

Masterthesis

De publieke waarde(n) van big, open en linked data

Marijn Detiger, s1395106

Begeleider: Dr. T. Kerkhoff

Tweede lezer: Prof. dr. Z. van der Wal

Management van de Publieke Sector

Universiteit Leiden

januari, 2018



VOORWOORD

Met het afronden van de scriptie 'De publieke waarde(n) van big, open en linked data' komt voor mij een einde aan de Master Management van de Publieke Sector van de Universiteit Leiden. De afgelopen jaren heb ik met plezier gestudeerd in Den Haag, het politiek-bestuurlijke hart van Nederland. Als afsluiting heb ik een afstudeeronderwerp gekozen dat me sinds mijn functie als data-analist bij het Centraal Bureau voor de Statistiek (CBS) heeft geïnteresseerd.

In dit onderzoek staat het gebruik van *big, open en linked data* (BOLD) door de overheid centraal. Hoewel de techniek en mogelijkheden van data al regelmatig onderwerp van onderzoek zijn, heb ik in mijn scriptie uiteindelijk de normatieve vraag gesteld wat het gebruik van BOLD kan betekenen voor de kwaliteit van overheidsbestuur. Daarvoor bedank ik voor hun input: Chantal Brakus en Carin Kruiskamp (van het CBS) en ook Thijmen Schep en Tom van de Wetering (die voor de organisatie SETUP onderzoeken welke impact nieuwe technologieën op de samenleving hebben).

Graag dank ik verder in het bijzonder Toon Kerkhoff, die tijdens het begeleiden van mijn scriptie vader is geworden. Ondanks de slapeloze nachten wees hij me met scherpe vragen telkens de juiste richting op. Zijn uitgebreide kennis van het onderwerp en onze periodieke bijeenkomsten hebben veel bijgedragen aan de vervolmaking van deze scriptie.

Ook bedank ik professor Zeger van der Wal voor zijn bereidheid om tweede lezer te zijn. Als autoriteit op het gebied van publieke waarden is zijn beoordeling waardevol. Het schrijven van deze scriptie was daarnaast niet mogelijk geweest zonder de medewerking van de respondenten. Hun behulpzaamheid droeg bij aan een goed verloop van het onderzoek. Tevens maakten de handreikingen en bemoedigende woorden van collega's, vrienden en familie het proces een stuk aangenamer.

Van het begin af aan was het mij duidelijk dat het schrijven van een scriptie gecombineerd met een fulltime baan een uitdaging zou zijn. Ook daarom ben ik trots op het eindresultaat. Tot slot bedank ik Stacey voor haar begrip en het geduld dat zij heeft moeten opbrengen tijdens dit onderzoek. Ik besef dat dit niet altijd even makkelijk is geweest!

Ik wens u veel leesplezier.

Marijn Detiger

Den Haag, januari 2018



MANAGEMENTSAMENVATTING

De afgelopen jaren hebben technologische en maatschappelijke vooruitgang het mogelijk gemaakt om steeds gemakkelijker data te verzamelen, op te slaan en te analyseren. De termen *big*, *open* en *linked data* (ook wel: BOLD) hebben daarmee steeds meer aandacht gekregen. De interesse vanuit de overheid is begrijpelijk: data spelen een grote rol bij overheidshandelen. Over het algemeen is men dan ook optimistisch over de mogelijkheden van het gebruik van BOLD. Veel minder aandacht gaat uit naar de bedreigingen die voortvloeien uit (her)gebruik van data, het consolideren van gegevens uit meerdere bronnen, het toepassen van *data-analytics* op gecombineerde datasets en het handelen op basis van de resultaten. Het doel van dit onderzoek is om gestructureerd en exploratief onderzoek te doen naar de normatieve consequenties van BOLD op de kwaliteit van de overheid. De centrale onderzoeksvraag is: “*Wat zijn de gevolgen van big, open en linked data voor goed openbaar bestuur?*”.

Goed openbaar bestuur kan gedefinieerd worden als bestuur dat belangrijke morele standaarden en principes bevordert en borgt (publieke waarden). Op basis van de bekendheid in de wetenschappelijke literatuur en de toepasbaarheid in de context van BOLD, zijn vervolgens 7 publieke waarden geselecteerd. Aan de hand hiervan zijn veronderstellingen opgesteld, die zijn getoetst aan de opvattingen en inschatting van 11 betrokken personen uit de praktijk. Uit de resultaten blijkt dat veel van het enthousiasme voor BOLD wordt ingegeven door de toegenomen technologische mogelijkheden om maatschappelijke problemen kwantificeerbaar en ‘objectief’ te maken. Daarmee zou BOLD tot beter, dat wil zeggen, efficiënter en effectiever bestuur kunnen leiden. Rechtmatigheid van overheidshandelen, gelijkheid tussen burgers en respect voor menselijke waardigheid zijn lastiger te borgen. Het risico van BOLD (hoofdzakelijk *big data*) is dat de privacy van burgers afneemt, de kans op discriminatie en vooroordelen toeneemt en het *big brother* probleem groter wordt. Dat hangt samen met de verantwoording over datagebruik en participatie van burgers, die volgens dit onderzoek vaak achterblijven.

Een van de aanbevelingen van deze studie is dat - ondanks de grote beloften van BOLD – de verdere uitwerking gelijktijdig plaatsvindt met het ontwikkelen van waarborgen om de mogelijke risico’s te beperken. Het is daarom goed om bij dataprojecten de normatieve uitgangspunten in beschouwing te nemen; data en ‘feiten’ zijn namelijk niet altijd zo objectief als het lijkt. Over de kwaliteit van de data, de definitie van concepten, de keuzes van methoden en de interpretatie van resultaten, is altijd discussie mogelijk. Door hier voldoende aandacht aan te schenken en publieke waarden te borgen en te bevorderen, kan bij burgers blijvend vertrouwen ontstaan in de manier waarop overheden bij de uitvoering van hun taken van BOLD gebruik maken.



INHOUD

VOORWOORD.....	2
MANAGEMENTSAMENVATTING	3
INHOUD	4
1 INLEIDING	6
Maatschappelijke relevantie	8
Wetenschappelijke relevantie	9
1.1 Leeswijzer	10
2 THEORETISCH KADER.....	11
2.1 Het publieke van de overheid	11
2.2 Goed openbaar bestuur	12
2.3 Een selectie van publieke waarden	15
2.4 Waardenbotsingen	18
2.5 BOLD	19
Big data	19
Open data	24
Linked data	25
2.6 Hypothesen	25
3 ONDERZOEKSOPZET	30
3.1 Onderzoekseenheid.....	30
3.2 Dataverzameling en -analyse	30
3.3 Operationalisering.....	32
Onafhankelijke variabele	32
Afhankelijke variabelen	33
3.4 Betrouwbaarheid en validiteit.....	34
4 ANALYSE	36
4.1 Het gebruik van BOLD door gemeenten	36
4.2 De invloed van BOLD op efficiëntie	39
4.3 De invloed van BOLD op effectiviteit.....	40
4.4 Rechtmatigheid bij het gebruik van BOLD.....	41
4.5 Verantwoording en het gebruik van BOLD.....	42
4.6 BOLD en gelijkheid.....	44
4.7 Participatie van burgers bij het gebruik van BOLD.....	45
4.8 BOLD en respect voor wenselijke waardigheid	46
4.9 Toetsing hypothesen	47



5	CONCLUSIE EN AANBEVELINGEN	50
5.1	Beantwoording onderzoeksvraag.....	50
5.2	Discussie	51
5.3	Praktische aanbevelingen.....	52
6	VERWIJZINGEN	54
7	BIJLAGEN	60
7.1	Bijlage 1: Interview vragenlijst	60
	Introductie	60
	Interviewvragen	61



1 INLEIDING

De afgelopen jaren hebben technologische en maatschappelijke vooruitgang het mogelijk gemaakt om steeds gemakkelijker data te verzamelen, op te slaan en te analyseren. Bijna elk elektronisch apparaat heeft toegang tot het internet en kan gekoppeld worden aan extra computerkracht en analytische mogelijkheden (ook wel het *Internet of Things* (IoT) (Gubbia, Buyyab, Marusic & Palaniswami, 2013). De explosie van gegevens belooft enorme gevolgen te hebben voor onder andere de economie, politiek, veiligheid, wetenschap, onderwijs, beleid, bestuur, gezondheidszorg en volksgezondheid (Metcalf, Keller & Boyd, 2017). Als belangrijk onderdeel van hun *smart city* agenda, zetten ook steden in op meer gebruik van data (Caragliu, Del Bo & Nijkamp, 2011). Een *smart city* verbetert de prestaties van een stad door gebruik te maken van (*big*) data en informatietechnologieën (IT) om onder andere efficiëntere dienstverlening te bieden aan burgers, effectiever te monitoren, de bestaande digitale infrastructuur te verbeteren, samenwerking tussen economische actoren te vergroten en innovatie aan te moedigen in de private en publieke sector (Marsal-Llacuna, Colomer-Llinàs & Meléndez-Frigola, 2014).

Het verzamelen, verwerken, opslaan en delen van data is al langere tijd een van de kerntaken van de publieke sector. Publieke organisaties zijn een van de grootste dataproducenten en -verzamelaars (Janssen, 2011). Adequate opslag van informatie over functies, taken, beleid, dataverzameling en -verwerking, beleidsuitkomsten en processen is essentieel voor het functioneren en de transparantie van een organisatie (Janssen & Van den Hoven, 2015). De interesse vanuit de overheid is dus begrijpelijk: data spelen een grote rol bij overheidshandelen. Er zijn bijna geen beleidsterreinen die niet (in)direct worden beïnvloed door inzichten op basis van gegevens. De termen *big*, *open* en *linked data* (ook wel: BOLD) krijgen daarmee steeds meer aandacht van onderzoekers, professionals en de media met betrekking tot zowel de private als publieke sector. Bij *big data* gaat het om grote volumes aan gegevens die vanuit meerdere bronnen kunnen worden verwerkt (McAfee & Brynjolfsson, 2012). *Open data* maakt toegang tot bepaalde gegevens mogelijk voor overheden, burgers en bedrijven zonder vooraf gedefinieerde beperkingen of gebruiksvoorwaarden (Janssen & van den Hoven, 2015). *Linked data* heeft betrekking op koppelen en verbinden van (online) data op basis waarvan analyses kunnen worden gedaan (Bizer, Heath & Berners-Lee, 2009). BOLD kan van grote invloed zijn op het functioneren van overheidsorganisaties en de interactie met de samenleving (Janssen & Van den Hoven, 2015).



Vaak worden in gemeenten al op kleinschalige en incidentele basis gegevens verzameld die worden aangevuld met inzichten uit het openbaar bestuur. Bronnen zijn onder andere interviews, eenmalige enquêtes en andere kwalitatieve data, die op geaggregeerd niveau worden geanalyseerd en daarmee een momentopname van de stad opleveren (Kitchin, 2016). Tegenwoordig kunnen datasets in toenemende mate worden aangevuld met (nieuwe) bronnen die andere kenmerken hebben dan de traditionele, kleine datasets. De gegevens worden bijna *real-time* gegenereerd en hebben betrekking op een veel groter deel van de stad (Kitchin, 2014). Dit soort data worden bijvoorbeeld mogelijk gemaakt door camera's, sensoren, meters en smartphone apps en heeft *datafication* tot gevolg: de radicale uitbreiding van het volume, bereik en detail van de data die iets over personen en plaatsen zeggen (Crawford & Schultz, 2014). Van de analyses op dit soort data wordt op lokaal niveau veel verwacht. Aandacht gaat bijvoorbeeld uit naar het oplossen van uitdagingen op het gebied van energie, transport en criminaliteit (Boer & van der Lans, 2014). Volgens Mitlin (2004) gaat daarmee samen dat stadsbesturen een steeds belangrijkere rol toebedeeld krijgen bij het aanpakken van deze stad-overstijgende problemen.

Het gebruik van BOLD wordt dan ook steeds populairder in de Nederlandse publieke sector. De Nationale Politie gebruikt *big data* bijvoorbeeld om diefstal en straatroven te voorspellen (Van Dijk, 2015). De Belastingdienst analyseert data om risicoprofielen op te stellen en belastingontduiking en -fraude op te sporen (Maurits, 2015). En ook op gemeentelijk niveau ontstaan meer en meer (*big*) dataprojecten, vaak als onderdeel van een *smart city* agenda. Kwaliteitsinstituut Nederlandse Gemeenten (KING) moedigt de afgelopen jaren al data-innovaties bij gemeenten aan (Kwaliteitsinstituut Nederlandse Gemeenten, 2016). Gemeenten staan voor de uitdaging om aan te sluiten op ingrijpende ontwikkelingen zoals bezuinigingen, decentralisaties, een complexere samenleving en stad-overstijgende problematiek. Steeds meer wordt samengewerkt om maatschappelijke problemen aan te pakken, ook op het gebied van data. Zo startten de gemeente Almere en onderzoeksbureau TNO in 2013 het Big Data Value Center, een platform waar verschillende partijen en organisaties experimenteren met *big data* (TNO, 2013). Ook wordt met private partijen samengewerkt. BOLD lijkt het publieke dienstverleningsdomein nog niet te domineren, maar kan in de toekomst niet meer weggedacht worden als onderdeel van beleidsvorming (Wetenschappelijke Raad voor Regeringsbeleid, 2016a).

Zowel de professionele als de wetenschappelijke literatuur zijn optimistisch over de mogelijkheden van het gebruik van BOLD. Volgens Clarke (2016) is het enthousiasme zelfs doorgeslagen naar een onbeheerste euforie. In veel mindere mate wordt aandacht besteed aan de consequenties van BOLD, bijvoorbeeld voor goed openbaar bestuur. De Wetenschappelijke Raad voor Regeringsbeleid (WRR)



waarschuwde in 2011 al dat de relatie tussen het openbaar bestuur en haar burgers verandert door de overgang naar een informatieoverheid. Politiek enthousiasme gaat vaak hand in hand met waarden als efficiëntie en doelmatigheid, terwijl waarden als transparantie en verantwoording veel minder aandacht krijgen (Wetenschappelijke Raad voor Regeringsbeleid, 2011). Utrecht Data School (2017) ontworpen de De Ethische Data Assistent (DEDA) om ethische problemen in dataprojecten, datamanagement en databeleid te herkennen. Ook andere maatschappelijke partijen roepen organisaties, bestuurders en lokale overheden op om zich bij *big data*-toepassingen bewust te zijn van ethiek en het behartigen van publieke belangen (Schep, 2016; Weerdt & Vries, 2014; Schäfer, 2017; Council for Big Data, Ethics, and Society, 2014). Desondanks lijkt de aandacht in de praktijk en de literatuur elders te liggen. Met het oog op publieke waarden als verantwoording, rechtmatigheid en gelijkheid, zouden overheidsorganisaties zich bewust moeten zijn van de normatieve gevolgen van BOLD-toepassingen. Het is van belang om de effecten van het toenemend gebruik van data en digitale intelligentie op de publieke sector te onderzoeken, omdat aantasting van publieke waarden kan leiden tot minder goed openbaar bestuur. Gezien het bovenstaande, staat de volgende onderzoeksvraag in deze studie centraal: *“Wat zijn de gevolgen van big, open en linked data voor goed openbaar bestuur?”*.

Het doel van de studie is om gestructureerd en exploratief onderzoek te doen naar de normatieve consequenties van BOLD op publieke waarden van goed openbaar bestuur. Allereerst wordt een literatuurverkenning gedaan op basis waarvan toetsbare veronderstellingen worden opgesteld. Vervolgens worden de verwachtingen aan de hand van de opvattingen en inschatting van betrokken personen uit de praktijk onderzocht (hierover volgt in latere hoofdstukken meer). Daarmee levert dit onderzoek grondstof voor volgende onderzoekers en doet het een toevoeging aan de schaarse literatuur. Op basis van de resultaten kan gereflecteerd worden op de stelling van Clarke (2016) dat de onbeheerste euforie over BOLD gematigd zou moeten worden omdat publieke waarden in het geding zijn. Daarbij kunnen aanbevelingen gegeven worden om het gebruik van BOLD in de toekomst zo in te richten dat publieke waarden zo veel mogelijk worden gewaarborgd en zo min mogelijk worden aangetast.

Maatschappelijke relevantie

De wetgeving die over (*big*) datagebruik bestaat richt zich voornamelijk op het verzamelen van gegevens en niet zozeer op wat met de data gebeurt (Van der Weerdt & De Vries, 2014). Om in grijze gebieden de juiste keuzes te maken, wordt ethiek steeds belangrijker. De overgang van kleinschalige en trage data naar uitgebreide en snelle data als gevolg van nieuwe digitale technologieën, ligt aan de basis van de ontwikkeling van *smart cities*. Een van de consequenties is dat stadssystemen en (digitale) infrastructuren meer geïntegreerd en gelinkt worden (een verdere verdieping volgt in hoofdstuk 2)



(Kitchin, 2016). Voor de samenleving is het belangrijk om zicht te krijgen op waarden die in het openbaar bestuur botsen en hoe bestuurders, managers en uitvoerders daarmee omgaan bij *big data*-analyses, profileringssoftware en het beïnvloeden van keuzegedrag op basis van data (*nudging*). Onderzoek naar de afweging van waarden bij datagebruik kan daarbij van pas komen.

Overheidsorganisaties en bedrijven werken daarnaast steeds meer samen aan innovatieve oplossingen om maatschappelijke problemen aan te pakken. Hierdoor vervagen de grenzen tussen publiek en privaat en is het de vraag welke (onbedoelde) effecten dit heeft op publieke waarden voor goed openbaar bestuur (hierover volgt het volgende hoofdstuk 2). Onderzoek naar de afweging en botsing van verschillende publieke waarden met betrekking tot BOLD kan bestuurders en (gemeente)ambtenaren helpen om bewustere keuzes te maken. Burgers zijn met dit onderzoek tevens beter op de hoogte van de mogelijke gevolgen van BOLD en de invloed die dit heeft op hun directe persoonlijke situatie. Dit kan het vertrouwen van burgers in de overheid en de politiek ten goede komen. Het vertrouwen is van belang omdat dit direct verband houdt met de legitimiteit van de overheid (Bovens & Wille, 2008).

Wetenschappelijke relevantie

De wetenschap is overwegend positief over BOLD. De belofte is dat met nieuwe *datascience*-technieken bredere, diepere en actuelere kennis vergaard kan worden in vergelijking met traditionele methoden (Kitchin, 2016). Veel minder aandacht gaat uit naar de bedreigingen die voortvloeien uit (her)gebruik van data, het consolideren van gegevens uit meerdere bronnen, het toepassen van *data-analytics* op de gecombineerde datasets en het handelen op basis van de resultaten (Clarke, 2016). Ook Janssen & Van den Hoven (2015) geven in hun studie aan dat er behoefte is aan fundamenteel en toegepast onderzoek naar de vraag hoe *big*, *open* en *linked data* publieke organisaties en de samenleving (zouden moeten) beïnvloeden. Ondanks deze suggesties, is het aantal onderzoeken over dit onderwerp gering. Een van de oorzaken hiervan kan zijn dat er naar publieke waarden en BOLD nog weinig gestructureerd onderzoek is gedaan, waardoor grondstof voor vervolgonderzoek ontbreekt. Tevens bestaat er veel conceptuele discussie over publieke waarden en kunnen methodologische uitdagingen het in de praktijk minder aantrekkelijk maken om onderzoek te doen naar publieke waarde(n).

Daarnaast is de bestaande aandacht voor de normatieve consequenties van BOLD beperkt omdat een klein aantal waarden wordt onderzocht (vaak zijn dit privacy en transparantie, zie bijvoorbeeld: Janssen & Van den Hoven, 2015). Het is overigens begrijpelijk dat deze waarden vaak genoemd worden. Door het IoT ontstaan steeds meer digitale data die met behulp van slimme algoritmes iets over personen kunnen zeggen. De *big data* analyses zijn echter voor gebruikers en zelfs programmeurs



vaak ongebruikelijk en onvoorspelbaar (Crawford & Schultz, 2014). Privacy en transparantie zijn daardoor lastig te waarborgen. Om een uitspraak te kunnen doen over de gevolgen van BOLD op de kwaliteit van de overheid, is het echter essentieel om onderzoek te doen naar meer publieke waarden van goed openbaar bestuur. Op basis van een breed scala aan publieke waarden wordt het overheidshandelen namelijk gelegitimeerd en de idee van de goede samenleving gevormd (een verdere verdieping volgt in hoofdstuk 2) (Rutgers, 2011).

1.1 Leeswijzer

In deze studie staat de vraag centraal welke gevolgen BOLD heeft op goed openbaar bestuur en welke spanning bij het gebruik hiervan tussen publieke waarden ontstaat. In het volgende hoofdstuk 'Theoretisch kader' wordt dieper ingegaan op de betekenis en ontwikkeling van publieke waarden van goed openbaar bestuur. Tevens wordt een uiteenzetting van BOLD gepresenteerd. Het hoofdstuk sluit af met enkele theoretische veronderstelling, in het bijzonder over de consequenties van *big data*-analyses, datadeling en het koppelen van datasets voor publieke waarden voor goed openbaar bestuur. Vervolgens wordt in het derde hoofdstuk een beschrijving gegeven van de methoden en technieken die aan de basis van het empirische werk liggen. Er wordt ingegaan op de casus, dataverzameling en data-analyse, validiteit en betrouwbaarheid van het onderzoek. De casus van dit onderzoek betreft in het bijzonder het gebruik van BOLD door de gemeente Den Haag rondom het UDC met het CBS. Het vierde hoofdstuk 'Analyse' presenteert de resultaten van het onderzoek door een uiteenzetting te geven over de resultaten van de dataverzameling en -analyse. In dit hoofdstuk worden de theoretische verwachtingen getoetst. Tot slot volgt de conclusie waarin de belangrijkste bevindingen van het onderzoek worden samengevat en de onderzoeksvraag wordt beantwoord. Tevens wordt op het onderzoek gereflecteerd en worden beperkingen van het onderzoek toegelicht. Tot slot volgen enkele praktische aanbevelingen.



2 THEORETISCH KADER

In dit hoofdstuk worden bevindingen uit de wetenschappelijke literatuur besproken die van belang zijn om de onderzoeksvraag te beantwoorden. Allereerst staat centraal wat overheidsorganisaties en private organisaties van elkaar onderscheidt in het nastreven van waarden. Dit is ten eerste van belang omdat met het woord ‘publieke’ in de term ‘publieke waarden’, vaak een scheiding met de private sector wordt gesuggereerd. Ten tweede is dit van belang omdat toenemende interactie tussen publieke en private partijen (*governance*) heeft geleid tot de vraag welke waarden voor goed overheidsbestuur van belang zijn en of deze verschillen van bedrijfswaarden. Ook op het gebied van BOLD is steeds meer samenwerking tussen publieke en private organisaties zichtbaar. Vervolgens wordt ingegaan op de definitie van waarden en welke waarden in relatie tot goed openbaar bestuur en BOLD belangrijk zijn. Hierop volgt een bespreking van de betekenis en mogelijkheden van het gebruik van *big, open* en *linked data*. Tot slot worden de verwachtingen van dit onderzoek opgesteld over de consequenties van BOLD op publieke waarden van goed openbaar bestuur.

2.1 Het publieke van de overheid

Op verschillende niveaus denken beleidsmakers na over de vraag of specifieke goederen of diensten het best geproduceerd kunnen worden door de overheid, de private sector, publiek-private samenwerking of hybride instituties. De vraag is onder welke voorwaarden het publiek het best gediend wordt en wat de publieke waarde van voorzieningen, goederen en diensten is. In toenemende mate wordt de allocatie van goederen en diensten toegewezen aan de private sector, bijvoorbeeld door middel van privatisering, *contracting-out* en samenwerking tussen publieke en private organisaties. Een van de argumenten is dat de markt in bepaalde opzichten efficiënter is en dat de overheid private denkbeelden over zou moeten nemen (Bozeman, 2002). Als gevolg hiervan zijn volgens Rainey (1997) en Frederickson (1991) de veronderstellingen over de aanbieders van goederen en diensten en met name de betekenis van ‘publiek’ meer en meer ambigu geworden (Bozeman, 2002).

Op basis van Rainey (1997), Bozeman & Bretschneider (1994) en Scott & Falcone (1998) zijn er meerdere benaderingen te onderscheiden die de demarcatie tussen publiek en privaat duiden (Pesch, 2008). In de private sfeer zou het bijvoorbeeld vooral gaan om het produceren van individuele goederen, keuzevrijheid voor personen en de mogelijkheid om het eigen belang na te streven (Pesch, 2008). Stereotyperend wordt de private sector dan ook geassocieerd met waarden als efficiëntie en marktwerking (Reynaers, 2013). De publieke sfeer zou daarentegen vooral betrekking hebben op het nastreven van publiek belang, het beïnvloeden van de samenleving, de focus op gemeenschappelijke



waarden en het betrekken van maatschappelijke actoren om organisatiedoelen te behalen. Waarden als solidariteit, verantwoording en rechtvaardigheid zouden volgens auteurs in deze sector centraal staan (Reynaers, 2013). Een combinatie van politieke en economische overwegingen bepaalt voor een belangrijk deel dus de houding en het handelen van een organisatie. De overwegingen hangen sterk samen met de waarden die een organisatie nastreeft (Bozeman, 2007). Het normatieve verschil tussen publiek en private organisaties is volgens Pesch (2008) afhankelijk van het 'publieke' gehalte van de geproduceerde goederen en het 'publieke' van het belang dat wordt gediend. Met normatief worden waarden bedoeld die nagestreefd 'zouden moeten worden' in bepaalde sectoren of organisaties. Dit in tegenstelling tot empirisch, dat ingaat op de vraag welke waarden daadwerkelijk in de praktijk nagestreefd worden.

Een strikte grens tussen de publieke en private sector wordt volgens verschillende auteurs betwist. Vaak spreken zij over *blurring boundaries* (Reynaers, 2013). Ook empirisch onderzoek toont aan dat de publieke en private sector bepaalde waarden gemeenschappelijk hebben (Van der Wal, Huberts, Van den Heuvel & Kolthoff, 2006). Een van de oorzaken is dat de private sfeer steeds meer voet aan de grond heeft gekregen. Onder meer komt dit door de veranderde kijk op de rol van het openbaar bestuur (van *government* naar *governance*). Vooral *New Public Management* (NPM) heeft volgens Osborne & Gaebler (1992) benadrukt dat de overheid moet sturen in plaats van roeien (Pesch, 2008). Van de overheid wordt niet langer verwacht dat deze centraal bepaalt welke richting de samenleving op gaat. In plaats daarvan is de overheid één van de spelers die samen met andere partijen beleid moet vormen en uitvoeren om maatschappelijke vraagstukken op te lossen. Volgens sommige auteurs (zie bijvoorbeeld Stone (1997), Brown (1994) en Zerbe & McCurdy (1999) lijdt de *publicness* van de publieke sector hieronder (Bozeman, 2002). De focus op effectiviteit en efficiëntie zou bijvoorbeeld ten koste gaan van de betrouwbaarheid en verantwoording.

2.2 Goed openbaar bestuur

Door de toenemende interactie tussen publieke en private organisaties en grotere complexiteit van de samenleving is er dus meer aandacht gekomen voor goed openbaar bestuur. Zo is er over de jaren heen veel discussie geweest over de vraag hoe de overheid of het openbaar bestuur zou moeten functioneren (De Graaf & Van der Wal, 2010). Een algemene definitie van goed openbaar bestuur is volgens Hendriks & Drosterij (2012) bestuur dat handelt in overeenstemming met essentiële kwaliteitsnormen (De Graaf, Huberts, & Smulders, 2013). Met betrekking tot goed openbaar bestuur of *good governance* is het vervolgens de vraag welke publieke waarden in de overheid dan van belang zijn en of deze anders zijn dan private waarden. Het idee van het bevorderen en bewaken van ethisch geïnspireerde waarden wordt steeds meer toegepast bij de inrichting van moderne overheden. Het



aantal empirische en normatieve onderzoeken naar belangrijke kernwaarden van publieke organisaties is daarmee sterk toegenomen (Van der Wal, De Graaf & Lasthuizen, 2008).

In dit onderzoek is het de vraag hoe publieke waarden – die door overheden gewaarborgd en bevorderd zouden moeten worden – volgens de perceptie en inschatting van respondenten door het gebruik van BOLD worden beïnvloed. In het verlengde daarvan, maar niet de focus van deze studie, ligt de (empirische) vraag welke positieve sociale en economische uitkomsten overheden door middel van BOLD genereren (publieke waarde). Een kanttekening is op zijn plaats. Over de concepten ‘publieke waarde’ en ‘publieke waarden’ is geen eenduidigheid. Publieke waarde(n) zijn moeilijk te duiden: in de bijna onbegrensde literatuur bestaat een meervoud aan definities (zie voor een uitgebreid overzicht Van der Wal, 2008). Rutgers (2008: 6) biedt de volgende algemene definitie voor publieke waarden: *“Publieke waarden zijn duurzame opvattingen over de vormgeving van en bezigheden in de samenleving die al dan niet nastrevenswaardig worden beschouwd voor het voortbestaan van die samenleving, het welbevinden van haar leden en in samenhang daarmee, de aard en het gedrag van degenen – de publieke functionarissen – die deze waarden moeten behartigen”*. Publieke waarden gaan volgens hem dus het idee van de ‘goede’ samenleving en het algemeen belang aan. Waarden die door een individu of groep voor zichzelf als belangrijk worden gezien, kunnen daartegenover worden gezet als private waarden (Rutgers, 2011).

Ondanks de verschillende conceptuele definities, zijn auteurs het eens dat waarden niet tastbaar zijn. Waarden kunnen niet gezien of gehoord worden, maar kunnen worden geobserveerd in de manier waarop ze zich manifesteren in gedragingen, voorkeuren, beslissingen en acties (Van der Wal et al., 2008). Op basis van waarden kan gesproken worden van goed, slecht, gedurfd of onaanvaardbaar openbaar bestuur. Het waarderen van informatie is volgens Self (1982) dan ook een cruciaal aspect in het werk van een ambtenaar (Rutgers, 2011). Een publieke functionaris schat de betekenis van beleid in door publieke waarden af te wegen. De afweging gaat bijvoorbeeld om de rechtvaardigheid, efficiëntie of het democratisch gehalte van overheidshandelen. In deze lijn definiëren van Van der Wal et al. (2008) publieke waarden als belangrijke (morele) standaarden en principes die een bepaald gewicht hebben bij het maken van keuzes.

Publieke waarden hebben dus een centrale plaats in en rondom het openbaar bestuur (Rutgers, 2011). Waarden zijn cultureel ingebed, worden sociaal bepaald en zijn daarmee voor een groot deel afhankelijk van de organisatie (Rutgers, 2011). In de literatuur wordt beargumenteerd dat organisaties een eigen, specifieke set waarden en normen (regels die voorschrijven wat het juiste gedrag in bepaalde situaties is) hebben die in de organisatiecultuur besloten zijn (Van der Wal et al., 2008). Medewerkers vervangen volgens Jackall (1988) de individuele normen en waarden voor de collectieve



moraal van de organisatie (Van der Wal et al., 2008). Dit maakt dat de keuzes en het gedrag van personen voor een groot deel worden bepaald door de doelen, waarden en kennis van de organisatie (Van der Wal et al., 2008). Meerdere auteurs gaan daarbij uit van waardenpluralisme (Graaf, Doeveeren, Reynaers & Van der Wal, 2011). Volgens Spicer (2010) staan twee kenmerken bij waardenpluralisme centraal. Enerzijds zijn waarden niet altijd te combineren (*incompatibility* van waarden). Dit betekent dat het nastreven van de ene waarde ten koste kan gaan van andere waarden. Het tweede kenmerk betreft de onvergelykbaarheid van verschillende waarden (*incommensurability* van waarden). Vanwege het feit dat waarden vaak niet vergelijkbaar zijn, kan niet altijd een rationale afweging gemaakt worden tussen twee verschillende opties. Dat een goede afweging van waarden echter van essentieel belang is, is in de inleiding al beargumenteerd.

Volgens Rutgers (2011) is in de politiek-bestuurlijke praktijk echter sprake van een 'feiten fetisjisme': acties worden het liefst in gedetailleerde urenverantwoordingen beschreven en de bestuurlijke praktijk gaat vaak over 'objectiveerbare' vereisten. De normatieve uitgangspunten blijven meestal buiten beschouwing. Hoewel de afgelopen jaren de aandacht voor de betekenis van waarden voor het openbaar bestuur is toegenomen, zou de waardengeladenheid of normativiteit van het openbaar bestuur volgens Rutgers (2011) nog veel meer centraal moeten worden gesteld. Ook het gebruik van BOLD is een treffend voorbeeld van het feiten fetisjisme. Het enthousiasme voor meer gebruik van data neemt op verschillende terreinen toe met het idee dat 'objectieve' *big, open en linked data* voor betere besluitvorming zorgen. Een van de oorzaken hiervan zijn de toegenomen technologische mogelijkheden. Vraagstukken over de leegstand van gebouwen in een gemeente kunnen bijvoorbeeld gemakkelijk gekwantificeerd en gevisualiseerd worden. Vervolgens kunnen bestuurders beargumenteerd een beslissing nemen. Door echter te schuiven met de assen van de grafiek, kan een totaal andere conclusie ontstaan met mogelijk grote consequenties voor de woonsituatie van bewoners. Data-analisten kunnen zodoende grote invloed uitoefenen op de beslissingen van bestuurders (Hartholt, 2017).

De onderliggende aanname is dat feitelijke data over personen zonder problemen op neutrale, waarde vrije en objectieve wijze verzameld en geanalyseerd zouden kunnen worden (Kitchin, 2016). Een valkuil is dat de resultaten voor waarheid worden aangenomen. Een feit is namelijk niet zo 'feitelijk' als het lijkt (Rutgers, 2011). Over de uitgangspunten waarop een feit wordt gebaseerd is bijna altijd discussie mogelijk, zeker wanneer het complexe maatschappelijke vraagstukken zoals armoede betreft. In de huidige politiek-bestuurlijke praktijk wordt bij het gebruik van BOLD desondanks een sterk feit-waarde onderscheid aangehouden. Bij de dataprojecten wordt vaak geclaimd dat het slim gebruik van data leidt tot "*meer inzicht in wat er feitelijk in de stad gebeurt op gebieden als zorg, werk, onderwijs en wonen*" (Gemeente Den Haag, 2017). De scheiding tussen feiten en waarden stelt ten



onrechte dat enkel waarden als irrationele en subjectieve voorkeuren kunnen worden verklaard. Echter, ook feiten kunnen moeilijk zuiver worden gehouden. Popper (1959) wees er al op dat waarnemingen altijd theoriegeladen zijn: er vindt altijd interpretatie en beoordeling plaats (Rutgers, 2011). Evenzo zijn datatoepassingen niet enkel neutraal of objectief, maar reflecteren de keuzes van onderzoekers, interpretaties en conclusies voor een specifiek doel. De ontwikkeling dat overheden normatieve vraagstukken met behulp van BOLD kwantitatief 'objectiveerbaar' willen maken, doet daarmee de normativiteit van het openbaar bestuur onrecht aan. Zeker als blijkt dat het gebruik van BOLD niet gepaard gaat met toenemend bewustzijn rondom het werken met data en het nastreven van publieke waarde(n).

2.3 Een selectie van publieke waarden

Met het toenemen van hervormingen in de publieke sector, hebben bestuurskundigen steeds meer belangstelling getoond voor de effecten van dergelijke hervormingen op publieke waarden (Reynaers, 2013). Op basis van de literatuur zijn omvangrijke lijsten met publieke waarden samen te stellen (zie voor een uitgebreid overzicht Van der Wal, 2008). Van der Wal et al. (2006) hebben bijvoorbeeld een uitgebreide verkenning gedaan in zowel de bestuurs- als bedrijfskundige literatuur. Na een analyse van honderden verschillende waarden, werd een set van de 20 belangrijkste organisatorische waarden samengesteld die zowel in de private als publieke sector getoetst kon worden. Van der Wal et al. (2008) hebben deze set vervolgens verwerkt in een digitale enquête en uitgezet bij een brede steekproef respondenten in zowel publieke als private organisaties. De resultaten van de studie wijzen uit dat de volgende 10 waarden volgens publieke managers het belangrijkste zouden moeten zijn in overheidsorganisaties: verantwoording, effectiviteit, onomkoopbaarheid, betrouwbaarheid, rechtmatigheid, professionaliteit, efficiëntie, transparantie, neutraliteit en dienstbaarheid.

Sommige auteurs onderschrijven het relatieve belang van waarden in relatie tot verschillende beleidsfasen van het openbaar bestuur (zie Tabel 1). De vraag wordt bijvoorbeeld gesteld of bepaalde waarden belangrijker zijn bij de agendavorming van beleid, ten opzichte van waarden die bij het vaststellen en uitvoeren van beleid juist belangrijk zijn. Dat kan voor dit onderzoek relevant zijn omdat bestuurders veelal bezig zijn met de inhoud en vaststelling van beleid, ambtenaren bij het vormen van beleid en uitvoerders met het uitvoeren van beleid. Henkriks & Drosterij (2012) en Huberts (2014) maken een indeling van waarden die in de verschillende fasen van de beleidscyclus van belang zijn (De Graaf, Huberts, & Smulders, 2013). Een aantal waarden is volgens de auteurs altijd van belang bij alle fasen van het beleid van het openbaar bestuur. De overheid zou bijvoorbeeld te allen tijde integriteit, behoorlijke contacten, openheid en professionaliteit centraal moeten stellen. In de fase van agendavorming en beleidsvoorbereiding wordt volgens de auteurs van het openbaar bestuur verwacht dat burgers en belangrijke partijen betrokken worden bij het vormen van beleid (participatie) en dat



er maatschappelijk draagvlak bestaat (legitimiteit). Rechtmatigheid wordt belangrijker bij het voorbereiden en vaststellen van beleid, omdat bekeken moet worden of beleid binnen bestaande wet- en regelgeving past. Ook bij de uitvoering van beleid is het van belang dat vastgestelde regels worden gevolgd.

Effectiviteit en efficiëntie gaan volgens De Graaf, Huberts, & Smulders (2013) in latere fasen een rol spelen. Verantwoorden staat volgens hen vooral centraal bij het vaststellen en uitvoeren van beleid. Hier moet de overheid verantwoording afleggen over gemaakte keuzes en eventuele problemen die zijn opgetreden (De Graaf, Huberts, & Smulders, 2013). Het gebruik van BOLD is een opkomend verschijnsel en is nog geen prominent onderdeel van de beleidsvorming in de publieke sector. Het is waarschijnlijk dat overheden met betrekking tot BOLD nog niet in de fase van beleidsvaststelling of -uitvoering zitten. Een onderscheid van waarden per beleidsfase kan daardoor minder toepasbaar lijken, maar kan anderszijds nuttig zijn om te verklaren waarom gemeenten al dan niet aandacht voor bepaalde waarden hebben.

Tabel 1.

Relatieve belang van waarden per beleidsfase

Agendavorming	Beleidsvoorbereiding	Beleidsvaststelling	Beleidsuitvoering
Integriteit			
Behoorlijke contacten			
Openheid			
Professionaliteit			
Participatie			
Legitimiteit			
	Rechtmatigheid		
	Effectiviteit		
	Efficiëntie		
	Verantwoording		

Noot. Herdrukt van De Graaf, Huberts, & Smulders (2013)

Om de hoeveelheid aan publieke waarden hanteerbaar te maken, hebben auteurs geprobeerd de waarden in te delen aan de hand van een aantal dimensies. Bij veel auteurs (zie bijvoorbeeld Huberts (2005) en Hendriks & Drosterij (2012) komt een aantal kernaspecten in verschillende vormen terug: democratisch gehalte, rechtmatigheid, effectiviteit en integriteit (De Graaf, Huberts, & Smulders, 2013). Bovens, 't Hart, Van Twist & Rosenthal (2001) beschrijven in hun handboek vier criteria die goed passen bij deze kernbegrippen om de kwaliteit van het openbaar bestuur vast te stellen. In dit onderzoek worden drie van deze waardenclusters gehanteerd:

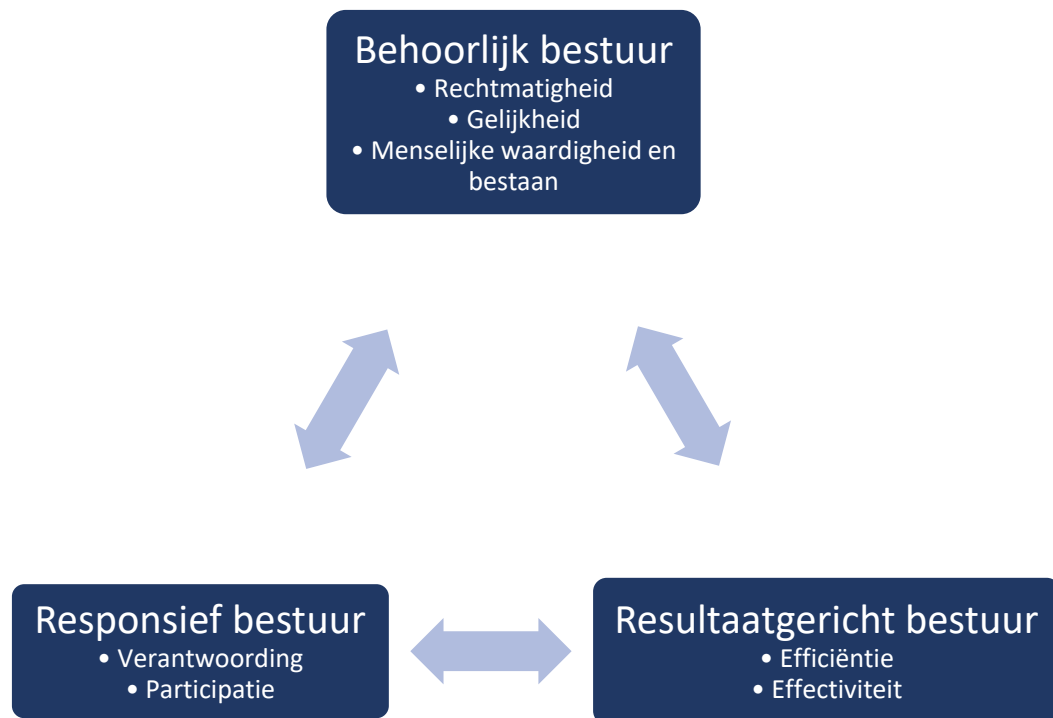


- behoorlijk bestuur: goed openbaar bestuur handelt in lijn met geschreven en ongeschreven regels. De overheid heeft een voorbeeldfunctie waar het gaat om omgangsnormen en het zijn van een betrouwbare partij.
- responsief bestuur: er is sprake van goed openbaar bestuur als de wensen en verlangens van burgers terugkomen in het functioneren van de overheid. Het democratisch gehalte is van belang, net als verantwoording, openheid en transparantie van het publieke handelen.
- resultaatgericht bestuur: de overheid handelt resultaatgericht door een duidelijke visie uit te dragen over maatschappelijke doelen en wat moet gebeuren om de doelstellingen te bereiken. Openbaar bestuur is goed wanneer middelen efficiënt en effectief worden ingezet.

Op basis van een breed scala aan mogelijke waarden, richt deze studie zich op verantwoording, effectiviteit, efficiëntie, rechtmatigheid, gelijkheid, participatie en menselijke waardigheid en bestaan. Deze selectie is niet gebaseerd op een uitgebreide primaire literatuurstudie of empirisch onderzoek van eerste hand (zoals Van der Wal et al. (2008)). De keuze is gebaseerd op de bekendheid van de waarden in de publieke waardenliteratuur en hun toepasbaarheid in de context van BOLD. Daarbij zijn de waarden ook (grotendeels) consistent met genoemde, belangrijke waarden in gedragscodes (zoals de Nederlandse code voor goed openbaar bestuur, BZK (2009)). Op geen manier wordt gesuggereerd dat deze waarden belangrijker zijn dan andere waarden. Over de toepasbaarheid in de context van BOLD volgt in de volgende paragrafen meer (met name paragraaf 2.6 Hypothesen en verwachtingen).

Efficiëntie betekent dat resultaten van de overheid met minimale middelen worden bereikt (doelmatigheid), terwijl **effectiviteit** betrekking heeft op het behalen van de gewenste resultaten (doeltreffendheid). Bij **rechtmatigheid** draait het vervolgens om openbaar bestuur dat in lijn met bestaande wet- en regelgeving handelt (Van der Wal et al., 2008). **Verantwoording** kent meerdere facetten waarbij het vaak gaat om een overheid die acties aan betrokken belanghebbenden rechtvaardigt en uitlegt (Van der Wal et al., 2008). Het gaat in deze studie ook om de formele verantwoordelijkheid voor overheidshandelen en de verantwoordelijkheid nemen bij fouten (Stoker, 2006). Deze waarde veronderstelt een bepaalde mate van openheid en transparantie van het openbaar bestuur. Van het openbaar bestuur wordt daarnaast **gelijkheid** verwacht, dat wil zeggen dat zonder arbitrair vooroordeel ten aanzien van specifieke groepsbelangen wordt gehandeld (Van der Wal et al., 2008). **Participatie** heeft betrekking op de activiteiten van de overheid waarmee personen, burgers, consumenten worden betrokken bij het vormgeven overheidsbeleid (Daniel, 1987). Tot slot volgt de waarde van **menselijke waardigheid en bestaan**. Openbaar bestuur draait om mensen en bestaat in het verlengde van een democratie uit personen die werken voor het bevorderen van het algemeen welzijn van de samenleving (Newland, 2012). Van de overheid wordt verwacht dat het fundamentele mensenrechten als privacy beschermt en levensonderhoud bevordert, bijvoorbeeld

door het recht op werk of op bijstand. De waarde duidt hiermee ook een mate van respect voor de waardigheid en emoties van burgers. Volgens de drie eerder genoemde waardenclusters kunnen de mogelijke botsingen tussen de publieke waarden van goed openbaar bestuur in relatie tot BOLD visueel worden weergegeven (zie figuur 1).



Figuur 1. *Botsingen tussen waardenclusters*

2.4 Waardenbotsingen

Spanningen tussen waarden zijn volgens Van der Wal, De Graaf & Lawton (2011) vaak toe te schrijven aan botsingen tussen klassieke bureaucratische waarden zoals rechtmatigheid en integriteit en (zogenaamd) private waarden als efficiëntie, effectiviteit en innovatie. Het is vervolgens de vraag hoe bestuurders, managers en uitvoerders omgaan met de waardenbotsingen die zij ervaren. Een vaak gebruikte indeling over copinggedrag beschrijft zes strategieën en is ontwikkeld door Thacher & Rein (2004). Huberts & Van Hout (2011) beargumenteren dat er op basis van deze indeling drie hoofdstrategieën zijn die bepalen hoe met conflicten tussen waarden omgegaan wordt (De Graaf, Huberts, & Smulders, 2013). Allereerst kunnen mensen duidelijke keuzes maken voor specifieke waarden (1). Dit is in lijn met wat Thacher & Rein (2004) *bias* en *firewalls* noemen. Met *bias* wordt bedoeld dat personen ervoor proberen te zorgen dat een enkele waarde de overhand krijgt. Volgens Steenhuisen (2009) kan dit bijvoorbeeld gerealiseerd worden door *technicization*: doordat mensen zaken meetbaar en rationeel willen benaderen, krijgt een waarde als efficiëntie de overhand boven procesgerichte waarden waarvan de uitkomsten moeilijker meetbaar zijn. *Firewalls* hebben betrekking



op het institutioneel scheiden van waarden, waardoor verschillende afdelingen/personen zich niet met elkaars keuze bemoeien (Thacher & Rein, 2008).

Personen kunnen volgens Huberts & Van Hout (2011) ook heen en weer bewegen tussen verschillende waarden, waarbij de ene waarde soms de voorkeur krijgt en op een ander moment een andere waarde (2). De nadruk kan volgens Thacher & Rein (2008) bijvoorbeeld een tijd lang op een dominante waarde liggen, totdat blijkt dat hierdoor andere waarden benadeeld worden en een verschuiving van de focus plaatsvindt (*cycling*). Een andere mogelijkheid is dat het belang van een waarde per casus besloten wordt, vaak op basis van eerdere ervaringen en persoonlijke inschattingen (*casuistry*). Waardenverschuivingen kunnen ook plaatsvinden door kleine incrementele stappen die genomen worden om zonder grote inspanning waardenconflicten op te lossen (*incrementalism*). Tot slot kunnen personen volgens Huberts & Van Hout (2008) waarden combineren om meerwaarde te creëren (3). Thacher & Rein (2008) spreken van *hybridization* waarbij waarden in de tijd aangevuld en vermengd worden met andere waarden.

2.5 BOLD

De beschikbare literatuur staat het toe om uitgebreider in te gaan op publieke waarde(n), goed openbaar bestuur en waardenbotsingen. In dit onderzoek staat daarnaast een ander concept centraal dat in volgende paragrafen wordt besproken. Het is in deze studie namelijk de vraag hoe publieke waarden van goed openbaar beïnvloed worden door BOLD.

Big data

Big data is onderdeel van BOLD en is uitgegroeid tot een veelgebruikt concept. Consensus over hoe dit begrip gedefinieerd kan worden is echter nog niet bereikt (Franke et al., 2016). Van *big data* wordt in de context van *smart cities* verwacht dat het met nieuwe technologieën gegronde kennis kan opleveren voor de stedelijke beleidsvorming (Giest, 2017). Johnson (2012) definieert *big data* als extreem grote datasets die informatie kunnen bevatten over consumentengedrag, sensoren, sociale media en meer. Fischer et al. (2012) leggen de lat *big* op het niveau wanneer datasets niet op een rechte doorzee en eenduidige manier verwerkt kunnen worden met de huidige technologische mogelijkheden. Een andere algemene definitie is dat *big data* geen technologieën op zich zijn, maar refereren aan data die zo groot, gevarieerd en dynamisch zijn, dat conventionele dataverwerkings-technologieën er niet mee om kunnen gaan (Klievink, Romijn & De Bruin, 2016).

Het gebruik van *big data* gebeurt in de huidige politiek-bestuurlijke praktijk voornamelijk op experimentele basis, maar zal in de toekomst een grote rol spelen binnen overheden (Wetenschappelijke Raad voor Regeringsbeleid, 2016b). Onderzoekers karakteriseren *big data* vaak



aan de hand van technische begrippen die met een 'V' beginnen. Het gaat volgens auteurs bijvoorbeeld om grote hoeveelheden data (*volume*), die in verscheidenheid (*variety*) en met een hoge snelheid worden gegenereerd (*velocity*) (Aggarwal, 2016). Sommige auteurs voegen een vierde en vijfde V toe: het gaat dan over de meerwaarde van *big data* (*value*) en waarheidsgetrouwheid (*veracity*) (Schroeck et al., 2012). Verschillende papers, artikelen en rapporten stellen manieren voor waarmee overheden *big data* kunnen gebruiken om de dienstverlening te verbeteren en nationale uitdagingen (zoals de stijgende zorgkosten, natuurrampen en terrorisme) te overwinnen (Manyika et al., 2011). Genoemd worden bijvoorbeeld profilering, voorspellende analyses, *social media* monitoring, semantisch web analyses en *process mining* (Van der Weerd & De Vries, 2014). Een van de grote beloftes is dat *big data* leidt tot grotere efficiëntie en effectiviteit (zie bijvoorbeeld de *Digital Agenda towards a Data Driven Economy* van de Europese Commissie, 2014).

In 2014 werd er op technologiebeurs CeBIT in Hannover door bondskanselier Angela Merkel al op gewezen dat aanvullende regels vereist zijn om data op een verantwoorde wijze te gebruiken (AG Connect, 2014). In een gerichte studie naar het gebruik van *big data* door de overheid, benadrukt ook de WRR dat burgerrechten in het licht van *big data* ontwikkelingen versterkt moeten worden (Wetenschappelijke Raad voor Regeringsbeleid, 2016). De wetgeving die over (*big*) datagebruik bestaat, richt zich voornamelijk op het verzamelen van gegevens en niet zozeer op wat met de data gebeurt (Van der Weerd & De Vries, 2014). De Wet bescherming persoonsgegevens (Wbp), de informatie-/toestemmingsplicht en Europese regelgeving voorzien nu voor een deel in wetgeving over het gebruik van data en bescherming van privacy. De legale principes in deze wetgeving staan echter op gespannen voet met *big data* (Wetenschappelijke Raad voor Regeringsbeleid, 2016b). *Big data* sluiten niet goed aan bij principes als anonimisering, doelbinding en minimalisatie van persoonsgegevens. Statistische analyses op *big data* kunnen allereerst zo'n precisie hebben dat datasets gede-anonimiseerd kunnen worden. Gegevens kunnen daarmee herleid worden tot identificeerbare natuurlijke personen. Ten tweede staat het koppelen en hergebruik van informatie in *big data*-analyses centraal, waarmee het in gaat tegen het doelbindingsprincipe (de regel dat data maar voor één doel gebruikt mogen worden). Ten derde draaien *big data*-analyses in tegenstelling tot dataminimalisatie, in beginsel om datamaximalisatie. Het is daarom een reëel risico dat personen die betrokken zijn bij *big data* projecten de principes van doelbinding, dataminimalisatie en anonimisering niet volgen (Wetenschappelijke Raad voor Regeringsbeleid, 2016b). Het gebruik van *big data* daagt daarom de rechtmatigheid van de overheid en de privacy van burgers uit.

In april 2016 is de Europese algemene verordening gegevensbescherming (AVG, 2016/679) aangenomen die de oude Europese privacyrichtlijn uit 1995 en de Wbp in 2018 (deels) zal vervangen. De toekomstige wetgeving bepaalt onder andere dat individuen bij het inzamelen van gegevens altijd



gewaarschuwd moeten worden welke data worden verzameld en waarom, de principes anonimisering, doelbinding en dataminimalisatie versterkt moeten worden, data alleen bewaard mogen worden voor zo lang het doel dit vereist en dat betrokken personen de mogelijkheid moeten hebben om te kunnen weigeren (Europa Decentraal, 2016). Tevens zijn er door inbreng van het Nederlands kabinet verschillende mogelijkheden bijgekomen om in de toekomst de data van burgers toch voor andere doelen te gebruiken dan waarvoor ze oorspronkelijk zijn verzameld (Van der Steur, 2016). Door een wet aan te nemen hoeft het doelbindingsprincipe niet altijd te worden gevolgd. Vervolgens kan brede koppeling van databronnen en profileringssoftware het bijvoorbeeld mogelijk maken om criminaliteit en huiselijk geweld statistisch te voorspellen. Gezien ook de toekomstige wetswijziging, en met name de bovengenoemde Nederlandse toevoeging, is een goede morele afweging van belang.

Big data projecten hebben vaak geen concreet doel; grote hoeveelheden data worden verzameld waarvan de waarde en het potentieel pas duidelijk wordt nadat ze verzameld zijn (Wetenschappelijke Raad voor Regeringsbeleid, 2016b). Het kan voor organisaties steeds minder belangrijk worden om te werken met correcte en nauwkeurige gegevens over specifieke personen, zolang een hoog percentage van de analyse maar een algemeen correct beeld oplevert. Het neigt dan naar het bekende *quantity over quality*, omdat organisaties gewend raken te werken met kwalitatief mindere data. Ook in de publieke sector lijkt het dat werken met onbetrouwbare gegevens of bronnen steeds meer voorkomt (Wetenschappelijke Raad voor Regeringsbeleid, 2016). Daarbij zijn *big data* analyses zijn in de huidige tijd voor gebruikers en zelfs programmeurs vaak nog ongebruikelijk en onvoorspelbaar (Crawford & Schultz, 2014).

Met betrekking tot *big data* wordt uitgegaan van de realistische (epistemologische) veronderstelling dat een externe realiteit bestaat die onafhankelijk van de waarnemer functioneert (Kitchin, 2016). Deze realiteit zou objectief en accuraat gemeten, gevolgd en geanalyseerd kunnen worden om de wereld te presenteren zoals deze daadwerkelijk is. Met andere woorden, stedelijke data zouden zonder problemen op neutrale, waardevrije en objectieve wijze verzameld en geanalyseerd kunnen worden zodat een representatief beeld ontstaat van wat gemeten wordt. Deze gedachtegang wordt in de sociale wetenschap bekritiseerd vanwege de positivistische, deterministische en mechanische insteek (Kitchin, 2016). Ook Boyd & Crawford (2012) verwerpen het blinde vertrouwen in kwantitatieve data en de claim dat data objectief en accuraat zijn. Daarbij gaan in tegenstelling tot enkele jaren geleden de inhoud en analyses van datasets steeds meer om mensen en minder om systemen (Metcalf, 2017).

Als *big data* bronnen vervolgens worden gelinkt en aangewend voor profileringssoftware, kan dit ten koste gaan van de onpartijdigheid van de overheid. Profileren is gebaseerd op het verzamelen, opslaan



en analyseren van de kenmerken van personen, zoals demografische karakteristieken of gedragsinformatie. Met profileringssoftware kan op basis van een profiel de meest relevante informatie naar de burger of klant worden gestuurd en beleid worden gevormd. *Big data* toepassingen draaien vaak niet om individuele profielen, maar om groepsprofielen. Niet om retrospectieve of historische analyses, maar om probabiliteitsanalyses en voorspellende algoritmen met een bepaalde foutmarge. Groepsprofielen zijn echter niet toepasbaar op ieder individu. Ook zijn algoritmen niet neutraal of objectief, maar reflecteren de keuzes van personen over data, koppelingen, interpretaties en conclusies voor een specifiek doel. Het risico is dat *big data* via profilering en selectief beleid kunnen leiden tot minder gelijkheid tussen burgers. Ook de kennisasymmetrie tussen de overheid en de burger kan groter worden (vaak wordt gesproken over *big brother*). Daarmee wordt bedoeld dat BOLD de onbalans tussen overheden en burgers kan versterken, doordat individuen steeds minder invloed op het functioneren van grote organisaties hebben (Wetenschappelijke Raad voor Regeringsbeleid, 2016b). *Big data* liggen in de handen van de overheid (en samenwerkende partijen) terwijl de burgers de gevolgen van de analyses ondervinden. De eerste groep ondervindt de toegevoegde waarde, maar de tweede groep wordt vaak niet geïnformeerd over de algoritmes die tot beslissingen hebben geleid.

De Wet op de Inlichtingen- en Veiligheidsdiensten, beter bekend als de controversiële ‘Sleepwet’, is hier een goed voorbeeld van. De wet treedt in 2018 in werking en vergroot de middelen die de Algemene Inlichtingen- en Veiligheidsdienst (AIVD) en de Militaire Inlichtingen- en Veiligheidsdienst (MIVD) kunnen inzetten. Met de wetswijziging kunnen de AIVD en MIVD onder andere een digitaal ‘sleepnet’ inzetten om online communicatie van verdachte en niet verdachte burgers af te luisteren. Ook maakt de wet het mogelijk om elektronische apparaten hacken, (ruwe) data met buitenlandse inlichtingendiensten te delen en een geheime DNA-databank aan te leggen. Het onderwerp leidde tot verontwaardiging onder burgers en heeft tot gevolg dat in 2018 een referendum gehouden wordt over de ‘Sleepwet’. In de praktijk heeft de politiek echter al aangegeven de uitslag te negeren (NRC, 2017). Ook bij de aanpak van fraude is de rode draad dat de overheid meer data wil koppelen en analyseren om betere ‘risicoclassificaties’ te kunnen vaststellen. Uit een Kamerbrief van toenmalig minister Opstelten van Veiligheid en Justitie blijkt dat de Rijksoverheid op basis van gedrag, omstandigheden en andere demografische data op zoek wil gaan naar potentiële fraudeurs (Opstelten, 2014). Het zou miljoenen opleveren aan het terugvorderen van uitkeringen en openstaande boetes.

De bijvangst aan data over onschuldige burgers wordt in deze brief niet genoemd. Daar komt bij dat burgers die het gevoel hebben dat zij gemonitord worden, hun gedrag aanpassen (zie bijvoorbeeld Schep (2017)). Wanneer algoritmen beoordelen wat burgers doen en handelingen worden herinnerd in *big data*, kan dit kan leiden tot maatschappelijke *chilling effects* op het uitoefenen en genieten van rechten. Personen kunnen het gevoel hebben dat hun recht op vrijheid van meningsuiting en privacy



in het geding zijn. Het Rathenau Instituut pleitte in een rapport op 10 mei 2017 dan ook dat het behoud van menselijke waardigheid in het digitale tijdperk nieuwe mensenrechten vereist (Rathenau Instituut, 2017). Ook op individueel niveau kan BOLD effect hebben op menselijke waardigheid en bestaan.

Verzekeraars en bedrijven verzamelen en verwerken bijvoorbeeld steeds meer data met persoonlijke gegevens om risico's in te schatten. Het is allereerst de vraag in hoeverre de privacy is gewaarborgd. Een voorbeeld waarbij het gebruik van data-analyses ook druk uitoefent op menselijke waardigheid en bestaan is het premiebeleid van zorgverzekeraars. De Consumentenbond deed hiernaar onderzoek en concludeerde dat premies van verzekeringen tot op huisnummer kunnen verschillen (Consumentenbond, 2015). Verzekeraars zouden verzekeringsaanvragen namelijk beoordelen door analyses uit te voeren op risicogebieden en gemiddelde inkomens via de postcode. Als een persoon in een wijk met sociale huurwoningen woont, kan het zijn dat diegene vanwege een risicoselectie een hogere premie betaalt. Discriminatie op basis van geslacht, ras en nationaliteit en indirect via woongegevens mag volgens de Algemene wet gelijke behandeling (AWBG) niet. Toch kan dit echter gebeuren, bijvoorbeeld als er voornamelijk mensen met een niet-Nederlandse afkomst in een gebied wonen. Deze vorm van indirecte discriminatie is moeilijk te bewijzen, mede omdat verzekeraars geen openheid geven over de opbouw van premies omdat dit 'concurrentiegevoelige informatie' is (Niewold, 2015). Met gebruik van BOLD is het mogelijk om ook overheidsbeleid op dezelfde vorm te geven. Op basis van analyses op inkomen, opleidingsniveau, leeftijd, geslacht, rookgedrag, gezondheid kunnen beslissingen worden genomen die direct verband houden met het levensonderhoud van personen (uitkeringen, voedsel, veiligheid). Daarmee kunnen botsingen optreden met menselijke waardigheid en bestaan. Gebieden waar inkomens al laag zijn kunnen daarmee in een neerwaartse spiraal terechtkomen.

Dat data ook op individueel niveau vergaande effecten kunnen hebben op menselijke waardigheid en bestaan, blijkt uit het lot van Aleksandr Dolmatov in 2013. Als activist was Dolmatov lid van een Russische oppositiegroep waarvoor hij werd gearresteerd in Moskou. Hij vluchtte naar Nederland en vroeg asiel aan. Zijn asielaanvraag werd echter geweigerd, waartegen hij in beroep ging. Hij pleegde uiteindelijk zelfmoord in een terugkeercentrum in Rotterdam. Uit onderzoek bleek dat door verschillende organisaties fouten waren gemaakt (Inspectie Veiligheid en Justitie, 2013). In computersystemen was namelijk per ongeluk ingevoerd dat Dolmatov uitzetbaar uit Nederland was. Later bleek dat er in 300 andere gevallen sprake was van dezelfde datafout.

Een ander voorbeeld uit de luchtvaart kan tevens belichten hoe data-analyses effect kunnen hebben op menselijke waardigheid. In april 2017 zat David Dao in een vliegtuig van United Airlines om van Chicago naar Louisville te vliegen. Vlak voor vertrek werd de passagier gemaand om zijn plek in het



vliegtuig af te staan aan crewleden van de luchtvaartmaatschappij. Dao weigerde dit en brak zijn neus toen hij hardhandig uit zijn stoel werd gewerkt. Andere passagiers filmde hoe Dao weggesleept werd en plaatste de video's op het internet (Back, 2017). De oorzaak waardoor Dao gesommeerd werd het vliegtuig te verlaten, was volgens ingewijden te wijten aan een algoritme dat getraind was om een klant van de 'laagste waarde' te vinden (Biryukov, 2017). In dit geval was dat Dao, een reiziger in de *economy-class*, die minder voor zijn ticket had betaald dan anderen en geen lid was van het spaarprogramma van United Airlines. Daarbij zou het algoritme berekenen dat voor Dao als alleenstaand reiziger geen grote additionele kosten gemaakt zouden hoeven worden (bijvoorbeeld voor overnachtingen in vergelijking met families). De video's leidden tot wereldwijde verontwaardiging, waarbij werd gesproken over inbreuk op de menselijke waardigheid (O'Neil, 2017).

Open data

Open data is het tweede onderdeel van het acroniem BOLD. Data zijn 'open' wanneer overheden, burgers en bedrijven er zonder beperkingen toegang toe hebben (Janssen & Van den Hoven, 2015). Veel data zijn via het internet al openbaar toegankelijk zoals weer- en transportinformatie. In essentie worden *open data* van de overheid geproduceerd met publieke middelen en bevatten geen data op persoonsniveau (Janssen & Van den Hoven, 2015). Ook berusten er geen auteursrechten op en zijn *open data* bij voorkeur computer-leesbaar en gereed voor analyses. Managers en ambtenaren zijn over het algemeen niet geneigd tot het openen van data, omdat het publiek daarmee tot nieuwe inzichten kan komen die op hun beurt tot kritische vragen leiden. Volgens Janssen & Van den Hoven (2015) zijn nieuwe mechanismen nodig om de publieke datavragen te monitoren en te beantwoorden.

Burgers kunnen *open data* gebruiken en door middel van samenwerkingsnetwerken nieuwe inzichten mogelijk maken (Chun, Shulman, Sandoval & Hovy, 2010). Burgers staan buiten de grenzen en controle van de organisatie, maar worden door datadeling onderdeel van de verwerkingssystemen. Zij kunnen bijvoorbeeld zelf data verrijken, combineren en verzamelen. Door data openbaar te maken vervagen de traditionele grenzen tussen de overheid en burgers en wordt het systeem meer opengesteld (Janssen et al., 2012). Het openen van systemen biedt burgers de mogelijkheid om feedback te geven op basis waarvan publieke organisaties kunnen leren. Traditioneel gezien is feedback dan ook van groot belang in open systemen (Jackson, 2003). De relatie tussen de overheid en de omgeving verandert door *open data* en leidt ertoe dat de traditionele plannings- en controlemechanismen niet altijd meer geschikt zijn (Janssen et al., 2012). Ondanks de significantie is er volgens Janssen et al. (2012) nog weinig systematisch en gestructureerd onderzoek verricht naar *open data*.



Linked data

Zoals hierboven beschreven zijn er verschillende argumenten voor het openbaar maken van overheidsdata. Het delen van data bevordert transparantie en nodigt burgers uit om de data collectief te verbeteren. Ook stimuleert het innovatieve voorzieningen en economische groei (Shadbolt & O'Hara, 2013). *Linked data*, het derde onderdeel van BOLD, kan deze effecten versterken. Met *linked data* wordt bedoeld op het koppelen van datasets op basis waarvan analyses kunnen worden gedaan (Bizer, Heath & Berners-Lee, 2009). Data van het web worden bijvoorbeeld aan andere gepubliceerde gegevens gelinkt, op dezelfde manier waarop webpagina's met hypertext worden gekoppeld. Deze data hebben ieder een label, ook wel *Uniform Resource Identifier* (URI), waarvan de *Uniform Resource Locators* (URL's) de bekendste zijn. In een *linked data* context worden URI's gebruikt om alles te identificeren, niet alleen webpagina's, afbeeldingen of video's, maar ook mensen, plaatsen, documenten, concepten, enzovoort. De betekenis van bronnen en onderlinge relaties kunnen worden vastgelegd in een *Resource Description Framework* (RDF).

Met *linked data* wordt informatie in de publieke sector ondubbelzinnig op het web gepubliceerd, op een manier die het voor andere partijen makkelijk maakt om gekoppelde data te gebruiken en te begrijpen. Het combineren van (*big*) databronnen maakt het mogelijk om digitale intelligentie en *data-analytics* in te zetten. Data kunnen daarnaast worden gebruikt om op persoonsniveau burgerprofielen op te stellen om fraude op te sporen en terrorisme te bestrijden (Holsapple, Lee-Post & Pakath, 2014). De aanname van *linked* (en *open*) *data* is dat het de mogelijkheden van *big data* versterkt. Het koppelen van datasets kan leiden tot nieuwe inzichten doordat het onverwachte relaties blootlegt. Interpersoonlijke connecties, die vroeger gevormd werden door middel van brieven en telefoongesprekken, worden sinds het IoT gerealiseerd in online (sociale) netwerken die grote hoeveelheid data accumuleren en personen met elkaar verbinden (Heath, 2011). De keerzijde van het koppelen en samenvoegen van datasets, uit verschillende databases en met verschillende typen data, is dat het oorspronkelijke doel van de dataverzameling verloren gaat (Wetenschappelijke Raad voor Regeringsbeleid, 2016).

2.6 Hypothesen

In de voorgaande paragrafen zijn de centrale concepten van deze studie besproken. Allereerst is ingegaan op de normatieve verschillen tussen de publieke en private sector. Omdat privaat meer voet aan de grond heeft gekregen (onder meer door *governance*) heeft het publieke daar volgens sommige auteurs onder geleden. Daarmee is er de afgelopen jaren meer aandacht gekomen voor publieke waarden en goed openbaar bestuur. Publieke waarden gaan over het maatschappelijk belang en spelen een grote rol in en rondom het openbaar bestuur. Voor deze studie zijn vervolgens zeven



waarden geselecteerd die van belang zijn voor goed openbaar bestuur en die toepasbaar zijn in de context van BOLD. Ook is ingegaan op de betekenis van BOLD, dat een grote invloed belooft te hebben op het functioneren van de overheid. In deze slotparagraaf worden de hypothesen van het onderzoek opgesteld en wordt een conceptueel model gepresenteerd. De hypothesen geven richting aan het empirisch onderzoek en zorgen voor verdieping in de beantwoording van de onderzoeksvraag.

Op basis van de theorie wordt duidelijk dat BOLD belooft voor meer efficiëntie in de overheid te zorgen. Met nieuwe technologische mogelijkheden kan de overheid gemakkelijker data verzamelen, opslaan en analyseren. Als onderdeel van een *smart city* agenda wordt vaak gesteld dat BOLD-toepassingen zorgen voor betere kostenbeheersing van de dienstverlening aan burgers. Dit leidt tot de eerste verwachting:

Hypothese 1. Het gebruik van BOLD leidt tot meer efficiëntie bij het uitvoeren van overheidsbeleid

Daarnaast wordt verondersteld dat BOLD – naast efficiënter handelen – leidt tot een effectievere overheid. De wetenschap en de praktijk stellen verschillende manieren voor waarmee de overheid met BOLD-analyses beleidsdoelen doelmatiger zou kunnen bereiken, bijvoorbeeld op het gebied van de economie, zorg, en veiligheid. Dit leidt tot de tweede hypothese:

Hypothese 2. Het gebruik van BOLD leidt tot meer effectiviteit bij het uitvoeren van overheidsbeleid

Uitgaande van de bestaande wetgeving is tevens gebleken dat BOLD (hoofdzakelijk *big data*) niet goed aansluit op legale principes over anonimiteit, doelbinding en dataminimalisatie. Ook met de intrede van de Europese AVG in 2018, is het de vraag in hoeverre overheden zich bewust zijn van de rechtmatigheid van BOLD-toepassingen. Dit bewustzijn is met de komende wetgeving en de toevoeging van het Nederlandse kabinet van belang. Het is een reëel risico dat BOLD-projecten op gespannen voet staan met huidige (en toekomstige) legale principes. De derde verwachting is als volgt geformuleerd:

Hypothese 3. Het gebruik van BOLD leidt tot minder rechtmatigheid van overheidsbeleid ten opzichte van burgers

Het is voor de legitimiteit van de overheid daarnaast van belang dat datagebruik en algoritmen worden verantwoord. Vanaf het begin van de beleidscyclus wordt van de overheid openheid en transparantie verwacht (deze waarden vallen in deze studie grotendeels ook onder verantwoording). Met betrekking tot BOLD zouden burgers dus op de hoogte gebracht moeten worden bij het maken van beleid. BOLD-analyses zijn echter hoogtechnologisch en de resultaten worden door gebruikers en zelfs programmeurs niet altijd begrepen. Verantwoordelijkheid bij het delen en koppelen en analyseren van



data is dus van belang. Het gaat hierbij om de vraag wie verantwoordelijk is voor de resultaten en eventuele tekortkomingen. Hoewel het openen en beschikbaar maken van data zou kunnen leiden tot meer transparantie, is het de vraag hoe overheden verantwoording afleggen over het eigen gebruik van data. Daar komt bij dat managers en ambtenaren volgens de literatuur niet geneigd zijn tot het verantwoorden van datagebruik omdat dat kritische vragen tot gevolg kan hebben. Het gebruik van BOLD kan daarmee tot problemen leiden in de verantwoording van het openbaar bestuur. Dit is vertaald naar de vierde verwachting:

Hypothese 4. Het gebruik van BOLD leidt tot minder verantwoording van de overheid over haar beleid ten opzichte van burgers

Een dataset is gebaseerd op een beperkt aantal kenmerken van een persoon en is daarmee een versimpelde weergave van de werkelijkheid. Normatieve vraagstukken worden met BOLD vertaald in kwantitatieve analyses, onder de veronderstelling dat de analyses leiden tot objectievere beleidsvorming. Analyses en algoritmen zijn echter niet neutraal, maar reflecteren de keuzes van personen over data, koppelingen en foutmarges. Een risico is dat de overheid, net als zorgverzekeraars, selectief beleid voert en mensen uitsluit op basis van BOLD analyses. Het gebruik van BOLD draagt dus de eventuele (onbewust of bewuste) menselijke vooroordelen. Deze verwachting is als volgt geformuleerd:

Hypothese 5. Het gebruik van BOLD leidt tot minder gelijkheid van burgers ten opzichte van de overheid

Als onderdeel van BOLD heeft vooral *open data* in theorie de intentie om de participatie van burgers en andere partijen te verhogen. Door data te publiceren komt informatie beschikbaar voor bedrijven en burgers en kunnen door samenwerkingsnetwerken nieuwe inzichten ontstaan. Dit ligt in lijn met de *governance* gedachte waarin van de overheid wordt verwacht dat met andere partijen wordt samengewerkt. Burgers en bedrijven kunnen bijvoorbeeld zelf data verrijken, combineren en verzamelen. De overheid wordt hiermee meer opengesteld en participatie wordt meer gestimuleerd. Dit leidt tot de zesde hypothese:

Hypothese 6. Het gebruik van BOLD leidt tot meer participatie van burgers in het overheidsfunctioneren

Het gebruik van BOLD kan druk leggen op mensenrechten en daaraan verbonden, menselijk(e) waardigheid en bestaan. Burgers kunnen allereerst het gevoel hebben dat hun recht op vrijheid van meningsuiting en privacy in het geding zijn omdat handelingen gemonitord en geanalyseerd worden

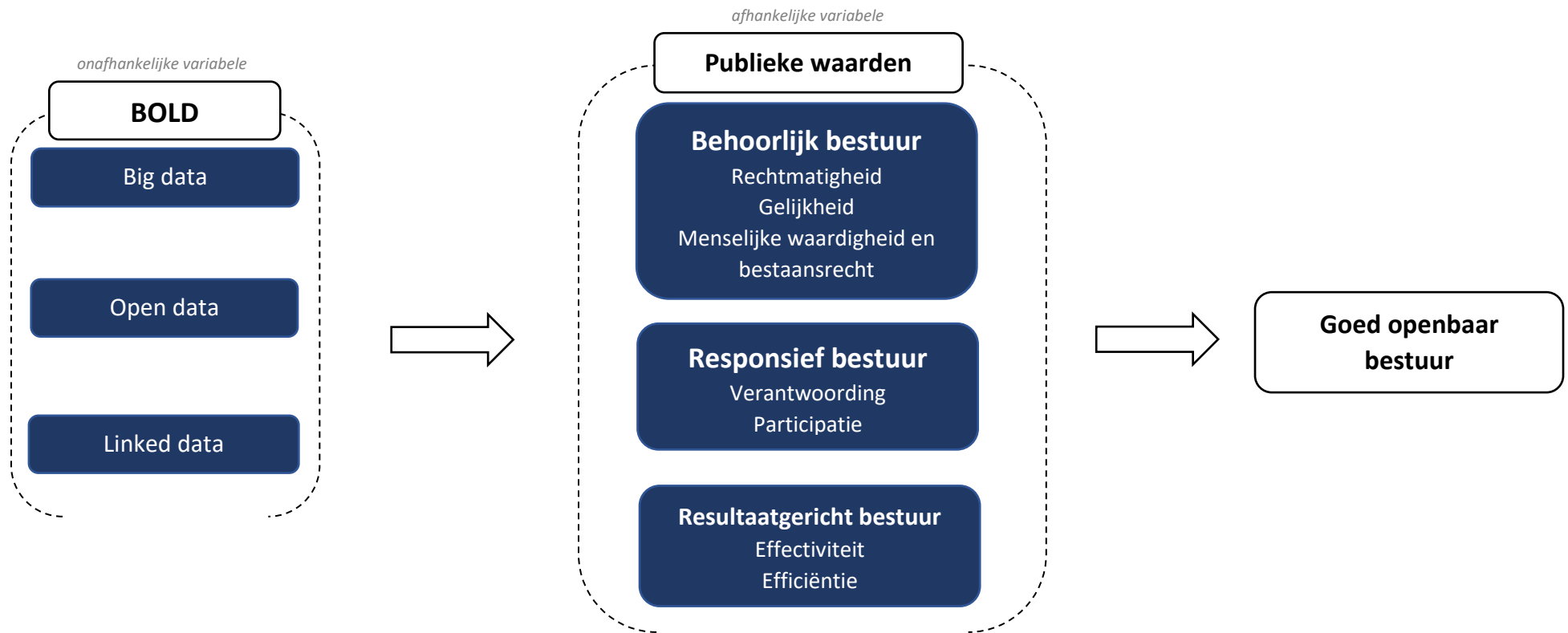


met *big data*. Ten tweede kunnen via profilering besluiten met rechtsgevolgen worden genomen zonder menselijke tussenkomst (hoewel de komende wetgeving dit beoogt te voorkomen). Beleid dat te maken heeft met levensonderhoud (uitkeringen, toeslagen, woningen, etc.) houdt direct verband met menselijke waardigheid en bestaan. Een combinatie van datastromen kan er bijvoorbeeld – bewust of onbewust – onterecht toe leiden dat een persoon bijvoorbeeld geen recht meer krijgt op een uitkering.

Hypothese 7. Het gebruik van BOLD in de overheid leidt tot minder respect voor de menselijke waardigheid van de burger

Op basis van de literatuurverkenning en de opgestelde hypothesen is een conceptueel raamwerk geschetst (zie figuur 2).

Figuur 2. Conceptueel model onderzoek



3 ONDERZOEKSOPZET

In het theoretisch kader zijn verschillende hypothesen en een conceptueel model opgesteld. Met behulp van empirisch onderzoek zijn de aannames getoetst om de hoofdvraag te beantwoorden. In dit hoofdstuk wordt de opzet van het onderzoek besproken. Allereerst wordt ingegaan op de eenheid van analyse en de focus van deze studie. Hier wordt kort de casus van deze studie besproken en worden de keuzes voor de methoden en technieken geëxpliceerd en toegelicht. Tenslotte volgen enkele opmerkingen over de validiteit en betrouwbaarheid van het onderzoek.

3.1 Onderzoekseenheid

In dit onderzoek is gekozen voor een enkelvoudige *casestudy*. Om de mogelijke normatieve consequenties van BOLD op goed openbaar bestuur empirisch te kunnen onderzoeken, is een casus uit de publieke sector geselecteerd waarbij organisaties bezig zijn met *big, open* en *linked data*. Een enkele casus maakt het mogelijk om de ervaringen en waarnemingen in een bepaalde omgeving gedetailleerd te beschrijven. De casus van deze studie is het Urban Data Center (UDC) van gemeente Den Haag en het CBS. Op 26 september 2017 hebben gemeente Den Haag en het CBS dit 7^e UDC geopend. Het doel van de samenwerking is het verbeteren, verbreden en verdiepen van lokale informatie door data te koppelen en kennis te delen. Naast de juiste data, biedt het CBS in het UDC een technisch platform waar analyses op laag niveau anoniem uitgevoerd kunnen worden (CBS, 2016). Hiermee kunnen de gemeentelijke besluitvorming en het beleid worden verbeterd (Van der Sangen, 2017). De samenwerking op het gebied van BOLD belooft grote gevolgen te hebben op het gebied van armoede, ondermijning, mobiliteit, onderwijs en (cyber) bedreigingen (Gemeente Den Haag, 2017). Dit type samenwerking is populair in de huidige praktijk; steeds meer gemeenten werken met andere partijen samen op het gebied van meer en beter gebruik van (gekoppelde) data.

3.2 Dataverzameling en -analyse

In dit onderzoek is voornamelijk gebruik gemaakt van kwalitatieve dataverzameling en -analyse en een interpretatieve benadering van onderzoek (Bryman, 2012). Om de centrale onderzoeksvraag een theoretisch kader te geven, zijn verschillende relevante wetenschappelijke artikelen, theorieën en onderzoeken verzameld en geanalyseerd. Op basis van de verzamelde theorie is een conceptueel model opgesteld. De opgestelde hypothesen zijn vervolgens in de praktijk getoetst aan de opvattingen en inschatting van personen betrokken bij de casus. De methode bestaat uit een documentenstudie en verschillende diepte-interviews. Aan de hand van een vooraf opgestelde vragenlijst zijn met 11 personen interviews gehouden. De respondenten waren werkzaam op verschillende niveaus: managers, beleidsadviseurs en uitvoerende medewerkers (zie Tabel 2).



Tabel 2.

Respondenten van het onderzoek

Naam	Functie	Organisatie
H. Miessen	Programmamanager ICT	Gemeente Den Haag
R. Duivesteijn	Concernadviseur / ICT architect	Gemeente Den Haag
L. Stroucken	Cöördinator UDC Den Haag en Leidschendam-Voorburg	CBS
C. Brakus	Accountmanager BZK en Big Data	CBS
C. Kruiskamp	Programmamanager UDC's	CBS
B. Braaksma	Innovatiemanager	CBS
R. Dood	Directeur Statistische Dienstverlening en Informatieverstrekking	CBS
R. Schoonhoven	Senior Relatiebeheerder Microdataservices	CBS
E. van Mierlo	Cöördinator Open Data	CBS
M. Das	Urban Statistics Professor / Statistisch Onderzoeker	Centre for BOLD Cities / CBS
L. Schokker	Executive manager	Centre for BOLD Cities

De respondenten waren medewerkers van gemeente Den Haag, het CBS en het Centre for BOLD Cities. Bijna alle respondenten zijn vanuit hun functie betrokken bij de ontwikkeling van het UDC in Den Haag. Sommige respondenten hadden daarnaast ervaring met andere UDC's of ervaring met het datagebruik van gemeenten in het algemeen. Een enkele respondent is vanwege specialistische kennis over *linked open data* tevens geïnterviewd, ondanks dat deze persoon in mindere mate betrokken is bij UDC's. Het Centre for BOLD Cities is een samenwerking tussen de Universiteit Leiden, Technische Universiteit Delft en Erasmus Universiteit Rotterdam en is sterk betrokken geweest bij de aanzet van het Haagse UDC. Het Centre organiseerde daartoe bijvoorbeeld een datakamp waarin onderzoekers van de gemeente Den Haag, het CBS en universiteiten onderzoek hebben gedaan naar het thema armoede. Verder onderhoudt het Centre een netwerkfunctie en draagt het bij aan oplossingen voor stedelijke datavraagstukken, waarbij het perspectief van de burger centraal staat (Centre for BOLD Cities, 2017).

De ratio achter de keuze van de respondenten is allereerst dat gegarandeerd is dat de respondenten ervaring hebben met de centrale thema's van dit onderzoek. Centraal staan dan ook de percepties van de respondenten over het gebruik van BOLD door gemeenten. Tevens is het mogelijk om antwoorden te vergelijken door respondenten van verschillende organisaties te spreken en overeenkomstige, tegenstrijdige en conflicterende dilemma's in kaart te brengen. Uit de interviews is tevens duidelijk geworden dat het merendeel van de respondenten vanuit hun functie ervaring en informatie hadden over soortgelijke casussen. Deze inzichten zijn aanvankelijk niet de focus van de studie geweest, maar zijn wel meegenomen in de analyse. Ook is gebleken dat het Haagse UDC voor sommige



interviewvragen te recentelijk was opgericht en de vragen daardoor niet altijd van toepassing waren. In die gevallen is uitgeweken naar andere dataprojecten of in meer algemene zin over datagebruik van gemeenten gesproken.

De vragenlijst is opgenomen in Bijlage 1. In beginsel zijn de vragen zo ontworpen dat iedere hypothese de revue zou passeren. Allereerst is het onderzoek toegelicht en zijn inleidende vragen gesteld. Vervolgens is aan de respondenten gevraagd welke beslissingen zij nemen, of genomen zien worden in de publieke sector in relatie tot BOLD. Omdat waarden abstract zijn, zijn soms concretere dilemma's en mogelijke waardenbotsingen besproken. In totaal duurden de interviews ongeveer een uur, waarbij de gesprekken zijn opgenomen om zorgvuldig uit te kunnen werken. Naast de interviews zijn andere bronnen over de casus geraadpleegd ter aanvulling van de (persoonlijke) interviews. De interviews zijn uitgewerkt en met het programma Atlas.ti gecodeerd in de wijze die aansluit op de hypothesen. Het codeerscherm is opgenomen in Bijlage 2. Respondenten hebben uitgewerkte teksten toegestuurd gekregen om het verslag op feitelijke onjuistheden te controleren. Vanwege gevoelige informatie zijn sommige interviews naar aanleiding van de terugkoppeling van respondenten aangepast. In Bijlagen 3 tot en met 13 zijn de woordelijke interviews opgenomen.

3.3 Operationalisering

Op basis van het literatuuronderzoek zijn de centrale concepten van dit onderzoek beschreven. De begrippen zijn vervolgens geoperationaliseerd, dat wil zeggen, bruikbaar gemaakt voor empirisch onderzoek. Uit de literatuur is gebleken dat publieke waarden abstract zijn en op verschillende manieren gedefinieerd kunnen worden. Ook BOLD kan op verschillende manieren worden opgevat, waardoor het nodig is om voor het empirisch onderzoek een definitie te kiezen. Omdat dit onderzoek een verklarende component bevat, is gewerkt met onafhankelijke en afhankelijke variabelen (zie ook Figuur 2).

Onafhankelijke variabele

De onafhankelijke variabele van dit onderzoek is BOLD, dat bestaat uit *big*, *open*, en *linked data*. Over de definitie van de concepten bleek uit de literatuurverkenning geen overeenstemming te zijn en is daarom bewust gekozen voor een brede, meer algemene definitie (zie Tabel 3). In dit onderzoek wordt met *big data* algemeen gerefereerd aan grote, gevarieerde en dynamische datasets die informatie kunnen bevatten over consumentengedrag, sensoren, sociale media en meer. *Open data* zijn vervolgens gedefinieerd als openbaar gepubliceerde gegevens waar overheden, burgers en bedrijven zonder beperkingen en op toegankelijke wijze (zonder speciale programma's) toegang tot hebben. Uit de theoretische verkenning is duidelijk geworden dat niet alle online data *open data* is. Zo is het op internet plaatsen van een document waarin een tabel staat wel openbaar, maar niet direct op



efficiënte wijze leesbaar door een computer of gereed voor analyses. Met *linked data* wordt bedoeld op de methode van het online publiceren van (gestructureerde) data, zodat deze beter beschikbaar zijn en vooral gemakkelijk gekoppeld kunnen worden. Op basis van het linken van gegevens kunnen analyses worden gedaan. De drie onderdelen kunnen ieder verschillende effecten hebben, maar zijn samengevoegd tot het acroniem BOLD om in meer algemene zin te verkennen welke normatieve consequenties het op publieke waarden van goed openbaar bestuur heeft.

Tabel 3.

De gehanteerde definities van big, open en linked data

Begrip	Betekenis
Big data	Grote, gevarieerde en dynamische datasets die informatie kunnen bevatten over consumentengedrag, sensoren, sociale media en meer
Open data	Openbaar gepubliceerde gegevens waar overheden, bedrijven en burgers zonder beperkingen en op toegankelijke wijze toegang tot hebben
Linked data	Digitale methode voor het online publiceren van gestructureerde data waarmee gegevensbronnen gemakkelijk onderling gelinkt kunnen worden

Afhankelijke variabelen

Publieke waarden zijn daarmee de afhankelijke variabelen van dit onderzoek. Uit de literatuur bleek ook over dit begrip veel conceptuele discussie te bestaan. Naarmate deze waarden meer of minder worden geborgd kan in normatieve zin gesproken worden van beter of slechter openbaar bestuur. Zoals uit de theorie duidelijk werd zijn waarden niet tastbaar, maar manifesteren ze zich in de meningen, houdingen, beslissingen en voorkeuren van personen. Waarden zijn dus verbonden aan een waardenmanifestatie en losstaand onzichtbaar. In dit onderzoek zijn waarden – in lijn met het theoretisch kader – gedefinieerd als belangrijke (morele) standaarden en principes die bij het maken van keuzes invloed hebben. Het gaat in dit onderzoek dan ook om de ervaringen in de praktijk, de beslissingen die respondenten nemen of genomen zien worden en wat daarvoor belangrijke drijfveren zijn (geweest). Op basis van het literatuuronderzoek zijn vervolgens waarden geselecteerd die toepasbaar zijn in de context van het gebruik van BOLD in de publieke sector. In Tabel 4 zijn de waarden opgesomd en gedefinieerd.

Tabel 4.

Selectie van publieke waarden voor goed openbaar bestuur

Publieke waarde	Betekenis
Efficiëntie	Het maximale rendement uit beschikbare middelen halen
Effectiviteit	Handelen om de gewenste doelen te behalen
Verantwoording	Open handelen door aan betrokken belanghebbenden acties te rechtvaardigen en uit te leggen. Verantwoordelijkheid over beleid nemen.
Rechtmatigheid	Handelen in overeenstemming met bestaande wet- en regelgeving
Gelijkheid	Zonder vooroordeel ten aanzien van specifieke groepsbelangen handelen.
Participatie	De omgeving en belanghebbenden betrekken bij vormgeven van overheidsbeleid
Menselijke waardigheid en bestaan	Handelen met respect voor fundamentele mensenrechten en het bestaan van burgers

3.4 Betrouwbaarheid en validiteit

Tot slot enkele opmerkingen over de betrouwbaarheid en validiteit van dit onderzoek. Als gesproken wordt over de betrouwbaarheid wordt vaak verwezen naar de samenhang en de herhaalbaarheid van een meting (Bryman, 2008). Onderscheid kan gemaakt worden tussen interne en externe betrouwbaarheid. Met interne betrouwbaarheid wordt de mogelijke vertekening van de resultaten bedoeld die aan de onderzoeker te wijten is. Met het oog op de interne betrouwbaarheid is voorafgaand aan de interviews de formulering van de vragen met een beleidsmedewerker besproken om te garanderen dat de vragenlijst begrijpelijk zou zijn. Naar aanleiding hiervan zijn enkele wijzigingen aangebracht zodat de respondenten de vragen op een juiste manier zouden interpreteren. Om verdere vertekening te voorkomen zijn door dezelfde onderzoeker – op zo objectief mogelijke wijze – aan alle respondenten identieke vragen en dilemma's voorgelegd. Externe betrouwbaarheid gaat in op de repliceerbaarheid van het onderzoek. Hoewel de gebruikte methoden zijn gedocumenteerd en andere onderzoekers het onderzoek daarmee kunnen repliceren, is het niet waarschijnlijk dat herhaling dezelfde resultaten oplevert; het gebruik van BOLD in de overheid is daarvoor te nieuw.

Validiteit heeft betrekking op de vraag of een meetinstrument daadwerkelijk meet wat het zou moeten meten en of de resultaten van een onderzoek kloppen met de werkelijkheid (Bryman, 2008). Tussen verschillende soorten validiteit kan onderscheid gemaakt worden, waarbij in dit onderzoek wordt ingegaan op interne en externe validiteit. Interne validiteit gaat over op de kwaliteit van de onderzoeksopzet en de gebruikte methoden. Het gebruik van verschillende methoden om informatie te verzamelen (ook wel triangulatie) draagt bij aan het verhogen van de interne validiteit (Bryman, 2012). In dit onderzoek betreffen het zoals eerder benoemd een documentenstudie van wetenschappelijke en niet-wetenschappelijke literatuur en empirisch werk in de vorm van interviews.



Met het oog op het wetenschappelijk gehalte en de kwaliteit van het onderzoek is in de literatuurverkenning hoofdzakelijk gebruik gemaakt van *peer reviewed* artikelen (Bryman, 2012). Naar aanleiding hiervan is een semigestructureerde vragenlijst opgesteld. Ondanks dat hoofdzakelijk open vragen zijn gesteld en respondenten de mogelijkheid hadden om anoniem te blijven, zijn sociaal wenselijke antwoorden niet geheel uit te sluiten. Door geen vaste volgorde van vragen te hanteren is in gesprekken flexibel omgegaan met vragen en antwoorden. De gedetailleerde antwoorden dragen bij aan de interne validiteit. Daarnaast is het belangrijk om nogmaals te benadrukken dat het gaat om de opvattingen en inschatting van personen die betrokken zijn bij het gebruik van BOLD door de overheid. Dat betekent dat het onderzoek iets anders kan meten dan de daadwerkelijke gevolgen voor goed openbaar bestuur.

Externe validiteit heeft betrekking op de generaliseerbaarheid van een onderzoek. Hoewel een enkele casus geschikt is om de ervaringen en waarnemingen in een bepaalde omgeving gedetailleerd te beschrijven, stuit het op problemen in de (on)mogelijkheid om de resultaten te generaliseren (Bryman, 2012). Omdat het gebruik van BOLD in opkomst is, is de doelstelling van deze studie dan ook niet om generaliseerbare uitkomsten te vinden die gelden voor alle BOLD projecten van overheden. Meerdere respondenten hadden vanuit hun functie wel informatie over soortgelijke *cases* en konden daar hun opvattingen en inschatting op baseren. Desondanks is de externe validiteit van deze studie laag. De resultaten leveren tegelijkertijd wel zaken op die vervolgens door andere onderzoekers getoetst kunnen worden in soortgelijke casussen. Het onderzoek heeft tevens aanbevelingen opgeleverd op basis waarvan het gebruik van BOLD in de toekomst zo ingericht kan worden dat goed openbaar bestuur zoveel mogelijk wordt gewaarborgd.

4 ANALYSE

Dit hoofdstuk presenteert de resultaten van het onderzoek naar aanleiding van de dataverzameling en -analyse. In de eerste paragraaf wordt in algemene zin besproken hoe de gemeente Den Haag (en gemeenten in het algemeen) volgens de respondenten omgaat met het gebruik van BOLD. De ervaringen en percepties van de respondenten staan hier centraal. Waar mogelijk worden ervaringen in soortgelijke casussen meegenomen of wordt in meer algemene zin gesproken over BOLD gebruik door gemeenten. Vervolgens wordt per publieke waarde ingegaan op de resultaten om uiteindelijk antwoord te kunnen geven op de centrale onderzoeksvraag. In de slotparagraaf worden de implicaties voor de opgestelde hypothesen besproken.

4.1 Het gebruik van BOLD door gemeenten

Op basis van de antwoorden op de algemene vragen wordt in deze paragraaf een algemene indruk gegeven van het gebruik van BOLD door gemeenten. In de gemeente Den Haag zijn op het gebied van data de afgelopen tijd verschillende experimenten gedaan en nieuwe dingen uitgetoet. Het UDC met het CBS is volgens de programmamanager ICT van de gemeente Den Haag de grootste oogst waar de gemeente mee verder gaat (zie Bijlage 12). Ook hogescholen, universiteiten en bedrijven worden hierbij betrokken. In het UDC Den Haag gaan de gemeente en het CBS de komende tijd met verschillende projecten aan de slag, zoals (Gemeente Den Haag, 2017):

- Armoede in Den Haag scherper in beeld brengen, in het bijzonder bij kinderen
- Inzichtelijk maken van de structuren en processen achter ondermijning (bijvoorbeeld witwassen en dubieuze vastgoedtransacties)
- In kaart brengen welke mensen uit Den Haag vertrekken en waarom (kenmerken en motieven)
- De monitoring verbeteren van kinderen/jongeren die hoog risico hebben op voortijdig schoolverlaten
- Mogelijke bedreigingen voor Den Haag scherper in beeld brengen, zowel in het fysiek als het cyberdomein, en maatregelen onderzoeken die de weerbaarheid van de stad kunnen versterken

De gemeente Den Haag is daarnaast bezig met het ontwikkelen van de datakennis door een *community* te bouwen met *business intelligence* (BI) specialisten, IT'ers, onderzoekers, statistici en beleidsambtenaren. Het datagebruik van de gemeente gaat zoals eerder genoemd breder dan enkel het UDC. Zo wordt bijvoorbeeld samenwerkt met de Haagse Hogeschool, andere steden, ministeries en overheidsinstellingen. Ook worden binnen de gemeente zelfstandig analyses gedaan op eigen,



gesloten datasets. Het gebruik van *big data* – in de zin van grote, ongestructureerde en gevarieerde datasets – staat bij de gemeente Den Haag nog in de kinderschoenen. In het UDC wordt wel met grote databestanden gewerkt, maar dit zijn bijna altijd gestructureerde datasets. Ook *linked data*, wat wil zeggen dat data onderling beter gelinkt kunnen worden, is nog een vrij nieuw concept. Hoewel datasets van de gemeente en het CBS wel worden gekoppeld in het UDC, betreft dit bijvoorbeeld nog niet online data via een RDF-standaard (zoals in het hoofdstuk literatuur beschreven). De gemeente heeft daarnaast de afgelopen tijd geïnvesteerd in een *open data* beheerunit, die structureel helpt bij het open maken van datasets.

De programmamanager ICT van gemeente Den Haag heeft ervaren dat in de gemeente vaak in silo's wordt gewerkt en dat dat moeilijk te doorbreken is. Ook zijn afdelingen niet gewend om data onderling te delen, datasets van goede beschrijvingen te voorzien en eigenaarschap goed vast te leggen (zie Bijlage 13). Middels experimenten wordt in Den Haag momenteel bekeken in hoeverre *big*, *open* en *linked data* bij kunnen dragen aan het functioneren en beleid van de gemeente. In het voorjaar van 2017 is bijvoorbeeld – als voorloper van het Haagse UDC – een datakamp georganiseerd waarin het CBS, het Centre for BOLD Cities, universitaire onderzoekers en de gemeente aan de slag gingen met data rondom het thema armoede. Omdat BOLD de beleidsvorming in gemeente Den Haag of andere gemeenten (nog) niet domineert, is soms in algemenere zin gesproken over het toenemend gebruik van data, koppeling van datasets en het openbaar publiceren van gegevens.

Op de vraag hoe en waarom *big*, *open* en *linked data* kunnen worden gebruikt door de gemeente Den Haag en gemeenten in het algemeen is verschillend geantwoord. De gemeente Den Haag en veel andere gemeenten maken ten eerste al geruime tijd gebruik van data; het is volgens één van de respondenten niet zo dat gemeenten nu vaak 'blind varen' (zie Bijlage 7). Aan de andere kant is het gebruik van data niet de *core-business* van een gemeente. Vaak worden veel data verzameld en opgeslagen, maar gebeurt dit niet altijd op een (her)bruikbare manier (zie Bijlage 13). Het gebruik van BOLD is niet alleen in Den Haag nog in opkomst; veel andere gemeenten zijn volgens de programmamanager UDC's van het CBS nog in een verkennende fase (zie Bijlage 4). Voorspellende algoritmes op basis van enorme hoeveelheden ongestructureerde en gekoppelde data worden door gemeenten nog niet veel toegepast. Het meten van de daadwerkelijke gevolgen van BOLD is dus (nog) niet goed mogelijk. Wat wel mogelijk is – en in deze studie centraal staat – is het meten van de opvattingen en verwachtingen van betrokkenen.

Een groot voordeel van het gebruik van BOLD is volgens meerdere respondenten dat gemeenten, in tegenstelling tot het verleden, hun beleid meer kunnen baseren op feitelijke cijfers. De respondenten



spreken onder andere van *evidence-based policy making* of *fact-based* beleid maken. Een van de respondenten verwoordt het als volgt (Bijlage 3):

“...een gemeente die dingen eerst op onderbuikgevoel deed, daar zetten we nu harde cijfers tegenaan en we zien nu al legio voorbeelden waarin ze door datagestuurd werken betere besluiten kunnen nemen en daardoor betere dingen kunnen doen voor hun burgers.”

Uit de interviews blijkt dus dat gemeenten inzetten op BOLD om maatschappelijke fenomenen ‘objectiveerbaar’ te maken. Het enthousiasme wordt vooral ingegeven door de toegenomen technische mogelijkheden. Gemeenten zien dat datagestuurd werken voordelen kan hebben en zetten daarom in op meer gebruik van data, soms zonder hiervoor een concreet doel te hebben. Een van respondenten zei hierover: *“Data wordt vaak als doel op zich gezien, in plaats van een middel om ergens te komen”* (zie Bijlage 7). Naast de juiste data, biedt het CBS in UDC’s een technisch platform aan waarbij analyses op laag niveau anoniem uitgevoerd kunnen worden. Meerdere respondenten ervaren dat gemeenten het liefst data op individueel niveau willen gebruiken om *“op de voordeur te kunnen kloppen”* (zie Bijlage 3, 4, 5 en 7). De onderzoeksvraag van gemeenten richt zich daarbij vaak op handhaving en persoonlijke contacten. Analyses op straatnaam of huisnummer zijn in een UDC – ondanks de wens van gemeenten – meestal niet mogelijk. Wel merkt een respondent op dat grenzen verschuiven en in sommige gevallen steeds meer mogelijk is. Meerdere respondenten zijn dan ook van mening dat het ethisch bewustzijn bij het gebruik van data beter ontwikkeld zou mogen worden. Een respondent zei hierover (zie Bijlage 3):

“...in je maatschappelijke rol kan het denk ik geen kwaad om ook het bewustzijn bij je personeel meer te ontwikkelen. Van joh, wat je doet, het zijn niet gewoon cijfertjes, het zijn niet alleen eentjes en nulletjes, het zijn geen respondenten, het zijn geen berichtgevers: het gaat om mensen. En dat bewustzijn zou ik meer willen meegeven.”

Een andere respondent vraagt zich hierbij af in hoeverre van een gemeente verwacht kan worden dat er genoeg *know-how* is om de beveiliging van data goed op orde te hebben (zie Bijlage 6). Daar kan de opmerking van een andere respondent mee samenhangen dat gemeenten door bezuinigingen en decentralisaties, naast overheidstaken, ook veel gevoelige data in hun schoot geworpen hebben gekregen. De aandacht voor databeveiliging en wetgeving kan daardoor een andere prioriteit krijgen (zie Bijlage 7). Op de algeme vraag of waarden weleens botsen, worden door bijna alle respondenten antwoorden gegeven gerelateerd aan privacy en wetgeving. Met nieuwe technieken is een hoop mogelijk, maar wetgeving sluit hier niet altijd op aan (hierover volgt in paragraaf 4.4 meer). Veel gemeenten, waaronder Den Haag, hebben een *privacy officer* in dienst die in de gaten houdt of wettelijke grenzen niet overschreden worden. Een van de respondenten ervaart daarbij dat



gemeenten snel in een privacyreflex schieten en dat privacy soms te veel als voornaamste zorg wordt gezien bij dataprojecten (zie Bijlage 10). Dat privacy de overhand kan krijgen boven deze waarden en daardoor een remmende factor kan zijn, is wat Thacher & Rein bedoelen met *bias*, zoals behandeld in hoofdstuk 2. Ook is te zien dat waarden, door het instellen van een *privacy officer*, institutioneel gescheiden worden. In het geval van gemeente Den Haag is dat vooral bedoeld om verschillende afdelingen en personen bewuster te maken van de juridische grenzen bij het datagebruik (zie Bijlage 11 en 13). Met betrekking tot BOLD is het relevant om te benoemen dat de afstand tussen politiek en uitvoering in de gemeente Den Haag (en andere gemeenten) door respondenten als kort wordt ervaren. Tussen wethouders en beleidsmedewerkers vindt veel interactie plaats. Sommige gemeenten hebben daarnaast een aparte onderzoeksafdeling, los van de beleidsafdelingen, terwijl andere gemeenten per beleidsafdeling onderzoekers in dienst hebben (zoals de gemeente Den Haag). Uit verschillende interviews is gebleken dat dit verschillende implicaties heeft voor het gebruik van BOLD, zoals voor het evalueren van beleid of het openbaar maken van gegevens (dit wordt behandeld in de volgende alinea's).

4.2 De invloed van BOLD op efficiëntie

Op verschillende manieren wordt verondersteld dat het gebruik van BOLD kan leiden tot meer efficiëntie in de overheid (ook wel resultaatgericht bestuur). Dat wil zeggen dat overheden door gebruik van BOLD meer rendement uit de beschikbare middelen kunnen halen (hypothese 1). Op de vraag welke gevolgen het gebruik van BOLD voor gemeenten kan hebben, hebben zeven respondenten een antwoord gegeven gerelateerd aan efficiëntie. Uit de interviews met medewerkers van de gemeente Den Haag blijkt de verwachting dat er door gebruik van *big*, *open* en *linked data* efficiëntieslagen te behalen zijn, bijvoorbeeld door hergebruik van beschikbare databronnen (zie Bijlage 13). Ook bij de verkennende gesprekken tussen het CBS en de gemeente werd uitgegaan van het idee van kostenbesparing door meer gebruik van data (zie Bijlage 5):

“Dat was aan de orde bij het UDC in Den Haag. Ik ben toen mee geweest met het gesprek bij gemeente Den Haag en toen zei iemand ‘met een beetje mazzel kun je met betere data drie procent op je uitkeringen besparen’. Nou toen begonnen de ogen van de wethouder te glimmen.”

Het gebruik van data onder de huidige omstandigheden blijkt uit de gesprekken nog niet erg gestructureerd te zijn. Het is vooral gebaseerd op vraag en aanbod; afdelingen zoeken elkaar op als data nodig is en maken daar individueel afspraken over (zie Bijlage 11 en 13). Door voorzieningen beter in te richten zodat meer gebruik gemaakt kan worden van BOLD, kunnen beleidsontwikkelaars en statistici sneller problemen oplossen. Volgens de derde geïnterviewde begint het algemene besef ook



door te dringen bij ministeries, provincies en gemeenten dat er al veel data beschikbaar is waarmee beleid gemaakt kan worden en geld efficiënter kan worden besteed (zie Bijlage 5). Door databronnen te hergebruiken en te koppelen, kunnen gemeenten doelgroepen en fenomenen beter in kaart brengen. Sommige gemeenten schrijven bijvoorbeeld kostbare enquêtes uit, die door gebruik van BOLD en nieuwe technieken vervangen kunnen worden door goedkopere opties (zie Bijlage 9). Een van de respondenten antwoordde dat iets op intuïtie roepen sneller en goedkoper is dan met data-analyses uitzoeken. Op de lange termijn is het volgens dezelfde respondent echter goedkoper en sneller om beleid te laten leiden door harde cijfers (zie Bijlage 6).

4.3 De invloed van BOLD op effectiviteit

De tweede veronderstelling van deze studie was dat de overheid op verschillende beleidsgebieden, door gebruik van BOLD, effectiever beleid zou kunnen voeren (hypothese 2). Dat wil zeggen dat het openbaar bestuur doeltreffender is in het behalen van gewenste doelen (resultaatgericht bestuur). Zeven respondenten gaven – met betrekking tot de gevolgen het gebruik van BOLD – een antwoord gerelateerd aan effectiviteit. De respondenten van gemeente Den Haag verwachten dat de effectiviteit van de gemeente door BOLD toeneemt (zie Bijlage 11 en 13). Door meer inzicht te hebben in wat er in de stad gebeurt, kan de effectiviteit en kwaliteit van dienstverlening worden verhoogd. De signalen die *big data-analyses* opleveren zijn volgens de programmamanager ICT van Den Haag waardevol, maar er moet altijd een realiteitscheck plaatsvinden. Een voorbeeld dat werd genoemd is dat de gemeente Den Haag onder andere beleid voert door in wijken beschikbare regelingen te promoten. Data-analyses zouden kunnen uitwijzen in welke wijk het bereik lager ligt, waardoor de gemeente effectiever beleid kan voeren (zie Bijlage 7).

Gemeenten zitten het dichtst op de burger en zouden ook volgens een respondent door data-analyses en hergebruik van bronnen effectiever doelgroepen en maatschappelijke problemen in kaart kunnen brengen (zie Bijlage 3). Het beleid zou daarmee beter gebaseerd zijn op de realiteit, in plaats van op een aanname van de realiteit. Door dit *'fact-based'* beleid neemt de effectiviteit van de maatregelen toe. Veel respondenten zien op dit vlak meerwaarde in het gebruik van BOLD (zoals aangegeven in paragraaf 4.1). Op de vraag hoe dit in de praktijk met behulp van BOLD mogelijk is, gaf een respondent het volgende voorbeeld (zie Bijlage 5):

“Door mensen gericht aan het werk te helpen bijvoorbeeld. Als je groepen op een goede manier kan karakteriseren en begeleiden naar werk, dan hebben ze geen uitkering nodig. Het werkt zowel aan de negatieve kant als fraudebestrijding als ook aan de positieve kant dat mensen geen uitkering meer nodig hebben.”



Het koppelen en hergebruiken data kan leiden tot nieuwe inzichten van oude databronnen. Basisregistraties kunnen bijvoorbeeld worden verrijkt, wat passender en effectiever beleid mogelijk maakt. Het inrichten van het gebruik van BOLD kan er tevens toe leiden dat cijfers beter beschikbaar komen voor beleidsmakers (zie Bijlage 7).

4.4 Rechtmatigheid bij het gebruik van BOLD

Uit het theoretisch kader is naar voren gekomen dat het gebruik van BOLD (hoofdzakelijk *big data*) niet goed aansluit op de huidige wettelijke regelingen over anonimiteit, doelbinding en dataminimalisatie. Tevens richt de huidige wetgeving zich vooral op het verzamelen van data en niet zo zeer op wat met de data gebeurt. Rechtmatigheid zou vooral van belang zijn bij de beleidsvoorbereiding, -vaststelling en -uitvoering. De derde hypothese was vervolgens opgesteld waarin wordt verwacht dat het gebruik van BOLD leidt tot overheidsbeleid dat minder in overeenstemming is met bestaande wet- en regelgeving (als onderdeel van behoorlijk bestuur).

De verwachting van de respondenten van de gemeente Den Haag is dat BOLD niet noodzakelijkerwijs leidt tot minder rechtmatigheid. Bij de gemeente Den Haag wordt het belang van wetgeving bij het gebruik van data hoog geacht. Dat wordt zowel door medewerkers van de gemeente als van het CBS beaamd (zie Bijlage 7, 9, 10 en 13). Een van de respondenten van de gemeente Den Haag zei hierover dat de aandacht voor rechtmatigheid in de cultuur van de gemeente gegoten zit, wat bevestigt dat het nastreven van waarden voor een groot deel bepaald wordt door de organisatiecultuur (zie Bijlage 13). Ook het instellen van een *privacy officer* bevestigt dat de gemeente over rechtmatigheid en privacy nadenkt. Op experimentele basis onderzoekt gemeente Den Haag momenteel in hoeverre het gebruik van BOLD en data-analyses voor de praktijk van belang kunnen zijn en in hoeverre dit geoorloofd is. In het Haagse UDC (en UDC's in het algemeen) hanteert het CBS de norm dat louter statistisch onderzoek wordt uitgevoerd dat beleidsondersteunend is. In beginsel worden geen onderzoeken uitgevoerd gericht op opsporing en handhaving en worden enkel analyses op wijk- en buurtniveau gedaan in plaats van op persoonsniveau. Het botst volgens verschillende respondenten weleens omdat gemeenten data het liefst op individueel niveau willen gebruiken en onderzoeksvragen vaak richten op handhaving en persoonlijke contacten (zie Bijlage 3 en 7).

Uit de interviews blijkt – ondanks het belang dat eraan gehecht wordt – dat de grenzen over wat wel en niet toegestaan is, niet altijd helder zijn. Een respondent heeft ervaren dat de ene gemeente aangeeft dat ze veel gegevens hebben maar denkt het niet te mogen koppelen, terwijl een andere gemeente dezelfde data, bijvoorbeeld persoonsdata en bijstandsgegevens, wel koppelt en gebruikt in data-analyses (zie Bijlage 7). Ook wordt toegevoegd: “...als ik vertel hoe het CBS met data omgaat en de gefronste wenkbrauwen zie, dan is mijn inschatting dat bij gemeenten lang niet zo strikt wordt



omgegaan met persoonsgegevens” (zie Bijlage 7). Een andere respondent bevestigt dat de kennis over wat kan en mag, zeker met de komende Europese wetgeving, lang niet bij iedereen aanwezig is (zie Bijlage 8). Volgens een geïnterviewde is het zonder meer belangrijk dat wetgeving en privacy hoog in het vaandel worden gedragen, maar kan het ook een remmende factor zijn als dit als grootste zorg wordt gezien bij dataprojecten (zie Bijlage 10). Ook andere respondenten hebben opmerkingen gemaakt waaruit blijkt dat het gebruik van BOLD kan wringen met wetgeving (zie Bijlage 3, 4, 5, 6, 8 en 12). Dat heeft er onder andere mee te maken dat de ruimte tussen de politiek en de uitvoering relatief kort is. Een respondent heeft hierover het volgende gezegd (zie Bijlage 6):

“...die statische afdelingen staan vaak onder druk van allerlei beleidsmedewerkers of wethouders. Dat zie je natuurlijk bij de Rijksoverheid ook. Er zijn best wel mensen die snappen dat privacy belangrijk is, maar als er een rapport voor de minister gemaakt moet worden dan komt het soms toch wel in een ander daglicht te staan.”

Dat heeft volgens deze respondent tot gevolg dat in het bestuurlijk krachtenveld beslissingen genomen kunnen worden die niet in lijn zijn met juridische normen. Hiermee kan de opmerking van een andere respondent samenhangen, waarin wordt gesteld dat een cultuurverschuiving gaande is waarin alles weten belangrijker wordt dan privacy. Deze is analoog aan ontwikkelingen rondom terrorisme en veiligheid (zie Bijlage 9).

4.5 Verantwoording en het gebruik van BOLD

Van een responsieve overheid wordt verwacht dat acties aan betrokken belanghebbenden worden uitgelegd en gerechtvaardigd. Dit veronderstelt een mate van openheid en transparantie. Het openen en beschikbaar maken van data (*open data*) leidt in theorie tot meer transparantie en openheid. Het is echter de vraag hoe overheden verantwoording afleggen over beleid dat tot stand komt door gebruik van BOLD. Uit de literatuur bleek dat managers en ambtenaren niet altijd geneigd zijn tot het verantwoorden en publiceren van uitkomsten en algoritmes van dataprojecten omdat dit tot kritische vragen kan leiden. De vierde verwachting van dit onderzoek was daarom dat het gebruik van BOLD tot minder verantwoording van overheidsbeleid leidt.

De verwachting van respondenten van de gemeente Den Haag is dat BOLD tot meer openheid leidt, hoewel dat in de praktijk soms anders uitpakt. De gemeente Den Haag voert sinds een aantal jaren een zogenaamd ‘open, tenzij’ beleid met als doel om meer data als *open data* beschikbaar te maken. Hiertoe is bijvoorbeeld een *open data* beheerunit opgesteld die technische ondersteuning geeft bij het openbaar maken van datasets. Uit de gesprekken blijkt dat er uiteindelijk meer opgesloten is in systemen dan men denkt en wil (zie Bijlage 7, 11 en 13). Een van de oorzaken – is zoals eerder gesteld – dat medewerkers al jaren lang gewend zijn in silo’s te werken en dat dat moeilijk te doorbreken is.



Ook is het gebruik, koppelen, analyseren en publiceren van data momenteel nog niet gestructureerd en professioneel genoeg ingericht (zie Bijlage 13). Een moeilijke afweging is dat het BOLD-analyses confronterende resultaten kunnen opleveren. Medewerkers kunnen volgens de programmamanager ICT van de gemeente Den Haag daardoor geneigd zijn data zelf te analyseren, om vervolgens eerst hun gedrag aan te passen voordat er openheid van zaken wordt gegeven (zie Bijlage 11).

Verschillende respondenten ervaren dat er op dit moment ook bij andere gemeenten te weinig verantwoording over datagebruik plaatsvindt en zouden dit graag meer zien. Een gemeente zoekt het volgens de programmamanager UDC's van het CBS vaak in *open data* platforms waarin data openbaar worden gemaakt. Dat zou voornamelijk zijn hoe gemeenten denken over verantwoording afleggen aan burgers (zie Bijlage 3). Verantwoording en transparantie houden volgens een andere respondent echter meer in dan alleen data op internet zetten (zie Bijlage 7). Het toelichten van het datagebruik en het publiceren van resultaten vindt volgens verschillende respondenten desondanks amper plaats; er zijn weinig partijen die bij de beleidsverantwoording vermelden welke databron ze gebruikt hebben en welke analyses ze hebben gedaan om beleid te maken (zie Bijlage 8). Het is daarbij wel de vraag in hoeverre dit van gemeenten verwacht zou moeten worden, gezien het gebruik van BOLD en nieuwe analysetechnieken nog in de kinderschoenen staat. Als vervolgens wordt samengewerkt met commerciële partijen, kan dit volgens een geïnterviewde tot problemen leiden in de transparantie van de overheid (zie Bijlage 5). Private partijen hebben belang bij een gebrek aan transparantie en zijn niet vaak geneigd tot het openbaar maken van dataprocessen, terwijl dit van de overheid juist verwacht wordt. Het risico van het inhuren van een derde partij is daarnaast dat het kennisniveau van de gemeente daarmee weinig toeneemt (zie Bijlage 10).

Meerdere respondenten herkennen dat gemeenten het als lastig ervaren om over het gebruik van data openbaarheid van zaken te geven. Dat kan in een UDC wringen met de verplichting van het CBS om resultaten en onderzoeken te allen tijde openbaar te maken. Een respondent heeft hierover het volgende gezegd (zie Bijlage 3):

“...wat we merken is dat niet iedereen heel erg staat te springen dat we alle resultaten openbaar maken. Maar het is onze wettelijke taak om melding te maken van het onderzoek dat we doen en ook van de resultaten.”

Hier hangt mee samen dat een gemeente veel wisselwerking kent tussen de politiek en beleidsmakers; het publiceren van onderzoeksresultaten kan politiek gevoelig liggen. Het naar buiten brengen kan bijvoorbeeld tot gevolg hebben dat een oppositie lastige vragen stelt, waardoor beleidsmedewerkers geen tijd hebben om beleid te maken omdat zij zich bezig moeten houden met het beantwoorden van de vragen (zie Bijlage 7).

4.6 BOLD en gelijkheid

Van behoorlijk bestuur wordt een gelijke behandeling van burgers verwacht. Analyses en algoritmen kunnen echter zo neutraal of selectief zijn als de personen die ze ontwikkelen. BOLD toepassingen dragen dus eventuele menselijke vooroordelen en kunnen via profilering en selectief beleid leiden tot minder gelijkheid tussen burgers. Ook kan de onbalans tussen burgers en de overheid toenemen als burgers steeds minder kennis van en invloed hebben op wat er binnen de overheid gebeurt. In de vijfde hypothese wordt gesteld dat het gebruik van BOLD leidt tot minder gelijkheid van burgers ten opzichte van de overheid.

Aan respondenten werd het voorbeeld voorgelegd van een gemeente die in kaart wil brengen hoe vaak inwoners van de stad boetes krijgen, omdat ze hier vaak klachten over krijgen. Het stadsbestuur wil liever niet eigen bewoners beboeten, maar heeft het beleid gericht op het ontmoedigen van forenzen. Ook bij de gemeente Den Haag worden op basis van profielen of groepen analyses gedaan, maar dat is vaak privacygevoelig en gebeurt met eigen, gesloten en gestructureerde data. *Big, open* en *linked data* worden hier niet voor gebruikt. Het betreffen bijvoorbeeld onderzoeken naar vroegtijdig schoolverlaters, uitkeringsgerechtigden en criminelen (zie Bijlage 11). Het risico kan volgens een respondent van gemeente Den Haag ontstaan dat resultaten van BOLD-analyses voor waar worden aangenomen en daarnaar gehandeld wordt. Het is dan de vraag of geanticipeerd wordt op de toekomst, of de toekomst wordt bepaald (zie Bijlage 13). Een van de andere respondenten zei over het voorbeeld (zie Bijlage 5):

“...die vragen leven wel degelijk inderdaad. En dat voorbeeld van eigen volk eerst, bewoners anders behandelen dan gasten in de gemeente, dat leeft ook. De behoefte aan profielen is er gewoon.”

Met meer gebruik van BOLD wordt stigmatisering vervolgens een risico (zie Bijlage 5). Een van de respondenten merkte hierbij op dat profileren niet nieuw is. Al jaren worden analyses gedaan en tabellen gepubliceerd op basis van profielen, bijvoorbeeld over het aantal bijstandsmensen in een bepaalde gemeente naar leeftijd en geslacht. Anekdotisch vertelt de respondent wel het gevoel te hebben dat de kennisasymmetrie tussen de overheid en burgers door datagebruik groter wordt omdat er weinig verantwoording en participatie plaatsvindt (zie Bijlage 3). Het risico speelt dat door gebruik van BOLD selectief beleid ontstaat, maar gemeenten zijn daar volgens sommige respondenten nog niet zo ver in (zie Bijlage 7, 11 en 13). In de ervaring van de respondenten is de intentie van een gemeente bij het gebruik van BOLD veelal het verbeteren van de dienstverlening, bijvoorbeeld door in kaart te brengen in hoeverre een regeling de doelgroep bereikt (zie Bijlage 3, 7 en 9).



Aan de andere kant neigen gemeenten naar opsporing en handhaving en worden op experimentele basis analyses uitgevoerd, dat een van de respondenten typeerde als grijs gebied (zie Bijlage 5). Een gemeente deed bijvoorbeeld onderzoek naar het detecteren van fraude en onderzocht wat de voorspellers van frauduleus gedrag zijn. Als iemand aan de opgestelde criteria voldoet, zou vervolgens een signaal afgegeven kunnen worden. De respondent gaf aan dat soort risico- en doelgroepanalyses riskant en stereotyperend te vinden. Etnisch profileren ligt in het verlengde hiervan en kan tot gevolg hebben dat men in een feedbackloop terecht komt (zie Bijlage 9):

“Gezien dat Marokkanen oververtegenwoordigd zijn in de criminaliteit kun je vervolgens selectief Marokkanen staande houden en bekijken of ze bijvoorbeeld drugs in de auto hebben. Af en toe zal je dat vinden, maar vervolgens voed je die statistiek weer met extra Marokkanen, omdat je die selectiever benadert. Voor je het weet loopt zo iets heel erg uit de hand, als je dat niet altijd ijkt aan een neutrale dataset die je niet met behulp van dat soort profileren hebt gemaakt.”

4.7 Participatie van burgers bij het gebruik van BOLD

Voor responsief bestuur is het van belang dat de omgeving, burgers en bedrijven worden betrokken bij beleid. Vooral *open data* heeft het potentieel om de participatie te verhogen (als ook de transparantie). Uit de gesprekken met respondenten, acties van de gemeente op internet en raadsinformatie komt naar voren gekomen dat de gemeente Den Haag het belang van *open data* groot acht. Zo is, zoals eerder benoemd, een *open data* beheerunit ingesteld die gespecialiseerd is in het openen van datasets. De data worden online gepresenteerd zodat gebruikers ze volgens een ‘open standaard’ en zonder restricties kunnen gebruiken. In samenwerking met de Haagse Hogeschool en andere partners wordt de aanpak op het gebied van *open data* verder ontwikkeld. Aan de samenleving wordt de ruimte gegeven om de data te gebruiken en (zakelijke) kansen te benutten. Zonder gegronde reden zal de gemeente bijvoorbeeld niet zelf applicaties ontwikkelen met *open data* (Engelshoven & Baldewsingh, 2013). Erkend wordt wel dat een slechts een select gezelschap in de stad nu gebruik maakt van een vorm van *open data* (zie Bijlage 13). Naast hogescholen, overheden en bedrijven, lijken burgers niet veel te worden betrokken. De programmamanager ICT van gemeente Den Haag geeft daarbij aan dat de wens om burgers te laten participeren in de gemeente groot is. Bij het betrekken van burgers bij data is het lastige volgens haar (zie Bijlage 11):

“...dat het moeilijk is om een representatieve steekproef te krijgen. Wij doen het bijvoorbeeld wel met studenten, die laten we de wijk in gaan om mensen te interviewen en te enquêteren. Maar daar krijgen we eigenlijk geen brede respons op. Nooit jongeren, maar heel vaak witte ouderen. Dus we doen wel pogingen, maar we zijn daar nog niet zo succesvol in.”



De programmamanager UDC's van het CBS zei het volgende over participatie (zie Bijlage 4):

“Op dit moment vind ik het voor zover ik kan zien nog te beperkt. Dus daarmee zie ik vooral wel de wens om de buitenwereld te betreden. Dus ondernemers, wetenschap en hogescholen en dat soort mensen, maar met burgers zie ik daarin nog helemaal geen stappen gezet worden.”

Een van de respondenten van het Centre for BOLD Cities beaamde dit en vertelde dat in het Center verschillende dataprojecten van de overheden zijn bekeken en geanalyseerd. De projecten bevatten vaak een element van burgerparticipatie, maar het wordt vaak gezien als iets “om af te vinken” en maar weinig mensen worden aan de voorkant bij beleid betrokken (zie Bijlage 10 en 12). Tussen meerdere respondenten bestaat overeenstemming over de constatering dat de overheid burgers over het algemeen niet voldoende betreft bij het gebruik van BOLD.

4.8 BOLD en respect voor wenselijke waardigheid

Uit de theoretische verkenning is tevens duidelijk geworden dat het gebruik van BOLD druk kan leggen op menselijke waardigheid en bestaan. Burgers kunnen het gevoel hebben dat fundamentele mensenrechten, zoals recht op vrijheid van meningsuiting en privacy, in het geding zijn omdat handelingen gemonitord en geanalyseerd worden met *big data* en direct gevolgen kunnen hebben. De verwachting van dit onderzoek was dat het gebruik van BOLD door de overheid leidt tot minder respect voor de menselijke waardigheid van de burger. Respondenten vonden het bij dit onderwerp lastig om de vinger op de (eventuele) wond te leggen. Aan respondenten werd daarom het voorbeeld van verzekeraars voorgelegd die verzekeringsaanvragen onder andere beoordelen door analyses te doen op risicogebieden op basis van gemiddelde inkomens via postcodes. Ook werd het *big brother* idee van de overheid genoemd. Vervolgens werd de vraag gesteld welke risico's er zijn als de overheid met behulp van dergelijke analyses beleid zou maken en bijvoorbeeld uitkeringen op basis hiervan zou toebedelen.

De respondenten van de gemeente Den Haag gaven allereerst aan dat er geen experimenten worden gedaan met dergelijke onderzoeken (zie Bijlage 11 en 13). De verwachting is dan ook niet dat BOLD voor dergelijk beleid zal worden gebruikt (zie Bijlage 13). Wel werd het risico van polarisatie genoemd en dat mensen vroegtijdig kunnen gaan afvallen bij gebruik van dit soort analyses (zie Bijlage 11). Op het gebied van veiligheid zouden de verschillen tussen voorstanders van veiligheid en voorstanders van privacy bijvoorbeeld versterkt kunnen worden. Ook verschillen tussen inkomensgroepen zouden benadrukt kunnen worden. De gemeente Den Haag moet nog een antwoord formuleren op de vraag tot hoever het acceptabel is om bij het gebruik van BOLD te gaan (zie Bijlage 11). Over het algemeen hebben respondenten ook hier aangegeven dat het uitgangspunt van gemeenten vooral is om de burger verder te helpen en de dienstverlening te verbeteren (zie Bijlage 3, 4, 7, 9, 10). Gemeenten



brengen het gebruik van regelingen bijvoorbeeld in kaart om vervolgens meer mensen te kunnen bereiken (en bestaan te bevorderen). Uit een van de interviews bleek wel dat burgers bij dataprojecten een wisselend sentiment kunnen hebben over de gemeentepolitiek. Bij een onderzoek naar uitkeringsgerechtigden werd aan cliëntenraden bijvoorbeeld om hun mening gevraagd over het gebruik van hun data. Zij reageerden volgens een respondent met het volgende: *“...’het ligt er ook maar net aan wie er op dat moment een gemeente bestuurt, hoe er wordt omgegaan met de data die daar beschikbaar zijn”* (zie Bijlage 11).

De gevolgen voor menselijke waardigheid en bestaan hangen volgens een respondent af van hoe BOLD wordt gebruikt en door overheden wordt geframed. Aan de ene kant kan het volgens de respondent voelen als *big brother*, maar als het vrijblijvend wordt aangeboden en de keuze bij de burger ligt, kan de overheid meer voor burgers betekenen en diensten laagdrempeliger aanbieden (zie Bijlage 3). Aan de andere kant benoemen twee respondenten dat de reikwijdte van de overheid al behoorlijk ver gaat, bijvoorbeeld op het gebied van terrorisme en met het intreden van de ‘Sleepwet’ (zie Bijlage 5 en 7). Op het voorbeeld reageerde een andere respondent met het volgende: *“Dan is 1984 dichtbij. Je hoort toch veel mensen praten die voor dat soort dingen niet veel schroom hebben”* (zie Bijlage 6). Het *big brother* probleem is volgens een geïnterviewde iets van alle tijden, maar de technische mogelijkheden zijn de afgelopen tijd flink verruimd. Daarmee is de kloof tussen de wereld van IT’ers en die van beleidsmakers en politici groter geworden (zie Bijlage 5).

4.9 Toetsing hypothesen

In de voorgaande paragrafen zijn de resultaten van de interviews per publieke waarde beschreven. Centraal stonden de percepties en ervaringen van respondenten in het kader van BOLD en publieke waarden. Naarmate deze waarden meer of minder nagestreefd kunnen worden, kan gesproken worden van beter of slechter openbaar bestuur. In deze slotparagraaf worden de implicaties voor de verwachtingen van deze studie beschreven. Per hypothese wordt kort benoemd of de resultaten van dit onderzoek de verwachting al dan niet ondersteunen. Met behulp van de percepties en inschatting van de respondenten kunnen de mogelijke normatieve consequenties van het gebruik van BOLD voor goed openbaar bestuur in kaart worden gebracht.

Over het algemeen zijn de respondenten het er over eens dat BOLD kan leiden tot een efficiëntere overheid. Door hergebruik van databronnen kunnen overheden met de beschikbare middelen meer rendement behalen. Ook maken data-analyses het mogelijk om bestaande werkprocessen te vervangen door goedkopere opties. De resultaten van de respondenten ondersteunen de eerste hypothese:

Hypothese 1. Het gebruik van BOLD leidt tot meer efficiëntie bij het uitvoeren van overheidsbeleid



Het gebruik van BOLD kan volgens meerdere respondenten ook tot effectiever overheidsbeleid leiden. De opmerking werd meermaals gemaakt dat overheidsbeleid met gebruik van BOLD op de realiteit gebaseerd kan worden in plaats van op intuïtie. Dat feiten en data niet altijd zuiver gehouden kunnen worden en altijd bediscussieerd kunnen worden, werd niet benoemd. Volgens de perceptie van respondenten kunnen beleidsdoelen met gebruik van BOLD beter worden behaald en neemt de effectiviteit van maatregelen toe. Deze resultaten ondersteunen de tweede hypothese:

Hypothese 2. Het gebruik van BOLD leidt tot meer effectiviteit bij het uitvoeren van overheidsbeleid

Hoewel de wetgeving over datagebruik volgens respondenten hoog in het vaandel wordt gedragen, bleek uit de resultaten dat de kennis over wat kan en mag lang niet bij iedereen aanwezig is. In de praktijk wordt volgens de respondenten niet altijd even strikt wordt omgaan met persoonsgegevens. Daar gaat mee gepaard dat de kennis over de data-opslag en het delen van gevoelige informatie niet altijd op orde is. Ook kan rechtmatigheid bij datagebruik in een ander daglicht komen te staan door politieke invloeden. Op basis van de opmerkingen van de respondenten wordt de derde hypothese ondersteund:

Hypothese 3. Het gebruik van BOLD leidt tot minder rechtmatigheid van overheidsbeleid ten opzichte van burgers

Met het oog op verantwoording ervaren respondenten dat gemeenten het vooral zoeken in *open data*, hoewel verantwoording meer inhoudt dan data op internet plaatsen. Weinig partijen vermelden bij de beleidsverantwoording welke databron ze gebruikt hebben en welke analyses hebben bijgedragen aan beleid. Ook aan de voorkant van de beleidsvorming vindt er volgens de respondenten te weinig verantwoording van datagebruik plaats. Dat personen niet staan te springen om resultaten openbaar te maken, hangt samen met de korte afstand tussen beleidsmakers en de gemeentelijke politiek. Ook kunnen analyses confronterende resultaten opleveren, waardoor de drempel voor het publiceren van de resultaten hoger wordt. De opmerkingen ondersteunen de vierde hypothese:

Hypothese 4. Het gebruik van BOLD leidt tot minder verantwoording van de overheid over haar beleid ten opzichte van burgers

Gemeenten maken al langere tijd bij data-analyses gebruik van profielen. Daar is volgens de respondenten niet direct iets mis mee. Aan de andere kant benoemden respondenten risico's van dataprojecten, die vooral op het vlak van stigmatisering en stereotypering liggen. De kans op



discriminatie en vooroordelen neemt daarmee toe. Stigma's en stereotypen kunnen volgens een van de respondenten als een *self fulfilling prophecy* werken; als de resultaten als waar worden beschouwd, wordt daarnaar gehandeld. Ook neemt de kennisasymetrie tussen burgers en overheid volgens de perceptie van verschillende respondenten toe. Deze opmerkingen ondersteunen de vijfde hypothese:

Hypothese 5. Het gebruik van BOLD leidt tot minder gelijkheid van burgers ten opzichte van de overheid

Vooraf *open data* stimuleert volgens respondenten inderdaad de participatie van de omgeving, burgers en bedrijven. In de praktijk is het echter een selecte groep die gebruik maakt van *open data*. Ondanks de wens van gemeenten, wordt door respondenten erkend dat vooral burgers bij dataprojecten te weinig bij de agendavorming en beleidsvoorbereiding worden betrokken. Daar hangt mee samen dat het volgens bepaalde respondenten lastig is om een representatieve steekproef te vinden die betrokken kan worden bij de beleidsvorming. In dit geval ondersteunen de opmerkingen van respondenten de zesde hypothese niet:

Hypothese 6. Het gebruik van BOLD leidt tot meer participatie van burgers in het overheidsfunctioneren

Aan de ene kant kan de overheid, door gebruik van BOLD, diensten passender en laagdrempelig aanbieden, waarbij de keuze bij de burger ligt. Aan de andere kant kan het analyseren van risicogroepen volgens respondenten de tegenstellingen tussen partijen of bevolkingsgroepen versterken (polarisatie). Door het toenemen van de technische mogelijkheden en uitblijven van verantwoording en participatie, wordt ook het *big brother* probleem groter. Daar komt bij dat door meerdere respondenten erkend wordt dat het ethisch bewustzijn van personen bij datagebruik beter ontwikkeld zou moeten worden. De opmerkingen van respondenten bieden ondersteuning voor de laatste hypothese:

Hypothese 7. Het gebruik van BOLD in de overheid leidt tot minder respect voor de menselijke waardigheid van de burger



5 CONCLUSIE EN AANBEVELINGEN

In dit slothoofdstuk wordt kort en bondig antwoord gegeven op de onderzoeksvraag. Hiervoor wordt geput uit voorgaande empirische en analytische hoofdstukken. Vervolgens wordt op het onderzoek gereflecteerd en ingegaan op de methodologische beperkingen. Tevens worden hier suggesties voor vervolgonderzoek gegeven. In de discussie worden de resultaten geïnterpreteerd en worden praktische aanbevelingen gedaan.

5.1 Beantwoording onderzoeksvraag

Gemeenten staan onder andere door bezuinigingen, decentralisaties, digitalisering en stad-overstijgende problematiek voor een uitdaging. *Big, open* en *linked data* kunnen een nieuwe impuls geven aan de rol van (lokale) overheden en het functioneren van de samenleving. Analyses en voorspelling van het gedrag van burgers kunnen aan de ene kant leiden tot passender en doelmatiger overheidsbeleid. Aan de andere kant kan het gebruik van BOLD ten koste gaan van fundamentele mensenrechten als privacy en vrijheid van meningsuiting. In deze studie is onderzocht welke mogelijke consequenties het gebruik van BOLD heeft op publieke waarden en daarmee de kwaliteit van het openbaar bestuur. Naarmate publieke waarden meer of minder worden geborgd en bevorderd, kan in normatieve zin gesproken worden van beter of slechter openbaar bestuur.

Het borgen en bevorderen van efficiëntie en effectiviteit is een belangrijk onderdeel van resultaatgericht bestuur. Het gebruik van BOLD biedt kansen voor het verhogen van de efficiëntie en effectiviteit in de uitvoering van taken. Data-analyses kunnen ertoe leiden dat de overheid doeltreffender is in het behalen van beleidsdoelen en meer rendement uit beschikbare middelen kan behalen. Datasets kunnen herbruikt worden in (*big*) data-analyses door deze toegankelijker te maken (als *open data*) en makkelijker te koppelen te maken (als *linked data*). Inzichten komen daarmee beter in de buurt van de realiteit waardoor de effectiviteit van beleid toeneemt. Veel van het enthousiasme voor BOLD is ingegeven door de toegenomen technologische mogelijkheden om maatschappelijke fenomenen kwantificeerbaar en 'objectief' te maken. Op dit gebied kan BOLD dus tot beter, dat wil zeggen, resultaatgerichter bestuur leiden.

Als onderdeel van behoorlijk bestuur zijn onder andere rechtmatigheid, gelijkheid en respect voor menselijke waardigheid en bestaan van belang. Uit de resultaten is gebleken dat BOLD in dit waardencluster tot botsingen kan leiden. De wetgeving over datagebruik heeft een normerend karakter en wordt in de praktijk hoog in het vaandel gedragen. Desondanks zijn de grenzen over wat wel en niet kan, niet altijd duidelijk. Daar hangt er mee samen dat ook de kennis en het ethisch bewustzijn over datagebruik, -opslag en analyse niet genoeg ontwikkeld zijn. De interactie in



gemeenten tussen politici en ambtenaren, maakt het lastig om rechtmatigheid te allen tijde te borgen en na te streven.

Het risico van het gebruik van BOLD (en vooral *big data*) is daarnaast dat de kans op discriminatie en vooroordelen toeneemt. Algoritmen reflecteren de keuzes en interpretaties van data-analisten en zijn niet neutraal of objectief en daarbij ook onderhevig aan mogelijke menselijke (denk)fouten. De resultaten bevestigen dat het gebruik van BOLD kan wringen met de gelijke behandeling van burgers. Ook de (kennis)kloof tussen burger en overheid neemt toe, vooral wanneer het toelichten van datagebruik uitblijft. Tot slot wordt, door toegenomen technische mogelijkheden aan de ene kant en minder verantwoording en participatie aan de andere kant, het *big brother* probleem groter. Uiteindelijk zou dit kunnen leiden tot maatschappelijk *chilling effects* op mensenrechten zoals vrijheid van meningsuiting en privacy. Wat betreft behoorlijk bestuur kleven er dus risico's aan het gebruik van BOLD.

Als laatste staan bij responsief bestuur verantwoording en participatie centraal. De drempel om resultaten van dataprojecten openbaar te maken en toe te lichten kan hoog zijn; de resultaten kunnen politiek gevoelig liggen en confronterend zijn. Verantwoording is daardoor lastig te bevorderen. Bij het gebruik van BOLD lijkt ook participatie een moeilijk na te streven wens. In de praktijk is het een selecte groep die betrokken wordt. Hoewel bedrijven de meerwaarde van dataparticipatie (en vooral *open data*) inzien, worden burgers bij de beleidsvorming te weinig betrokken (hoewel dit probleem niet per se de oorzaak van het gebruik van BOLD is). Ook hier blijkt uit de resultaten dat BOLD in responsief bestuur mogelijke waardenbotsingen tot gevolg heeft.

5.2 Discussie

De gebruikte theorie in dit onderzoek, op basis waarvan toetsbare veronderstelling zijn opgesteld, is gebaseerd op de omvangrijke publieke waardenliteratuur en de beschikbare literatuur over BOLD. Met betrekking tot publieke waarden komt allereerst naar voren dat waarden die als vanzelfsprekend kunnen worden ervaren, in de context van BOLD weer relevant kunnen worden. Het gaat bijvoorbeeld om menselijke waardigheid, vrijheid van meningsuiting en privacy. Het is een belangrijke vraag hoe deze waarden gewaarborgd kunnen worden in het licht van BOLD. Ook is gebleken dat de interactie tussen politici en ambtenaren in een gemeente een belangrijke factor is in het borgen en bevorderen van publieke waarden. In de gebruikte theorie is daarentegen weinig ingegaan op de frictie die tussen de twee groepen kan ontstaan. Vervolgonderzoek wordt aangeraden dat zich buigt over de vraag hoe uitvoerende ambtenaren bij dataprojecten met politieke druk om kunnen gaan.

Deze studie heeft verder weinig onderzoek gedaan naar de samenhang tussen waarden, terwijl de onderlinge relatie tussen waarden van belang kan zijn (waardenpluralisme). Daarom wordt



vervolgonderzoek aangemoedigd dat kijkt naar de interactie en conflicten tussen bijvoorbeeld privacy en transparantie, effectiviteit en gelijkheid en participatie en verantwoording. Met betrekking tot publieke waarden wordt tevens onderzoek aangeraden dat zich richt op mechanismen om publieke dataparticipatie te stimuleren en vragen te monitoren en te beantwoorden. In de fase waar overheden zich nu in bevinden, zijn juist participatie en verantwoording van belang.

Ook is duidelijk geworden dat BOLD verschillende effecten heeft. *Open data* heeft vooral het potentieel om voor meer participatie en transparantie te zorgen door data klaar voor gebruik op internet te plaatsen. *Linked data* hangt daar mee samen en kan de effecten versterken; het gaat namelijk vooral om het koppelen van (openbare) data. In theorie zijn *open* en *linked data* dus niet privacygevoelig. De data kunnen wel in *big data*-analyses worden gebruikt en gekoppeld worden aan gesloten data. Op het vlak van voorspellende software en *machine-learning* kunnen dan ook vooral botsingen met waarden als privacy en transparantie verwacht worden.

Een methodologische uitdaging is geweest dat het gebruik van BOLD ten tijde van deze studie hoofdzakelijk in experimentele setting plaatsvindt. Daardoor zou gesteld kunnen worden dat bepaalde verwachtingen van dit onderzoek te vroeg getoetst zijn. Aan de andere kant heeft het onderzoek grondstof voor vervolgonderzoek opgeleverd en doet het daarmee een bijdrage aan de wetenschap en praktijk voor de inrichting van goed openbaar bestuur bij het gebruik van BOLD. De lezer moet er daarnaast rekening mee houden dat opvattingen en inschatting van respondenten zijn gebruikt om de veronderstellingen te toetsen. Daarmee kan het onderzoek iets anders meten dan de daadwerkelijke gevolgen van BOLD. Sociaalwenselijke antwoorden en *confirmation bias* zijn daarbij niet uit te sluiten. Tot slot is de studie gebaseerd op een enkele casus en kan niet zonder meer gegeneraliseerd worden naar andere *cases*, ondanks dat meerdere respondenten hun antwoorden op basis van ervaring met meerdere gemeenten konden baseren. Generalisatie over alle dataprojecten ging dan ook voorbij aan het doel van deze studie. Een grotere steekproef had daarentegen voor meer verzadiging kunnen zorgen en wordt voor vervolgonderzoek aangeraden.

5.3 Praktische aanbevelingen

Veel overheden bevinden zich in een verkennende fase voor het gebruik van BOLD. Juist nu zijn participatie van en verantwoording naar burgers van belang. Hoewel dit als een lastig streven wordt ervaren, zou hier meer aandacht voor moeten komen; het vertrouwen van burgers is essentieel voor de legimiteit van de overheid. Wellicht kunnen overheden dit organiseren op een manier die bij de huidige tijd past. De meeste burgers maken gebruik van *apps* en *social media* om te laten weten wat zij van iets denken. Door middel van dit soort digitale platformen zouden burgers kunnen worden geraadpleegd en ingelicht over het datagebruik en beleid van overheden. Daarbij moet er wel rekening



gehouden worden met de (kleine groep) burgers die hier niet de digitale vaardigheden of mogelijkheden voor heeft. Daarnaast is het een risico dat de kloof tussen data-analisten, ambtenaren en politici groter wordt. Data-analisten kunnen vanuit hun functie grote invloed uitoefenen op de beslissingen van bestuurders. Het is daarom belangrijk om BOLD-toepassingen altijd te ijken aan neutrale datasets of zelfs door gewoon op straat verwachtingen te checken. Daar hoort bij dat resultaten niet direct voor waarheid worden aangenomen. Niet iedereen is namelijk te vatten in profielen en etnisch profileren loert om de hoek. Dat is niet alleen onrechtvaardig, maar ook ineffectief en vertekend.

Afsluitend is het – ondanks de grote beloften van BOLD – belangrijk dat de verdere uitwerking gelijktijdig plaatsvindt met de ontwikkeling van doeltreffende waarborgen om de mogelijke risico's te beperken. Met het oog op goed openbaar bestuur moet voorkomen worden dat data-analyses leiden tot discriminatie, inbreuk op privacy en minder transparantie. Het is daarom goed om bij dataprojecten de normatieve uitgangspunten in beschouwing te nemen; data en 'feiten' zijn namelijk niet altijd zo objectief als het lijkt. Over de kwaliteit van de data, de definitie van concepten, de keuzes van methoden en de interpretatie van resultaten, is altijd discussie mogelijk. Het is belangrijk dat overheden, bedrijven en burgers deze discussie aangaan en samen nadenken over de vraag hoever de overheid met data-analyses zou moeten gaan. Met behulp van experimenten en lerend vermogen, kunnen kennis en ethisch bewustzijn over BOLD en data-analyses ten behoeve van de praktijk worden opgedaan. De Ethische Data Assistent (DEDA) van Utrecht Data School (2017) is hierbij een passend instrument. Door voldoende aandacht aan de hierboven genoemde aspecten te schenken en publieke waarden daardoor op de juiste wijze te waarborgen en te bevorderen, kan bij burgers blijvend vertrouwen ontstaan in de manier waarop overheden bij de uitvoering van hun taken van BOLD gebruik maken.



6 VERWIJZINGEN

- AG Connect. (2014, maart 10). *Databescherming centraal in opening Cebit*. Opgehaald van <https://agconnect.nl/artikel/databescherming-centraal-in-opening-cebit>
- Aggarwal, A. (2016). *Managing Big Data Integration in the Public Sector*. Pennsylvania: IGI Global Information Science Reference.
- Albino, V., Berardi, U., & Dangelico, R. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*(22:1), 3-21.
- Back, M. (2017, april 27). *United Airlines schikt met weggesleepte passagier*. Opgehaald van NRC: <https://www.nrc.nl/nieuws/2017/04/27/united-airlines-schikt-met-weggesleepte-passagier-a1556347>
- Biryukov, N. (2017, april 13). *United Fiasco: How Do Airlines Select Who to Remove From Overbooked Flights?* Opgehaald van NBC News: <https://www.nbcnews.com/storyline/airplane-mode/united-fiasco-how-do-airlines-select-who-remove-overbooked-flights-n746331>
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data — The story so far. International. *International Journal on Semantic Web*(5), 1-22.
- Boer, N., & van der Lans, J. (2014). *DEcentraal: De stad als sociaal laboratorium*. Amsterdam: Atlas Contact.
- Bovens, M., & Wille, A. (2008). Politiek vertrouwen langs de meetlat. *Socialisme en Democratie*(65), 32-43.
- Bovens, M., Hart, P., Twist, M., & Rosenthal, U. (2001). *Openbaar bestuur. Beleid, organisatie en politiek*. Alphen aan den Rijn: Kluwer.
- boyd, d., & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication & Society*(15), 662-679.
- Bozeman, B. (2002). Public-Value Failure: When Efficient Markets May Not Do. *Public Administration Review*(62), 145-161.
- Bryman, A. (2012). *Social research methods*. Oxford: Oxford University Press.
- BZK. (2009). *Nederlandse code voor goed openbaar bestuur*. Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.
- Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of Urban Technology*(18:2), 65-82.
- CBS. (2016, september). *CBS Urban Data Centers*. Opgehaald van CBS: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/wegwijzer/cbs-urban-data-centers>
- Centre for BOLD Cities. (2017). *Over ons*. Opgehaald van Leiden-Delft-Erasmus: <http://www.leiden-delft-erasmus.nl/nl/centres/centre-for-bold-cities>
- Chun, S. A., Shulman, S., Sandoval, R., & Hovy, E. (2010). Government 2.0: Making connections between citizens, data and government. *Information Polity*(15:1/2), 1-9.



- Clarke, R. (2016). Big data, big risks. *Information Systems Journal*(26), 77-90.
- Consumentenbond. (2015, augustus 26). *Premies verzekeringen verschillen tot op huisnummer*. Opgehaald van Consumentenbond: <https://www.consumentenbond.nl/autoverzekering/verzekeringspremies-verschillen-tot-op-huisnummer>
- Council for Big Data, Ethics, and Society. (2014). *Homepage*. Opgehaald van <http://bdes.datasociety.net/>
- Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Review*, 93-128.
- Daniel, H. (1987). Wilderness politics: Public participation and values. *Environmental Management*(11:3), 283-293.
- Dijk, B. van. (2015, september 4). *Big data: nieuwe software voorspelt diefstal en straatroof*. Opgehaald van Het Financieele Dagblad: <https://fd.nl/economie-politiek/1091395/politie-laat-computer-criminaliteit-voorspellen>
- Engelshoven, I., & Baldewingsh, R. (2013, december 3). *Den Haag Raadsinformatie*. Opgehaald van Gemeente Den Haag: <https://denhaag.raadsinformatie.nl/document/3378923/1/RIS268722%20Voortgang%20Open%20Data>
- Europa Decentraal. (2016, mei 26). *PRIVACY: DE ALGEMENE VERORDENING GEGEVENSBESCHERMING*. Opgehaald van Kenniscentrum voor Europees recht en beleid: <https://europadecentraal.nl/onderwerp/informatiemaatschappij/gegevensbescherming-en-de-avg/>
- Fischer, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012). Interactions with big data analytics. *Magazine Interactions*(19:3), 50-59.
- Franke, B., Plante, J. F., Roscher, R., Lee, E. S., Smyth, C., Hatefi, A., . . . Ried, N. (2016). Statistical Inference, Learning and Models in Big Data. *International Statistical Review*(83:3), 371-389.
- Gemeente Den Haag. (2017, september 28). *CBS Urban Data Center Den Haag*. Opgehaald van Gemeente Den Haag: <https://www.denhaag.nl/nl/bestuur-en-organisatie/feiten-en-cijfers/cbs-urban-data-center-den-haag-.htm>
- Graaf, G. de, & van der Wal, Z. (2010). Managing conflicting public values: governing with integrity and effectiveness. *The American Review of Public Administration*, 623-630.
- Graaf, G. de, Reynaers, A., Wal, Z. van der, & Doeveren, V. van. (2011). Goed bestuur als management van spanningen tussen verschillende publieke waarden. *Bestuurskunde*(20:2), 5-11.
- Graaf, G., Huberts, L., & Smulders, R. (2013). *Publieke waarden. De beginselen van goed bestuur in de dagelijkse praktijk en gemeente*. Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties in samenwerking met Vrije Universiteit Amsterdam.
- Gray, C. (1999). Reforming Public Institutions and Stengthening Governance: A World Bank Strategy. *Word Bank Public Sector Board*, 41.



- Gubbia, J., Buyyab, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*(29), 1645–1660.
- Hartholt, S. (2017, maart 3). *Binnenlands Bestuur*. Opgehaald van DE DUBIEUZE BEWIJZEN VAN BIG DATA: <http://www.binnenlandsbestuur.nl/digitaal/achtergrond/achtergrond/de-dubieuze-bewijzen-van-big-data.9570366.lynkx>
- Heath, T. (2011). Linked Data - Welcome to the Data Network. *IEEE Internet Computing*(15), 70-73.
- Hendriks, F., & Drosterij, G. (2012). *De zucht naar goed bestuur in de stad. Lessen uit een weerbarstige werkelijkheid*. Den Haag: Boom Lemma uitgevers.
- Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems*(64), 130-141.
- Hordijk, M., & Baud, I. (2009). Policy analytical capacity and evidence-based policy-making: Lessons from Canada. *Canadian Public Administration*(52:2), 153-175.
- Huberts, L. (2005). *Integriteit en integritisme in bestuur en samenleving*. Amsterdam: Vrije Universiteit.
- Huberts, L., & Van Hout, E. (2011). Goed bestuur: kiezen of delen? *Bestuurskunde*(20), 53-62.
- Jackson, M. C. (2003). *Systems thinking: Creative holism for managers*. Chichester, Engeland: John Wiley.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*(28:4), 446-456.
- Janssen, M., & Hoven, J. van den. (2015). Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly*(32), 363–368.
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*(29:4), 258-268.
- Johnson, J. E. (2012). Big data + big analytics = big opportunity. *Financial & Executive*, 51-53.
- Jongejan, W. (2016, april 23). *Big data analyse ongeoorloofd voor gemeenten in het sociale domein*. Opgehaald van Zorg-ICT zorgen: <https://www.zorgictzorgen.nl/big-data-analyse-ongeorloofd-voor-gemeenten-in-het-sociale-domein/>
- Justitie, I. V. (2013). *Het overlijden van Alexander Dolmatov*. Den Haag: Inspectie Veiligheid en Justitie.
- KING. (2016, april 19). *Big data op kleine schaal: data en gemeenten*. Opgehaald van Kwaliteits Instituut Nederlandse Gemeenten: <http://magazine.kinggemeenten.nl/data-en-gemeenten/#!/titelpagina>
- Kitchin, R. (2014). *The data revolution: big data, open data, data infrastructures and their consequences*. London: Sage.
- Kitchin, R. (2016). The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society A*, 1-15.
- Klievink, B., Romijn, B., S., C., & Bruin, H. de. (2016). *Big data in the public sector: Uncertainties and readiness*. Delft: Springer.



- Lavertu, S. (2015). We All Need Help: “Big Data” and the Mismeasure of Public Administration. *Public Administration Review*, 864-872.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.
- Marsal-Llacuna, M., Colomer-Llinàs, J., & Meléndez-Frigola, J. (2014). Lessons in urban monitoring taken from sustainable and livable cities to better address the Smart Cities initiative. *Technological Forecasting and Social Change*(90), 611–622.
- Maurits, M. (2015, april 21). *Baas Belastingdienst over Big Data: 'Mijn missie is gedragsverandering'*. Opgehaald van De Correspondent: <https://decorrespondent.nl/2720/baas-belastingdienst-over-big-data-mijn-missie-is-gedragsverandering/83656320-f6e78aaf>
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution that Will Transform How We Live, Work and Think*. Londen: John Murray.
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*(90), 60–66.
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*(90), 60–66.
- Metcalf, J., Keller, E., & boyd, d. (2017, mei 23). *Perspectives on Big Data, Ethics, and Society*. Opgehaald van THE COUNCIL FOR BIG DATA, ETHICS, AND SOCIETY: <http://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/>
- Mitlin, D. (2004). Reshaping local democracy. *Environment for Urbanization*(16:1), 3-8.
- Mitlin, D. (2004). Reshaping local democracy. *Environment and Urbanization*(16:1), 3-8.
- Newland, C. (2012). Values and Virtues in Public Administration: Post-NPM Global Fracture and Search for Human Dignity and Reasonableness. *Public Administration Review*(72), 293-302.
- Niewold, M. (2015, augustus 26). *Letter achter je huisnummer? Verzekering 100 euro duurder*. Opgehaald van RTL Nieuws: <https://www.rtlnieuws.nl/economie/home/letter-achter-je-huisnummer-verzekering-100-euro-duurder>
- NRC. (2017, oktober 28). *Buma: kabinet zal referendum over de sleepwet negeren*. Opgehaald van NRC.nl: <https://www.nrc.nl/nieuws/2017/10/28/buma-kabinet-zal-referendum-over-de-sleepwet-negeren-a1579085>
- O'Neil, C. (2017, 24 april). *In the age of 'Big Data,' United Airlines shows how human dignity suffers (Commentary)*. Opgehaald van Syracuse: http://www.syracuse.com/opinion/index.ssf/2017/04/big_data_hurts_humans.html
- Opstelten, I. (2014, december 12). *Kamerbrief over voortgang rijksbrede aanpak van fraude*. Opgehaald van Rijksoverheid: <https://www.rijksoverheid.nl/documenten/kamerstukken/2014/12/20/tk-voortgang-rijksbrede-aanpak-van-fraude>
- Pesch, U. (2008). The Publicness of Public Administration. *Administration & Society*(40), 170-193.
- Rathenau Instituut. (2017). *Human rights in the robot age*. Den Haag: Rathenau Instituut.



- Reynaers, A. (2013). Public Values in Public–Private Partnerships. *Public Administration Review*(74:1), 41-50.
- Rutgers, M. (2011). *Het Pantheon van de Publieke Waarden*. Rede uitgesproken bij de aanvaarding van het ambt van hoogleraar Filosofie van het Openbaar Bestuur aan de Universiteit van Amsterdam op vrijdag 11 februari 2011.
- Sangen, M. (2017, juli 2017). *CBS Urban Data Center/Den Haag van start*. Opgehaald van CBS: <https://www.cbs.nl/nl-nl/corporate/2017/26/cbs-urban-data-center-den-haag-van-start>
- Schäfer, M. (2017, maart 24). Meer aandacht voor ethiek nodig bij data-projecten. *Binnenlands bestuur*(6).
- Schep, T. (2016, februari 10). *Data is the new gold, who are the new thieves?* Opgehaald van Youtube (TEDx Talk): <https://www.youtube.com/watch?v=XNF-rGiGb50>
- Schep, T. (2017, juni 30). *Data Leads to Social Cooling*. Opgehaald van Social Cooling: <http://onlinelibrary.wiley.com.ezproxy.leidenuniv.nl:2048/doi/10.1111/j.1540-6210.2011.02536.x/epdf>
- Schroeck, M., Schockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012, oktober). *Analytics: the real world use of big data*. Opgehaald van IBM Institute for Business Value / Saïd Business School: http://www.ibm.com/smarterplanet/global/files/se__sv_se__intelligence__Analytics_-_The_real-world_use_of_big_data.pdf
- Shadbolt, N., & O'Hara, K. (2013). Linked Data in Government. *IEEE Internet Computing*(17), 72-77.
- Spicer, M. (2010). *In defense of politics in public administration. A value pluralist perspective*. Alabama: The University of Alabama Press.
- Steenhuisen, B. (2009). *Competing public values. Coping strategies in heavily regulated utility industries (proefschrift)*. Opgehaald van <https://repository.tudelft.nl/islandora/object/uuid:3257cd11-13ad-4636-8074-32b2dd5276e2/datastream/OBJ/download>
- Steur, A. van der. (2016, januari 7). *Brief van de minister van Veiligheid en Justitie inzake de verwerking en bescherming persoonsgegevens*. Opgehaald van Overheid.nl: <https://zoek.officielebekendmakingen.nl/dossier/32761/kst-32761-91>
- Stoker, G. (2006). A New Narrative for Networked Governance. *American Review of Public Administration*(36:1), 41-57.
- Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business economies, societies and nations*. New York, NY: Doubleday.
- Thacher, D., & Rein, R. (2004). Managing value conflict in public policy. *Governance*(17:4), 457-486.
- TNO. (2013, april). *Big Data Value Center*. Opgehaald van TNO: <https://www.tno.nl/nl/aandachtsgebieden/industrie/networked-information/information-creation-van-data-naar-informatie/big-data-value-center/>
- Utrecht Data School. (2017, maart). *De Ethische Data Assistent (DEDA)*. Opgehaald van Utrecht Data School: <https://dataschool.nl/deda/>



- Wal, Z. van der. (2008). *Value Solidity: Differences, Similarities and Conflicts between the Organizational Values of Government and Business*. PhD diss.: Amsterdam Free University.
- Wal, Z. van der. (2012). Elite Ethics: Comparing Public Values Prioritization Between Administrative Elites and Political Elites. *International Journal of Public Administration*(37), 1030-1043.
- Wal, Z. van der, Graaf, G. de, & Lasthuizen, K. (2008). What's valued most? Similarities and differences between the organizational values of the public and private sector. *Public Administration*(86), 465-482.
- Wal, Z. van der, Graaf, G. de, & Lawton, A. (2011). Competing values in public management. *Public Management Review*(13), 331-341.
- Wal, Z. van der, Huberts, L., Heuvel, J. van den, & Kolthoff, E. (2006). Central Values of Government and Business: Differences, Similarities and Conflicts. *Public Administration Quarterly*(30), 314–364.
- Weerdt, C. van der, & Vries, A. de. (2014). *Dienstverlening verbeteren met big data*. Den Haag: TNO.
- Wetenschappelijke Raad voor het Regeringsbeleid. (2011). *iOverheid*. Amsterdam: Amsterdam University Press.
- Wetenschappelijke Raad voor het Regeringsbeleid. (2016a). *Big Data in een vrije en en veilige samenleving*. Amsterdam: Amsterdam University Press.
- Wetenschappelijke Raad voor het Regeringsbeleid. (2016b). *Exploring the boundaries of big data*. Amsterdam: Amsterdam University Press.
- World Bank. (1989). *Sub-saharan Africa. From crisis to sustainable growth. A long-term perspective study*. Washington, D.C.: The World Bank.



7 BIJLAGEN

7.1 Bijlage 1: Interview vragenlijst

Introductie

Allereerst hartelijk bedankt voor uw medewerking aan dit onderzoek. Dit interview maakt onderdeel uit van mijn onderzoek naar publieke waarden binnen het openbaar bestuur in relatie tot *big*¹, *open*² en *linked*³ data (BOLD) in het kader van mijn masterscriptie. Het onderzoek voer ik uit onder begeleiding van T. Kerkhoff van de Universiteit Leiden.

Het interview gaat voornamelijk over de ontwikkeling van een data gedreven overheid, het gebruik van *big data* analyses, het delen van data en het koppelen van datasets. Daarbij gaat het allereerst over uw ervaringen in de praktijk, de beslissingen die u neemt of genomen ziet worden en wat daarvoor belangrijk is. Daarna richten we ons wat meer op de waarden die daarbij spelen.

Ik ben met name geïnteresseerd in vragen als welke publieke waarden u van groot belang acht in relatie tot BOLD, of verschillende publieke waarden wel eens met elkaar botsen, en zo ja, hoe u daar dan mee omgaat. Het onderzoek kan bestuurders, managers en ambtenaren helpen om bewustere keuzes te maken bij het gebruik van BOLD en in de overheid.

In totaal zal het interview ongeveer een uur duren. Als u daar geen bezwaar tegen heeft, wil ik het interview graag opnemen om het zorgvuldig te kunnen uitwerken. U krijgt die tekst toegestuurd om het verslag op feitelijke onjuistheden te controleren. Als u op dit moment verder geen vragen heeft, dan zou ik graag beginnen.

¹ Voor het gemak verstaan we onder *big data*: grote, gevarieerde en dynamische datasets die informatie kunnen bevatten over consumentengedrag, sensoren, sociale media en meer.

² *Open data* zijn openbaar gepubliceerde gegevens waar overheden, bedrijven en burgers zonder beperkingen en op toegankelijke wijze toegang tot hebben.

³ Met *linked data* wordt bedoeld op de digitale methode voor het online publiceren van gestructureerde data waarmee gegevensbronnen makkelijk onderling gelinkt kunnen worden. Op basis van het koppelen van datasets kunnen analyses worden gedaan.



Interviewvragen

1. Kunt u uw werk kort typeren? Wat is uw taak en functie?
2. Hoe wordt u geconfronteerd met concepten als *big*, *open* en *linked data*?
3. Steeds meer ontstaan samenwerkingsverbanden gericht op het gebruik van data door de overheid. Wat zijn volgens u moeilijke afwegingen als het gaat om de ontwikkeling van een data-gedreven overheid?
4. Gaat het daarbij in uw beleving ook om waarden⁴ en zo ja, welke waarden botsen er volgens u in de genoemde keuzes?
5. Hoe komen partijen volgens u tot een beslissing bij die waarden(botsingen)?
6. Wat zijn volgens u de grootste voordelen van het gebruik *big*, *open* en *linked data* (voor de overheid/gemeente Den Haag)?
7. Leidt het gebruik van *big*, *open* en *linked data* volgens u tot meer efficiëntie en effectiviteit (van de overheid/gemeente Den Haag), en zo ja, hoe?
8. Hoe wordt er – in uw beleving – bij het datagebruik van de overheid/gemeente Den Haag/UDC's rekening gehouden met relevante wetgeving?
9. Hoe wordt verantwoording afgelegd over het beleid dat in de overheid/gemeente Den Haag/UDC's met *big*, *open* en *linked data* wordt gevormd?
10. Wie is er verantwoordelijk voor de resultaten van de analyses van *big*, *open* en *linked data* in de overheid/gemeente Den Haag/UDC's?

Dilemma: BOLD kan in de overheid gebruikt worden om beleid te maken dat onderscheid maakt in burgerprofielen. Het voordeel is dat beleid afgestemd kan worden op de burger.

11. Hoe wordt hier mee omgegaan in de overheid/gemeente Den Haag/ UDC's?

Dilemma: Verzekeraars beoordelen verzekeringsaanvragen door analyses uit te voeren op risicogebieden en gemiddelde inkomens via de postcode. Premies van verzekeringen kunnen daarmee tot op huisnummer verschillen.

12. Wat zijn volgens u de mogelijke risico's als de overheid/gemeente Den Haag beleid baseert op dit soort toepassingen?
13. Hoe worden de omgeving, bedrijven, burgers en relevante belanghebbenden betrokken bij het gebruik van *big*, *open* en *linked data* in de overheid/gemeente Den Haag/in UDC's?

⁴ In dit onderzoek zijn waarden gedefinieerd als belangrijke (morele) standaarden en principes die bij het maken van keuzes invloed hebben.



14. Hoe worden gegevens van de overheid/gemeente Den Haag/UDC's openbaar toegankelijk gemaakt?

Dilemma: Burgers kunnen het gevoel hebben dat hun recht op vrijheid van meningsuiting en privacy in het geding zijn als handelingen gemonitord en geanalyseerd worden met big data.

15. Hoe wordt hier in de overheid/gemeente Den Haag/in UDC's mee omgegaan?