeker. De bespaarde transactiekosten zijn ca. €75 (1 uur junior onderzoeker) per onderzoeker.
- *Consumentensurplus.* Het consumentensurplus zal afhankelijk zijn van de indirecte effecten.

Indirecte effecten

- *Bijdrage aan wetenschappelijke kennis.* De baten van het hergebruiken van open onderzoeksdata en het creëren van mogelijkheden voor nieuwe types (multidisciplinair) onderzoek in andere gebieden zijn moeilijk te kwantificeren omdat de waarde van wetenschappelijk onderzoek van vele factoren afhankelijk is.
- *Betrouwbaarheid.* Als de dataset als 3-sterren open data beschikbaar en beter vindbaar is, zal de transparantie van wetenschappelijke resultaten toenemen omdat de kwaliteit van de onderzoeksoutput door buitenstaanders kan worden gecontroleerd. Daarmee neemt de betrouwbaarheid van de resultaten toe. Ook hier geldt dat die effecten moeilijk te kwantificeren zijn. Wel kan hier ook gedacht worden aan vermeden reputatieschade door onderzoeksfraude.
- *Nieuwe producten en bedrijven.* Er is tot nu toe geen bewijs van nieuwe producten op basis van de Kindermishandeling-dataset.

Overige effecten

- *Ontwikkelen van beleid.* Wetenschappelijk onderzoek kan aantonen wat de verwachte langetermijntontwikkelingen kunnen zijn, en deze laatste kunnen van groot belang zijn voor het ontwikkelen of aanpassen van beleid. Voor dit onderzoek zijn deze effecten niet te kwantificeren.

¹¹⁹ KNAW [Koninklijke Nederlandse Akademie van Wetenschappen] (2013). Publieke Kennisinvesteringen en de waarde van wetenschap. Amsterdam, Koninklijke Nederlandse Akademie van Wetenschappen: 68, p.59.

- *Bijdragen aan maatschappelijke oplossingen.* Door effectiever beleid zal de maatschappij daarvan profiteren, maar zijn voor dit onderzoek niet te kwantificeren.

Tabel 6.2: Gemonetariseerd overzicht van verschillen effecten in euro's

Kindermishandeling in NL anno 2005	0 (open data 2-sterren via DANS)	-1 (alleen op verzoek beschikbaar)	+1 (open data 3-sterren, via meerdere platforms)	+2 ('linked data' 4-sterren, via meerdere platforms)
Investeringskosten				
-Kosten om data open te maken			500	500
-Kosten om data 'linked' te maken				0-400
Directe effecten				
-Serverkosten		marginaal	marginaal	marginaal
-Baten voor onderzoekers		pm	pm	pm
-Consumentensurplus*		pm	pm	pm
Indirecte effecten				
-Wetenschappelijke kennis		pm	pm	pm
-Betrouwbaarheid		pm	pm	pm
-Nieuwe producten en bedrijven		pm	pm	pm
Overige effecten				
-Maatschappelijke oplossingen		pm	pm	pm
-Beleidsontwikkeling		pm	pm	pm
Totaal		pm	500+ pm	500-900+pm

* Zie indirecte kosten.

6.6 Conclusies

WODC heeft een aantal onderzoekdatasets, waaronder de 'Kindermishandeling in Nederland anno 2005' beschikbaar gesteld als open data via DANS, een platform voor wetenschappelijke data, voornamelijk in de geesteswetenschappen. WODC was een van de eerste organisaties in Nederland die data als open data beschikbaar stelde, en ondervindt nu de gevolgen van de remmende voorsprong. De dataset voldoet niet aan de principes voor open data vastgesteld in de NODA. Het grootste probleem is echter niet dat de dataset niet beschikbaar is in een open machine-verwerkbaar formaat, maar dat de dataset moeilijk te vinden is voor buitenstaanders die niet op de hoogte zijn van het bestaan dat deze specifieke dataset. En daarmee niet aan de motieven voor het beschikbaar stellen als open data voldoet, namelijk het faciliteren van hergebruik. Het zal zeker de moeite waard zijn om de dataset als 'linked data' beschikbaar te stellen via meerdere opendataplatforms. Met een relatief minimale inspanning zal de dataset aan kwaliteit en vindbaarheid winnen. Het zal efficiënter zijn om de dataset beschikbaar te stellen als 'linked open data' volgens het 4-sterren niveau op de TBL-index omdat de meerkosten van het linken van data marginaal zullen zijn ten opzichte van het geschikt maken volgens het 3-sterren niveau op de TBL-index. De dataset zal als 'linked data' meer relevantie hebben, en makkelijker te koppelen zijn aan andere data. Het is aanbevelingswaardig om de eigendomsmatige versie van de dataset naast een open-formaatversie te handhaven om te voldoen aan de wensen van gebruikers die een voorkeur hebben voor eigendomsmatige software.

7 Polisbus: Banen en lonen van werknemers in Nederland

7.1 Inleiding

In de Polisbus-dataset van CBS zijn kwantitatieve en kwalitatieve gegevens opgenomen over banen en lonen van werknemers bij Nederlandse bedrijven in een bepaalde verslagmaand. De gegevens zijn afkomstig uit de Polisadministratie van UWV, gebaseerd op de loonaangiften bij de Belastingdienst en omvatten onder meer alle banen van werknemers in Nederland, maar ook uitkeringen en pensioenen. Het UWV levert de Polisbus-dataset in ruwe vorm aan CBS.¹²⁰ Polisbus bevat gegevens op microniveau (koppelbare gegevens op persoons-, bedrijfs- en adresniveau). Polisbus bevat dus niet alleen informatie die herleidbaar zou kunnen zijn tot natuurlijke personen, maar ook bedrijfsgevoelige informatie. Daarom is de Polisbus-dataset van CBS niet openbaar. Alleen gemachtigde gebruikers, voornamelijk onderzoekers, kunnen – onder strikte voorwaarden – via een beveiligde omgeving toegang krijgen tot een geanonimiseerde versie van Polisbus. Polisbus is één van de meest gewilde microdatabestanden van CBS.

Open data van CBS

CBS hanteert een opendatabeleid en heeft sinds 18 juli 2014 zijn volledige databank met statistische gegevens als open data beschikbaar gesteld. De reden hiervoor was om statistische informatie beter bereikbaar en makkelijker (zonder beperkingen die verbonden zijn aan het rechtstreeks toegang hebben tot registerdata) herbruikbaar te maken, waardoor niet alleen wetenschappers maar ook het bedrijfsleven tijd en geld besparen en makkelijker CBS-data kunnen gebruiken om nieuwe toepassingen te maken.¹²¹ Echter, de Polisbusgegevens zijn (nog) geen open data. Ten eerste is CBS niet de data-eigenaar en kan alleen UWV die beslissing nemen. Ten tweede is, door de gevoelige aard van de Polisbusgegevens, dit voor de huidige versie van het bestand niet mogelijk. Om Polisbus tóch in een opendataversie beschikbaar te maken zullen de ruwe Polisbusgegevens eerst geaggregeerd moeten worden en daar is een extra inspanning voor nodig.

7.2 Huidige situatie: Polisbus alleen voor gemachtigde gebruikers

De ruwe Polisbusgegevens verstrekt door UWV worden door CBS bewerkt en geharmoniseerd om voor eigen gebruik geschikt te zijn. Daarbij worden direct identificeerbare variabelen (naam, adres, BSN) vervangen door een zogenaamde RIN (Record Identification Number). Dit bewerkte bestand wordt beschikbaar gesteld aan externe onderzoekers. Echter, de data in de bewerkte versie zijn nog wel koppelbaar – op persoonsniveau – met andere registraties en registers.

CBS bepaalt aan de hand van een aantal voorwaarden of een gebruiker gemachtigd kan worden om een bepaalde microdataset te gebruiken. Er zijn op dit moment ca. 400 onderzoekers die de Polisbus-dataset van CBS gebruiken.

De microdata van CBS zijn alleen beschikbaar via een beveiligde omgeving. Gemachtigde gebruikers kunnen naar de CBS-vestiging in Den Haag of Heerlen gaan, of kunnen toegang krijgen via een

¹²⁰ Wet van 20 november 2003, houdende vaststelling van een wet op het Centraal bureau voor de statistiek (Wet op het Centraal bureau voor de statistiek), art.33.

¹²¹ <https://www.cbs.nl/nl-nl/nieuws/2014/29/cbs-lanceert-open-data-portal-volledige-databank-nu-beschikbaar>

zogenaamde remote-accessverbinding (RA-verbinding). Binnen de remote-accessomgeving kan een onderzoeker eigen bestanden uploaden en koppelen aan de microdata. De onderzoeker krijgt een afgeschermd omgeving ter beschikking om te werken. De microdata zelf blijven binnen de beveiligde omgeving van CBS en kunnen niet worden gedownload. Voor het opstarten van een nieuw onderzoek geldt dat er opstartkosten in rekening worden gebracht, bijvoorbeeld voor een intake-overleg, het opstellen van een contract en geheimhoudingsverklaring, het inrichten van een projectomgeving en het maken van een vingerafdrukpasje voor toegang tot de remote-accessomgeving. Verder zijn er kosten tijdens de duur van het onderzoek, zoals voor een verplicht serviceabonnement of extra opslagcapaciteit. Ook zijn er kosten na afloop van een onderzoeksproject, zoals voor opslag (archivering). Als een onderzoeker (tussen)resultaten buiten de beveiligde omgeving wil brengen, dan exporteert CBS die resultaten na een controlecheck om te zien dat de tussenbestanden geen tot personen of bedrijven herleidbare gegevens bevatten. Voor die controles worden ook kosten in rekening gebracht. Het is verplicht voor onderzoekers om publicaties gebaseerd op CBS-microdata te voorzien van een bronvermelding die verwijst naar het feit dat er gebruik is gemaakt van deze data. Voor grotere onderzoeksprojecten met meerdere onderzoekers kunnen de kosten in de tienduizenden euro's lopen.¹²²

7.3 Projectalternatieven Polisbusdata

Vanwege de gevoelige aard van de Polisbus-dataset is de onderzoeksversie niet openbaar en dus niet beschikbaar voor als open data. In het kader van dit onderzoek beschouwen we de theoretische mogelijkheid van een opendataversie van de Polisbus-dataset. De dataset zal dan eerst bewerkt moeten worden door de data te aggregeren. De ruwe dataset bestaat uit circa 20 miljoen records op individueel niveau (persoon/adres). De dataset kunnen 'ingedikt' worden naar bijvoorbeeld grotere geografische eenheden (bijvoorbeeld wijken of buurten) maar ook naar inkomensklassen. Er moet goed worden nagedacht over welke definities worden gebruikt: wordt de dataset te veel ingedikt, dan verliest de dataset betekenis en wordt deze minder waardevol voor gebruikers. Maar als de dataset te weinig wordt ingedikt, dan zouden de gegevens herleidbaar kunnen zijn tot personen, zeker als de data gecombineerd worden met andere (open) datasets.

Statistische datasets hebben waarde voor hergebruikers. Er is daarbij een hogere potentiële waarde voor statistische datasets met een hoge updatefrequentie, die daardoor actueler zijn, datasets die gedetailleerd zijn en datasets die relevant zijn voor economische of financiële kwesties. Echter, dergelijke hoogwaardige statistische datasets kosten ook meer om in te winnen en te beheren.¹²³ Volgens de Bank of England zullen de meeste economische statistische data gebruikt worden door beleidsmakers, onderzoekers, analisten en journalisten. Die gebruikers zullen verschillende behoeften hebben wat betreft de data: de een zal behoefte hebben aan actuele data om huidige trends snel te kunnen detecteren, een ander zal juist behoefte hebben aan tijdseries om langetermijnrelaties te bepalen.

Echter, de Polisbus-dataset wordt voornamelijk gebruikt door wetenschappelijke onderzoekers en is niet beschikbaar voor andere doelgroepen. Beleidsmakers kunnen gebruik maken van Polisbusgegevens via maatwerkproducten geproduceerd door CBS. Maar andere doelgroepen hebben nu nog geen toegang tot de Polisbus-dataset van CBS.

¹²² Gebaseerd op gebruik van CBS-microdata via Remote Access op de Technische Universiteit Delft, waarbij opgemerkt moet worden dat de TU Delft een atypische gebruiker is vanwege de omvang van de meerjarige onderzoeksprojecten, het aantal onderzoekers verbonden aan de projecten en de gedegen statistische voorkennis van die onderzoekers.

¹²³ Bank of England (2006). Cost-benefit analysis of monetary and financial statistics. A practical guide, p.10.

Er is een groeiende behoefte onder wetenschappers om onderzoeksresultaten en de onderliggende onderzoeksgegevens openbaar te maken. Volgens de Adviesraad voor Wetenschap, Technologie en Innovatie (AWTI) en de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) creëert hergebruik van open onderzoeksdata mogelijkheden voor nieuwe types (multidisciplinair) onderzoek in andere gebieden. Bovendien zullen door verbeterde transparantie van wetenschappelijke resultaten de kwaliteit en de betrouwbaarheid van het onderzoek toenemen omdat de resultaten door buitenstaanders kunnen worden gecontroleerd. Daarnaast kan de efficiëntie van wetenschappelijk onderzoek omhoog gaan omdat onnodig onderzoek of dubbel inwinnen van gegevens kan worden voorkomen. Ook de maatschappij en het bedrijfsleven kunnen baat hebben bij toegang tot onderzoeksdata, hoewel dat wel sterk zal afhangen van de aard van de data.¹²⁴ Onderzoekdata kunnen het bedrijfsleven kansen bieden om innovatieve producten te ontwikkelen.¹²⁵ Wetenschappelijke publicaties zijn van belang voor het delen van kennis onder onderzoekers, maar het bedrijfsleven heeft waarschijnlijk meer aan de onderliggende data voor het daadwerkelijk ontwikkelen van innovatieve producten en diensten. Er zijn ook maatschappelijke baten: doordat onderzoeksresultaten gedeeld worden, zullen de afgeleide producten en toepassingen die voor de maatschappij nuttig zijn, bijvoorbeeld medicijnen of nieuwe technologieën, eerder beschikbaar komen.¹²⁶ De verwachte baten zijn echter van meerdere factoren afhankelijk en het kan soms jaren duren voordat ze zichtbaar worden. Op grond van deze meer algemene ervaringen en overwegingen kan ook verwacht worden dat er positieve maatschappelijke effecten zijn van het open aanbieden van data gebaseerd op Polisbus-gegevens.

Gegeven de beleidscontext zullen de volgende alternatieven met elkaar worden vergeleken:

- 0-alternatief: de Polisbus-dataset is niet openbaar en alleen beschikbaar voor gemachtigde gebruikers onder strikte voorwaarden en tegen een vergoeding. Dit alternatief scoort 0 sterren op de TBL (Tim Berners-Lee)-index.
- Min 1 alternatief: de Polisbus-dataset is niet openbaar en ook niet beschikbaar voor onderzoekers. Ook dit alternatief scoort 0 sterren op de TBL-index.
- Plus 1 alternatief: de Polisbus-dataset wordt, in aanvulling op het 0-alternatief, ook in een geaggregeerde vorm beschikbaar gesteld als open data volgens 3 sterren op de TBL-index.
- Plus 2 alternatief: de Polisbus-dataset wordt, in aanvulling op het 0-alternatief, ook in een geaggregeerde vorm beschikbaar gesteld als 'linked open data' volgens 4 sterren op de TBL-index.

Het 0-alternatief is het huidige beleid en zal in de nabije toekomst het meest waarschijnlijk zijn. CBS heeft geen plannen om de Polisbus-dataset in een opendataversie beschikbaar te stellen.

Het Min 1 alternatief is niet realistisch, de Polisbus-dataset is een gewilde dataset onder onderzoekers en er zijn geen signalen dat het huidig beleid zal worden bijgesteld.

Het Plus 1 alternatief is als technische stap meegenomen om de dataset te converteren naar een versie geschikt als open data, conform het opendatabeleid van de overheid. Voor dit alternatief zullen keuzes moeten worden gemaakt rond de benodigde definities voor het aggregeren van de gegevens, om de mogelijkheid van herkenning of herleidbaarheid tot bijvoorbeeld individuen te

¹²⁴ AWTI [Adviesraad voor Wetenschap, Technologie en Innovatie] (2016). Durven Delen. Op weg naar een toegankelijke wetenschap. Den Haag, p.12, p.112-113.

¹²⁵ KNAW [Koninklijke Nederlandse Akademie van Wetenschappen] (2013). Publieke Kennisinvesteringen en de waarde van wetenschap. Amsterdam, Koninklijke Nederlandse Akademie van Wetenschappen, p.40.

¹²⁶ AWTI [Adviesraad voor Wetenschap, Technologie en Innovatie] (2016). Durven Delen. Op weg naar een toegankelijke wetenschap. Den Haag, p. 103.

voorkomen. Verder moeten er rekenmodellen worden ontwikkeld en toegepast en moeten er handleidingen worden geschreven voor de gebruikers.

Het realiseren van het Plus 2 alternatief vergt een extra investering die het mogelijk maakt externe betekenis toe te voegen aan de data. In dit alternatief is gekozen voor de (relatief) simpele versie van 'linked data' en is er bij deze versie geen extra semantiek nodig.¹²⁷ Hoewel dit alternatief een extra inspanning zal vergen ten opzichte van het Plus 1 alternatief, is het de moeite waard om dit te onderzoeken aangezien er voor het Plus 1 alternatief al ook nagedacht moet worden over definities en rekenmodellen. De extra stap voor het Plus 2 alternatief zou in deze fase waarschijnlijk minder groot zijn dan wanneer de dataset reeds beschikbaar zou zijn als open data.

7.4 Kwalitatieve omschrijving van investeringskosten en effecten

De belangrijke kwalitatieve verschillen in de effecten tussen de alternatieven zitten in de volgende aspecten:

Investeringskosten

Kosten om data open te maken: Het gaat hierbij om de kosten die nodig zijn om een opendataversie van de Polisbus-dataset te maken volgens het Plus 1 alternatief. De meeste kosten zullen zitten in het opzetten, testen en toepassen van definities en rekenmodellen. De initiële kosten worden geschat op circa €50.000. Daarnaast zullen er extra kosten zijn voor ondersteuning van opendatagebruikers. Naast de eenmalige kosten zijn er ook doorlopende kosten omdat voor elke nieuwe versie van de dataset (dataset wordt eens per maand geüpdatet), een opendataversie opnieuw zal moeten worden gemaakt.

Kosten om data 'linked' te maken: Het gaat hierbij om kosten die verder gaan dan het structureren van de data conform het huidig opendatabeleid van de Rijksoverheid omdat er ook moet worden nagedacht over het toepassen van RDF-standaarden. Als dit een onderdeel is van het proces om de data geschikt te maken als open data, zullen de extra kosten van het Plus 2 alternatief in verhouding tot het Plus 1 alternatief marginaal zijn.

Directe effecten

Serverkosten: Meer gebruik door ('linked') open data (Plus 1 en Plus 2 alternatieven vergeleken met de andere twee alternatieven) kan ervoor zorgen dat een hogere capaciteit van computerservers nodig is vergeleken met geen data beschikbaar stellen volgens het Min 1 alternatief. CBS heeft al een serverpark in gebruik voor de open data. De extra serverkosten ten gevolge van een opendataversie van de Polisbus-dataset zullen waarschijnlijk marginaal zijn in verhouding tot de totale serverkosten.

Administratieve lasten: In theorie zouden minder financiële transacties moeten leiden tot besparingen op administratieve lasten (Min 1, Plus1 en Plus 2 alternatieven vergeleken met 0-alternatief). Het verwerken van betalingen en de handhaving van betalingen in het 0-alternatief kosten geld. In het Plus1 en Plus 2 alternatief blijven deze kosten ongewijzigd. Alleen in het Min 1 alternatief zullen deze verdwijnen. Deze kosten zijn echter marginaal ten opzichte van de totale beheerkosten van microdata.

Baten voor onderzoekers: Open statistische data kunnen voor meer onderzoek worden hergebruikt en onderzoeksgelden kunnen efficiënter worden besteed.

¹²⁷ In deze versie worden de data omgezet naar RDF (Resource Description Framework) en met URI's (Uniform Resource Identifiers), volgens standaarden van het World Wide Web Consortium (W3C) beschikbaar gesteld. Zie hiervoor ook Bizer, C., R. Cyganiak & T. Heath (2007), How to Publish Linked Data on the Web (<http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>, bezocht: 29/11/2016) of Stappenplan Linked (Open) Data voor Archieven (<https://informatie2020.pleio.nl/file/download/45253222>, bezocht 29/11/2016).

Consumentensurplus: Het gratis beschikbaar stellen van de Polisbus-dataset kan leiden tot een consumentensurplus. Dit geldt met name waar het extra gebruikt wordt. De Polisbus-dataset is één van de meest gewilde microdatasets van CBS. Gezien die gewilde status zal het consumentensurplus toenemen, als het bedrijfsleven de data zullen gaan hergebruiken en de indirecte effecten vruchten gaan afwerpen.

Indirecte effecten

Bijdrage aan kennis met wetenschappelijke en maatschappelijke relevantie: Open data op het gebied van inkomens en banen hebben een hoge maatschappelijke en wetenschappelijke relevantie. Deze data kunnen worden hergebruikt en creëren mogelijkheden voor het ontwikkelen van kennis met zowel maatschappelijke als wetenschappelijke relevantie.

Betrouwbaarheid: Een opendataversie van de Polisbus-dataset maakt dat de toegang tot de data laagdrempeliger wordt voor onderzoekers die gebruik kunnen maken van deze geaggregeerde data alsmede voor beleidsmakers en analisten die nu geen toegang hebben. Voor het bedrijfsleven kan het beschikbaar zijn van betrouwbare kennis zorgen voor meer gerichte (en daardoor meer efficiënte) marketing van bestaande producten.

Nieuwe producten en bedrijven: De onderzoeksresultaten bieden het bedrijfsleven die nu geen gemachtigde gebruiker zijn, kansen om de opendataversie van de Polisbus-dataset zelf te gebruiken voor bijvoorbeeld eigen R&D en marketing.

Tabel 7.1: Kwalitatief overzicht van verschillen effecten (ordinaal)

Polisbus-dataset van CBS	0 (alleen voor gemachtigde gebruikers beschikbaar als Remote Access-versie)	-1 (niet beschikbaar)	+1 (geaggregeerde versie als open data 3-sterren naast Remote Access-versie)	+2 (geaggregeerde versie als 'linked data' 4-sterren naast Remote Access-versie)
Investeringskosten -Kosten om data open te maken -Kosten om data 'linked' te maken			-	-
Directe effecten -Serverkosten -Administratieve lasten -Baten voor onderzoekers -Consumentensurplus		+ + - -	0 - + +	0 - + +
Indirecte effecten -Wetenschappelijke kennis -Betrouwbaarheid -Nieuwe producten en bedrijven		- - -	+ + +	+ + +
Totaal		-	+	+

+ = meer baten

- = meer kosten

Tabel 7.1 geeft een overzicht van de hypothetische verschillen tussen de alternatieven. Dit is per effect ordinaal gebeurd. De serverkosten in het Min 1 alternatief zijn bijvoorbeeld lager dan in het 0-alternatief. In de Plus 1 en Plus 2 alternatieven zijn de serverkosten ongeveer gelijk aan die in het 0-alternatief. Dit zegt niets over de omvang van deze kosten, noch over de verschillen tussen deze

stappen. Toch kunnen op grond van de tabel (in vergelijking met het 0-alternatief) de volgende conclusies worden getroffen. We moeten hierbij opmerken dat dit hypothetische alternatieven zijn: CBS heeft geen plannen om een (linked) opendataversie van de Polisbus-dataset te produceren.

Min 1 alternatief: Polisbus-dataset niet meer beschikbaar, ook niet via remote access

- *Investeringskosten.* Geen.
- *Directe effecten.* Door de Polisbus-dataset niet meer beschikbaar te stellen aan onderzoekers zal de dataset niet meer gebruikt worden buiten CBS. Dit leidt tot slechts marginale besparingen op de serverkosten voor de verwerking van de data. De administratieve lasten zullen dalen. Onderzoekers zullen de data zelf moeten inwinnen, waardoor hun projectkosten hoger zullen zijn en de data incompleet (geen registerdata). Dat zal weer een impact hebben op de kwaliteit van de onderzoeksresultaten. Door het niet beschikbaar zijn van de data verdampt een deel van het consumentensurplus wat niet volledig zal worden gecompenseerd door het gestegen producentensurplus.
- *Indirecte effecten.* Doordat er minder en/of kwalitatief minder goede onderzoeksresultaten zijn, zal dit een impact hebben op de bijdrage aan de wetenschappelijke kennis. De output zal minder transparant zijn en niet controleerbaar voor buitenstaanders. De spillovereffecten naar het bedrijfsleven zullen lager zijn, waardoor er minder innovatieve toepassingen zullen worden gecreëerd.
- *Overige effecten.* Het bedrijfsleven zal minder toepassingen ontwikkelen die gebaseerd zijn op de onderzoeksoutput, waardoor de maatschappelijke baten lager zullen zijn. De overheid zal voor onderbouwing van beleid in beperktere mate gebruik kunnen maken van wetenschappelijke onderzoeksresultaten.

Plus 1 alternatief: Polisbus-opendataversie naast de versie voor gemachtigde gebruikers

- *Investeringskosten.* Er zullen structureel kosten moeten worden gemaakt om de opendataversie van de Polisbus-dataset te maken.
- *Directe effecten.* Door het beschikbaar stellen van een opendataversie van de Polisbus-dataset zullen de open data meer worden hergebruikt. Het intensiever gebruik zal waarschijnlijk een marginaal effect hebben op de serverkosten omdat CBS al een serverpark heeft ingericht voor het hosten van open data. De administratieve lasten voor het verwerken van betalingen, zullen naar verwachting marginaal dalen omdat een aantal van de huidige onderzoekers voldoende zullen hebben aan de open data, maar niet verdwijnen. De microdata blijven in de remote access-omgeving beschikbaar en blijven aanzienlijke meerwaarde bieden voor onderzoek (denk aan koppeling met andere bestanden). Naast de Polisbus-dataset zijn er nog andere microdatabestanden die alleen via een remote access-omgeving beschikbaar blijven. De transactiekosten van de huidige onderzoekers die de opendataversie gaan gebruiken zullen lager zijn omdat de kosten en beperkingen van de remote access-omgeving wegvallen. Onderzoeksfondsen kunnen efficiënter worden gebruikt indien de opendataversie kan worden gebruikt in plaats van de huidige microdata versie. Maar daar staat tegenover dat de datakwaliteit van de opendataversie minder zal zijn dan van de onderzoeksversie omdat de data 'ingedikt' zijn. Het bedrijfsleven kan direct gebruik maken van Polisbus-opendataversie voor eigen R&D, waardoor ze niet (meer) afhankelijk zijn van onderzoeksoutput. Door het extra gebruik van de data zal het consumentensurplus stijgen en groter zijn dan het verlies aan producentensurplus.
- *Indirecte effecten.* Als de Polisbus-dataset meer gebruikt zal worden, ook direct door het bedrijfsleven, zullen de spillovereffecten groter zijn en zullen meer innovatieve toepassingen worden gecreëerd en mogelijk zelfs meer in statistische toepassingen gespecialiseerde bedrijven ontstaan. Indien huidig onderzoek kan volstaan met de open dataversie dan is er een positief effect op de betrouwbaarheid van onderzoeksoutput omdat de onderliggende data voor iedereen controleerbaar zijn.

- *Overige effecten.* De ontwikkelde innovatieve toepassingen zullen sneller beschikbaar komen en daarmee een nuttigere bijdrage leveren aan de maatschappij. Bovendien zal overheidsbeleid voor de lange termijn een betere kwantitatieve onderbouwing hebben omdat er meer en multidisciplinaire onderzoeksoutput zal komen.¹²⁸

Plus 2 alternatief: Polisbus-opendataversie als 'linked open data' naast de versie voor gemachtigde gebruikers

- *Investeringskosten.* Er zullen extra kosten moeten worden gemaakt om de data 'linked' te maken, maar die zullen marginaal extra zijn ten opzichte van het Plus 1 alternatief als CBS er voor kiest om dit tegelijk te doen met de data geschikt maken als open data.
- *Directe effecten.* De directe effecten van het Plus 2 alternatief zijn vergelijkbaar met het Plus 1 alternatief. Door het linken van de data wint de dataset aan relevantie. Extra gebruik zal ertoe leiden dat meer serverkosten moeten worden gemaakt. Als CBS ervoor kiest op alle open data (op termijn) te linken, zullen de serverkosten van het opendata-serverpark toenemen. Maar voor de Polisbus 'linked data' in isolatie zullen de meerkosten marginaal zijn. Het consumentensurplus zal verder stijgen. Door extra gebruik overstijgt dit het verlies aan producentensurplus.
- *Indirecte effecten.* Deze zijn vergelijkbaar met het Plus 1 alternatief, maar door de betere vindbaarheid en relevantie van 'linked' Polisbus-dataset zal dit leiden tot een grotere betrouwbaarheid van onderzoeksoutput en spillovereffecten.
- *Overige effecten.* Deze zijn vergelijkbaar met het Plus 1 alternatief. Het linken van data zal mogelijk resulteren in verdere verbetering van betrouwbaarheid van onderzoeksoutput en snellere beschikbaarheid van innovatieve toepassingen.

7.5 Monetarisering

De Polisbus-dataset wordt (nog) voornamelijk gebruikt voor eigen gebruik, voor maatwerkproducten in opdracht van ministeries en andere overheden en voor wetenschappelijk onderzoek. Echter, het meten van de economische en maatschappelijke waarde van het gebruik door onderzoekers ten behoeve van wetenschap is moeilijker.¹²⁹ De kosten van de alternatieven kunnen nog wel bepaald worden, maar de baten zijn niet eenvoudig te monetariseren.

Investeringskosten

- *Kosten om data open te maken.* Deze kosten zijn afhankelijk van de keuzes die moeten worden gemaakt en zullen eenmalig ca. €50.000 zijn en de beheerkosten ca. €5.000 per jaar. Totaal gaat het dan om een contante waarde van €140.000.
- *Kosten om data 'linked' te maken.* Deze kosten zijn afhankelijk van in welke fase de data gelinkt zullen worden. Als dit in het conversieproces om data open te maken zal zijn, zullen deze kosten marginaal zijn. Als dit in een latere fase zou gebeuren, dan zouden de kosten ca. €25.000 zijn.

Directe effecten

- *Serverkosten.* Afhankelijk van de keuzes die gemaakt worden zullen de extra serverkosten voor het Plus 1 en Plus 2 alternatief marginaal zijn ten opzichte van het 0-alternatief.

¹²⁸ AWTI [Adviesraad voor Wetenschap, Technologie en Innovatie] (2016). Durven Delen. Op weg naar een toegankelijke wetenschap. Den Haag, p.12, p.112-113.

¹²⁹ KNAW [Koninklijke Nederlandse Akademie van Wetenschappen] (2013). Publieke Kennisinvesteringen en de waarde van wetenschap. Amsterdam, Koninklijke Nederlandse Akademie van Wetenschappen: 68, p.59.

- *Administratieve lasten.* Deze lasten zijn afhankelijk van de keuzes die gemaakt worden. Als er zowel een remote access-onderzoeksversie voor gemachtigde gebruikers gehandhaafd blijft als een opendataversie, zullen de administratieve lasten toenemen omdat er ook ondersteuning voor opendatagebruikers geboden zal moeten worden. De conversiekosten om de Polisbus-dataset via Remote Access beschikbaar te stellen aan onderzoekers zijn moeilijk te distilleren van de andere taken, en zijn ca. €20.000 per jaar (contant tegen 5,5% is dit ca €360.000). De beheerkosten voor Remote Access Polisbus-dataset zijn ca. €7.000 per jaar (Contant tegen 5,5% is dit €130.000). Die kosten zouden in het Min 1 alternatief bespaard worden.
- *Baten voor onderzoekers.* Deze kosten zijn afhankelijk van het soort onderzoek, het type onderzoeker, het aantal onderzoekers per project en het aantal bestandsonderwerpen. De baten zullen waarschijnlijk te vinden zijn bij nieuwe groepen onderzoekers waarvoor de opendataversie voldoende zal zijn. Voor de bestaande onderzoekers, die de data op microniveau nodig hebben, zullen de baten marginaal zijn. De kosten die de onderzoekers nu maken (en die ze niet meer maken in alternatief min 1) geven een ondergrens aan de baten die zij verwachten te hebben van de onderzoeksdata. Op grond van de Dienstencatalogus microdatenservices van CBS en de afschrijvingstermijn van een computer van vier jaar en het gegeven dat deze computer in een afgeschermd ruimte moet staan (uitgegaan is van een cel van 4 m² kantooruimte met tafel en stoel), de veronderstelling dat toegang door 4 gebruikers wordt gedeeld en gebruikt wordt door 400 onderzoekers zijn deze kosten €7,3 miljoen.¹³⁰ Wanneer de verhouding tussen kosten en baten voor de Remote Access faciliteit gemiddeld gesproken 1:2 is, komt dit neer op bruto onderzoeksbatens van €14,6 miljoen (netto €7,3 miljoen)
- *Consumentensurplus.* Het consumentensurplus voor het extra beschikbaar komen van open onderzoeksdata zal afhankelijk zijn van de indirecte effecten van de spillovers van wetenschappelijk onderzoek voor het bedrijfsleven. Het is onbekend hoeveel extra gebruikers deze data zullen trekken.

Indirecte effecten

- *Bijdrage aan wetenschappelijke kennis.* De baten van het hergebruiken van open onderzoeksdata en het creëren van mogelijkheden voor nieuwe soorten (multidisciplinair) onderzoek in andere gebieden zijn moeilijk te kwantificeren omdat de waarde van wetenschappelijk onderzoek van vele factoren afhankelijk is.
- *Betrouwbaarheid.* Als statistische data als open data beschikbaar zijn, zal de transparantie van wetenschappelijke resultaten toenemen omdat de kwaliteit van de onderzoeksoutput door buitenstaanders kan worden gecontroleerd. Daarmee neemt de betrouwbaarheid van de resultaten toe. Ook hier geldt dat die effecten moeilijk te kwantificeren zijn. Wel kan hier ook gedacht worden aan vermeden reputatieschade door onderzoeksfraude.
- *Nieuwe producten en bedrijven.* Om van kennisuitwisseling naar innovatieve toepassingen te komen is een interactief proces en verloopt niet lineair. Om een benadering te kunnen maken, wordt in de literatuur vaak gebruik gemaakt van de zogenaamde paybackmethode,

¹³⁰ Deze berekening is gebaseerd op 100 geïnstalleerde Remote Access computers, waarbij een dergelijke computer (die met veel data ingewikkelde statistische bewerking aan moet kunnen) in vier jaar wordt afgeschreven. De kosten voor de installatie zijn daarmee per jaar 1/4 van €1500,- en de kosten van de computer zelf 1/4 van €1200,-. Het servicecontract is €165 per maand (€1980 per jaar). Uitgegaan wordt van vier geautoriseerde gebruikers die dit elk voor gemiddeld vier jaar doen. Dit betekent dat jaarlijks €375 aan autorisatiekosten worden betaald. De ruimte is begroot op €800 per jaar (4 m² a €200), meubilair op €100 per jaar. Als deze jaarlijkse kosten tegen 5,5% worden gekapitaliseerd komen deze op €71.455. Verder zijn er eenmalige kosten: de basisopstartkosten (€1600) en de kosten per bestandsonderwerp (€160). Het totaal is dan €73.215. Voor 100 installaties zijn de kosten dan €7,3 miljoen.

die indrukwekkende rendementen laat zien.¹³¹ De ontwikkeling van innovatieve toepassingen is ook afhankelijk van andere factoren. Het betreft hier niet alleen het vervolg op wetenschappelijke doorbraken¹³², maar er zal vooral sprake zijn van tal van praktische toepassingen.

Overige effecten

- *Bijdragen aan maatschappelijke oplossingen.* Het delen van onderzoeksresultaten zal resulteren in het eerder beschikbaar komen van resultaten en toepassingen die voor de maatschappij nuttig zijn, maar deze zijn voor dit onderzoek niet te kwantificeren.
- *Ontwikkelen van beleid.* Wetenschappelijk onderzoek kan aantonen wat de verwachte langetermijnonontwikkelingen kunnen zijn. Deze kunnen van groot belang zijn voor het ontwikkelen of aanpassen van beleid. Voor dit onderzoek zijn deze effecten niet te kwantificeren.

Tabel 7.2: Gemonetariseerd overzicht van verschillen effecten in €1.000

Polisbus-dataset van CBS	0 (alleen voor gemachtigde gebruikers beschikbaar als Remote Access-versie)	-1 (niet beschikbaar)	+1 (geaggregeerde versie als open data 3-sterren naast Remote Access-versie)	+2 (geaggregeerde versie als 'linked data' 4-sterren naast Remote Access-versie)
Investeringskosten -Kosten om data open te maken -Kosten om data 'linked' te maken			- 140	- 140 0 tot -25
Directe effecten -Serverkosten -Baten voor onderzoekers -Consumentensurplus*		Marginaal -7.321 pm	pm pm	0 pm pm
Indirecte effecten -Wetenschappelijke kennis -Betrouwbaarheid -Nieuwe producten en bedrijven		pm pm pm	pm pm pm	pm pm pm
Overige effecten				
-Maatschappelijke oplossingen		pm	pm	pm
-Beleidsontwikkeling		pm	pm	pm
Totaal		-6.831 + pm	-140 + pm	-165 + pm

* Zie indirecte kosten.

¹³¹ Er is ook al een poging gedaan om het 'payback framework' toe te passen op de Nederlandse situatie, zie Oortwijn *et al.* (2008). Assessing the impact of health technology assessment in the Netherlands. *International Journal of Technology Assessment in Health Care*, 24(3), 259-269. Bredere toepassing van de 'payback'-methode op publiek gefinancierde onderzoeksprojecten zou volgens de onderzoekers op lange termijn tot interessante nieuwe inzichten kunnen leiden en de bepaling van de impact van wetenschap kunnen verbeteren.

¹³² KNAW [Koninklijke Nederlandse Akademie van Wetenschappen] (2013). Publieke Kennisinvesteringen en de waarde van wetenschap. Amsterdam, Koninklijke Nederlandse Akademie van Wetenschappen: 68, p.41.

7.6 Conclusies

Er zijn circa 400 onderzoekers die gebruik maken van de Polisbus-dataset van CBS via beveiligde Remote Access-faciliteiten. Daarmee is het een van de meest gewilde microdatasets van CBS. CBS hanteert een opendatabeleid en gegeven de inhoud van de dataset kan het de moeite waard zijn om de Polisbus-dataset in een geaggregeerde opendataversie beschikbaar te stellen. Hierbij zullen kosten gemaakt moeten worden om de data te anonimiseren. Het zal in dat geval efficiënter zijn om dan meteen een opendataversie van de Polisbus-dataset geschikt te maken om als 'linked open data' beschikbaar te stellen; de meerkosten van het linken van data zullen marginaal zijn ten opzichte van het als open data beschikbaar stellen. De dataset zal als 'linked data' meer relevantie hebben en makkelijker te koppelen zijn aan andere data.

8 Conclusies

Dit onderzoek geeft inzicht in de (potentiële) maatschappelijke kosten en baten van open data. Achtergrond vormt de ambitie die de Minister van Binnenlandse Zaken en Koninkrijksrelaties in de Nationale Open Data Agenda (NODA) heeft geformuleerd om zoveel mogelijk overheidsgegevens die zich daarvoor lenen, als open data beschikbaar te stellen. Daarbij geeft het kabinet prioriteit aan 'high value'-datasets: datasets die met voorrang beschikbaar worden gesteld, omdat zij een hoge sociaal-economische meerwaarde hebben of in hoge mate bijdragen aan een transparante en beter controleerbare overheid. In dit onderzoek wordt de verhouding onderzocht tussen de kosten van het beschikbaar stellen van overheidsdata als open data en de maatschappelijke baten daarvan. Het onderzoek bestaat uit twee delen: (1) een internationaal vergelijkend literatuuronderzoek naar de kosten en maatschappelijke baten van het openen van overheidsdata, en (2) een maatschappelijke kosten-batenanalyse van vijf overheidsdatasets.

Literatuurstudie naar kwalitatieve en kwantitatieve effecten van open data

Alle onderzochte studies in de internationale literatuurstudie geven een positief beeld van de verhouding tussen kosten en baten van open data. De kwalitatieve effecten zijn divers en overtuigend. De kwantitatieve studies laten een verhouding tussen kosten en baten zien van 1:1,6 tot 1:70. Dat wil zeggen dat de baten van het openstellen van overheidsdata ruim 1,5 tot 70 keer zo groot zijn als de kosten daarvan. Uit de internationale literatuurstudie blijkt echter ook dat het lastig is om de baten methodologisch goed te kwantificeren en uit te drukken in geld. De meeste studies waarin de baten gemonetariseerd zijn, hebben een zwakke onderbouwing. De resultaten zijn hierdoor lastig op hun merites te beoordelen. Daarentegen zijn veel niet-monetaire effecten zoals meer raadplegingen en downloads van de datasets door het bedrijfsleven, het gebruik van de data in andere domeinen, kwaliteitsverbeteringen van de datasets door terugmeldingen, en efficiëntieslagen binnen de overheid ex-post, 'hard', vastgesteld. De literatuurstudie laat verder zien dat de kosten van het verstrekken van beschikbare data als open data marginaal zijn ten opzichte van de totale kosten van informatievoorziening van een organisatie. Wat deze literatuurstudie ook laat zien, is dat de (meer)kosten van open data moeilijk te onderscheiden zijn van de kosten die een dataverstrekker sowieso moet maken om zijn kerntaken uit te voeren.

Naast de literatuurstudie zijn vijf datasets onderzocht waarvan de overheid een potentieel grote toegevoegde waarde verwacht – zogenaamde high value-datasets. De onderzochte cases zijn de Basisregistraties Adressen en Gebouwen (BAG), de Nationale Databank Wegverkeersgegevens (NDW), de landelijke voorziening WOZ-waarden (LV WOZ), de dataset Kindermishandeling in Nederland anno 2005 van WODC en de Polisbus-dataset van CBS (met data over banen en lonen van werknemers in Nederland). De BAG, NDW-gegevens en de 'Kindermishandeling in NL 2005' dataset zijn reeds als open data beschikbaar. Voor de andere datasets zou dat mogelijk op termijn het geval kunnen worden. In de cases zijn voor drie alternatieven de maatschappelijke kosten en baten onderzocht: (1) van het aanbieden van de data als open data, (2) het niet als open data aanbieden, en (3) het aanbieden van de data als 'linked open data'.

Conclusies individuele cases

De belangrijkste conclusies over de individuele cases zijn als volgt.

Basisregistraties Adressen en Gebouwen (BAG)

De BAG is sinds 2012 beschikbaar voor een ieder als open data, zij het tegen marginale verstrekingskosten voor het bedrijfsleven. Het beschikbaar stellen van de BAG heeft een efficiëntere

en effectievere overheid opgeleverd. Dubbel inwinnen van data vindt minder vaak plaats en er wordt meer gebruik gemaakt van BAG. De authentieke gegevens uit de BAG worden als betrouwbare brongegevens beschouwd. Zo gebruikt de Landelijke Voorziening WOZ de BAG voor het recent geopende WOZ-Waardeloket. De maatschappelijke baten zijn ook groter dan voorheen: bedrijven gebruiken de BAG steeds vaker als een bron voor het verbeteren van hun dienstverlening en voor het creëren van nieuwe producten en diensten. Hergebruik van de BAG versterkt het imago van de overheid als een aanbieder van transparante en betrouwbare data. Het handhaven van de huidige financiële drempel lijkt geen meerwaarde te bieden ten opzichte van het gratis aanbieden van de data. Pilots met het verwijderen van de laatste financiële drempel tonen aan dat open data een positief effect hebben op de bedrijvigheid, vooral in de vorm van nieuwe kleine bedrijfjes. Het aanbieden van open data is op organisatorisch gebied slechts een kleine stap binnen het gehele programma van het Kadaster om te komen tot betere ruimtelijke informatie. Het effect van open data is daardoor lastig af te zonderen van de totale innovatie. Zowel qua kosten als baten is het effect van open data bescheiden ten opzichte van de totale begroting. De stap naar 'linked data' brengt extra kosten met zich mee om de data op een kwalitatief hoger niveau te tillen. Maar gezien de potentiële baten voor de overheid (betere uitvoering van een complexe Omgevingswet) zullen de uiteindelijke baten hoger zijn dan de kosten. Daardoor zullen de maatschappelijke baten ook hoger zijn: kwalitatief hoogwaardige data, beter geïnformeerde burgers, betere vindbaarheid van de data en betere koppelingen voor nieuwe toepassingen.

Nationale Databank Wegverkeersgegevens (NDW)

NDW biedt sinds september 2013 actuele verkeersgegevens beschikbaar als open data aan een ieder. Open data aanbieden is op organisatorisch gebied slechts een kleine stap binnen het gehele programma van NDW om betere gegevens over wegen en verkeer te verzamelen en beschikbaar te stellen. Het effect van open data is daardoor lastig af te zonderen van de totale innovatie. Zowel qua kosten als baten is het effect van open data bescheiden ten opzichte van de gehele inspanning die wordt geleverd op het gebied van wegverkeersgegevens. Hoewel de baten van actuele verkeersgegevens als gevolg van open data moeilijk te isoleren zijn van de baten als gevolg van andere data-overeenkomsten, is de maatschappelijke kosten-batenbalans van open data overigens positief en het kan daardoor als een sluitstuk worden beschouwd van data-innovaties bij NDW.

De WOZ-Waardeloket van de Landelijke Voorziening WOZ

Vanaf oktober 2016 worden WOZ-waarden uit de Landelijke Voorziening WOZ openbaar gemaakt via het WOZ-Waardeloket. Het betreft dus een dienst om een ieder inzage te geven in WOZ-waarden van woningen. Het WOZ-Waardeloket is nog te jong om harde conclusies te trekken over de maatschappelijke baten; de eerste signalen zijn reden voor een voorzichtig positief oordeel. De beheerkosten en infrastructuurkosten van de Landelijke Voorziening WOZ moeten in ieder geval gemaakt worden om de primaire processen te kunnen uitvoeren. Ook het openbaar maken van WOZ-waarden via het WOZ-Waardeloket is onderdeel van het primaire proces. Pilots met WOZ-voormeldingen in een aantal gemeenten hebben aangetoond dat meer transparantie van WOZ-waarden leidt tot minder bezwaren en daarmee lagere administratieve lasten voor de gemeenten. Uit eerdere pilots met voormeldingen van WOZ-waarden bleek bovendien dat de transparantie heeft geleid tot meer draagvlak voor de belastingheffingen. Of de openbaarheid van de volledige WOZ-waarden dezelfde effecten zullen hebben, zal in 2017 duidelijk worden wanneer de volgende beschikkingen zullen worden verstuurd. De WOZ-waarden zijn nu alleen openbaar en niet herbruikbaar voor het bedrijfsleven omdat de WOZ-waarden gerelateerd kunnen worden aan personen. Als er een aangepaste, geaggregeerde, versie beschikbaar komt als open data, zullen de maatschappelijke baten waarschijnlijk nog verder stijgen als het bedrijfsleven de data kan hergebruiken, onder andere voor de ontwikkeling van nieuwe producten. De waarde voor het bedrijfsleven zal afhankelijk zijn van de mate van aggregatie en de manier waarop de open data beschikbaar worden gesteld.

Kindermishandeling in Nederland anno 2005

Het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het Ministerie van Veiligheid en Justitie heeft een aantal onderzoeksdatasets, waaronder de 'Kindermishandeling in Nederland anno 2005,' dataset beschikbaar gesteld als open data. De dataset is tot op heden nog niet veel gedownload, wat veroorzaakt zou kunnen worden doordat de dataset moeilijk te vinden voor buitenstaanders. Het vindbaar zijn van data is een belangrijke voorwaarde voor hergebruik. Met een relatief kleine inspanning (en daardoor relatief lage kosten) zal de vindbaarheid van de dataset toenemen wanneer deze als 'linked open data' beschikbaar wordt gesteld.

Polisbus: Banen en lonen van werknemers in Nederland

De Polisbus-dataset bevatten gegevens over lonen van individuele personen, privacygevoelige gegevens dus. Omdat de Polisbus-dataset persoonsgegevens bevatten, kan de dataset niet openbaar worden gemaakt, laat staan als open data beschikbaar worden gesteld. De circa 400 onderzoekers die gebruik maken van Polisbus-dataset krijgen daarom, na ondertekening van een contract, via beveiligde Remote Access-faciliteiten toegang tot de dataset. CBS hanteert een open databeleid, en gegeven de waardevolle inhoud van de dataset kan het de moeite waard zijn om de Polisbus-dataset in een geaggregeerde versie als een opendataversie beschikbaar te stellen. Het zal in dat geval efficiënter zijn om dan meteen de data beschikbaar te stellen als 'linked open data'; de meerkosten van het linken van data zullen marginaal zijn ten opzichte van het als open data beschikbaar stellen. De dataset zal als 'linked data' meer relevantie hebben en makkelijker te koppelen zijn aan andere data.

Onze studie laat zien dat de maatschappelijke baten groter zullen zijn dan de lasten bij het verstrekken van overheidsdata als open data. Het al dan niet beschikbaar stellen van overheidsgegevens als open data is meestal slechts een kleine stap in het databeleid in de cases. Daar waar de data vrij zijn van persoonsgegevens, zijn de additionele kosten van het verstrekken als open data marginaal, en lager dan de verwachte baten. In veel van de cases is er sprake van belangrijke databeleidsvernieuwingen: er worden al grote stappen gemaakt met betrekking tot investeren in infrastructuur en technologie als onderdeel van de primaire taken van de overheid. De maatschappelijke kosten en baten van het als open data beschikbaar stellen van overheidsdata zijn nu nog gering ten opzichte van de algehele investering in de ontwikkeling van deze data, maar – waar waarneembaar – laten een positieve balans zien.

De cases laten ook zien dat er sprake is van verschillende gradaties van openheid: van volledig open data via open geaggregeerde data en openbaar toegankelijke data tot gesloten data. De vraagstelling is dan ook niet zozeer gericht op open data versus gesloten data, maar op data die in verschillende gradaties van openheid ter beschikking worden gesteld.

Maatschappelijke kosten en baten van de onderzochte datasets

Uit de cases blijkt dat de kosten van het als open data beschikbaar stellen van overheidsdata relatief laag zijn, mits de data met een geringe bewerkingsslag gepubliceerd kunnen worden als open data. Met name voor data die nu door de overheid tegen betaling worden aangeboden zijn de kosten voor de data als open data beschikbaar stellen marginaal. Deze overheidsdata blijven geproduceerd worden, ook als hier geen inkomsten uit verkoop meer tegenover staan. Het kosteloos toegankelijk maken van de data trekt een nieuwe groep gebruikers aan. Tegenover het wegvallen van de inkomsten staat dezelfde maatschappelijke baat bij de afnemers die eerst betaalden. Ook verdwijnen administratieve lasten rond betalingen. De directe maatschappelijke baten zijn hierdoor groter dan de kosten. Uit de cases blijkt dat ook het hergebruik door de overheid zelf door het verstrekken van de data als open data toeneemt. Dit draagt bij aan de efficiëntie van de overheid.

Daar waar de data niet geschikt zijn om zonder bewerking te publiceren, zijn de verhoudingen binnen de cases anders. Indien er bijvoorbeeld sprake is van persoonsgegevens in een dataset moeten de data worden geanonimiseerd om als open data beschikbaar gesteld te kunnen worden (i.c. de WOZ-waarden, en de Polisbus-dataset van CBS). Het anonimiseren van data brengt extra kosten met zich mee en resulteert in een dataset met een mogelijk lagere utiliteitswaarde voor hergebruikers. De algemene positieve kosten-baten balans van open data zal hierdoor kleiner worden of mogelijk omslaan naarmate de oorspronkelijke datasets meer persoonsgegevens bevatten. Het laatste is aannemelijk bij de Polisbus-dataset van CBS.

Zelfs datasets die geen directe persoonsgegevens bevatten kunnen nog herleidbaar zijn tot personen. Het is sterk afhankelijk van de dataverzameling zelf – en ook de technische karakteristieken van deze dataverzameling – welke kosten er gemoeid zijn met de bewerking van de data om deze niet meer herleidbaar tot personen te maken. Desondanks is de verwachting dat de beschikbaarheid als open data van met name high value-datasets op een iets hoger aggregatieniveau dan de brondata nog steeds een positieve maatschappelijke kosten-batenverhouding zal hebben.

De kosten van het alternatief om de data ‘linked’ te maken, zijn – op grond van ervaringen in enkele van de cases – geraamd en betreffen slechts een fractie van de totale kosten van het ontwikkelen en beheren van deze data. De effecten zijn lastig te moneteriseren omdat linked (open) data nog niet breed is geïmplementeerd. De verwachting is dat veel van de verwachte positieve effecten van open data veel gemakkelijker bereikt kunnen worden bij het aanbieden van dezelfde data als ‘linked data’, omdat hierdoor veel betere aansluitingen tussen databestanden binnen de data-infrastructuur mogelijk zijn. Indien de kosten van het ‘linked’ maken overeenkomstig de geraamde bedragen blijven, verwachten we dat de maatschappelijke baten van high-value datasets hoger zijn dan de maatschappelijke kosten.

Het al dan niet beschikbaar stellen van overheidsgegevens als open data vergt in de onderzochte casussen meestal slechts een kleine aanpassing in het databeleid. In veel van de casussen is er sprake van belangrijke databeleidsvernieuwingen: er wordt al veel geïnvesteerd in infrastructuur en technologie als onderdeel van de primaire taken van de overheid. De maatschappelijke kosten en baten van het als open data beschikbaar stellen van overheidsdata lijken nu nog misschien gering ten opzichte van de algehele investering in de ontwikkeling van deze data, maar – waar waarneembaar – laten een positieve balans zien, zeker ten opzichte van het *niet* beschikbaar stellen van de data.

In, in ieder geval, twee casussen bevat de dataset persoonsgegevens (WOZ-waarden, en de Polisbus-dataset van CBS). Voor deze casussen is het vanwege de eisen die de Wet bescherming persoonsgegevens¹³³ stelt ondenkbaar dat ze in de toekomst als open data beschikbaar zullen komen. Daarom hebben we voor deze casussen bij het open data en open linked data alternatief gekozen voor een geaggregeerde, dat wil zeggen anonieme, versie van de oorspronkelijke dataset. Deze aanpassing heeft ertoe geleid dat er in de beleidsvarianten een vergelijking wordt gemaakt tussen verschillende alternatieven met verschillende datasets. Dit betekent tevens dat er voor het maken van de opendataversie extra bewerkingen moeten plaatsvinden waarvan de kosten niet goed zijn in te schatten. Hier speelt tevens de vraag wanneer een dataset nu redelijkerwijs niet meer te herleiden is tot een natuurlijke persoon. De mate van aggregeren heeft ook zijn weerslag op de utiliteitswaarde van de dataset. Deze onzekerheden zijn in het onderzoek onderkend maar bleken lastig te kwantificeren. In het geval van de Polisbus-dataset van CBS hebben we ook geen gebruik kunnen maken van ervaringen met vergelijkbare datasets in het buitenland.

Dit betekent dat de uitkomsten van het onderhavige onderzoek een indicatie geven van de maatschappelijke kosten en baten maar niet als absolute waarheid mogen worden geïnterpreteerd.

¹³³ En vanaf 25 mei 2018 de Algemene Verordening Gegevensbescherming.

Dankwoord

Wij willen graag alle mensen bedanken die hebben meegewerkt aan het tot stand komen van dit rapport. Allereerst alle mensen die bereid waren om geïnterviewd te worden of gegevens hebben aangeleverd voor onze case studies. Met hun input hebben ze waardevolle bijdrage geleverd aan de totstandkoming van dit rapport. Ook willen we met name de leden van de begeleidingscommissie: Jaron Haas en Arie Scheer van Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Rien Bout van Ministerie van Infrastructuur en Milieu, Ramon de Louw van Ministerie van Economische Zaken, Fanny Wallebroek van Ministerie van Veiligheid en Justitie, Hainy Laken en Hugo Haastert van Ministerie van Financiën, Magdaleen de Schepper en Arjen Wilms van Ministerie van Sociale Zaken en Werkgelegenheid bedanken. Ook willen we de Stuurgroep Open Data bedanken voor hun feedback. Verder willen we Niels Peters van Ecorys bedanken voor het delen van de voorlopige resultaten van de evaluatie onder NDW Data+Diensten-gebruikers. Ten slotte willen we de volgende medewerkers van het OTB, TU Delft bedanken: Marianne de Vries voor haar analyses van de BAG IP-data, Sylvia Janssen met haar hulp in het opzetten van de vragenlijsten, en Dirk Dubbeling voor het proeflezen en de redactie van dit rapport.

Lijst van geïnterviewden

Naam	Organisatie	Datum interview
Ruben Dood	CBS	30-09-2016
Henk Harmsen	DANS	23-11-2016
Erwin Folmer	Kadaster	13-10-2016
Hugo Haastert	Ministerie van Financiën	22-11-2016
Peter van der Heuij	Ministerie van Financiën	11-10-2016 en 22-11-2016
Fanny Wallebroek	Ministerie van Veiligheid en Justitie	6-10-2016
Ernst Scheerder	NDW	4-10-2016
Ruud Kathmann	Waarderingskamer	19-10-2016
Ronald Meijer	WODC	6-10-2016 en 14-11-2016
Andreas Pelke	Oberer Gutachterausschuss für Grundstückswerte in NRW	14-10-2016

Geraadpleegde literatuur

- Amsterdam Mobiel (2015). Eindrapport praktijkproef Amsterdam. Amsterdam: 63. <http://docplayer.nl/25460318-Eindrapport-praktijkproef-amsterdam.html>, (opgeroepen 28/11/2016)
- Australian Bureau of Statistics (2014). Annual Report 2013-14. Canberra.
- Bank of England (2006). Cost-benefit analysis of monetary and financial statistics. A practical guide: 36. <http://www.bankofengland.co.uk/statistics/Documents/about/cba.pdf>, (opgeroepen 28/11/2016).
- Bizer, C., R. Cyganiak and T. Heath (2007). How to Publish Linked Data on the Web. Berlin, Freie Universität Berlin <http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>, (opgeroepen 28/11/2016).
- Brandt, U. (2013) Die Gutachterausschüsse für Grundstückswerte in Niedersachsen in Beschwerdeverfahren und gerichtlichen Auseinandersetzungen, Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung 63(2+3), 3-14.
- Carolan L., 2014, Estimating the cost of a government open data initiative, *ODI News* (Open Data Institute), <https://theodi.org/blog/estimating-the-cost-of-a-government-open-data-initiative>.
- CBS [Centraal Bureau voor de Statistiek] (2015a). Documentatierapport Banen en lonen van werknemers in Nederland (POLISBUS). Den Haag, Centrum voor Beleidsstatistiek en Microdata Services: 24.
- CBS [Centraal Bureau voor de Statistiek] (2015b). Dienstencatalogus Microdataservices (On Site / Remote Access) 2015/2016. Den Haag, Centrum voor Beleidsstatistiek en Microdata Services: 43.
- CPB/PBL (2015) Maatschappelijke kosten en baten prijsbeleid personenauto's, Den Haag <https://www.cpb.nl/sites/default/files/publicaties/download/cpb-pbl-achtergronddocument-13apr2015-mkba-prijsbeleid-personenautoos.pdf> (opgeroepen 07/11/2016).
- Davies, T. (2013). Open Data Barometer. 2013 Global Report Web Foundation and Open Data Institute: 45.
- Eijgenraam, C.J.J., C.C. Koopmans, P.J.G. Tang & A.C.P. Verster (2000). Evaluatie van Infrastructuurprojecten; Leidraad voor kosten-batenanalyse. Deel I: Hoofdrapport. Den Haag, Centraal Planbureau: 84. http://www.mkba-informatie.nl/index.php/download_file/force/107/195/. (opgeroepen 28/11/2016)
- Farrow, S. (2013). *How (not) to lie with Benefit-Cost Analysis*. The Economists' Voice 10 (1 (Dec. 2013)):45-50. doi: 10.1515/ev-2013-0012.
- Fornefeld, M. (2009). The value to Industry of PSI: the business sector perspective. In P.F. Uhler (Ed.) The Socioeconomic Effects of Public Sector Information on Digital Networks: Towards a Better Understanding of Different Access and Reuse Policies. Workshop Summary. Washington, The National Academies Press: 10-16, ISBN 978-0-309-13968-7.
- Friedt, M. & T. Luckhardt (2014) Open Data: Zukunftsorientierte Bereitstellung von amtlichen Geodaten im Land Berlin, zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement 139(5), 269-77.
- Grimmelikhuijsen, S.G. (2012). *Transparency and trust. An experimental study of online disclosure and trust in government*. Ph.D. Dissertation, Utrecht: Utrecht University, ISBN 978-90-8570-971-2
- Häggquist, E. & P. Söderholm (2015). *The economic value of geological information: Synthesis and directions for future research*. Resources Policy 43 (March 2015):91-100. doi: 10.1016/j.resourpol.2014.11.001.

- Heusser, F.I. (2012). Understanding Open Government Data and addressing its impact (Draft Version).
- IRM [Independent Reporting Mechanism] (2015). Aligning supply and demand for better governance. Washington, Independent Reporting Mechanism (IRM): 40. <http://www.opengovpartnership.org/sites/default/files/IRMReport-OpenData.pdf>, (opgeroepen 19/10/2015)
- Kamerstukken (2013). Mobiliteitsbeleid; Brief regering; Beter geïnformeerd op weg: Koersbepaling Reisinformatie & Verkeersmanagement. Ministerie van Infrastructuur en Milieu. Den Haag. kst II 2012/13, 31 305, nr. 202.
- Kamerstukken (2015). Mobiliteitsbeleid; Brief regering; Samenhang en activiteiten Smart Mobility. Ministerie van Infrastructuur en Milieu. Den Haag. kst II 2015/16, 31 305, nr. 214.
- Kamerstukken (2015). Nationale open data agenda 2016 (NODA). Den Haag, Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. kst II 2015/16, 32 802, nr. 20
- Kertscher, D. (2009) Transparency with the Online-Real-Estate-Price-Calculator in Lower Saxony, FIG Working Week 2009 Surveyors Key Role in Accelerated Development, Eilat, https://www.fig.net/resources/proceedings/fig_proceedings/fig2009/papers/ts01e/ts01e_kertscher_3221.pdf (opgeroepen 11/10/2016)
- KIM (2014) Mobiliteitsbeeld 2014, Kennisinstituut voor Mobiliteitsbeleid, Den Haag, <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2014/10/23/mobiliteitsbeeld-2014-kennisinstituut-voor-mobiliteitsbeeld/mobiliteitsbeeld-2014-kennisinstituut-voor-mobiliteitsbeeld.pdf> (opgeroepen 07/11/2016).
- KIM (2015) Trendprognose wegverkeer 2015 - 2020 voor RWS, Kennisinstituut voor Mobiliteitsbeleid, Den Haag, http://www.kimnet.nl/binaries/kimnet/documenten/rapporten/2015/06/29/trendprognose-wegverkeer-2015_2020-voor-rws/trendprognose-wegverkeer-2015-2020-voor-rws.pdf (opgeroepen 07/11/2016).
- KIM (2016) Trendprognose wegverkeer 2016-2021 voor RWS, Kennisinstituut voor Mobiliteitsbeleid, Den Haag, <http://www.kimnet.nl/publicaties/rapporten/2016/06/06/trendprognose-wegverkeer-2016-2021-voor-rws> (opgeroepen 07/11/2016).
- Kronenburg, T., T. Monasso, E. Boschker & M. Thaens (2012). De waarde van open data. Keuzes en effecten van open-datastrategieën voor publieke organisaties. Den Haag, ZENC i.o.v. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties: 109
- Liebig, S. (2012) BORIS.NI – das neue Bodenrichtwertinformationssystem: VBORIS – ein nutzerorientierter Online-Dienst der AdV, Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung 62(3+4), 3-11.
- Lind, M. (2014). *Addresses and Address Data. Socio-economic benefits of Open Address Data experiences in Denmark*. State of the Map France Conference, Paris, 4-6 April 2014. OpenStreetMap France.
- Mouter, N. (2012). Voordelen en nadelen van de Maatschappelijke Kosten- en Baten analyse nader uitgewerkt: 18. <http://www.mkba-informatie.nl/mkba-voor-gevorderden/publicaties-over-mkba/mouter-2012-voordelen-en-nadelen-van-de-maatschappelijke-kos/>, (opgeroepen 28/1/2017).
- NWD [Nationale Databank Weggegevens] (2016). NDW Jaaroverzicht 2015: 26. <http://www.ndw.nu/downloaddocument/72a2ffef1ed01dde3ce7ab041f384a96/NDW%20Jaaroverzicht%202015%20versie%201.0.pdf>, (opgeroepen 1/11/2016).

- Onderzoekcentrum Drechtsteden (2014). Doorbraak WOZ 2014. Ervaringen en effecten van een nieuwe ronde: 71.
http://www.onderzoekcentrumdrechtsteden.nl/dds/up/ZixzvkdJG_def_rapportage_doorbraak_WOZ_2014.pdf,
- Oortwijn, W.J., S.R. Hanney, A. Ligtoet, S. Hoorens, S. Wooding, J. Grant, M.J. Buxton & L.M. Bouter (2008). Assessing the impact of health technology assessment in The Netherlands. International Journal of Technology Assessment in Health Care vol. **24**(3): p.259-69 doi: 10.1017/s0266462308080355
- Platform Linked Data Nederland (2014). Linked data in beeld. Amersfoort: 24.
http://www.pilod.nl/w/images/3/32/Linked_Data_in_bbeeld_2014.pdf, (opgeroepen 28/11/2016)
- Prest, A.R. & R. Turvey (1965) Cost-Benefit Analysis: a Survey, *The Economic Journal*, 75(300), 683-735.
- Richtlijn 2010/40/EU van het Europees Parlement en de Raad van 7 juli 2010 betreffende het kader voor het invoeren van intelligente vervoerssystemen op het gebied van wegvervoer en voor interfaces met andere vervoerswijzen (ITS Richtlijn.PbEU L 207/1-13.
- Romijn, G & G. Renes (2013) Algemene leidraad voor maatschappelijke kosten baten analyse, CPB/PBL, Den Haag.
- te Velde, R. (2009). Public Sector Information: Why bother? In P.F. Uhler (Ed.) *The Socioeconomic Effects of Public Sector Information on Digital Networks: Toward a Better Understanding of Different Access and Reuse Policies: Workshop Summary* (pp. 25-28). Washington: The National Academies Press.
- TNO Stappenplan <http://www.pilod.nl/wiki/BoekTNO/stappenplan>, (opgeroepen 27/11/2016).
- van Loenen, B. & F. Welle Donker (2014). De stand in opendataland. Delft: OTB - Research for the Built Environment, doi:uuid:7b425db6-2f9a-4be6-b5f0-3492055268d8
- van Loenen, B., F. Welle Donker & R. Braggaaar (2016). De stand in opendataland 2016. Delft: Kenniscentrum Open Data, doi:uuid:b76cb0d8-e9b5-4125-8d7c-b452de1da0d0.
- van Loenen, B., J. Zevenbergen, G. Giff & J. Cromptvoets (2007). Open toegankelijkheidsbeleid voor geo-informatie vergeleken: het gras leek groener dan het was (pp. 116): Technische Universiteit Delft.
- Voss, W. (2011) Transparency of Property Markets – The 1. National Level Market Report, FIG Working Week 2011: Bridging the Gap between Cultures, FIG, Marrakech.
- Walm, A. (2013) Informations- und Einführungsveranstaltung VBORIS 2.0: Informationsveranstaltung der Projektgruppe VBORIS, des Arbeitskreises Liegenschaftskataster der AdV am 08.03.2013 auf der CeBIT Hannover, Nachrichten der Niedersächsischen Vermessungs- und Katasterverwaltung 63(2+3), 30-32.
- Wanzke, H. (2009) Internet presentation of real estate market data in NRW (North Rhine-Westphalia), Germany, International Workshop on Spatial Information for Sustainable Management of Urban Areas, FIG Commission 3 Workshop 2009, FIG, Mainz.
- Welle Donker, F., & B. van Loenen (2014). Open data beoordelingsraamwerk. Deel: review kosten baten analyses (pp. 32). Delft: Kenniscentrum Open Data
- Wenzler, P. and S. Kropp (2015) Real Estate Reference Values for a Better Market Transparency, FIG Working Week 2015: From the Wisdom of the Ages to the Challenges of the Modern World, Sofia,
https://www.fig.net/resources/proceedings/fig_proceedings/fig2015/papers/ts01j/TS01j_wenzler_kropp_7584.pdf (opgeroepen 21/10/2016)
- Werkgroep Discontovoet 2015 (2015). Rapport Werkgroep Discontovoet 2015: 95.

<https://www.rijksoverheid.nl/ministeries/ministerie-van-financien/documenten/rapporten/2015/11/13/rapport-werkgroep-discontovoet-2015-bijlage>, (opgeroepen 28/11/2016).

Wildavsky, A. (1966) The Political Economy of Efficiency: Cost-Benefit Analysis, Systems Analysis and Program Budgeting, *Public Administration Review*, 26(4), 292-310.

Witteveen + Bos (2011) MKBA-kengetallen voor omgevingskwaliteiten: aanvulling en actualisering Rotterdam, <http://www.omgevingseconomie.nl/wp-content/uploads/2012/03/Kengetallen-Omgevingskwaliteiten.pdf> (opgeroepen 07/11/2016).

Onderzochte kosten-batenstudies

- ACIL Tasman (2008). The value of Spatial Information. The impact of modern spatial information technologies on the Australian Economy: 221.
- ACIL Tasman (2009). Spatial Information in the New Zealand Economy. Realising productivity gains: 153.
- Baudel, T., L. Dablanc, P. Aguiar Melgarejo & J. Ashton (2016). *Optimizing urban freight deliveries: from designing and testing a prototype system to addressing real life challenges*. Tenth International Conference on City Logistics 17-19 June 2015, Tenerife, Spain, Elsevier, pp 170-180.
- Booz Allen Hamilton Inc. (2012). Measuring Socioeconomic Impacts of Earth Observations. A Primer. National Aeronautics and Space Administration (NASA): 64.
- Bregt, A.K., W. Castelein, L. Grus & D. Eertink (2013). De effecten van een open basisregistratie topografie (BRT). Wageningen: 40.
- Bregt, A., L. Grus & D. Eertink (2014). Wat zijn de effecten van een open basisregistratie topografie na twee jaar? Wageningen, Wageningen University: 49.
- Bregt, A.K., L. Grus, T. van Beuningen & H. van Meijeren (2016). Wat zijn de effecten van een open Actueel Hoogtebestand Nederland (AHN)? Onderzoek uitgevoerd in opdracht van het Ministerie van Economische Zaken. Wageningen, Wageningen University & Research: 53.
- Capgemini Consulting (2015). Creating value through open data. Study on the impact of re-use of public data resources. Digital Agenda for Europe. European Commission: 112.
- Carpenter, J. & P. Watts (2013). Assessing the value of OS OpenData™ to the economy of Great Britain -Synopsis. A synopsis of the Full Study Interim Report prepared for Department of Business, Innovation and Skill by Consulting Where Limited & ACIL Tasman. Ordnance Survey: 32.
- Castelein, W., A. Bregt & Y. Pluijmers (2010). *The economic value of the Dutch geo-information sector*. International Journal of Spatial Data Infrastructures Research vol. 5. pp. 58-76.
- ConsultingWhere Ltd & ACIL Tasman (2012). Assessing the value of Ordnance Survey OpenData™ to the economy of Great Britain. Full study interim report: 107.
- Coote, A. & A. Smart (2010). The Value of Geospatial Information to Local Public Service Delivery in England and Wales. Final Report, Consulting Where Ltd and ACIL Tasman: 128.
- Danish Government / Local Government Denmark (2012). Basic Data for improved efficiency and growth. The eGovernment strategy 2011-2015. Albertslund, Rosendahls-Schultz Distribution: 40.
- datos.gob.es (2012). Characterization Study of the Infomediary Sector. 2012 Edition: 120.
- datos.gob.es (2013). Re-use of Public Sector Information. Free Access, Public Fees and Prices. Topic Report. ePSI Platform. Topic Report No. 2013/04: 26.
- de Vries, M. (2014). Wah kos'dah dan? Onderzoek naar de incrementele kosten van aan Open Data doen: 60.
- de Vries, M., L. Kapff, M. Negreiro Achiaga, P. Wauters, D. Osimo, P. Foley, K. Szkuta, J. O'Connor & D. Whitehouse (2011). Pricing of Public Sector Information Study (POPSIS) - Summary Report. Brussels, Deloitte Consulting CVBA.
- DECA [Danish Enterprise and Construction Authority] (2010). The value of Danish address data: Social benefits from the 2002 agreement on procuring address data etc. free of charge: 8.
- Dekkers, M., F. Polman, R. te Velde & M. de Vries (2006). Measuring European Public Sector Information Resources (MEPSIR) Study. Final report of study on exploitation of public sector information - benchmarking of EU framework conditions, HELM Group of Companies of Moira,

Northern Ireland & ZENC, the Netherlands: 94.

- Deloitte (2013). Market Assessment of Public Sector Information. London, Department for Business Innovation & Skills: 235.
- Deloitte (2014). The impact of open geographical data. English summary of Danish report to the Danish Geodata Agency: 17.
- Deloitte LLP (2012). Open Growth: Stimulating demand for open data in the UK. A briefing note from Deloitte Analytics. London, Deloitte Touche Tohmatsu Limited: 12.
- DotEcon (2015). Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI). A Report by DotEcon for the CMA.: 234.
- Eaves D. (2010). Case Study: How Open data saved Canada \$3.2 Billion, <https://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>
- Fornefeld, M., G. Boele-Keimer, S. Recher & M. Fanning (2008). Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report. Study commissioned by EC in 2007. Dusseldorf, MICUS: 101.
- Freebairn, J.W. & J.W. Zillman (2002). *Economic benefits of meteorological services*. Meteorological Applications 9(01): 33-44.
- Genovese, E., G. Cotteret, S. Roche, C. Caron & R. Feick (2010). *The EcoGeo Cookbook for the Assessment of Geographic Information Value*. International Journal of Spatial Data Infrastructures Research 5(2010): 120-144.
- Grus, L., A. Bregt & D. Eertink (2015). De effecten van open data BRT na 3 jaar, Wageningen University & het Kadaster: 7.
- Hägglund, E. and P. Söderholm (2015). *The economic value of geological information: Synthesis and directions for future research*. Resources Policy 43(March 2015): 91-100.
- Heusser, F.I. (2012). Understanding Open Government Data and addressing its impact (Draft Version).
- Houghton, J. (2011). Costs and benefits of data provision. Report to the Australian National Data Service. Melbourne, Centre for Strategic Economic Studies, Victoria University: 66.
- Jetzek, T. (2013). *The value of Open Government Data*. Geoforum Perspectiv 12(23 (2013)): 47-56.
- Khalil, S., W. Saffar & S. Trabelsi (2015). *Disclosure Standards, Auditing Infrastructure, and Bribery Mitigation*. Journal of Business Ethics, 132(2), 379-399. doi: 10.1007/s10551-014-2321-6
- Koski, H. (2011). *Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth?* Keskusteluaiheita Discussion Papers, 28 September 2011 no. 1260. Helsinki, ETLA - The Research Institute of the Finnish Economy: 20.
- Krishnamurthy, R. & Y. Awazu (2016). *Liberating data for public value: The case of Data.gov*. International Journal of Information Management 36(4): 668-672.
- Kronenburg, T., T. Monasso, E. Boschker & M. Thaens (2012). De waarde van open data. Keuzes en effecten van open-datastrategieën voor publieke organisaties. Den Haag, ZENC i.o.v. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties: 109.
- Lazo, J K., R.E. Morss & J.L. Demuth (2009). *300 Billion Served. Sources, Perceptions, Uses, and Values of Weather Forecasts*. Bulletin of the American Meteorological Society 90(6): 785-798.
- Mäkelä, J. (2012). *Model for Assessing GIS Maturity of an Organization*. GSDI 13 World Conference. Global Geospatial Conference 2012 - Spatially Enabling Government, Industry and Citizens, Quebec, GSDI Association Press.

- Manyika, J., M. Chui, P. Groves, D. Farrell, S. Van Kuiken & E. Almasi Doshi (2013). Open Data: Unlocking innovation and performance with liquid information, McKinsey Global Institute: 116.
- Newbery, D., L. Bently & R. Pollock (2008). Models of Public Sector Information Provision via Trading Funds. Study commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury in July 2007. Cambridge, Cambridge University: 154.
- Nilsen, K. (2007). Economic theory as it applies to Statistics Canada: A review of the literature, Statistics Canada: 114.
- OECD [Organisation for Economic Co-operation and Development] (2006). Digital Broadband Content: Public Sector Information and Content, Directorate for Science, Technology and Industry. Committee for Information, Computer and Communication Policy: 82.
- OFT [Office of Fair Trading] (2006). The commercial use of public information (CUPI): 183.
- Omidyar Network (2014). Open for Business: How open data can help achieve the G20 growth target. A Lateral Economics report commissioned by Omidyar Network: 84.
- Oslo Economics (2011). Markedspotensial ved økt tilgjengeliggjøring av offentlig data: 123.
- OSTP [Ocean Science and Technology Partnership] (2011). Lessons learned from OOS in Canada: Preliminary Assessment of OOS Value: 19.
- Oxera [Oxford Economic Research Associates Ltd] (2013). What is the economic impact of geo services? Prepared for Google. Oxford, Oxera Consulting Ltd.
- Pettifer, R.E.W. (2009). *PSI in European Meteorology - an unfulfilled potential*. 7th Eastern European eGov Days: eGovernment & eBusiness Ecosystem & eJustice Conference. Prague, European Projects and Management (EPMA) (CZ).
- Pham, N.D. (2011). The Economic Benefits of Commercial GPS Use in the U.S. and The Costs of Potential Disruption, ndp consulting: 19.
- PIRA International Ltd, University of East Anglia & KnowledgeView Ltd (2000). Commercial exploitation of Europe's public sector information - Final report. Pira International Ltd, European Commission Directorate General for the Information Society: 132.
- Ploos van Amstel, W. (2016) Gebruik actuele verkeersdata levert mogelijk 1 miljard op. De Laatste Meter.
- Pollock, R. (2008). The Economics of Public Sector Information, University of Cambridge: 50.
- Pollock, R. (2011). Welfare gains from opening up Public Sector Information in the UK, University of Cambridge: 4.
- PriceWaterhouseCoopers Australia (PwC) (2014). Deciding with data. How data-driven innovation is fuelling Australia's economic growth: 68.
- Proyecto Aporta (2011). Characterization Study of the Infomediary Sector: 88.
- PwC & Uscreates (2015). Nesta and the Open Data Institute Open Data Challenge Series. Final Report: 32.
- PwC (2014). Deciding with data. How data-driven innovation is fuelling Australia's economic growth: 68.
- RDW (2013). Eindevaluatierapport - PoC Open Data Voertuigen: 24.
- SCGOE [Standing Committee on Government Operations and Estimates] (2014). Open Data: The way of the Future. Report of the Standing Committee on Government Operations and Estimates: 54.
- Shakespeare, S. (2013). Shakespeare Review: An independent review of Public Sector Information: 71.
- The World Bank (2014). Open data for economic growth: 20.

- Ubaldi, B. (2013). Open Government Data, OECD Publishing: 61.
- Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Paris, Information Economics: 44.
- Weiss, P.N. & Y. Pluijmers (2004). Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impact. *In* Georg Aichholzer & Herbert Burkert (eds), Public Sector Information in the Digital Age. Between markets, public management and citizens' rights. Cheltenham, UK; Northampton, MA, US, Edward Elgar Publishing Ltd: 137-159.
- Welle Donker, F. & B. van Loenen (2015). Duurzame businessmodellen voor open data. Delft, Kenniscentrum Open Data: 60.
- Welle Donker, F., A. Bregt, M. Zweistra & B. van Loenen (2015). Effecten open data Liander 2014. Delft, Kenniscentrum Open Data: 18.
- Welle Donker, F., B. van Loenen & A. Bregt (2014). Open Data and Beyond II: Measuring the impact and modelling the data. Delft, Kenniscentrum Open Data, TU Delft: 40.

Lijst van afkortingen

ABS	Australian Bureau of Statistics
AHN	Actueel Hoogtebestand Nederland
API	Application Programming Interface
BAG	Basisregistratie Adressen en Gebouwen
BBP	Bruto Binnenlands Product
BIG Register	Beroepen in de Individuele Gezondheidszorg Register
BNP	Bruto nationaal product
BRT	Basisregistratie Topografie
CBS	Centraal Bureau voor de Statistiek
CH	Companies House
CIBG	Centraal Informatiepunt Beroepen Gezondheidszorg
DANS	Data Archiving and Networked Services
DECA	Danish Enterprise and Construction Authority
DUO	Dienst Uitvoerende Onderwijs
EU	Europese Unie
FCD	Floating Car Data
HR	Hogeschool Rotterdam
LOD	Linked open data
LV	Landelijke Voorziening
MBBL	Ministeriet for By, Bolig og Landdistrikter (Ministerie van Volkshuisvesting, Stedelijke en Plattelandszaken Denemarken)
Min OCW	Ministerie van Onderwijs, Cultuur en Wetenschap
Min. BIS	Ministry of Business, Innovation and Skills (Verenigd Koninkrijk)
Min. BZK	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties
Min. IenM	Ministerie van Infrastructuur en Milieu
Min. VenJ	Ministerie van Veiligheid en Justitie
MKBA	Maatschappelijke kosten-batenanalyse
NDW	Nationale Databank Weggegevens
NMB	National Meteorological Bureau
NODA	Nationale Open Data Agenda
PDOK	Publieke Dienstverlening op de Kaart
PGR	Provinciaal GeoRegister
PSI	Publieke sector informatie
RCE	Rijksdienst voor het Cultureel Erfgoed
RDF	Resource Description Framework
RDW	Dienst Wegverkeer
RIVM	Rijksinstituut voor Volksgezondheid en Milieu
TBL	Tim Berners-Lee, in dit geval een schaalverdeling van (open) data ontsluitingsniveaus
URI	Uniform Resource Identifier
VI	Verkeersinformatie
VK	Verenigd Koninkrijk
VM	Verkeersmanagement
VS	Verenigde Staten
Wob	Wet openbaarheid van bestuur
WODC	Wetenschappelijk Onderzoek- en Documentatiecentrum
WOZ	Waardering Onroerende Zaken

Bijlage 1 High value-datasets (lijst BZK)

Dataset	Status	Departement
IATI Dataset met transactiegegevens Ontwikkelingssamenwerking	Beschikbaar	BUZA
Uitgaven subsidies	Beschikbaar	BUZA
Maandelijke realisatie op jaarniveau, van de volledige financiële administratie van BUZA	Beschikbaar	BUZA
Rijksbegroting hoofdstuk XVII BHOS	Beschikbaar	BUZA
Wapenexportvergunningen	Gepland	BUZA
Resultaatfiches Ontwikkelingssamenwerking	In onderzoek	BUZA
Landenfiches Ontwikkelingssamenwerking	In onderzoek	BUZA
Wereldkaart Postennet	Gepland	BUZA
Lijst met alle posten van Nederland in het buitenland	Gepland	BUZA
Landelijke voorziening WOZ-waarden	In onderzoek	Fin
Begrotingsuitgaven en -ontvangsten (FIN/SG/FEZ)	Gepland	Fin
Jaargemiddelde luchtkwaliteit per postcode	In onderzoek	IenM
Actuele NO2-concentratie in Nederland	Beschikbaar	IenM
Actuele fijnstofconcentratie (pm10) in Nederland	Beschikbaar	IenM
Actuele Luchtkwaliteitsindex (LKI) in Nederland	Beschikbaar	IenM
Actuele ozonconcentratie in Nederland	Beschikbaar	IenM
Voertuig Open data	Beschikbaar	IenM
Actuele verkeersgegevens	Beschikbaar	IenM
Status verkeersgegevens	Beschikbaar	IenM
BRK (Kadastrale kaart)	Beschikbaar	IenM
Basisregistratie Topografie	Beschikbaar	IenM
NETPOS	In onderzoek	IenM
BAG (Basisregistratie Adressen Gebouwen)	Beschikbaar	IenM
Ruimtelijke plannen Wro: Ruimtelijkeplannen.nl	Beschikbaar	IenM
BGT (Basisregistratie Grootchalige Topografie)	Beschikbaar	IenM
Autoriteit Woningcorporaties	Beschikbaar	IenM
Toelating gewasbeschermingsmiddelen	Gepland	EZ

Bijlage 2 Analyse van alle onderzochte studies

Internationale studies

Studie	% BNP / Verhouding kosten : baten	Reflectie
ACIL Aus 2008	Baten: AUD 6,43-12,57mld/jr (0,6-1,2% BNP)	Geen kosten meegenomen. Onderzoek gebaseerd op extrapoleren van enkele casestudy's in 1 deelstaat naar meerdere sectoren en naar het gehele land.
ACIL NZ 2009	Baten: NZD 1,2 mln/jr (0,6% BNP)	Bouwt verder op ACIL 2008. Moet eerder als een beleidsadvies worden gezien dan als een robuuste KBA.
Cambridge Report 2008	Verschuiving van een kostendekkend regime naar een marginaalkostenregime voor basisdata zullen een positief effect op de welvaart hebben omdat het instellen van prijzen op kostendekking het vereiste heeft van een hoge mark-up, sterke en groeiende vraag naar data, en de kans op spillovers uit het gebruik van basisdata.	
Castelein 2010	Economische waarde NL geosector EUR 1,4 mld, 0,25% BNP, met een verwachte groei van 7% per jaar. De economische waarde van de private geosector werd op EUR 900 mln geraamd.	Geen kosten meegenomen. Economische waarde van publieke sector en onderzoekinstellingen = aantal fte's * EUR 100.000 per jaar. Voor de private sector is de omzet van de geosector bepaald via een enquête onder de leden van GeoBusiness Nederland door een extern bureau. GeoBusiness Nederland heeft later toegegeven dat het externe bureau de data niet goed had opgeschoond (o.m. een aantal bedrijven dubbel geteld).
Coote Value of GI to Local PS Delivery 2010	Ratio: 2,5-3,75:1	Baten uitgedrukt in % toename in productiviteit door betere toegang tot decentrale PSI.
CUPI UK 2006	Potentiële baten: GBP 1,1 mld/jr als er geen verlies van potentiële economische waarde zou zijn door: te hoge prijzen (GBP 20 mln/jr); oneerlijke concurrentie door datahouders (GBP 140 mln/jr), en falende exploitatie van PSI (GBP 350 mln/jr). Kosten: inkomstenverlies van datahouders (GBP 400 mln/jr); extra reguleringskosten (max. GBP 0,7 mld/jr). ratio: 1.100:401 ≈ 2,74:1	Ca. 50% van inkomsten van datahouders afkomstig van het bedrijfsleven, 45% van de publieke sector en 5% van de burgers. Veel van de gegenereerde PSI is resultaat van een wettelijke publieke taak (bijv. met Office), van proactieve taken (bijv. OS), van verplichte registers (bijv. CH), of een bijproduct (bijv. NHS).
DECA Address Data 2010	Baten: Directe economische baten EUR 62 mln (2005-2009), en EUR 14 mln in 2010; maatschappelijke baten ca. EUR 57 mln, incl. EUR 5 mln in bespaarde transactiekosten. Kosten: EUR 2 mln om data op niveau te krijgen + EUR 0,2 mln /jaar voor distributie. Ratio: 1 (EUR 0,2mln):70 (EUR14 mln)	EUR 1.3 mln van de eenmalige kosten door gemeenten en de rest door DECA betaald.
Deloitte BIS 2013	Baten: schatting maatschappelijke baten GBP 5 mld; lagere transactiekosten publieke sector GBP 50 mln/jr. Kosten: GBP 413 mln/jr verlies aan inkomsten. Ratio: 143:50 ≈ 2.86:1	Evaluatie van OFT (2006) voor het CUPI rapport, werd tegelijk met Shakespeare Review (2013) gepubliceerd. Baten zijn gebaseerd op ruwe schattingen. Rapport bevat anekdotische voorbeelden van maatschappelijke baten in de gezondheidszorg en de transportsector.
Deloitte Danish Geodata Agency Summary 2014	Baten: DKK 1,6 mld	Geen kosten meegenomen. Rapport beschrijft de nulmeting van de Deense Geodata Agency.
Denmark Base Data 2012	Baten: Overheid bespaarde DKK 38 mln in transactiekosten; bredere maatschappij DKK471 (2005-2009).	Er moeten kosten gemaakt worden om hoogwaardige basisregistratie op te zetten. De meerkosten van open data zijn niet gekwantificeerd.
EcoGeo Cookbook	Beste methode om netto waarde te bepalen is	Taxonomie van baten en kosten per doelgroep en

2010	betalingsbereidheid van eindgebruiker minus kosten van het product.	bijbehorende evaluatiemethode toe te passen op Quebec.
Freebairn Meteo 2002	Baten: Veelal in de landbouwsectoren voor besluitvorming. 1 rapport uit 1996: AUD 24 per huishouden per jaar/AUD 2,5 per ha. Kosten: Meeste nationale meteorologische bureaus berekenden tarieven die tot 40% van hun totale kosten dekken.	Artikel berekende de optimale hoeveelheid meteodiensten (Q) als functie van marginale kosten (MC) en baten (MB).
Houghton CBA of Data Provision 2011	Baten: ABS: direct AUD 4,97 mln/jr ; bredere impact ca. AUD 25 mln/jr. GA: direct transactiekostenbesparingen ca. AUD 375.000/jr. Bredere impact ca. AUD 15,5 mln/jr. Schatting totale baten: AUD 17,5 mln/jr. Kosten: ABS: netto AUD 3.53 mln/jr; GA: netto 1,3 mln /jr. Ratio: ABS: 25:3.53 ≈ 7:1; GA: 17,5:1,3 ≈ 13:1	ABS = Australian Bureau of Statistics GA = Geoscience Australia (topografische data) Baten/kosten = (kostenbesparingen voor aanbieder en voor gebruikers + verhoogd rendement op jaarlijkse uitgaven voor dataproductie) / (aanbieder- en gebruikerskosten)
Indecon Value Irish Geospatial Industry 2014	Baten: EUR 69,3-120 mld, incl. EUR 82 mln/jr aan kostenbesparingen binnen de overheid.	Geen kosten meegenomen. De kosten moeten toch worden gemaakt om aan eigen normen te voldoen.
Jetzek 2013	Baten open basisdata: Publieke sector: DKK 52 mln in 2015 tot DKK 260 mln in 2020; Private sector: DKK 500 mln/jr.	Geen kosten van open data meegenomen.
Krishnamurthy Liberating Data 2016	Kosten: > 640 competities met > USD 220 mln aan prijzengeld; data transformatiekosten, incl. digitalisatie van historische data en anonimiseren + ontwikkelen van tools.	Kosten die nodig zijn om open data te faciliteren en te promoten. Enige baten tot nu toe: er kunnen vergelijkingsanalyses worden uitgevoerd.
Lazo Value Weather Forecasts 2009	Baten: Geaggregeerde waarde USD 31,5 mld/jr. Kosten: USD 5,1 mld/jr (door publieke en private sector). Ratio: 31,5: 5,1 ≈ 6,2:1	Er is een nationale enquête met meer dan 1.500 respondenten gebruikt om te bepalen waar, wanneer en hoe vaak weersverwachtingen werden opgevraagd, en wat hun percepties, hun gebruik en hun waardering van de informatie waren.
McKinsey Open Data 2013	Baten: USD 3 triljoen wereldwijd, waarvan USD 900 mld in EU27. Kosten: Genoemd maar niet gekwantificeerd.	De enige studie die de mondiale waarde van open data tracht te kwantificeren. Deze studie neemt de potentiële waarde van alle open data (data uit de publieke en de private sector, en onderzoeksdata) in beschouwing.
MEPSIR 2006	PSI markt is gem. EUR 27 mld en draagt ca. 0,25% bij aan BNP van EU25+NO.	Resultaten waren waarschijnlijk te optimistisch geschat, en waren eerder EUR 3-5 mld i.p.v. 27 mld (te Velde 2009).
NASA 2013	Baten: Vermeden inkomstenverliezen en schade aan vliegtuigen na Eyjafjallajökull vulkaanuitbarsting USD 72 mln. Malaria Early Warning System in Afrika: 105 minder slachtoffers in Botswana, en potentieel 660.000 minder slachtoffers in 28 Sub-Saharalanden.	Had USD 200 mln kunnen zijn indien data direct beschikbaar waren geweest i.p.v. een week later. Kosten worden niet in beschouwing genomen.
Nilsen Economic Theory 2010	Prijsbepaling voor PSI kan nooit objectief worden bepaald, en iedere prijs boven de marginale kostprijs is economisch inefficiënt.	Uitgebreide literatuurstudie van economische theorieën die ten grondslag liggen aan PSI.
ODI Cost of OGD 2014		Opsomming van kosten benodigd om open data te faciliteren en capaciteit te bouwen. Geen kwantificatie van de kosten en geen baten genoemd.
OECD Broadband PSI & Content 2006	Verkoop van commerciële data door PSI-aanbieders heeft tot oneerlijke concurrentiepraktijken geleid in een aantal sectoren.	Dit rapport vormde de theoretische basis voor het Vickery rapport (2011) en beschrijft de PSI waardeketen die in vele volgende publicaties is hergebruikt. De PSI waardeketen geeft inzicht.

Omidyar OD for G20 targets 2014	Baten: 1,0% BNP van Australië AUD 15,4 mld/jr en potentiële economische waarde van alle OD: tot AUD 64 mld/jr. Geaggregeerde directe en indirecte waarde AUD 25 mld/jr (samenvatting van eerdere studies). Kosten: AUD 8-10 mld/jr Ratio: 15,4:8-10 \approx 1,9-1,54:1 (dit rapport) 25:8-10 \approx 3,13-2,5:1 (andere studies).	De opstellers van dit rapport denken dat de uitkomsten van McKinsey (2013) zelfs aan de conservatieve kant zijn. De totale baten van alle Open Data (data uit de publieke en de private sector ((bijv. seismologische data uit mijnbouw)), en wetenschappelijke data) worden hoger geraamd.
Open Data combats Tax Fraud	Baten: Ca. CAD 1 miljard/jaar door open data van Canadese Belastingdienst.	Analyse van nationaal donatiesregister van 2005 toonde aan dat er bij frauduleuze goede doelenorganisaties ca. CAN 1 mld/jr, en totaal CAN 3,7 mld aan donaties waren ondergebracht.
OS Open Data 2013	Baten: GBP 10,2-24,1 mln BNP per jaar. Kosten: verlies van inkomsten (OS licentiekosten) + GBP 3,72 mln/jr negatief effect op exports. Ratio: 10,2-24,1:3,72 \approx 2.74-6,48:1.	Data van 1 data-aanbieder. De grootste hergebruiker van OS data is Google. Onmogelijk om de economische waarde van dergelijke partijen te bepalen omdat deze geen inzage in financiën geven en geen belasting betalen in binnenland. Maatschappelijke baten zijn niet meegenomen in dit onderzoek.
Oslo Economics 2011	Baten: Topo: NOK 70 mln voor bedrijfsleven op de korte termijn, en consumentensurplus mogelijk tot NOK 73,5 mln; OG: NOK 75 mln korte termijn en consumentensurplus mogelijk tot NOK 84 mln; HR: geen korte termijn impact (geen huidige hergebruikers) en consumentensurplus mogelijk tot NOK 44 mln; Juris: maatschappelijke baten van gratis toegang. Kosten: Topo: NOK 70 mln (gederfde inkomsten); OG: NOK 75 mln (gederfde inkomsten); HR: NOK 35 mln (gederfde inkomsten); Juris: transformatiekosten (niet gekwantificeerd). Ratio: 1,12:1 (consumentensurplus: gederfde inkomsten).	Open Topografische Data + open Onroerend Goed (OG) data + open Handelsregister (HR) + open jurisprudentie data (Juris). Bij de kosten zijn alleen de gederfde inkomsten gekwantificeerd. Ook genoemd maar niet gekwantificeerd zijn extra kosten om met vragen om te gaan, ondermijning van samenwerking in Norge Digitalt, data transformatiekosten voor anonimiseren, etc. + gederfde inkomsten van huidige intermediair van juridische data.
OSTP Canada 2011	Baten Smart Bay: CAD 2,225/jr Kosten Smart Bay: CAD 7,1mln Ratio: 3,19:1 na 1 jaar; 1:1,57 na 5 jaar.	Overheidsinvestering van CAD 2 mln in Smart Bay was eenmalig met een extra subsidie van CAD 5 mln over 5 jaar. Er is geen financiering voor de lange termijn.
Oxera 2013	Baten: Directe effecten van geodiensten USD 150-270 mld mondiaal (0,2% van mondiaal BNP), bijv. kortere reistijden en brandstofbesparingen USD 22 mld/jr. GIS voor besluitvorming USD 12 mld/jr; GPS bespaart ca. USD10 mld/jr in transport & logistiek.	Geen kosten meegenomen.
Pettifer 2009	Kosten: EUR 290 mln/jr aan gederfde inkomsten voor de nationale overheden als gevolg van restrictieve datalicenties van nationale meteorologische bureaus.	De VS-markt voor meteo groeit 14x sneller dan de EU-markt.
Pham GPS 2011	Baten: Economische waarde van GPS (commerciële producten/diensten, productiviteitswinst en kostenbesparingen) USD 68-122 mld/jr of 0,5-0,9% van US BNP. Kosten: Investeringskosten in GPS constellation door US overheid: USD 35 mld + structurele kosten USD 1 mld/jr. Ratio: 1,88-3,88:1	Rapport beschrijft de potentiële kosten veroorzaakt door LightSquared, een bedrijf dat van plan is om een 4G-LTE wifi breedband netwerk op te zetten, waardoor de bestaande GPS-signalen waarschijnlijk zullen worden verstoord. Indien dit zou gebeuren, dan zouden de economische kosten: USD 96 mld/jr zijn aan gederfde inkomsten uit GPS-producten, hardware en verzonken investeringskosten en R&D. Bij 50% uitval van GPS zou dit USD 48 mld/jr zijn.
PIRA 2000	Baten: EUR 68 mld/jr Kosten: EUR 9,5 mld/jr Ratio: 9,5:68 \approx 1:7,16	Gebaseerd op extrapoleren van enkele casestudy's voor de EU15, en het includeren van de filmindustrie in de US economische waarde.
Pollock economics of PSI 2009	Baten: GBP 1,6-2,0 mld/jr (middenpunt) tot 4,5-6 mld (bovengrens)	Baten=(2/5)F λ Waar F= inkomsten uit kostendekkend regime,

Pollock Welfare Gains 2011	Kosten: GBP 400-550 mln/jr Ratio: 3,37:1 tot 12,6:1	λ =de 'multiplier' en ϵ =elasticiteit v.d. vraag. Bovengrens: $\lambda=8$, $\epsilon=3.5$ Middelpunt: $\lambda=5$, $\epsilon=2$
POPSIS 2011	Baten: Hogere aantallen hergebruikers en uit nieuwe categorieën vooral uit MKB; efficiëntieslagen bij datahouders; lagere transactiekosten. Kosten: (niet gekwantificeerd): verlies aan inkomsten voor datahouders (relatief laag t.o.v. totale begroting); verandercosten om bestaande processen en mindsets te veranderen.	Anekdotisch / op dataset niveau baten/kosten. Markt voor smartphone apps werden in 2011 geschat op USD 35 miljard in 2015, waarbij PSI worden hergebruikt, vooral voor weer, reis en transport apps. Kosten voor Open Data portalen varieerden van USD 9,2 miljoen voor data.gov tot een gemiddelde van EUR 0,1 miljoen voor Europese nationale en regionale initiatieven.
PWC Aus data-driven economic growth 2014	Baten uit data gedreven innovatie AUD 12,6 mld voor de overheid; transactiekosten besparingen t.g.v. real-time OV-data: 5 min/dag per reiziger (ca. AUD 400 mln/jr in NSW).	Geen focus op open data. Baten zouden hoger zijn als er beter opgeleid personeel zou zijn.
PwC ODCS impact 2015	Baten (verwacht): over komende 3 jaar: 17 tot 141 extra banen; GBP 5,3-10,8 mln in Bruto Toegevoegde Waarde (in Netto Contante Waarde-termen); GBP 161-302 mln bredere maatschappelijke en economische impact (in NCW-termen). Kosten: GBP 1,2 mln, incl. prijzengeld en ondersteuning. Ratio: 4,42-9:1 (bruto toegevoegde waarde); 134-251:1 (bredere impact).	Kosten-batenanalyse van de Open Data Challenge Series in de UK.
SCGOE Canada 2014	Baten: CAD 100 mld/jr (volgens McKinsey 2013).	Geen kosten meegenomen.
Shakespeare Report 2013		Geeft aan dat de baten en kosten niet exact te bepalen zijn vanwege de complexiteit.
Ubaldi OGD Initiatives 2013	Maatschappelijke baten: Worden gerealiseerd wanneer er voldoende interesse in en capaciteit voor hergebruik aanwezig is. Kosten: Infrastructuur + transformatiekosten voor open data + verandercosten/training + stimuleren + capaciteit bouwen. Veel van die kosten moeten in ieder geval gemaakt worden om te voldoen aan openbaarheidsverplichtingen (Wob) en/of delen tussen overheden.	Geen kwantiteit.
Vickery Review 2011	Baten: EUR 40 mld voor EU27 Kosten: gederfde inkomsten < 1% van PSI-houder begroting te compenseren door centrale overheid i.c.m. 'updater' model.	Gebaseerd op extrapoleren van Castelein <i>et al.</i> (2010) en ACIL Tasman (2008; 2010). En die rapporten waren ook al gebaseerd op extrapoleren van enkele casestudy's.
Weiss Borders in Cyberspace 2002	Geen baten en kosten gekwantificeerd.	Kostendekkende regimes werken op de lange termijn niet omdat de private afnemersmarkt te klein is, het inefficiënt is om andere overheden te factureren, de werkelijke inwinkosten toch niet kunnen worden gedekt en hoge prijzen voor informatie tot oneerlijke handelspraktijken leiden.
World Bank OD for Econ. Growth 2014	Anekdotisch voorbeeld van baten van open metadata voor geodata (als INSPIRE verplichting) in Catalonië: 4 jaar ontwikkelingskosten terugverdiend in 6 maanden door efficiëntieslagen binnen de overheid.	Overzicht van bestaande literatuur geeft aan dat de economische effecten van open data significant tot zeer significant zullen zijn.

Nationale Studies

de Vries Wah Kos'dah dan 2014	Baten: Kostenbesparingen KNMI ca. EUR 80.000/jr vanaf 2015/17; oostenbesparingen Kadaster: ca. EUR 115.000/jr vanaf 2014. Kosten: Infrastructurele kosten: ca. EUR 5-20.000/jr; veranderkosten: ca. EUR 50.000 (eenmalig); beheerkosten: EUR 10-45.000/jr.	Er zijn verschillen tussen organisaties wier kerntaak het is data te verstrekken en bijv. gemeenten. Bij de 1 ^e categorie betreft het terabytebestanden, maar worden de baten eerder zichtbaar. Bij gemeenten zijn de baten voornamelijk onbekend.
KC-OD Duurzame Businessmodellen 2015	Baten zijn vooral te vinden in interne efficiëntieslagen, en leiden in een aantal gevallen tot meer vraag naar maatwerk/hoogwaardigere data via betaaldiensten. De kosten van gedeelde inkomsten zijn relatief laag in verhouding tot totale begroting.	
Laatste Meter Actueel Verkeersdata 2016	Baten: Ca. EUR 1 mld aan externe kosten, effectiever gebruik van infrastructuur en schoner milieu. Kosten (voor de transportsector): Ca EUR 5,5 mln Ratio: 181:1	Kosten van beschikbaar maken van real-time verkeersgegevens zijn niet meegenomen.
Liander OD 2014	Baten: Nog te vroeg om waar te nemen. Kosten: Infrastructurele kosten: 1x extra server; verander/implementatiekosten: 3,3 FTE; beheerkosten: geen extra kosten.	Was een nulmeting voordat Liander open data beschikbaar stelde.
Liander OD 2015	Baten: Lagere transactiekosten bij aanbieder (kan ook gevolg zijn van andere efficiëntieslagen; ,meer samenwerking met andere partijen; beter imago	De kosten zijn een integraal onderdeel van de operationele kosten.
Open BRT Effecten 2013	Baten: 7.000 manuren bespaard aan uitleveren van data. Kosten: EUR 250.000 netto Ratio: 3,5:1	Uurloon is gebaseerd op aanname in vervolgmetingen (EUR 1.000/dag = EUR 125/uur).
Open BRT Effecten 2014	Baten: Efficiëntieslagen, ca. EUR 1.500 MH*EUR 125/u = EUR 187.500; toename in extern gebruik drievoudig; economisch effect: EUR 9 mln in 2013 Kosten: Hogere beheerkosten (PDOK) (niet gekwantificeerd).	Gebaseerd op een enquête. De 58 respondenten besteedden totaal 1105,2 dagen aan managen en toepassen van BRT * 1000/dag = 1,1 mln investeringswaarde. Als deze waarde wordt geëxtrapoleerd naar heel Nederland, dan zou de totale investeringswaarde EUR 11.570-14.467 mln zijn minus de bedrijven die de BRT al eerder gebruikten, dan komt de netto minimum investeringswaarde op EUR 9,2 mln t.g.v. open data.
Open BRT effecten 2015	Baten: Kwaliteitsslag door terugmeldingen, vooral van buiten de overheid. Toegenomen economisch effect: + EUR 4,4 mln t.o.v. 2013, nu ca. EUR 13,6 mln.	Aantal zakelijke gebruikers lijkt te stabiliseren. Kosten worden niet meegenomen.
RDW OD PoC 2013	Baten: Kwaliteitsslag door meer terugmeldingen; meer (her)gebruik van open data. Kosten: Projectkosten EUR 92.000; Investeringskosten EUR 90.000; Exploitatiekosten EUR 100.000/jr.	Open data heeft nauwelijks een effect gehad op de data betaaldiensten. De baten zijn niet gekwantificeerd.

Overzicht van (inter)nationale studies met percentages als uitkomst

Studie	Kwantitatieve uitkomst (% BNP)	Bereik
ACIL Tasman 2008	Baten: AUD 6,43-12,57mld/jr. (0,6-1,2% BNP) Geen kosten meegenomen.	Australische economie.
ACIL Tasman 2009	Baten: NZD 1,2 mln/jr. (0,6% BNP) Geen kosten meegenomen.	Nieuw-Zeelandse economie
Baudel 2016	Baten: Direct: potentiële productiviteitswinst 18% afstand- en 11% tijdsbesparing Indirect: betere luchtkwaliteit, etc.	Logistiek binnenstedelijk transport
Castelein 2010	Economische waarde EUR 1,4 mld, 0,25% BNP, waarvan	Nederlandse geosector.

	bedrijfsleven EUR 900 mln Geen kosten meegenomen.	
Denmark Geodata Agency 2014	Baten: DKK 1,6 mld Geen kosten meegenomen.	OD nulmeting Deense Geodata Agency.
Denmark Base Data 2012	Baten: Overheid: DKK 38 mln in transactiekosten en maatschappij DKK 471 mln. Meerkosten van open data zijn niet gekwantificeerd.	Deense Basisregistraties
Indecon Value Irish GI Industry 2014	Baten: EUR 69,3-120 mld, incl. EUR 82 mln/jr. aan kostenbesparingen binnen de overheid.	Ierse geosector.
Jetzek 2013	Baten open basisdata: Publieke sector: DKK 52 tot 260 mln (2015 tot 2020). Private sector: DKK 500 mln/jr.	Deense Basisregistraties.
Koski 2011	Baten: bedrijvengroei 15% hoger in landen waar marginale kostenregime wordt gehanteerd i.p.v. kostendekkend regime, vooral voor MKB.	Econometrische analyse van 15 landen 2000-2007.
McKinsey Report 2013	Baten: USD 3 triljoen wereldwijd, waarvan USD 900 mld in EU27. Datatransformatie kosten worden genoemd maar niet gekwantificeerd.	Potentiële waarde van alle open data (publieke en private sector data, en onderzoeksdata).
MEPSIR Study 2006	Waarde PSI markt gem. EUR27 mld, ca. 0,25% BNP van EU25+NO.	Economische waarde PSI markt EU25+NO.
NASA 2013	Baten: Vermeden inkomstenverliezen en vliegtuigschade na Eyjafjallajökull vulkaanuitbarsting USD 72 mln, had 200 mln kunnen zijn indien data direct beschikbaar waren geweest i.p.v. een week later. Malaria Early Warning System in Afrika: 105 minder slachtoffers in Botswana, en potentieel 660.000 minder slachtoffers in 28 Sub-Saharalanden.	Aardobservatie data.
Open Data combats Tax Fraud	Baten: Ca. CAD 1 mld/jr., totaal CAD 3,7 mld door open data van Canadese Belastingdienst.	Onderzoeksjournalistiek met open belastingdata.
Oxera 2013	Baten: Directe effecten USD 150-270 mld (0,2% van mondiale BNP) en Indirecte effecten: bijv. voor GPS-data USD 10 bjd/jr. kostenbesparingen voor logistieke sector en USD 22 bjn/jr. voor consumenten.	Mondiale impact van geodiensten.
Pettifer 2009	Kosten: EUR 290 mln/jr. gederfde inkomsten voor nationale overheden t.g.v. restrictieve licenties van NMB's.	Vergelijking private meteo sector EU en VS.
Vickery Review 2011	Baten: EUR 40 mld voor EU27 Kosten: gederfde inkomsten < 1% van PSI-houder begroting en te compenseren door centrale overheid.	Literatuur review KBA studies m.b.t. EU27.
de Vries Wah Kos'dah dan 2014	Baten: Kostenbesparingen KNMI ca. EUR 80.000/jr. vanaf 2015/17; Kadaster: ca. EUR 115.000/jr. vanaf 2014. Voor andere zijn baten niet zichtbaar. Kosten: Infrastructurele kosten: ca. EUR 5-20.000/jr; veranderkosten: ca. EUR 50.000 (eenmalig); beheerkosten: EUR 10-45.000/jr.	Kosten-batenanalyse voor KNMI, Kadaster, RCE, gemeente Rotterdam en Enschede.

Overzicht van (inter)nationale studies met ratio als uitkomst

Studie	Verhouding baten : kosten	Bereik
Coote GI Value to Local PS 2010	Ratio: 2,5-3,75:1	Geografische data decentrale overheden in VK.
CUPI UK 2006	Potentiële baten: GBP 1,1 mld/jr.) Kosten: gederfde inkomsten van datahouders GBP 400 mln/jr. + extra handhavingskosten GBP 0,7 mld/jr. Ratio: 1.100:401 ≈ 2,74:1	Economische waarde van PSI in VK.
DECA Address Data 2010	Baten: Directe economische baten EUR 62 mln (2010) en maatschappelijke baten ca. EUR 14 mln/jr. Kosten: EUR 2mln eenmalig en EUR 0,2 mln/jr. structureel. Ratio: 1 (EUR 0,2mln):70 (EUR14mln)	Evaluatie na 8 jaar open adresdata in Denemarken.
Deloitte BIS 2013	Baten: schatting maatschappelijke baten GBP 5 mld	Evaluatie van OFT (2006) CUPI

	lagere transactiekosten publieke sector GBP 50 mln/jr. Kosten: GBP 413 mln/jr. verlies aan inkomsten Ratio: 143:50 \approx 2.86:1	rapport, VK.
Houghton CBA of Data Provision 2011	Baten: ABS: direct AUD 4,97 mln/jr. + bredere impact ca. AUD 25 mln/jr. GA: direct transactiekosten besparingen ca. AUD 375.000/jr. Bredere impact ca. AUD 15,5 mln/jr. Kosten: ABS: netto AUD 3.53 mln/jr.; GA: netto 1,3 mln/jr. Ratio: ABS: 25:3.53 \approx 7:1 GA: 17,5:1,3 \approx 13:1	Kosten-batenanalyse Australian Bureau of Statistics (ABS) en Geoscience Australia (GA) topografische data.
Lazo Value Weather Forecasts 2009	Baten: Geaggregeerde waarde USD 31,5 mld/jr. Kosten: USD 5,1 mld/jr. (publieke en private sector) Ratio: 31,5: 5,1 \approx 6,2:1	Meteodiensten in US.
Omidyar OD for G20 targets 2014	Baten: 1,0% BNP (AUD 15,4 mld/jr.) en potentiële economische waarde van alle open data: tot AUD 64 mld/jr. Kosten: AUD 8-10 mld/jr. Ratio: 15,4:8-10 \approx 1,9-1,54:1	G20 Open Data (publieke en private sector data, en wetenschapsdata).
OS Open Data 2013	Ratio: 10,2-24,1:3,72 \approx 2.74-6,48:1	Kosten-batenanalyse open topografische data in VK.
Oslo Economics 2011	Baten: direct NOK 145 mln voor bedrijfsleven korte termijn, consumentensurplus mogelijk tot NOK 201,5 mln + maatschappelijke baten van gratis toegang. Kosten: NOK 180 mln (gederfde inkomsten) Ratio: 1,12:1	Kosten-batenanalyse open data in Noorwegen.
OSTP Canada 2011	Baten Smart Bay: CAD 2,225/jr. Kosten Smart Bay: CAD 7,1mln Ratio: 3,19:1 na 1 jaar en 1:1,57 na 5 jaar	Casestudy open oceanografische observatiedata Canada.
Pham GPS 2011	Baten: Economische waarde van GPS USD 68-122 mld/jr., 0,5-0,9% van US BNP Kosten: Investeringskosten door US overheid: USD 35 mld + structurele kosten USD 1 mld/jr. Ratio: 1,88-3,88:1	Kosten-batenanalyse GPS industrie VS, incl. potentiële schade door storingen.
PIRA 2000	Baten: EUR 68 miljard/jaar Kosten: EUR 9,5 miljard/jaar Ratio: 9,5:68 \approx 1:7,16	Economische waarde PSI in EU15.
Pollock economics of PSI 2009 & Pollock Welfare Gains 2011	Baten: GBP 1,6-2,0 mld/jr. tot 4,5-6 mld Kosten: GBP 400 tot 550 mln/jr. Ratio: 3,37:1 tot 12,6:1	Kosten-batenanalyse PSI in VK.
PwC ODCS impact 2015	(verwachte) baten over komende 3 jaar: 17 tot 141 extra banen, direct: GBP 5,3-10,8 mln; GBP 161-302 mln bredere maatschappelijke impact Kosten: GBP 1,2 mln (prijzengeld en ondersteuning) Ratio: 4,42-9:1 (direct) en 134-251:1 (bredere impact)	Kosten-batenanalyse van de Open Data Challenge Series in de VK.
Laatste Meter Actueel Verkeersdata 2016	Baten: Ca EUR 1 mld aan externe kosten, effectiever gebruik van infrastructuur en schoner milieu. Kosten (voor de transportsector): Ca EUR 5,5 mln Ratio: 181:1	Open data voor transport sector.
Open BRT Effecten 2013	Baten: 2.200 manuren bespaard \approx EUR 87,5000 Kosten: EUR 250.000 netto Ratio: 87.500:250.000 \approx 0,35:1 (negatief voor Kadaster)	Open topografische data in Nederland.

Bijlage 3 Casestudy-effecten van open data op microniveau

In deze bijlage zijn een zestal casestudy's opgenomen waarin de effecten van open data op datasetniveau zijn beschreven.

B.1. Effecten basisregistratie topografie open data

De basisregistratie topografie (BRT) is sinds 1 januari 2012 als open data beschikbaar. Bregt *et al.* (2013, 2014) en Grus *et al.* (2015) hebben tussen 2013 en 2015 de effecten in kaart gebracht.

Het onderzoek uit 2013 en 2014 is gebaseerd op respectievelijk 56 (in 2013) en 148 (in 2014) responses op een enquête beschikbaar gesteld via de website van de basisregistratie topografie (voor de meting van 2015 zijn geen aantallen respondenten genoemd). De enquêteresultaten lieten zien dat het aandeel zakelijke en particuliere gebruikers zeer sterk is toegenomen sinds de BRT als open data beschikbaar is. Van het aandeel 17% zakelijk en 3% particulier in 2012, tot 41% zakelijk en 5% particulier in 2013, tot 41% zakelijk en 22% particulier in 2014. In 2015 leken deze percentages te stabiliseren, met een lichte afname in % particulier. Ook zijn de toepassingsgebieden in 2014 sterk verruimd. Het aantal zakelijke gebruikers is twee jaar na het openstellen van de BRT ongeveer vervijfvoudigd (Bregt *et al.* 2014).

“Door het open databeleid is een bedrag van circa €0,5 miljoen aan omzet bij het Kadaster weggevalen. In 2012 is dit deels (€0,25 miljoen) gecompenseerd in het budget van [het Ministerie van] IenM. Het budget is echter in 2013 weer gedaald ten gevolge van de algemene bezuinigingen. De totale kosten van de BRT (productie, levering, beheer en ontwikkeling) zijn in 2013 licht gestegen en ook de kosten van PDOK zijn toegenomen. Daarmee zijn de kosten boven budget uitgekomen.” (Bregt *et al.* 2014, p.28)

De capaciteit die bedrijven in het jaar 2013 besteden aan verwerking van de BRT wordt op basis van een online enquête geschat op €11,5 tot 14,5 miljoen. Het opendataeffect werd in 2014 geschat op ten minste €9 miljoen en in 2015 zelfs op €13,6M. Op het gebied van type toepassingen is er tussen 2013 en 2015 een lichte afname in Ruimtelijke Ordening (van 50% naar 46%) en Milieu, Bodem en Natuur (van 48% naar 42%). De toepassing Cultuur en Recreatie is gestegen van 18% in 2013 naar 32% in 2015, mede omdat particulieren de BRT steeds vaker gebruiken. De BRT wordt steeds vaker gebruikt om deze voor analyses te koppelen met andere gegevens, en wordt steeds minder vaak als onbewerkt bestand doorgeleverd (van 30% in 2013 naar 10% in 2015. Voor de meerderheid van de afnemers zijn de kosten van het verkrijgen, beheer en gebruik in de afgelopen periode afgenomen.

Bij de wijze waarop het 'effect' van ten minste €9 miljoen (2014) en 13,6 miljoen (2015) wordt berekend kunnen vraagtekens worden gezet. Op basis van de antwoorden op de vraag: "Hoeveel dagen heeft u het afgelopen jaar besteed aan het verkrijgen, beheer en toepassen (ook voor klanten) van de open basisregistratie topografie?" heeft men een inschatting gemaakt. De zakelijke gebruikers die de enquête hebben ingevuld (in 2013 48 = 41% van 148 respondenten en in 2014 40 zakelijke gebruikers) besteden gemiddeld 19,06 dagen per jaar om BRT bij te houden en te verwerken (in totaal 1105,5 dagen). Men doet de volgende aannames om een inschatting van de 'waarde' van open data te berekenen:

1. Het aantal overheidsgebruikers is in de periode tussen het openstellen van de BRT in 2012 en de meting in het eerste kwartaal van 2014 niet veranderd;
2. De steekproef in de enquête is 'unbiased' wat betreft de percentageverhouding tussen overheden en bedrijven;
3. De verdeling van het aantal bestede dagen door bedrijven is in de totale populatie gelijk aan de verdeling in de steekproef.

Op basis daarvan redeneert men als volgt:

1. Het aantal overheidsorganisaties dat in 2012 gebruik maakt van de BRT bedraagt 400-500. Op basis van aanname (1) bedraagt het aantal overheidsorganisaties in 2013 dus 400-500.
2. Binnen de steekproef vormen de overheidsorganisaties 27% van de respondenten (i.c. 40 overheden). Het percentage bedrijven bedraagt 41% in de steekproef. Dit betekent een aantal van $607-759 (= 41/27 * 400-500)$ bedrijven in de gehele populatie die de BRT gebruiken.
3. Op basis van aanname (3) is het aantal dagen besteed door bedrijven in de gehele populatie $11570 - 14467 (= 19,06 * 607-759)$. Op basis van €1.000 per dag komt men op een investering door bedrijven van €11,5 - 14,5 miljoen. De onderzoekers merken hierbij op dat dit bedrag niet volledig kan worden toegeschreven aan het effect van open data omdat voor 2012 er al bedrijven waren die investeerden in de BRT (ongeveer 100-150 bedrijven, wat overeenkomt met 20% van de bedrijven). Hiervoor corrigerend komt men op een opendataeffect van ten minste €9 miljoen voor 2013.

Voor 2014 wordt het opendataeffect van de BRT voor het bedrijfsleven op ten minste €13,6 miljoen geraamd, wat overeenkomt met 80% van de €17-21 miljoen geïnvesteerd in de BRT (Grus *et al.* 2015, p.5).

De onderzoekers concluderen terecht dat nader onderzoek gewenst is. Zo gaat men er ten eerste vanuit dat er gerekend mag worden met de verhouding geënquêteerden afkomstig uit de zakelijke markt en de publieke sector. Dat deze verhouding 41%/27% een juiste weerspiegeling is van de gebruikers van de open basisregistratie topografie is op geen enkele wijze hard gemaakt. Verder is het niet uitgesloten dat de respondenten de juiste categorie gebruiker hebben aangevinkt en/of dat de antwoorden onjuist zijn ingevuld. Verder gaat men ervan uit dat per bedrijf slechts één gebruiker de enquête heeft ingevuld. Ook gaat men ervan uit dat als men gegevens heeft gezocht en deze niet geschikt bleken, dit als een positief effect van open data kan worden beschouwd (want anders had men toch niet gezocht).

Ten slotte wordt er gerekend met een dagtarief van €1.000 en wordt het gebruik per definitie als een baat gezien. Als we uitgaan van een situatie waarbij een junior medewerker tegen een kostprijs (uurloon + overhead) van €50 per uur (€400 per dag), dan wordt de baat 250% lager. Zelfs bij een kostendekkend uurtarief van €100 voor vergelijkbare werkzaamheden door een senior medewerker¹³⁴ zal de baat 125% lager zijn. Als deze juniormedewerker 20 dagen de open data heeft getest en tot de conclusie komt dat hij de open data niet kan gebruiken, dan is zijn investering in tijd toch in de berekening meegenomen.

Ondanks dat op basis van deze onderzoeken de conclusie kan zijn dat de effecten van open-basisregistratie positief zijn, moet de inschatting van de investeringskosten niet zonder meer als reëel worden bestempeld. Door het openstellen van de BRT is de gebruikerscommunity sterk gegroeid, en onderneemt het Kadaster initiatieven om deze community actief te betrekken. Begin

¹³⁴ CBS hanteert een uurtarief van €100 voor dergelijke werkzaamheden, zie Sectie 7.3. Voor kostendekkende uurtarieven van de Rijksoverheid zie <http://wettenpocket.overheid.nl/portal/add57ad6-77e3-4aca-9666-6d13d992f9ae/document/Handleiding%20Overheidstarieven%202016.pdf>.

mei 2012 is het Kadaster gestart met een BRT LinkedIngroep, en in het vierde kwartaal van 2015 was het ledenaantal gegroeid naar 663 (Grus *et al.* 2015).

B.2. Effecten Actueel Hoogtebestand Nederland (AHN)

Het Actueel Hoogtebestand Nederland is sinds maart 2014 beschikbaar als open data. Bregt, *et al.* (2016) hebben de effecten van open data middels literatuurstudie, een enquête onder gebruikers en interviews met gebruikers en aanbieders onderzocht en concluderen:

“Het AHN levert een grote bijdrage aan het realiseren van organisatiedoelen voor ongeveer de helft (45%) van de enquête-respondenten. Daarbij wordt door meer dan de helft van de zakelijke gebruikers (56%) aangegeven dat het openstellen een positief effect heeft op de omzet van de organisatie. Bovendien is een toenemende diversiteit aan toepassingen van het AHN waar te nemen. Waar in het verleden waterbeheersing, archeologie en wetenschappelijk onderzoek de boventoon voerden, komen nu ook de domeinen Bouw en infrastructuur, Milieu Bodem en Natuur en Ruimtelijke Ordening naar voren.

Het aantal zakelijke gebruikers is sterk gegroeid na het openstellen van de dataset, namelijk een stijging van 80%. De oorzaak hiervan zijn de lage huidige transactiekosten. De dagen die de geënquêteerden aan het AHN besteden in 2015 is omgerekend naar geld ongeveer €2 miljoen, waarvan de zakelijke sector verantwoordelijk is voor €639.000. Geëxtrapoleerd naar alle AHN-gebruikers in Nederland is dit een totale directe financiële investering van ca. €5.5 miljoen. Zakelijke gebruikers investeren in totaal rond de €1,8 miljoen in AHN gerelateerde producten en diensten.” (p.6)

Daarnaast worden effecten als een positieve invloed op de reputatie van de aanbieder van de open data waargenomen, blijkt het aantal meldingen bij de helpdesk in aantal (minder) en van aard te zijn veranderd met betrekking tot bestellingen en producten naar informatievragen over het AHN.

De wijze waarop de €5,5 miljoen aan directe financiële investering tot stand is gekomen verdient enige aandacht. De onderzoekers nemen het aantal zakelijke gebruikers van het AHN voordat het bestand als open data beschikbaar werd gesteld als startpunt (54 ‘oude’ zakelijke gebruikers). Dit aantal delen de onderzoekers door het aantal ‘oude’ zakelijke gebruikers dat de enquête hebben ingevuld (19 ‘oude’ zakelijke gebruikers hebben dit gedaan). Deze verhouding (54/19) wordt gebruikt als representatief voor de gehele populatie gebruikers (incl. overheid, burgers en nieuwe zakelijke gebruikers). Vervolgens is in de enquête gevraagd naar het aantal dagen dat een gebruiker heeft besteed aan het werken met het AHN. Dit totaal aan bestede dagen door overheid, zakelijke gebruikers en burgers is vermenigvuldigd met de verhouding 54/19. De uitkomst is vermenigvuldigd met “het gemiddeld dagtarief in deze sector” à €1.000 per dag. Zo komen de onderzoekers op een bedrag van €5,5 miljoen aan directe financiële investering door gebruikers van het AHN. Ten slotte zou dit met een 30% belastingtarief tot minimaal €1,65 miljoen per jaar aan belastinginkomsten leiden (Bregt *et al.* 2016, p.27).

Naast de vraag of het aantal bestede dagen kan worden vertaald in een directe financiële investering, is er onzekerheid omtrent de mate van representativiteit van het aantal ingevulde enquêtes in verhouding tot de oude zakelijke gebruikers en de mate waarin deze verhouding kan worden toegepast op andere gebruikersgroepen. Indien het tarief van €1000 per dag, (€125/uur) het gemiddelde dagtarief in deze sector is voor zakelijke gebruikers, is dit zeker niet zonder meer het geval voor gebruikers binnen de overheid en/of burgers. Kronenburg *et al.* (2012, p.85) hanteert bijvoorbeeld een uurtarief van €50/uur binnen de provinciale overheid voor vergelijkbare taken. Een

andere verhouding, onderscheid in gebruikersgroepen en het hanteren van andere tarieven kunnen de conclusies van de onderzoekers zeer sterk beïnvloeden. Dit los van het feit dat in ieder geval burgers geen belasting betalen over de tijd die zij hebben besteed aan het AHN.

B.3. Effecten Adresdata Denemarken

In 2002 besloot de Deense overheid dat vrije en onbeperkte toegang tot hoogwaardige adresdata van belang was voor de samenleving en voordelen zou opleveren voor de publieke sector en het bedrijfsleven. In het voorjaar van 2010 heeft de Danish Enterprise and Construction Authority (DECA) door een onafhankelijk consultancybedrijf laten onderzoeken wat de baten 8 jaar na dato waren van het vrij geven van adresdata. Uit dit onderzoek bleek dat 70% van de hergebruikers uit het bedrijfsleven kwamen, 20% uit de centrale overheid en 10% uit de gemeenten (DECA 2010). Sinds 2009 worden de adresdata beschikbaar gesteld via een webservice, en worden de data nu in een formaat aangeboden dat direct in navigatiesystemen kan worden gebruikt (Lind 2014).

Bredere effecten van open adresdata:

Volgens DECA (2010) en Lind (2014) waren de indirecte effecten van het vrijgeven van adresdata:

- Eliminatie van meervoudig inwinnen door andere overheidsorganisaties.
- Hergebruik van de referentiedata door onder meer OpenStreetMap, wat geleid heeft tot een kwaliteitsslag door meer terugmeldingen.
- Hogere datakwaliteit en standaardisatie: doordat het terugmelden van fouten is vereenvoudigd en nog maar door één organisatie wordt afgehandeld. Open adresdata hebben geleid tot standaardisatie van en bekendheid met één adresformaat voor Denemarken.
- Verbeterde dienstverlening-coördinatie: er is nu meer vertrouwen dat nooddiensten dezelfde referentiedata gebruiken, en dat gebruikers van de referentiedata meer kunnen rekenen op de juistheid van de adressen.

Financiële effecten van open adresdata

Het onderzoek voor DECA bepaalde dat de directe en indirecte baten van gratis adresdata €63 miljoen bedroegen in de periode 2005-2009. Dit bedrag was gebaseerd op het schatten van de bredere maatschappelijke baten door bijvoorbeeld betere navigatiesystemen en logistieke planning of OV-planners. In 2009 lag het in de lijn der verwachting dat dit bedrag zou stijgen met €14 miljoen in 2010. De totale kosten van het openstellen van adresdata was tot 2009 ongeveer €2,6 miljoen, waarvan €1,3 miljoen compensatie aan de decentrale overheden (afkopen van lopende dataverstrekkingsovereenkomsten) en €1,3 miljoen investeringen in een distributieplatform. De distributiekosten werden verwacht te stijgen met €0,2 miljoen in 2010 (DECA 2010).

Het DECA-rapport gaf aan dat de IT-infrastructuur nog te wensen overliet. Daarom werd door het (toenmalige) Ministerie van Volkshuisvesting, Stedelijke en Plattelandszaken (MBBL) besloten om verdere investeringen te plegen om de IT-infrastructuur te verbeteren. Tussen 2010 en 2014 werden nogmaals €5 miljoen (eenmalig) geïnvesteerd om de IT-infrastructuur voor een nieuw register en distributiedienst in te richten, met 85 extra fte bij de gemeenten en het MBBL. De operationele kosten zijn nu 40 fte extra per jaar bij de gemeenten en het MBBL. De baten voor de overheid worden geschat op ca. €30 miljoen per jaar vanaf 2015 door besparingen in transactiekosten, efficiëntie- en effectiviteitsslagen (Lind 2014).

B.4. Effecten Australian Bureau of Statistics¹³⁵

De Australian Bureau of Statistics (ABS) is een voorloper op het gebied van het implementeren van een opendatabeleid. Het aanbod van ABS bevat Australische nationale statistieken, de belangrijkste economische indicatoren, censusdata, indexcijfers en de effecten van de maatregelen op de algehele ontwikkeling van Australië. Daarnaast biedt ABS regionale statistieken en producten zoals sociale trends, het Australische statistisch jaarboek en wetenschappelijke publicaties. Daarnaast biedt ABS ook een aantal diensten aan voor het genereren van statistische tabellen, vergelijkbaar met Statline van CBS, en handleidingen. Ook biedt ABS een aantal diensten en producten tegen een vergoeding zoals boeken, de zogenaamde Confidentialised unit record files (CURFs) voor onderzoekers, en geselecteerde censusproducten. Verder heeft ABS de app 'Run that Town' ontwikkeld en gratis beschikbaar gesteld.

Sinds de implementatie van open data in 2005 is een aantal effecten waargenomen zoals een significante toename in het aantal downloads, en meer uiteenlopende hergebruikers. Om aan de vraag te kunnen voldoen, zijn er meer servers in gebruik genomen. Een ander effect is dat er minder vragen binnenkomen via de e-mail, en dat de aard van de vragen veranderd is van data- en prijsverzoeken naar vragen over de data zelf en de gebruiksvriendelijkheid ervan. De kwaliteit van de data en metadata is toegenomen.

ABS wordt door de centrale overheid gefinancierd en hoeft geen eigen inkomsten te genereren. Voor het aanbieden van maatwerkproducten worden marktconforme tarieven in rekening gebracht. Sinds het opendata-aanbod is er een lichte toename van betaaldiensten (consultancyopdrachten voor ABS) waarneembaar omdat niet alle gebruikers statistische analyses kunnen of willen uitvoeren.

Om het vertrouwen in ABS te kunnen meten, houdt het enquêtes (in 2010 voor het laatst) en monitort het de berichtgeving in de media. Zowel de enquête als de constante positieve berichtgeving in de media suggereren dat het steeds maar toenemende gebruik van ABS-statistieken een bewijs van blijvend vertrouwen is (Australian Bureau of Statistics 2014).

ABS-statistieken en het merk ABS worden via Facebook en Twitter gepromoot. Gedurende 2013–2014 werd het bereik van de ABS Facebook newsfeed meer dan verdrievoudigd, waardoor het merk ABS een groot publiek bereikt tegen verwaarloosbare kosten. Het aantal volgers van de ABS Twitteraccount is in een jaar gegroeid van 11.000 tot 18.000 (Australian Bureau of Statistics 2014, p. 35).

B.5. Effecten Companies House¹³⁶

Companies House (CH) in de VK is de uitvoerende overheidsorganisatie van het Ministerie van Business, Innovation and Skills (BIS) met als publieke taak het registreren van bedrijven en bedrijfsgegevens (incl. jaarverslagen) in een bedrijvenregister, en deze publiekelijk toegankelijk maken. CH is een zogenaamd Trading Fund, dat wil zeggen een overheidsorganisatie verantwoordelijk voor het genereren van eigen inkomsten om onkosten te dekken, en zelfs jaarlijks een dividend aan het Ministerie van BIS te betalen. CH hanteert een integraal kostenregime voor aangeboden diensten. De meeste inkomsten van CH komen uit verplichte registratiebijdragen, zowel

¹³⁵ Dit deel is afkomstig uit Welle Donker en Van Loenen 2015 p. 44 en verder.

¹³⁶ Dit deel is afkomstig uit Welle Donker en Van Loenen 2015 p. 47 en verder.

voor inschrijvingen en mutaties als jaarlijkse contributies. Daarnaast heeft CH in de afgelopen twee jaar al een aantal datadiensten voor het verstrekken van data als open data ontwikkeld. CH biedt een aantal diensten aan, waarvan er in 2014 twee als open data beschikbaar werden gemaakt. CH is nieuwe diensten aan het ontwikkelen en/of aan het verbeteren om geheel over te kunnen gaan op het aanbieden van uitsluitend open data.

In november 2013 is een gratis Accounts Data Product gelanceerd waarmee klanten toegang krijgen tot financiële bedrijfsgegevens die digitaal zijn aangeleverd. De Accounts Data Product biedt een dagelijks en een maandelijks download ZIP-bestand, dat een snapshot geeft van basis-bedrijfsgegevens in het handelsregister in cvs-formaat.

De WebCheck dienst is een zoekfunctie voor het bedrijvenregister, waarmee gratis op naam of op nummer kan worden gezocht naar basis-bedrijfsgegevens. De WebCheck-dienst kan ook als monitor gebruikt worden met een gratis emailmelding wanneer een geselecteerd bedrijf nieuwe documenten heeft aangeleverd.

CH heeft ook een app ontwikkeld die, zonder vooraf registreren, toegang geeft tot basis-bedrijfsgegevens, registerstatistieken en waarschuwingen wanneer bepaalde documenten aangeleverd moeten worden. Verder heeft CH application programming interfaces (API's) ontwikkeld die direct toegang geven tot het register van huidige en uitgesloten bedrijfsfunctionarissen.

Het grootste effect direct na het als open data beschikbaar stellen (effect in 2014) is een grote toename in het aantal online zoekopdrachten in het bedrijvenregister. De verbeterde dienstverlening heeft een positief effect gehad op de efficiëntie, voor zowel CH als voor de bedrijven. De kosten per bedrijf zijn van £24.01 in 2005/2006 naar £15.88 in 2012/2013 gedaald en CH verwacht dat dit nog verder zal dalen. Er zijn sinds het aanbod van open data minder verzoeken per telefoon binnengekomen in 2014 (7,5% minder), maar dat kan ook komen doordat steeds meer bedrijven via e-mail communiceren (9,9% toename).

B.6. Kosten-baten van open data: BZK-onderzoeken

De Vries heeft in opdracht van het Ministerie van BZK in 2014 een onderzoek uitgevoerd naar de eventuele kosten van het beschikbaar stellen van open data. Daarvoor zijn een vijftal casestudy's uitgevoerd, te weten KNMI, Kadaster, de Rijksdienst voor het Cultureel Erfgoed (RCE), de gemeente Rotterdam en de gemeente Enschede. Kronenburg *et al.* hebben in 2012 in opdracht van het Ministerie van BZK een soortgelijk onderzoek uitgevoerd, met als casestudy's RCE, Centraal Informatiepunt Beroepen Gezondheidszorg (CIBG), Ministerie van BZK, gemeente Rotterdam, Provinciaal GeoRegister (PGR) en Dienst Uitvoerend Onderwijs (DUO).

Aspecten van incrementele kosten

De Vries maakt een onderscheid tussen drie aspecten van de incrementele kosten, te weten de eenmalige infrastructurele kosten (extra servers, software), de veranderkosten en de structurele beheerkosten. De infrastructurele kosten verschillen per organisatie en per type van dataset. Voor KNMI waren die infrastructurele kosten van €20.000 beduidend hoger dan bijvoorbeeld voor de gemeentes. De eenmalige infrastructurele kosten bedragen gemiddeld €5.000.

De veranderkosten vormen de grootste kostenpost van het beschikbaar stellen van open data. De veranderkosten hebben betrekking op de personele kosten gemoeid met het voorbereiden en uitvoeren van een opendataproject, zoals het onderhandelen met data-eigenaren, het wegnemen van bezwaren en het faciliteren van duurzame ontsluiting en promoten van open data. Hoewel deze kosten ook verschilden per onderzochte organisaties, concludeert De Vries (2014) dat er wel rekening gehouden moet worden met circa 1 fte voor een jaar, ongeveer €50.000.

De beheerkosten van open data betreffen het duurzaam ontsluiten van open data, publiceren van updates, structurele kosten van de infrastructuur en kosten voor extra helpdesk medewerkers.

Beheerkosten variëren tussen de €10.000 en €15.000 per jaar, afhankelijk van organisatie, hoeveelheid en diversiteit van datasets.

Kronenburg *et al.* (2012) heeft voor de incrementele kosten alleen een inschatting gemaakt van de investeringskosten (infrastructuur en voorbereidingskosten), maar heeft de beheerkosten niet in alle cases meegenomen. Voor de door hun onderzochte zeven casestudy's liepen de incrementele kosten uiteen van €15.000 voor CIBG voor de extra server benodigd voor het beschikbaar stellen van het BIG Register tot €7.5 miljoen voor RCE. RIVM gaf aan dat er geen extra kosten zijn gemaakt om luchtkwaliteit- en emissiedata beschikbaar te stellen, maar dat open data een verschuiving van kosten binnen de eigen organisatie inhielden. De provincies hebben €450.000 en 800 uren bijgedragen om het Provinciaal GeoRegister (PGR) op te zetten. PGR is geen dataset maar een metadata-register van provinciale data, zodat provinciale (open) data beter te benutten en te delen worden. Het Ministerie van BZK heeft een extra €300.000 geïnvesteerd om de Wettenbank op te zetten. Het Ministerie van OCW heeft €500.000 bijgedragen om onderwijsdata openbaar te maken. RIVM, RCE en DUO gaven aan dat de meeste verandkosten gemaakt zijn voor het geschikt maken van open data (datakwaliteit verhogen, anonimiseren/aggregeren, etc.). Bij gemeente Rotterdam is €130.000 geïnvesteerd, waarbij het grootste gedeelte toe te schrijven is aan samenwerkingsprojecten met de Hogeschool Rotterdam (HR) om open data te faciliteren en te promoten.

Overigens konden in geen van de onderzochte cases de incrementele kosten exact worden bepaald aangezien in geen van de cases *ex-ante* een duidelijke budgettering heeft plaatsgevonden, er niet bijgehouden is wat de kosten *ex-post* precies waren. Bovendien is het moeilijk voor de organisaties om de kosten van het uitvoeren van publieke taken te scheiden van de kosten van open data.

Verschillen in onderzochte organisaties

De Vries (2014) concludeert uit de vijf casestudy's dat er grote verschillen zijn tussen het type data-aanbieder, aan de ene kant van het spectrum bevinden zich de organisaties van wie het een kerntaak is om data te verstrekken (KNMI, Kadaster) en aan de andere kant de andere organisaties zoals gemeentes en RCE. Bij de professionele dataverstrekkers is er al vaak een infrastructuur aanwezig om grote hoeveelheden dynamische data te verwerken. Bij andere organisaties ontbreekt vaak een dergelijke infrastructuur. Voor de professionele dataverstrekkers zit het grootste knelpunt voor open data in (het omleggen van) financiën omdat hun huidig verdienmodel onder druk wordt gezet. Voor gemeentes zit het grootste knelpunt eerder in het verduurzamen van open data, en het vasthouden van de energie. Open data worden bij gemeentes vaak door één persoon aangejaagd, en als die persoon wegvalt (ziekte, doorschuiven naar een andere positie of project), valt het momentum weg, zoals bij de gemeente Rotterdam. Een ander probleem voor gemeentes is dat de kosten vaak door verschillende onderdelen van de organisatie moeten worden betaald, en dat de data weer bij andere onderdelen worden beheerd. Daardoor kan het moeilijk zijn om de verschillende onderdelen van de organisatie te overtuigen van de benodigde kosten. In Tabellen 2.5 en 2.6 van het rapport wordt een samenvattend overzicht van de baten en kosten van de BZK onderzoeken gegeven.

Samenvatting onderzochte cases

Tabel B3.1: Samenvattend overzicht van de bevindingen in de 5 cases in euro's (de Vries 2014, p.25)

Case	Begin uitvoering	Data	Kosten in euro's			Dekking	Baten
			Infra-structuur-kosten	Verander-kosten	Beheer-kosten-		
Rotterdam	2011	Zeer divers, 137 sets en API's	20.000/jr.	50.000 1x	45.000/jr.	Extern (2011), daarna budget CIO	Onbekend
Enschede	2012	Zeer divers, 23 sets	10.000 1x	8.000 1x	Minimaal	Budget afdeling IM	Onbekend
KNMI	2013/14	Real-time, klimaat, seismologisch	4.000/jr.	50.000 1x	10.000/jr.	Vooralsnog eigen middelen	80.000/jr. vanaf 2015/17
Kadaster	2012	Basisregistratie Topografie	0, lift mee in PDOK	75.000 1x	25.000/jr.	250.000 van Min. lenM, rest eigen middelen, incl. inkomstenderving 300.00/jr.	115.000/jr. vanaf 2014
RCE	2010	Cultureel erfgoed, zoals monumenten	0	0	0	Ministerie OCW, subsidies, eigen middelen	Onbekend

Tabel B3.2: Samenvatting overzicht van de bevindingen in de 7 cases in euro's (Kronenburg et al. 2012) (samengevat door de auteurs)

Case	Begin invoering	Data	Kosten in euro's		Baten in euro's
			Infrastructureel	Veranderkosten	
RCE	2011	Cultureel erfgoed	60.000/jaar voor de Beeldbank	7.5mln	Intern: 7 fte/jaar (500.000/jr) minder + onderhoudscontracten (300.000/jr) extern: gemeenten besparen ca. 7 min. per opvraging van monumentendossiers
BZK	2001	Wettenbank		300.000	Lagere transactiekosten.
CIBG	2012	BIG Register	15.000	Moesten toch al gemaakt worden	Intern: 80.000 meer bevragingen per maand zonder extra personeel
RIVM	-*	Luchtkwaliteit & -emissiedata	100.000	250.000	Extern: waarschijnlijk lagere transactiekosten, bijv. Milieudefensie
Rotterdam	2012	Divers		130.000, waarvan 100.00 faciliteren HR	Extern: HR: bijna EUR 700.000 aan subsidies + EUR 200.000-300.000 aan publicitaire waarde
DUO	2011	Onderwijsdata		500.000	4.5 fte besparen in 2 jaar
Provincies	2007	Opzetten PGR	60.000	450.000 + 800 metingen in 2 jaar	Intern: Transactiekosten besparing provincies 100-500 metingen extern: bijv. PBL 50.000/jaar

* Invoerdatum niet vermeld in rapport.

De infrastructuurkosten voor KNMI en Kadaster zijn relatief laag, ondanks de grote hoeveelheid data die ze aanbieden (terabytes per dag), omdat deze organisaties meeliften op infrastructuur die al uitgerold was. Voor gemeente Rotterdam waren deze kosten relatief hoog omdat deze uitbesteed zijn. Voor RCE waren er geen cijfers beschikbaar omdat RCE geen onderscheid maakt tussen gebruikers en hergebruikers. Volgens De Vries (2014, p.28) is ontsluiten van open data bij RCE een kerntaak, en is er daarvoor geen aparte post in de begroting opgenomen, en is het daarom onmogelijk om de incrementele kosten vast te stellen. RCE beschouwde alle kosten die gemaakt moeten worden voor het digitaliseren van informatie en het verbeteren van de informatievoorziening, in totaal €7,5 miljoen, in zijn geheel als investeringen voor open data

(Kronenberg *et al.* 2012, p.31). RCE verwacht dit bedrag binnen 10 jaar terug te verdienen door besparingen op balie- en archiefmedewerkers (Kronenberg *et al.* 2012, p.33).

Baten van open data

De Vries (2014) heeft alleen kosten in beschouwing genomen, en daar waar genoemd, anekdotische baten meegenomen. Kronenburg *et al.* (2012) hebben wel de verwachte baten in beschouwing genomen. Een aantal van die baten worden verwacht binnen de organisatie te vallen. Zowel RCE als DUO verwachten op personele kosten te besparen. Bij de andere organisaties worden de baten verwacht buiten de eigen organisatie te vallen, zoals lagere transactiekosten voor gebruikers (Wettenbank van BZK, en de provincies). De gemeente Rotterdam ziet zelf nog geen directe baten. Wel heeft de samenwerking met HR geleid tot baten voor de hogeschool. HR gebruikt de open data als lesmateriaal en heeft bijna €700.000 aan nationale en Europese subsidies binnengehaald voor projecten. Ook gaf HR aan voor €200.000 à €300.000 aan publicitaire waarde gegenereerd te hebben met publicaties over (de diverse onderdelen van) Rotterdam Open Data platform (Kronenburg *et al.* 2012, p.80).

Bijlage 4 ITS Richtlijn 2010/40/EU

In een Europese context is voor de internationale vergelijking van de casus NDW de Europese Richtlijn 2010/40/EU van 7 juli 2010 betreffende het kader voor het invoeren van intelligente vervoerssystemen op het gebied van wegvervoer en voor interfaces met andere vervoerswijzen (ITS-richtlijn) van belang.¹³⁷ Met deze richtlijn wordt gepoogd de ontwikkeling van innovatieve vervoerssystemen te versnellen. De richtlijn is vanaf 1 juli 2015 van kracht.

“Intelligente vervoerssystemen (ITS) zijn geavanceerde toepassingen die, zonder intelligentie als zodanig te belichamen, gericht zijn op het aanbieden van innovatieve diensten inzake verschillende vervoerswijzen en verkeersbeheer en die uiteenlopende gebruikers in staat stellen zich beter te informeren en veiliger, meer gecoördineerd en “slimmer” gebruik te maken van vervoersnetwerken.” (Richtlijn 2010/40/EU, punt 3 van de preambule)

De gedachte achter de richtlijn is dat er grote druk staat op verkeers- en vervoerssystemen, dat traditionele oplossingen (zoals aanleg nieuwe wegen) onvoldoende oplossingen bieden en er innovaties op het gebied van informatie- en communicatietechnologie (ICT) nodig zijn. Nationale toepassingen van ICT “... blijven evenwel gefragmenteerd en ongecoördineerd en kunnen de geografische continuïteit van ITS-diensten in de hele Unie en aan haar buitengrenzen niet waarborgen.” (Richtlijn 2010/40/EU, punt 6 van de preambule) Daarom moet er een Europese ‘ITS-infrastructuur’ worden ontwikkeld:

“ITS moeten voortbouwen op interoperabele systemen die zijn gebaseerd op open en publieke normen en zonder onderscheid beschikbaar zijn voor alle leveranciers en gebruikers van toepassingen en diensten.” (Richtlijn 2010/40/EU, punt 11 van de preambule)

De hier geformuleerde ambitie is anders dan open data. Het gaat er hierbij om dat de data toegankelijk zijn voor leveranciers en gebruikers en dat de systemen kunnen worden gebruikt. Open data kunnen (naast andere alternatieven) echter wel een middel zijn om invulling te geven aan deze ambitie.

De richtlijn (artikel 2) kent de volgende prioritaire gebieden:

1. optimaal gebruik van weg-, verkeers- en reisgegevens;
2. continuïteit van ITS-diensten voor verkeers- en vrachtbeheer;
3. ITS-toepassingen voor verkeersveiligheid en -beveiliging;
4. koppeling van het voertuig aan de vervoersinfrastructuur.

Het optimaal gebruik van NDW-gegevens valt in eerste instantie onder prioriteit 1. Bijlage 1 [van de richtlijn] bevat specificaties van deze prioriteit. Deze specificatie bevat onder meer “de omschrijving van de eisen inzake de accuratesse en het over de grenzen heen ter beschikking stellen van multimodale reisinformatiediensten [...en...] real-time verkeersinformatie in de gehele Unie voor ITS-gebruikers, op basis van:

- de beschikbaarheid en toegankelijkheid van bestaande, accurate weg- en real-time verkeersgegevens die worden gebruikt voor multimodale reisinformatie voor ITS-dienstaanbieders, onverminderd de eisen qua veiligheid en verkeersbeheer;
- het faciliteren van grensoverschrijdende elektronische gegevensuitwisseling tussen de bevoegde overheden en belanghebbenden en de bevoegde ITS-dienstaanbieders;

¹³⁷ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:0001:0013:nl:PDF>

- het tijdig actualiseren van beschikbare weg- en verkeersgegevens die worden gebruikt voor multimodale reisinformatie door de bevoegde overheid en belanghebbenden;
- het tijdig actualiseren van multimodale reisinformatie door de ITS-dienstaanbieders.” (Richtlijn 2010/40/EU, Bijlage 1)

In bijlage 2 van de richtlijn worden de volgende criteria geformuleerd op grond waarvan maatregelen kunnen worden getoetst. Deze maatregelen moeten:

- a) effectief zijn;
- b) kosteneffectief zijn;
- c) proportioneel zijn;
- d) continuïteit van dienstverlening ondersteunen;
- e) interoperabiliteit bieden;
- f) achterwaartse compatibiliteit ondersteunen;
- g) de kenmerken van de nationale infrastructuren en netwerken in acht nemen;
- h) gelijke toegang bevorderen;
- i) maturiteit ondersteunen;
- j) kwaliteit van tijds- en positiebepaling bieden;
- k) intermodaliteit vergemakkelijken;
- l) samenhang in acht nemen.

Volgens een verslag van de Europese Commissie aan het Europees Parlement en de Europese Raad (van 21.10.2014, COM(2014) 642 final) over de invoering van de richtlijn verliep dit traag. Slechts zeven lidstaten hadden hun wetgeving op tijd omgezet. Ook het door de Europese commissie opstellen van specificaties over multimodale reisinformatiediensten en real-time verkeersinformatie in de gehele Unie was vertraagd. Omtrent het laatste gaf de Commissie daarbij aan dat deze vertraging gedeeltelijk was te verklaren door de ‘zeer grote reikwijdte van deze specificaties’ en door ‘de noodzaak te komen tot onderlinge overeenstemming tussen de lidstaten over het toepassingsgebied van de specificaties’ (COM(2014) 642 final, blz. 8). Tsjechië heeft zelfs een zaak aangespannen bij het Hof van Justitie, omdat de EU buiten haar mandaat zou zijn getreden met deze regels. Het gerecht heeft op 8 oktober 2015 (zaken T-659/13 en T-660/13) deze claim afgewezen. Deze situatie van complexiteit en onzekerheid over bevoegdheden is niet bevorderlijk geweest voor de snelheid van doorwerking van deze richtlijn.

In het kader van deze richtlijn heeft de Europese Commissies een aantal consultaties uitgevoerd. Een van deze consultaties (gehouden tussen 20/12/2013 en 14/03/2014) betrof de EU-brede voorziening van real-time verkeersinformatie. De vragenlijst betrof echter vooral het type gebruik van informatie en de rol van de EU daarin en minder de wijze waarop de data beschikbaar worden gesteld.

Op het gebied van real-time verkeersinformatiediensten heeft de Commissie op 18 december 2014 een Gedelegeerde verordening (2015/962) ter aanvulling van Richtlijn 2010/40/EU vastgesteld (Publicatieblad van de Europese Unie, 23.6.2015, L 157/21). In de preambule wordt het volgende over toegankelijkheid aangegeven:

“De toegankelijkheid van nauwkeurige en actuele statische weggegevens, dynamische weggegevens en verkeersgegevens is van essentieel belang voor het verlenen van real-time verkeersinformatiediensten in de hele Unie. De relevante gegevens worden verzameld en bewaard door wegenautoriteiten, wegenexploitanten en aanbieders van real-time verkeersinformatie. Om de uitwisseling en het hergebruik van deze gegevens te vergemakkelijken met het oog op de verlening van de diensten in kwestie, moeten wegenautoriteiten,

wegenexploitanten en verleners van real-time verkeersinformatiediensten de gegevens, de overeenkomstige metagegevens en de informatie over de kwaliteit van de gegevens ter beschikking stellen van andere wegenautoriteiten, wegenexploitanten, verleners van real-time verkeersinformatiediensten en producenten van digitale kaarten via een nationaal of gemeenschappelijk toegangspunt. Het toegangspunt kan de vorm aannemen van een databank, een register, een internetportaal of een vergelijkbare vorm, afhankelijk van het gegevenstype. De lidstaten moeten de bestaande openbare en particuliere toegangspunten samenvoegen tot één toegangspunt dat de toegang mogelijk maakt tot alle soorten relevante beschikbare gegevens die binnen het toepassingsgebied van deze specificaties vallen. Het moet de lidstaten worden toegestaan om met elkaar samen te werken om een gemeenschappelijk toegangspunt op te zetten dat betrekking heeft op de beschikbare gegevens van de deelnemende lidstaten. Het moet de lidstaten vrij staan om de toegangspunten die zijn vastgesteld uit hoofde van andere krachtens Richtlijn 2010/40/EU goedgekeurde gedelegeerde handelingen te gebruiken als nationale toegangspunten voor de gegevens die onder deze verordening vallen.” (Publicatieblad van de Europese Unie, 23.6.2015, L 157/21, preambule 14)

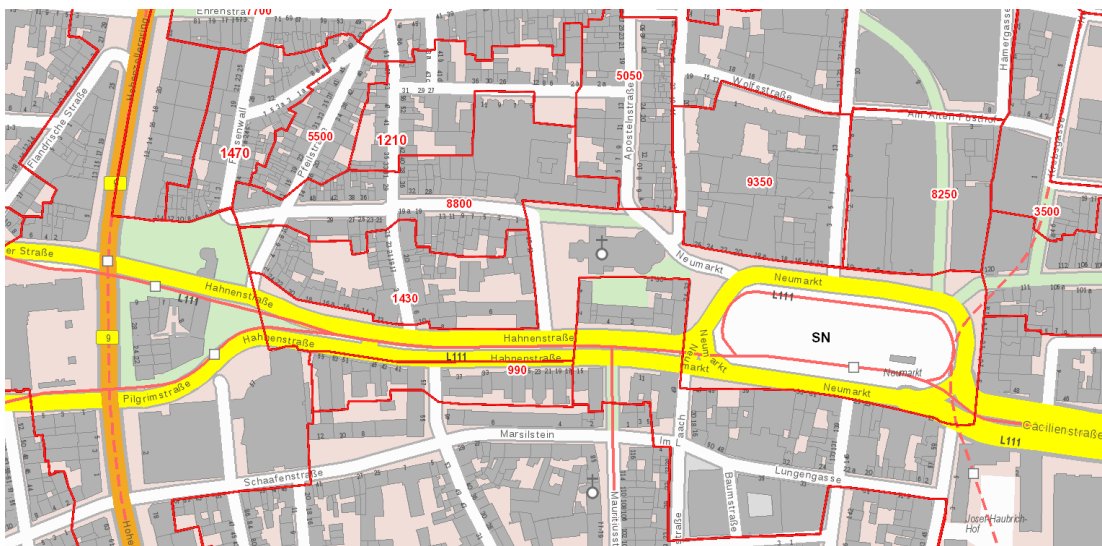
De regeling betreft slechts bestaande data die beschikbaar zijn in machine-leesbaar formaat. De regeling regelt niet dat lidstaten data moeten gaan verzamelen of machine-leesbaar moeten maken. Als de data er zijn, dan moet deze (op een niet-discriminerende wijze) toegankelijk worden gemaakt. De specifieke voorwaarden die gelden voor het hergebruik van die gegevens, worden overgelaten aan de betrokken partijen (natuurlijk binnen reeds eerder gestelde kaders zoals Richtlijn 2003/98/EG).¹³⁸

Op Europees niveau is er ook de DATEX II standaard (CEN/TC 278; zie ook <http://www.datex2.eu/>) ontwikkeld om uitwisseling van de data mogelijk te maken.

Er moet dus een online toegangspunt zijn om de data in te voeren en op te vragen. In Nederland beheert NDW het Nationaal Toegangspunt ITS. De Dienst Wegverkeer (RDW) is aangewezen als toezichthouder.

¹³⁸ <http://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:32003L0098&qid=1478274482091>

Figuur B5.2: Kaart met referentiegrondwaardes (variërend van €990 tot €9350 per m²) in een deel van Keulen (<http://www.boris.nrw.de/borisplus/>)



Deze referentiegrondwaarde wordt vastgesteld door een lokale commissie van deskundigen op het gebied van waardering van onroerende zaken (Gutachterausschüsse für Grundstückswerte) en staat in relatie tot de marktwaarde van het vastgoed. Deze deskundigen krijgen ten behoeve van de vaststelling van deze waarde via het notariaat de beschikking over alle gegevens van alle vastgoedtransacties. Met behulp van de vergelijkende methode worden op grond hiervan referentiegrondwaardes bepaald (Brandt 2013).

Het systeem van deze referentiegrondwaardes is sinds 1960 opgenomen in de wet en aanvankelijk bestond dit uit papieren kaartmateriaal dat tegen betaling werd verstrekt aan belangstellenden. Deelstaten gaan over de manier waarop deze verstrekking plaatsvindt. De deelstaat Noordrijn-Westfalen heeft in 2003 de elektronische toegang tot deze gegevens ontwikkeld via een Bodenrichtwerte Informationssystem (afgekort BORIS) dat in 2008 werd verbeterd tot BORIS-Plus (Kertscher, 2009; Wanzke, 2009). Ook in andere deelstaten is dit BORIS ontwikkeld, waarbij ook relaties zijn gelegd via het 'Vernetzte Bodenrichtwertinformationssystem' (VBORIS) dat deel uitmaakt van de Duitse geografische informatie infrastructuur (Liebig 2012; Walm 2013).

Deze informatie-infrastructuur kent een voorkant (de referentiegrondwaardes die worden gepresenteerd in BORIS) en een achterkant (de ordening van de informatie over de transacties, kenmerken van de bebouwing, planologische situatie, cartografische ondergrond) (Liebig 2012). Van deze achterkant biedt met name de verzameling van transactiepreizen relevante informatie. Deze informatie wordt ook beschikbaar gesteld in bijvoorbeeld de deelstaat Hamburg waar informatie over de 'Kaufpreissammlung' over vergelijkingsprijzen van bebouwde en onbebouwde grond kan worden verkregen ten behoeve van het door deskundigen taxeren van een object. Hierbij kunnen voor €410 maximaal 30 voor dat object vergelijkbare transacties worden opgevraagd. Hier is dus allerm minst sprake van open data. Een en ander hangt ook samen met het feit dat veel informatie van Kadaster en openbare registers, anders dan in Nederland, alleen toegankelijk zijn voor belanghebbenden.

Ook wat betreft de voorkant bestaan er tussen de deelstaten grote verschillen in het toegankelijkheidsregime van deze data. In Nedersaksen kan bijvoorbeeld voor €5 voor 15 minuten toegang worden verkregen tot het BORIS van de deelstaat wat een van de goedkoopste opties is uit

de 6 bladzijden lange legesverordening (Gebührenordnung) die op dit gebied is vastgesteld. Door een abonnement te nemen kan er elektronisch worden betaald.¹³⁹ In andere deelstaten wordt dit bestand gratis ter inzage gegeven (al dan niet aangevuld met betaalde diensten) of wordt het zelfs als open data aangeboden.

Berlijn levert de data niet als open data, maar biedt inzage in de kaart via een Web Mapping Service, en vervult bij het opendatabeleid een voortrekkersrol (Friedt & Luckhardt 2014). In het coalitieakkoord voor 2011-2016 oarmde de deelstaat Berlijn open data, waarbij geo-informatie expliciet als voorbeeld werd genoemd. Vervolgens werd in januari 2012 een *Berliner Open Data-Strategie* gepresenteerd en zijn sinds oktober 2013 veel geodata als open data beschikbaar gesteld. Rond 2010 waren de inkomsten uit de verkoop van deze geodata nog ongeveer 1,2 miljoen per jaar (Friedt & Luckhardt 2014). Een deel van deze inkomsten (€0,4 miljoen) betroffen interne verrekeningen ('In-Sich-Geschäfte') binnen de deelstaat zelf (de ene dienst 'kocht' data van de andere dienst). De externe inkomsten die daarmee op het spel stonden betroffen €0,8 miljoen. Daarnaast zou dit negatieve effecten kunnen hebben op andere aanbieders van deze data en zouden mogelijk extra investeringen gedaan moeten worden in IT om het toegenomen gebruik van open data te faciliteren. Zaken als privacy en databescherming worden ook als aspecten genoemd, maar in de randvoorwaardelijke sfeer. Volgens de Berlijnse deelstaatregering stonden tegenover deze aspecten de volgende maatschappelijke baten (Friedt & Luckhardt 2014).

1. Open data werden beschouwd als een wezenlijk bestanddeel van 'open government': het openen van het bestuur naar bevolking en economie, leidend tot grotere transparantie, participatie en een intensere samenwerking om gemeenschappelijke belangen te versterken.
2. Open data zouden leiden tot innovaties die de economie ten goede zouden komen en zo – op termijn –leiden tot extra belastinginkomsten.
3. De data waren reeds met belastinggelden (zo een €34 miljoen) gefinancierd.
4. Een deel van de data worden reeds door de veroorzakers (en niet de gebruiker) betaald.
5. Betaling heeft een afschrikwekkende werking voor kleine bedrijven (start-ups) en particulieren. Het gebruik van data zou zo verveelvoudigd kunnen worden.
6. Het sluit aan bij de ontwikkeling op het gebied van verbeterde toegankelijkheid van data in Europa, de Bond en de deelstaat.
7. Open Data maken deel uit van het eveneens in het coalitieakkoord opgenomen moderniseringsprogramma van de dienstestad Berlijn (er is zelfs een logo van ServiceStadt Berlin).
8. Het *Leitprinzip* van het bestuur is het tegemoetkomen aan maatschappelijke opgaven en niet winstmaximering.
9. Juridische datalicenties kunnen zo worden vereenvoudigd.
10. Door hoger gebruik komt er betere feedback waardoor gegevens sneller worden bijgewerkt en fouten worden gecorrigeerd. De kwaliteit van de data gaat daardoor omhoog.

In Berlijn zijn de gegevens van de referentiegrondwaarde echter alleen als Web Map Service (WMS) beschikbaar. Hiermee worden de bovengenoemde voordelen voor een belangrijk deel niet gehaald omdat een dergelijke kaart op het web belangrijke beperkingen oplegt aan het hergebruik.

In Noordrijn-Westfalen (NRW) zijn deze data per januari 2016 beschikbaar als open data.¹⁴⁰ Er is sprake van een atom-feed waarmee de gegevens in een ESRI shapefile (twee sterren volgens de TBL-classificatie) kunnen worden gedownload. De gehanteerde licentie is die van 'Datalizenz Deutschland

¹³⁹ <http://www.gag.niedersachsen.de/bodenrichtwerte/bodenrichtwerte-88091.html>

¹⁴⁰ <http://www.wms.nrw.de/rssfeeds/content/geoportal/html/1014.html>

Namensnennung 2.0' (ofwel dl-de-by-2.0), zie kader "Datenlizenz Deutschland – Namensnennung – Version 2.0".

Hiermee loopt Noordrijn-Westfalen vooruit op andere deelstaten. Daarnaast is door de deelstaat zelf ook een App ontwikkeld. Op 10 oktober 2016 heeft deze BORISplus.NRW App een best practice award gekregen van het Gesellschaft für Geodäsie, Geoinformation und Landmanagement. In deze app kun je ter plekke zien wat de referentiegrond- en vastgoedwaardes zijn in de directe omgeving (binnen de grenzen van de deelstaat). In veel andere deelstaten (Sleeswijk-Holstein, Hamburg, Mecklenburg-Vorpommern, Brandenburg, Berlijn, Saksen-Anhalt, Saksen, Thüringen, Saarland, Hessen en Rijnland Pfalz) is er gratis toegang tot de Web Mapping Service, waarbij het veelal mogelijk is betaald aanvullende diensten te verkrijgen. In sommige deelstaten kunnen de onderliggende marktanalyses gratis als pdf worden gedownload; in andere kan dat niet. In Mecklenburg-Vorpommern is deze toegang alleen te verkrijgen via de raadpleging van acht verschillende websites van lokale organisaties (waar niet naar wordt gelinkt, maar die wel met adres en telefoonnummer zijn vermeld) te raadplegen. In al deze acht districten is de informatie op het web in te zien. In een aantal deelstaten (zoals Nedersaksen en Beieren) is er alleen betaald toegang tot de website. De kleinste deelstaat Bremen geeft aan dat vanwege personeelsproblemen elektronische toegang nog niet mogelijk is en kan de kaart gratis worden ingezien bij een servicecentrum en kan deze ook (op papier) tegen betaling worden besteld. In Baden-Württemberg ten slotte is er iets mis met de beveiligingscertificaten van de website waarnaar VBORIS door- linkt. De gegevens van Stuttgart (en een paar kleinere districten) kunnen op het web worden ingezien. Op niveau van de deelstaat is er alleen betaalde toegang mogelijk. Daar waar in de deelstaten betaling wordt gevraagd (als extra dienst of als enige toegangsmogelijkheid) zijn de bedragen hoger dan de materiële kosten voor papier en drukinkt en de bij een efficiënte reproductie benodigde personeelslasten (bijvoorbeeld €150 voor een A4-rapport van 80 blz.; €70 voor een kaartblad); maar zullen deze gezien de verwachte oplage niet de kosten van dataverzameling kunnen dekken.

Kader: Datenlizenz Deutschland – Namensnennung – Version 2.0

(1) Jede Nutzung ist unter den Bedingungen dieser „Datenlizenz Deutschland – Namensnennung – Version 2.0“ zulässig.

Die bereitgestellten Daten und Metadaten dürfen für die kommerzielle und nicht kommerzielle Nutzung insbesondere

- vervielfältigt, ausgedruckt, präsentiert, verändert, bearbeitet sowie an Dritte übermittelt werden;
- mit eigenen Daten und Daten Anderer zusammengeführt und zu selbständigen neuen Datensätzen verbunden werden;
- in interne und externe Geschäftsprozesse, Produkte und Anwendungen in öffentlichen und nicht öffentlichen elektronischen Netzwerken eingebunden werden.

(2) Bei der Nutzung ist sicherzustellen, dass folgende Angaben als Quellenvermerk enthalten sind:

1. Bezeichnung des Bereitstellers nach dessen Maßgabe,
2. der Vermerk „Datenlizenz Deutschland – Namensnennung – Version 2.0“ oder „dl-de/by-2-0“ mit Verweis auf den Lizenztext unter www.govdata.de/dl-de/by-2-0 sowie
3. einen Verweis auf den Datensatz (URI).

Dies gilt nur soweit die datenhaltende Stelle die Angaben 1. bis 3. zum Quellenvermerk bereitstellt.

(3) Veränderungen, Bearbeitungen, neue Gestaltungen oder sonstige Abwandlungen sind im Quellenvermerk mit dem Hinweis zu versehen, dass die Daten geändert wurden.

B.2. Vastgoedwaarde als open data in Noordrijn-Westfalen

Zoals al eerder is aangegeven zijn in Noordrijn-Westfalen recent de gegevens als open data ter beschikking gesteld. Er is een telefonisch interview gevoerd met Andreas Pelke van de Obere Gutachterausschuss für Grundstückswerte in NRW – het coördinerend orgaan van de deskundigencommissies op het gebied van vastgoedwaarde in Noordrijn-Westfalen. Deze gaf aan dat voordat (per 1-1-2016) de referentiegrondwaarde als open data beschikbaar waren het bestand op CD geleverd werd voor een prijs van €80.000. Deze prijs was te hoog. Er waren wel gemeenten, kerken, energiebedrijven en grotere bedrijven die informatie inwonen over de mogelijke levering, maar de daadwerkelijke vraag was (gezien de prijs) beperkt en er waren slechts 2 tot 3 klanten die jaarlijks dit bestand afnamen. Deze klanten bouwden dit vervolgens in het eigen systeem in. Een bank had bijvoorbeeld een dergelijk systeem ingebouwd.

Tussen 1-1-2016 en 1-10-2016 (9 maanden) zijn de gegevens (als opendatabestand) 213 maal gedownload. Daarnaast zijn ook oudere gegevens vanaf 2011 beschikbaar. Deze zijn (als opendatabestand) 464 maal gedownload. Er is sprake van een sterk toegenomen gebruik.

Naast de beschikbaarstelling van het hele databestand zijn nu op www.boris.nrw.de ook andere producten gratis geworden. Het gaat dan om:

- de Immobilenrichtwert (de gemiddelde waarde per m² van een normvastgoedobject in een gebied);
- de grondmarktrapportage (per gemeente/district beschikbaar) en
- het afdrukken van de gegevens.

De prijzen van individuele afname van een dergelijk product was voor 1/1/2016 respectievelijk €28,-, €40,- en €5,-. Betaling geschiedde niet online (geen internetbetalingen of per creditcard) maar op basis van rekeningen die werden verstuurd en waarvoor ook ambtelijke capaciteit was vrijgemaakt om bij te houden of er wel betaald was.

De overstap naar open data betekende een bezuiniging op het gebied van dit betalingsverkeer (1 fte), maar ook in het nieuwe informatiesysteem van BORIS dat per 1-1-2016 operationeel is, was het niet meer nodig om een betaalmodule in te bouwen.

Een ander belangrijk verschil zit in de herpublicatie van deze grondwaarden. Als men voor 1-1-2016 een deskundige een rapportage (zoals een taxatie) op het web wilde plaatsen, kon dat niet als daarin de referentiegrondwaarde was opgenomen. Dit was niet toegestaan (en dit werd niet gedoogd).

In andere deelstaten wordt met belangstelling deze ontwikkeling gevolgd en men ziet de meerwaarde. De transparantie van de vastgoedmarkt neemt hierdoor toe.

Na veel aandringen noemde de heer Pelke een nadeel dat hij persoonlijk ervaarde: vroeger werd bij elke bestelling gevraagd waarvoor de data werden aangevraagd. Nu is het onbekend waarom data worden gedownload. Het is in Duitsland wettelijk niet toegestaan de IP-adressen van partijen die de data opvragen te bewaren. Analyse op grond van deze IP-adressen is daarmee niet mogelijk.

Het is wel bekend dat de data gebruikt worden door (universitaire) onderzoeksinstituten om waardeontwikkelingen te analyseren. Dit kan bijdragen aan een verdere vergroting van de transparantie van de grondmarkt in Noordrijn-Westfalen en bijdragen aan het doel van het verzamelen en beschikbaar stellen van referentiegrondwaardes.

Belangrijk bij de afweging van kosten en baten van open data op dit gebied is het maatschappelijk doel dat wordt nagestreefd met de vervaardiging van referentiegrondwaarde. Doel van de

deskundigencommissie is het verschaffen van transparantie over vastgoedwaardes om te voorkomen dat foutieve keuzes worden gemaakt (Wanzke 2009).

B.2.1 Overzicht maatschappelijke kosten en baten Noordrijn-Westfalen

Directe effecten

Geen kosten meer om betalingen te verwerken.

Deze kosten worden door Pelke geschat op 1 fte administratieve ondersteuning. Volgens de 'Gebührenberechnung nach dem Zeitaufwand' van het Landesbetrieb Information und Technik van de deelstaat Noordrijn-Westfalen waren personeelskosten voor een personeelslid in 'Mittlere Dienst' €70.465 per jaar. Dit is exclusief 'Arbeitsplatzkosten' (€26.235 per jaar, op grond van €15,90 per uur * 1650 uur). Verondersteld mag worden dat deze kosten ook op termijn tot bezuiniging kunnen leiden. Daarnaast was er sprake van een besparing op het niet hoeven te ontwerpen van een betaalmodule in het automatiseringsprogramma. Dit is geen jaarlijks terugkerende besparing en deze zal beperkt qua omvang zijn. Al met al betekent het voorgaande dat een totale besparing op verstrekingskosten geschat kan worden op €100.000 per jaar. Echter slechts een klein deel van deze kosten zullen zijn gemaakt in relatie tot de drie transacties per jaar waarmee de gehele dataset werd verkocht. Het merendeel van deze administratieve lasten zullen besteed zijn aan de veelheid van kleinere transacties rond het kunnen afdrukken van gegevens, het inzien van marktberichten en vastgoedreferentiewaarden. Wat betreft de grotere transacties van het gehele bestand worden deze geschat op 5% van de omzet hiervan, €12.000 per jaar.

Tabel B5.1: Gebührenberechnung nach dem Zeitaufwand, stand 2014.

Landesbetrieb Information und Technik Nordrhein-Westfalen											Stand: 2014
Gebührenberechnung nach dem Zeitaufwand											
Laufbahn- gruppen	Durch- schnitt- liche Dienst- bezüge 2014	Ver- sorgungs- zuschlag (30 %)	Personalnebenkosten		Zuschlag für Hilfs- personal	Zwischen- summe (Sp. 2-5)	Zuschläge für Ver- waltung und Leitung (15 %)	Gesamt- summe (Sp. 6+7)	Kosten je Arbeitsstunde		
			Beihilfen	Trennungs- entsch.. Umzugs- kostenverg. (0,5 %)					Personal- kosten (Sp.8 / 1652 durchschnittl. Jahres- arbeitsstd.)	Sach- kosten (Arbeitsplatz- kosten)	Gesamt- kosten (Sp.9+10) - gerundet -
Beträge in Euro											
1	2	3	4a	4b	5	6	7	8	9	10	11
Höherer Dienst	63.339	19.002	1.878	317	7.748	92.284	13.843	106.127	64.24	15.90	80
Gehobener Dienst	48.967	14.690	1.878	245	7.748	73.528	11.029	84.557	51.18	15.90	67
Mittlerer Dienst	39.577	11.873	1.878	198	7.748	61.274	9.191	70.465	42.65	15.90	59
Einfacher Dienst	28.247	8.474	1.878	141	-	38.740	5.811	44.551	26.97	15.90	43

Transferkosten en baten

Het verschil in betaling voor bestaande gebruikers zelf is vanuit een maatschappelijke kosten-batenanalyse irrelevant, aangezien het transferkosten betreft. Tegenover de maatschappelijke baat voor de gebruiker staan maatschappelijke kosten voor de overheid.

Consumentensurplus van nieuwe gebruikers van open data.

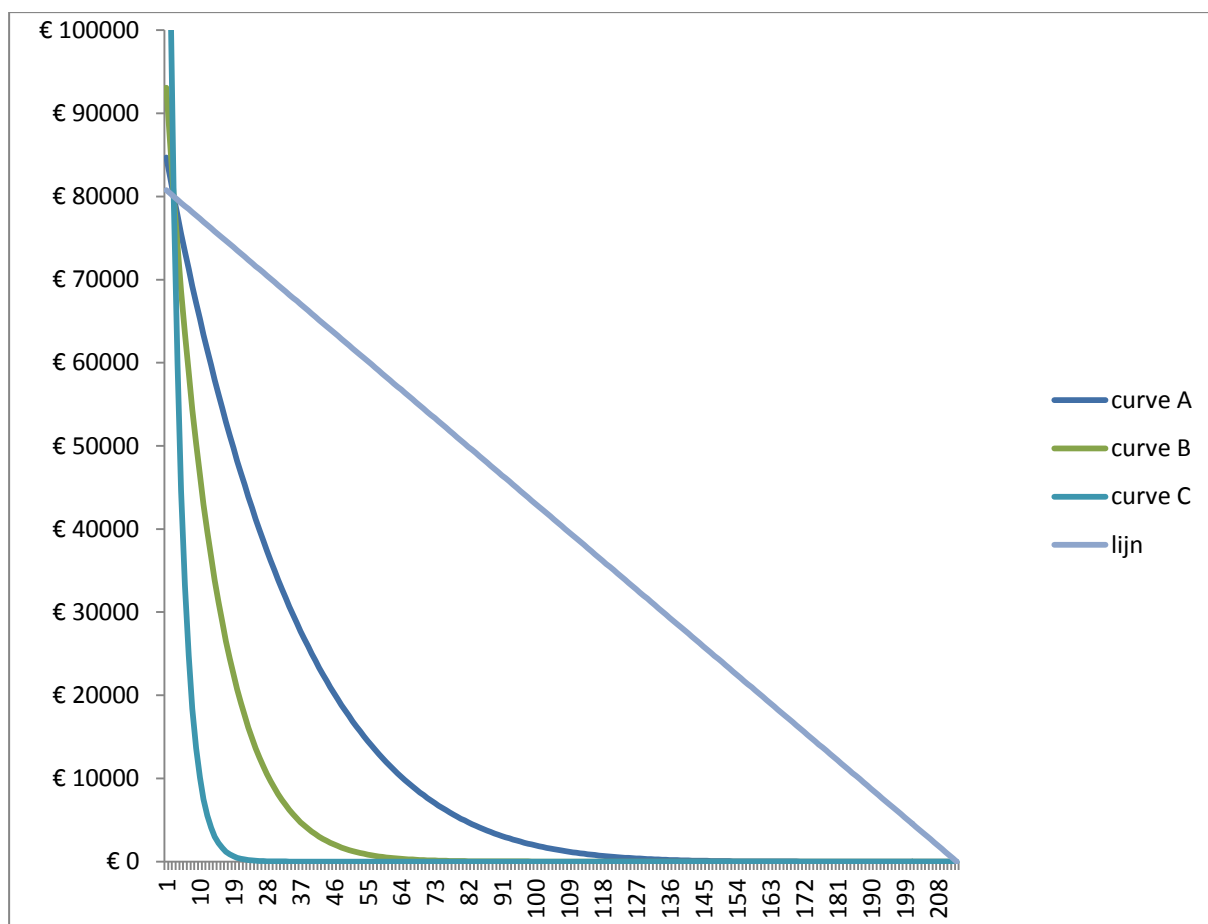
Veel gebruikers van de open data zouden hier wel een bedrag voor over hebben gehad dat hoger is dan €0. De leidraad hanteert hierbij de 'rule of half' die onzes inziens niet de juiste benadering is om in deze context het consumentensurplus te berekenen. Volgens de berekeningswijze van de rule of

half zou dit consumentensurplus de helft zijn van het product van het verschil in gebruik en het verschil in prijs tussen de alternatieven. In dit geval (uitgaande van de fictie dat gebruik over de eerste negen maanden gelijk zou zijn aan gebruik over het hele eerste jaar)

- Aantal gebruikers in nieuwe situatie: 213;
- Aantal gebruikers in oude situatie: 3 (2 à 3 naar boven afgerond)
- Verschil 210 consumenten.
- Oude prijs €80.000
- Nieuwe prijs €0
- Verschil is €80.000
- Daarmee is het consumentensurplus $0.5 * 210 * €80.000 = €8,4$ miljoen per jaar.

De mogelijke fout in deze berekening is dat er vanuit wordt gegaan dat de vraagcurve geen curve, maar een lijn is. De methode berekent de oppervlakte van de ruimte onder deze lijn/curve. Bij een curve is dit oppervlak veel kleiner dan bij een lijn. Deze berekeningswijze komt er op neer dat de helft van deze 210 extra gebruikers meer dan €40.000 per jaar over heeft voor deze data, maar slechts werd teruggeschrikt door het bedrag van €80.000. De gedachte achter open data is juist dat gratis toegankelijke data veel extra gebruik trekken. Bij het hanteren van een vraagcurve zal het leeuwendeel van deze extra gebruikers pas de data kopen als deze laag zijn geprijsd. Dat is een veel realistischere aanname. Echter, tussen twee punten kan slechts één lijn worden getrokken, maar er kunnen door twee punten oneindig veel curves worden getrokken die ieder leiden tot een verschil in consumentensurplus. Ten behoeve van dit onderzoek is een drietal curves uitgezet in Figuur 5.3.

Figuur B5.3: Alternatieven verloop vraagcurves en 'vraaglijn' (volume vraag en verkoopprijs)



Van deze alternatieven kunnen een aantal gegevens op een rij worden gezet.

Tabel B5.2: Alternatieve curve-gegevens

	Maximale opbrengst data			Consumentensurplus open data
		Bij prijs in euro's	Aantal sets	
Lijn	€4.320.762	€40.762	106	€8.400.000
Curve A	€1.050.998	€35.033	30	€2.400.190
Curve B	€476.436	€36.649	13	€988.743
Curve C	€240.000	€80.000	3	€272.902

Indien wordt aangenomen dat de oude prijs rationeel was bepaald en er met een prijs van €80.000 per dataset inderdaad een maximale opbrengst van de data kon worden gegenereerd (het zoveel mogelijk ter beschikking stellen van deze data sluit bovendien aan bij de publieke doelstelling) dan is curve C de juiste vraagcurve en is het consumentensurplus daarmee €272.902 per jaar.

Naast het consumentensurplus van het als open data beschikbaar stellen van het gehele databestand is er ook het consumentensurplus van het gratis beschikbaar stellen van de mogelijkheid om data af te drukken, de mogelijkheid om vastgoedwaarden te bekijken (en af te drukken) en de mogelijkheid het grondmarktbericht te bekijken.

Indirecte effecten

Een volgende maatschappelijke baat ligt in de doelstelling van de referentiegrondwaarde. De wetgever heeft besloten dat er een maatschappelijke baat is wanneer er transparantie is over vastgoedmarkten (Voss 2011). Onderzoek binnen een deelregio (Rijn-Sieg) heeft uitgewezen dat deze referentiegrondwaarden bijdragen aan transparantie, met name in stedelijke gebieden, waar er voldoende vastgoedtransacties zijn om te gebruiken bij het construeren van deze referentiegrondwaarden (Wenzler & Kropp 2015). In beginsel zit de waarde die deze transparantie heeft voor partijen die gebruikmaken van de data in het consumentensurplus. Hiermee zijn de directe effecten reeds geadresseerd. Dit betreft niet de indirecte effecten van de waarde die een transparante vastgoedmarkt heeft voor het ontstaan van minder fricties tussen vraag en aanbod op vastgoedmarkten en daarmee op het bieden van adequate ruimte voor economisch ontwikkeling. Gebrek aan transparantie zou kunnen bijdragen aan grotere verschillen in kennisniveaus tussen actoren op de markt, meer volatiliteit (grotere verschillen in prijzen) waardoor de optiewaarde van een herontwikkelingsoptie zou stijgen en de drempel hoger wordt voordat een optie zal worden uitgeoefend (een afwachtende houding wat betreft investeringen). Door al deze ontwikkelingen kan de tijd worden verlengd waarop het vastgoedaanbod (wat door bouwtijd altijd al relatief traag reageert) zich aanpast aan een veranderende vraag. Dit kan ook weer gevolgen hebben voor de financiële markt – veel vastgoed wordt deels met vreemd vermogen gefinancierd. Hoe groot deze positieve effecten zijn is onbekend.

Overige effecten

Economische activiteiten hebben een ruimtelijke component. Alle economische activiteiten vinden ergens plaats. Een goede transparante marktwerking op het gebied van de vastgoedmarkt kan een bijdrage leveren aan het beter functioneren van allerlei economische activiteiten. Daar waar een transparante vastgoedmarkt leidt tot positieve gevolgen voor de woningmarkt kan dit bijdragen aan betere huisvesting. Dit kan leiden tot het meer mobiel maken van de beroepsbevolking. Het is mogelijk om te verhuizen naar plekken waar werk is. Een goed woon- en leefklimaat kan positieve

effecten hebben voor de volksgezondheid. Het is onbekend hoe groot de bijdrage van open data op het gebied van grondprijzen aan dit complex aan effecten is.

B.2.2 Conclusie

Daar waar er sprake van jaarlijkse voordelen, is het mogelijk deze te disconteren. Voorwaarde hiervoor is wel dat ook jaarlijks up-to-date data worden geleverd. De waarde van het leveren van het huidige opendatabestand zal natuurlijk door veroudering snel afnemen. Op grond van de MKBA-leidraad is in het onderstaande een discontovoet van 5,5% (2,5% reëel risicovrij + 3% opslag voor macro-economisch risico) gehanteerd.

Tabel B5.3: Overzicht MKBA open data referentiegrondwaarde in Noordrijn-Westfalen, in euro's

in €	Per jaar	Totaal
Kostenbesparing	12.000	218.000
Consumentensurplus	273.000	4.962.000
Indirecte effecten	Pm	pm
Overige effecten	Pm	pm
Totaal	285.000 + pm	5.180.000 +pm

Het als open data aanbieden van de referentiegrondwaarde zorgt voor een maatschappelijke baat van ruim €5 miljoen plus pro memorie indirecte en overige effecten. Deze baat wordt voor een belangrijk deel gevormd door het consumentensurplus. Dit is echter een schatting die niet volledig hard te maken is. Veel belangrijker zijn de potentiële indirecte en overige effecten: een goed functionerende vastgoedmarkt is van groot belang voor de economische ontwikkeling. Het als open data beschikbaar stellen van referentiegrondwaardes levert een bijdrage aan deze transparantie.

Bijlage 6 Internationale vergelijking inkomen/belastingdata

In Nederland zijn werk- en inkomstendata geen openbare informatie, en daarom is een internationale vergelijking moeilijk te maken. In deze bijlage hebben we daarom gekozen om inzage in belastinggegevens in twee landen te beschrijven omdat de Polisbusgegevens die in Hoofdstuk 7 zijn beschreven, van oorsprong gegevens van de Belastingdienst zijn.

B.1. Noorwegen

Noorwegen kent van oudsher de mogelijkheid om belastinggegevens van andere burgers ten kantore van de belastingdienst in te zien. Met de informatisering van belastingaangiftes en gegevensverstreking is deze informatie nu online beschikbaar. Door de pers worden deze mogelijkheden benut om Quote500-achtige lijstjes van Noren die het meest verdienen op te stellen. Ook biedt het burgers de mogelijkheid om zelf informatie te verzamelen over inkomen en belastingafdracht van bijvoorbeeld de burens, familieleden, de potentiële en aanstaande schoonfamilie en potentiële klanten.

De informatie bevat de volgende gegevens

- Naam
- Geboortjaar (organisatienummer bij bedrijven)
- postcode
- woonplaats
- belasting-gemeente
- netto inkomen (algemeen inkomen voor bijzondere toeslagen)
- netto vermogen
- berekende belasting.

Voor personen worden de volgende gegevens niet opgenomen:

- informatie over personen met een geblokkeerd adres (als dit is vastgelegd in het bevolkingsregister)
- informatie over personen waar deze informatie kan onthullen over een klantrelatie (de Noorse term is 'klientforhold' en deze wordt bijvoorbeeld gebruikt voor de relatie tussen advocaten en hun cliënten)
- informatie over minderjarige personen
- informatie over mensen zonder een permanente verblijfsvergunning
- informatie over overledenen.

Ook voorlopige gegevens kunnen worden gepubliceerd en deze worden drie keer per jaar herzien. Vanaf het belastingjaar 2013 zijn er door het Noorse Parlement extra eisen opgesteld aan het raadplegen van deze informatie. Deze behelzen dat de belastingdienst moet bijhouden wie zoekopdrachten uitvoert op deze belastinggegevens. Volgens de Engelstalige site van de belastingdienst betekent dit het volgende:

“This will enable you to find out who has searched for you in the tax lists. When you have logged in to search the tax lists, there is a separate tab where you can view statistics concerning who has searched for you. You can see the name, year of birth, postcode and postal town of a person who has searched for you. The list of who has searched for and

viewed your information is updated hourly.” (<http://www.skatteetaten.no/en/person/Tax-settlement/Search-in-the-tax-lists/>; geraadpleegd op 5/10/2016)

Naast deze gegevens wordt ook inzicht gegeven in het tijdstip waarop deze zoekopdracht is verstrekt. Deze informatievoorziening zet een rem op het zomaar raadplegen van gegevens van bekenden, omdat deze dan weten dat deze informatie (achter hun rug om) is opgevraagd.

Bovendien zijn er nog twee restricties. Allereerst moet je minimaal 16 jaar oud zijn voordat je mag zoeken in deze gegevens. Voorts is het opvragen van belastinggegevens beperkt tot 500 mensen per maand.

Al eerder (vanaf 2011) waren er restricties opgelegd aan het opnieuw publiceren van deze data. Er waren daarvoor onlinebestanden waarin deze data werden gedupliceerd.

“The possibility of searching tax lists via online newspapers ceased with effect from 2011. The Parliament introduced this restriction. The aim is to prevent the information from being used for commercial purposes. However, the Directorate of Taxes cannot compel the press to delete tax lists from previous years. The quality of these lists will inevitably deteriorate as they become older.” (<http://www.skatteetaten.no/en/person/Tax-settlement/Search-in-the-tax-lists/>; geraadpleegd op 5/10/2016)

Voor journalistieke doeleinden heeft de pers wel toegang tot de data; daarvoor dient een overeenkomst te worden gesloten die het gebruik van deze data reguleert:

“By signing the agreement, the press undertakes to not to publish online or disclose to others any part of a tax list. However, the press can still refer to private individuals in newspaper articles and publish 'rich lists'.”¹⁴¹

Kortom, Quote500-achtige lijstje met gegevens over netto inkomen, netto vermogen en betaalde belasting blijven mogelijk en het is ook mogelijk om over personen te publiceren die bijvoorbeeld actief zijn in het bedrijfsleven, oud-politici en directeuren van organisaties die publieke dienstverlening verzorgen.

Het betreft dus geen open data, maar het gaat om het aan de nieuwe technologische mogelijkheden aanpassen van de toegankelijkheid van deze data. Hierbij zijn er een aantal maatregelen genomen om de verdere verspreiding die het internet mogelijk maakt, in te perken. Ook in de oude situatie is het mogelijk om naar het belastingkantoor te gaan en daar deze gegevens in te zien. In Noorwegen wordt het van oudsher van publiek belang geacht dat een belastingplichtige inzage kan hebben in belastinggegevens van andere burgers. In *The Guardian*¹⁴² stelt de directeur communicatie van Skatteetaten, Mariken Holter: “It has been in place a very long time and is generally accepted by the people of Norway”. In dit artikel wordt ook een relatie gelegd met de waarden en normen (en kleine mate van inkomensongelijkheid) in de Noorse samenleving die afwijken van die in de Britse samenleving.

Voor de Nederlandse discussie is het uitgangspunt anders. Hier wordt juist uitgegaan van het gegeven dat het in het publiek belang is dat belastinggegevens vertrouwelijk zijn. Hierbij is er voor bepaalde functionarissen (in de publieke sector) wel regelgeving opgesteld rond maximumsalarissen

¹⁴¹ <http://www.skatteetaten.no/en/person/Tax-settlement/Search-in-the-tax-lists/>

¹⁴² <https://www.theguardian.com/money/blog/2016/apr/11/when-it-comes-to-tax-transparency-norway-leads-the-field>

en is er sprake van een publicatieplicht rond bepaalde gegevens, maar deze plicht is duidelijk gericht op slechts bepaalde functies. De discussie over het publiceren van gegevens over inkomens, vermogen en betaalde belasting gaat veel verder dan de kosten en baten van open data. Het betreft een fundamentele wijziging – een politieke keuze – over hoe Nederland informatie over inkomstengegevens toegankelijk wil maken en geen technische exercitie van kosten en baten.

De toegang tot de data wordt verzorgd via een ICT systeem. De IT van de Noorse belastingdienst wordt verzorgd door *Skatteetatens IT- og servicepartner* (SITS), waar 880 medewerkers werken op kantoren te Oslo, Grimstad en Lillehammer. De kosten van de belastingadministratie was in 2015 €85,93 per belastingbetaler.¹⁴³ Hiervan was €26,75 voor ICT. Dit laatste bedrag is hoog vergeleken met Denemarken, Zweden en Finland. Door de Noorse belastingdienst wordt dit verklaard door het relatief (vergeleken met andere Noordse landen) hoge percentage van digitalisering. Er is hier echter geen uitsplitsing voor de kosten van de verstrekking van data aan anderen.

B.2. Verenigd Koninkrijk

In het VK mogen belastinggegevens in beginsel niet worden gedeeld met anderen dan de belastingbetaler zelf. Uitzonderingen hierop zijn specifiek omschreven in de wet en er zijn ook nog richtlijnen opgesteld. De *Commissioners for Revenue and Customs Act 2005* (CRCA, artikel 18) regelt dat informatie van HM Revenue & Customs (HMRC) niet mag worden verstrekt aan een ander tenzij hier een juridische basis voor is, bijvoorbeeld aan andere overheidsdiensten, lokale overheden en de politie.

Commissioners for Revenue and Customs Act 2005 artikel 18 lid 1 en 2

- (1) Revenue and Customs officials may not disclose information which is held by the Revenue and Customs in connection with a function of the Revenue and Customs.
- (2) But subsection (1) does not apply to a disclosure—
 - (a) which—
 - (i) is made for the purposes of a function of the Revenue and Customs, and
 - (ii) does not contravene any restriction imposed by the Commissioners,
 - (b) which is made in accordance with section 20 or 21,
 - (c) which is made for the purposes of civil proceedings (whether or not within the United Kingdom) relating to a matter in respect of which the Revenue and Customs have functions,
 - (d) which is made for the purposes of a criminal investigation or criminal proceedings (whether or not within the United Kingdom) relating to a matter in respect of which the Revenue and Customs have functions,
 - (e) which is made in pursuance of an order of a court,
 - (f) which is made to Her Majesty's Inspectors of Constabulary, the Scottish inspectors or the Northern Ireland inspectors for the purpose of an inspection by virtue of section 27,
 - (g) which is made to the Independent Police Complaints Commission, or a person acting on its behalf, for the purpose of the exercise of a function by virtue of section 28, or
 - (h) which is made with the consent of each person to whom the information relates."

Er wordt verwezen naar de artikelen 20 en 21 in lid 2. Artikel 21 gaat over de 'Disclosure to prosecuting authority' (verstrekking aan vervolgende autoriteit) en artikel 20 gaat over 'public interest disclosures'. Het publiek belang werd hier echter niet ruim omschreven. Er is een richtlijn (IDG60100) gepubliceerd over Public Interest Disclosures. Deze zijn gedefinieerd als zaken "where

¹⁴³ Årsrapport 2015 for Skatteetaten [Jaarrapportage 2015 van de Belastingdienst], <http://www.skatteetaten.no/globalassets/pdf/arsrapport2015skatteetaten.pdf>, blz. 15.

there is no appropriate statutory gateway, but where the public could reasonably expect HMRC to disclose information (for example, in order to meet legitimate public concern, or to prevent harm to the public)". Dit wordt strikt uitgelegd zie IDG60232. Verstrekking van gegevens over inkomstenbelasting aan derden belangstellenden (waaronder de tabloids) valt hier niet onder. Deze casus lijkt daarom niet relevant om verder te onderzoeken.

Kenniscentrum open data

Faculteit Bouwkunde, TU Delft

Postbus 5030, 2600 GA Delft

Telefoon: +31 (0)15 278 30 05

E-mail: opendata-bk@tudelft.nl

www.opendata.bk.tudelft.nl