



Essaybundel
**Behoorlijk datagebruik in
de openbare ruimte**

Van dialoogtafels naar voorstellen voor nieuw beleid

> digitaleoverheid.nl/behoorlijk-datagebruik

Essaybundel

Behoorlijk datagebruik in de openbare ruimte

Van dialogotafels naar voorstellen voor nieuw beleid

> digitaleoverheid.nl/behoorlijk-datagebruik

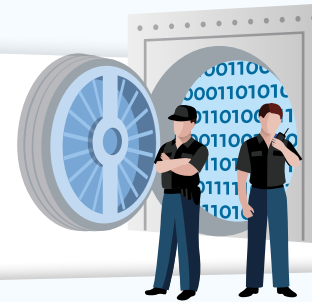
Essays

1 Hergebruik van gegevens in smart cities



2 Digitalisering, gedragsbeïnvloeding en de overheid

3 Bestuurlijke regie over experimentele data en algoritmen



4 De noodzaak van een digitale omgevingsvisie

5 AI in de digitale samenleving



6 Data makes the world go round

Inhoudsopgave

1/2

Voorwoord _____ 6

Inleiding _____ 7

Essay 1: Hergebruik van gegevens in smart cities _____ 9

Juridische en ethische kaders voor big data in de openbare ruimte

Prof. mr. dr. ir. Bart Custers

1. Inleiding _____ 11

2. Context en terminologie _____ 14

3. Hergebruik van gegevens _____ 17

4. Wettelijke kaders _____ 18

5. Ethisch perspectief - afwegingskaders _____ 27

6. Conclusies _____ 29

Essay 2: Digitalisering, gedragsbeïnvloeding en de overheid _____ 36

Het gladde ijs tussen nudging en overveillance

Prof. dr. Wijnand IJsselsteijn

1. Inleiding _____ 38

2. Digitale gedragsbeïnvloeding: persuasive technology _____ 40

3. Nudging: een duwtje in de goede richting? _____ 42

4. De kunstmatig intelligente overheid? _____ 47

5. Big Brother is Watching You: het spook van overveillance _____ 49

6. Naar participatoir ontwerp van de digitale openbare ruimte _____ 52

7. Digitale geletterdheid, mondige burgers en een eerlijke overheid _____ 54

8. Conclusie en aanbevelingen _____ 56

Essay 3: De bestuurlijke regie over experimentele data en algoritmen _____ 61

Aanbevelingen voor een nationaal algoritmeregister en rapportagekader

Prof. dr. Gerd Kortuem

1. Inleiding _____ 63

2. Data en algoritmen in de samenleving _____ 64

3. Algoritmen als experimentele technologie _____ 69

4. Bestuurlijke modellen voor experimenten met algoritmen _____ 77

5. Aanbevelingen _____ 82

Essay 4: De noodzaak van een digitale omgevingsvisie _____ 95

Van principes naar praktijk

Prof. dr. Liesbet van Zoonen en dr. Jiska Engelbert

1. Inleiding _____ 97

2. Een weids landschap aan principes _____ 97

3. De opgave voor de verschillende overheden _____ 101

4. Digitale omgevingsvisie _____ 103

Essay 5: AI in de digitale samenleving _____ 108

Kwaliteit van algoritmen en besluitvorming

Prof. dr. Eric Postma

1. Inleiding _____ 110

2. Deterministische algoritmen versus probabilistische algoritmen _____ 111

3. Machinelearning-algoritmen _____ 112

4. De AI-revolutie _____ 115

5. Observaties _____ 120

6. Beleidsvoorstellen _____ 125



Inhoudsopgave

2/2

Essay 6: Data makes the world go round _____ 128

*Voorstel voor onderzoek naar drie instrumenten ter versterking
van de (digitale) autonomie*

Theo Veltman en Rob van Kranenburg

1. Data, eeuwenoude bron van kennis, macht en geld _____ 130
2. De digitale platforms rijk en dominant, riskant voor (digitale)
autonomie _____ 135
3. Toenemende druk op (digitale) autonomie, een 'tipping point' 136
4. (Inter)nationale initiatieven voor houvast (digitale) autonomie 138
5. Wat nodig is voor versterking (digitale) autonomie _____ 139
6. Conclusie _____ 141

Overzicht beleidsaanbevelingen uit essays _____ 147

Epiloog: Beleidsuitdagingen bij datagebruik _____ 150

Voorstel voor beleidsmaatregelen en acties _____ 162

Colofon _____ 165



Voorwoord



Er wordt in Nederland volop geëxperimenteerd met het inzetten van data en digitale technologieën om problemen in de openbare ruimte op te lossen. Door deze experimenten komen verschillende vragen naar boven, bij de overheid en bij inwoners. Van wie zijn de data die in de publieke ruimte verzameld worden? Hoe kun je als inwoner of bezoeker meer zeggenschap krijgen over data en technologie? Hoe kunnen we als overheid data en technologie gebruiken en tegelijkertijd maatschappelijke waarden waarborgen?

Uit deze vragen spreekt een noodzaak om met elkaar, als overheid en inwoner, te komen tot spelregels voor een maatschappelijk verantwoorde inzet van data en technologie in de openbare ruimte. Daarom is in NL DIGITAAL: Data Agenda Overheid het traject *Behoorlijk datagebruik in de openbare ruimte* opgenomen. Met als doel: mogelijkheden verkennen om te kunnen innoveren, stappen zetten in het verantwoord gebruik van data in de openbare ruimte en het gesprek aangaan met elkaar, want dit is een dialoog die we steeds met elkaar moeten blijven voeren. Ik vind het belangrijk dat iedereen aan deze dialoog mee kan doen en niemand wordt buitengesloten.

In 2018 woonde ik in Eindhoven de eerste dialoogtafel over behoorlijk datagebruik in de openbare ruimte bij. De vragen van inwoners aan deze tafel troffen mij. “Hoe voorkomen we dat op basis van data gediscrimineerd wordt?” en “Hoe zit het eigenlijk met de verhouding tussen veiligheid en privacy?”. Terechte vragen, waar we direct mee aan de slag zijn gegaan.

Van publieke dialoogtafels via wetenschappelijke essays naar voorstellen voor nieuw beleid. Met praktijkvoorbeelden die als inspiratie dienen. Deze bundel, en meer nog de weg ernaartoe, is een combinatie van degelijk en innovatief. Ik ben ervan overtuigd dat deze manier van werken nodig is in deze tijd, waarin we geconfronteerd worden met razendsnelle technologische ontwikkelingen. Tegelijkertijd hebben we als overheid instrumenten tot onze beschikking, zoals beleid en wetgeving, waarbij zorgvuldigheid voorop staat in plaats van snelheid. Tussen deze werelden slaat dit traject een brug. Om met elkaar in gesprek te blijven over wat er kan, wat er mag, en wat geoorloofd is als het gaat om datagebruik in de openbare ruimte.

Raymond Knops

Minister van Binnenlandse Zaken en Koninkrijksrelaties

Inleiding

In maart 2019 heeft het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) NL DIGITAAL: Data Agenda Overheid aan de Tweede Kamer aangeboden. Deze agenda gaat over datagebruik in de samenleving en richt zich met name op goed en verantwoord datagebruik door de overheid. Het ministerie van BZK heeft zich geëngageerd aan deze agenda en is van plan om de resultaten van de in de agenda beschreven acties verder te brengen. In Data Agenda Overheid zijn voor de komende drie jaar verschillende actielijnen uitgezet. Eén van de actielijnen is het project 'Behoorlijk datagebruik in de openbare ruimte', dat past bij het huidige debat over publieke waarden en het gebruik van data in de openbare ruimte.

Het doel van dit project is om te komen tot een gezamenlijk beeld van behoorlijk datagebruik in de openbare ruimte en bijbehorend handelingsperspectief. Het middel hiervoor is een reeks thematische essays met concrete beleidsaanbevelingen¹. Dit project bouwt voort op de succesvol afgeronde proeftuin² waarmee een inventarisatie is gemaakt van knelpunten rond behoorlijk datagebruik in de openbare ruimte. Deze knelpunteninventarisatie heeft als basis gediend voor de thematische essays. De knelpunten zijn onder andere naar voren gebracht tijdens dialogotafels met inwoners. De geïdentificeerde knelpunten zijn:

1. *De spanning tussen privacy en techniek; wat kan, wat mag, wat moet je (niet) willen.*
2. *A. De spanning tussen bedrijven en burgers; de belangen van private commerciële partijen en B. De spanning tussen het individu en het collectieve, publieke belang.*
3. *De mate van transparantie en uitlegbaarheid van datagebruik en dataverzameling in de publieke ruimte.*
4. *Doelbinding versus hergebruik; de gewenste mate van doelmatigheid.*
5. *Publiek-private samenwerking.*
6. *Wat is de rol van de overheid binnen de smart society?*

In het proces om tot de essays te komen is nadrukkelijk de verbinding gezocht tussen wetenschap en de praktijk van het openbaar bestuur. Er is een begeleidingscommissie gevormd met leden vanuit de wetenschap en met ervaringsdeskundigen die werkzaam zijn bij verschillende overheidsorganisaties. Deze commissie heeft de hierboven beschreven knelpunten erkend en herkend en een belangrijke rol gespeeld bij de totstandkoming van de thematische essays. Ze heeft als opdrachtgever opgetreden door 'trekkers' van essays te benoemen. Deze trekkers hebben op hun beurt een beroep gedaan op hun netwerk om auteurs voor de essays te benaderen. Uiteindelijk heeft de commissie de keuze voor auteurs geëscordeerd, waarna het schrijven van de essays kon beginnen. Binnen het schrijfproces hebben de auteurs zoveel mogelijk ruimte gekregen om hun denkbeelden weer te geven op grond van hun kennis, ervaring en inzichten. Hoewel wetenschappers altijd verbonden zijn aan een universiteit, is die universiteit natuurlijk niet gebonden aan hun standpunt. Voor de overige auteurs betekent dit dat de essays op persoonlijke titel zijn geschreven. De commissie is enkele keren bij elkaar gekomen om de voortgang van het project en de essays te bespreken en de tussenresultaten van feedback te voorzien.

Het zestal essays, geschreven vanuit de wetenschap en de praktijk, heeft uiteindelijk geleid tot 24 beleidsaanbevelingen. De laatste fase van het project betrof de doorvertaling van deze beleidsaanbevelingen naar concrete voorstellen voor beleidsmaatregelen en acties³. Deze beleidsmaatregelen en acties zijn van een hoog abstractieniveau omdat er rekening gehouden moet worden met verschillende contextfactoren, waaronder politiek klimaat en (bestuurlijk) draagvlak. Om een en ander ten opzichte van elkaar te positioneren, heeft het Instituut Bestuurskunde van de Universiteit Leiden een kader ontwikkeld (als onderdeel van een epiloog bij de zes essays). Dit kader is gebruikt bij het proces van doorvertaling en geeft enkele suggesties voor beleidsontwikkeling in relatie tot datagebruik bij de overheid.

Het is belangrijk om te vermelden dat de beleidsaanbevelingen en de voorstellen voor beleidsmaatregelen en acties de uitdagingen van het databeleid in Nederland niet volledig afdekken. Wel dekken ze de geconstateerde knelpunten af. Daarmee kan de inhoud van deze essaybundel gezien worden als een eerste aanzet, en hopelijk een inspiratiebron, om de uitdagingen aan te gaan die voortkomen uit datagebruik in de openbare ruimte.

Leeswijzer:

Het eerste deel van deze essaybundel bestaat uit zes thematische essays (hoofdstuk 1 t/m 6). Alle essays worden ingeleid door een voorwoord dat geschreven is door de trekker van het desbetreffende essay. Na de essays volgt een overzicht van alle beleidsaanbevelingen en de epiloog: beleidsuitdagingen bij datagebruik, geschreven door het Instituut Bestuurskunde van Universiteit Leiden. De bundel wordt afgesloten met de doorvertaling van de aanbevelingen van de essays naar concrete voorstellen voor beleidsmaatregelen en acties.

Referenties

- 1 Eén van de beleidsaanbevelingen wordt inmiddels al gerealiseerd in de vorm van een project in een gezamenlijke opdracht met VNG waarin wordt gewerkt aan een modelovereenkomst 'Slimme Toepassingen'.
- 2 Aanleiding voor deze proeftuin was de oproep van toenmalige wethouders Ollongren (Amsterdam) en Depla (Eindhoven) voor samenwerking rond datagebruik in de openbare ruimte. Zij boden hun 'living labs' omgeving aan als experimenteerruimte om nieuwe technologieën te leren begrijpen en samen de nieuwe spelregels op te bouwen.
- 3 De concrete voorstellen voor beleidsmaatregelen en acties zijn op een longlist geplaatst ter voorbereiding voor de Data Agenda Overheid 2020. Deze lijst zal ter beoordeling worden voorgelegd aan de klankbord- en stuurgroep Data Agenda Overheid.

Hergebruik van gegevens in smart cities

Juridische en ethische kaders voor big data in de openbare ruimte

Prof. mr. dr. ir. Bart Custers
Hoogleraar Law and Science
Universiteit Leiden



Voorwoord

Door prof. mr. dr. Sofia Ranchordas

Steden staan traditioneel bekend als anonieme ruimtes waar individuen kunnen leven en werken met weinig (sociale) controle. Niets is tegenwoordig minder waar in de zogeheten smart cities, waar dagelijks grote hoeveelheden gegevens worden verzameld. Terwijl een groot deel van de verzamelde gegevens betrekking heeft op bijvoorbeeld het verkeer of de luchtkwaliteit van de stad, worden ook regelmatig persoonsgegevens verzameld, zowel door private als door publieke partijen.

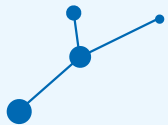
Deze gegevens worden tevens uitgewisseld en hergebruikt. In de laatste jaren zijn deze gegevens hergebruikt om bijvoorbeeld burgers te 'nudgen' om duurzamer of gezonder te zijn. Dit (her)gebruik van data is vaak gemotiveerd aan de hand van technocratische redeneringen. Toch blijft het hergebruik van (persoons)gegevens problematisch om juridische en ethische redenen. Kan het (her)gebruik van data als behoorlijk worden beschouwd als de autonomie van burgers onder druk wordt gezet?

Hebben lokale overheden concreet zicht op het (her)gebruik van data? Is het hergebruik van data in smart cities altijd noodzakelijk om goede publieke diensten te leveren? Hoe worden publieke waarden gewaarborgd als gegevens worden hergebruikt door private partijen?

In het essay van Bart Custers worden de juridische en ethische kaders van het (her)gebruik van gegevens in smart cities geanalyseerd. Custers, hoogleraar Law & Data Science (Universiteit Leiden), is een vooraanstaande expert op het gebied van privacy en cybercrime. Custers doet belangrijke aanbevelingen voor lokale overheden met het oog op een behoorlijk hergebruik van gegevens in smart cities: lokale overheden moeten goed op de hoogte zijn van welke gegevens zijn verzameld op een rechtmatige wijze, burgers actief betrekken bij digitale initiatieven en vaker gebruik maken van pilots.

Als rechtswetenschapper op het gebied van publiekrecht en digitale technologie, vond ik het bijzonder verrijkend om deel te mogen nemen aan de begeleidingscommissie en te reflecteren op zowel de kansen als risico's van gegevensverzameling en datahergebruik in smart cities.

Prof. mr. dr. Sofia Ranchordas
Adjunct-hoogleraar Rechtsgeleerdheid, Rijksuniversiteit Groningen
Lid begeleidingscommissie



Samenvatting

Slimme steden (*smart cities*) zijn stedelijke gebieden waar met gebruik van sensoren grote hoeveelheden gegevens worden verzameld om uiteenlopende processen in deze steden soepel te laten verlopen. Het gebruik van gegevens is juridisch en ethisch echter alleen toegestaan wanneer de gegevens op een behoorlijke manier worden verzameld en verwerkt. Voor veel steden is niet duidelijk welke (persoons)gegevens over burgers mogen worden verzameld en verwerkt en onder welke voorwaarden. In dit essay staat centraal de vraag: *in hoeverre mogen gegevens van burgers worden hergebruikt in de context van smart cities?* De nadruk ligt daarbij op het hergebruiken van gegevens.

Aan bod komt wat wordt verstaan onder smart cities, *Internet of Things*, *big data* en *nudging*. Verder worden verschillende vormen van hergebruik van gegevens onderscheiden via een typologie die bijdraagt aan de verdere duiding van hergebruik en helpt bij het beoordelen van de wenselijkheid van hergebruik van gegevens. De kern van dit essay is onderzoek naar de meest relevante juridische en ethische kaders voor hergebruik van gegevens.

De meest relevante juridische kaders zijn respectievelijk privacy en mensenrechten, bescherming van persoonsgegevens en bestuursrecht (in het bijzonder de algemene beginselen van behoorlijk bestuur). De meest relevante ethische kaders zijn deontologie, utilitarisme en deugdenethiek. De ethische perspectieven bieden afwegingskaders die kunnen worden gehanteerd binnen de juridische kaders, bijvoorbeeld voor het opstellen van gedragscodes en andere vormen van zelfregulering. Wanneer de juridische en ethische kaders in dit essay worden aangehouden, is er zeer waarschijnlijk sprake van behoorlijk (her)gebruik van gegevens. Wanneer deze kaders niet worden aangehouden, is er geen sprake van behoorlijk (her)gebruik.

Op basis van deze conclusies worden vier aanbevelingen gedaan. Gemeentelijke overheden moeten bij smart cities inzetten op behoorlijk hergebruik van gegevens via publiek-private samenwerkingen, burgers actief betrekken bij afwegingen, zorgen voor transparantie over gegevens en afwegingen, en via pilots smart cities geleidelijk verder ontwikkelen.



1. Inleiding

Door snelle opeenvolgende technologische ontwikkelingen is het mogelijk geworden steeds meer gegevens te verzamelen over burgers in de openbare ruimte. Via allerlei apparaten die zijn aangesloten op het *Internet of Things* (zoals auto's, camera's, vuilcontainers, elektriciteitsmeters en huishoudelijke apparatuur) kunnen datastromen worden gegenereerd. Door handig gebruik te maken van de beschikbare gegevens, worden steden steeds slimmer en duurzamer, ook in Nederland. Camera's en sensoren die agressie herkennen of met voorbijgangers communiceren kunnen bijvoorbeeld helpen om de openbare ruimte veiliger te maken. Via chips in vuilcontainers die aangeven of ze vol zijn, kan het ophalen van huisvuil worden geoptimaliseerd. Door middel van inductielussen in het wegdek en navigatiesystemen kunnen verkeersstromen *realtime* worden aangepast. Dit alles is mogelijk door enorme hoeveelheden data uit verschillende bronnen die realtime beschikbaar zijn (ook wel big data genoemd) en geavanceerde analysemethoden, zoals *machine learning*, *data mining* en kunstmatige intelligentie (*Artificial Intelligence*, AI) die geautomatiseerd patronen herkennen en soms zelfstandig beslissingen nemen. Steden die op deze manier slim gebruik maken van big data worden ook wel *smart cities* genoemd, of, in goed Nederlands, slimme steden. Het gebruik van gegevens kan steden niet alleen duurzamer, innovatiever en leefbaarder maken, maar draagt ook bij aan het vertrouwen van burgers in overheden en aan burgerparticipatie in de samenleving.¹

Binnen en buiten Nederland zijn de afgelopen jaren slimme steden steeds belangrijker geworden. Dat is niet verwonderlijk, want meer dan de helft van de wereldbevolking woont in stedelijke gebieden, zoals Tokyo, Delhi en Mexico City.² Daar komt nog bij dat de meeste van deze steden enorm groeien, zowel door bevolkingsgroei als toenemende urbanisatie. In reactie op de problemen die daaruit voortkomen, zoals groeiende ongelijkheid, klimaatverandering en onveiligheid, zoeken steden naar technologische oplossingen.³ Ook in Nederland experimenteren steden met vergaande inzet van datagedreven oplossingen. In Den Haag worden slimme lantaarnpalen ingezet. Die kunnen niet alleen licht geven, maar ook worden gebruikt voor het meten van luchtkwaliteit, geluids-overlast en verkeersdrukke.⁴ In Eindhoven wordt onderzocht of verlichting kan worden gebruikt om hulpdiensten naar ongevallen te geleiden en of de verlichting kan worden aangepast aan weersomstandigheden en de drukke op straat.⁵

In Groningen is het zo druk met fietsers dat wordt getracht verkeersstromen te monitoren en oplossingen in kaart te brengen, zoals het omleiden van verkeersstromen.⁶

In Nederland experimenteren steden met vergaande inzet van datagedreven oplossingen

Het gebruik van gegevens is vanuit juridisch en ethisch perspectief echter alleen toegestaan als de gegevens ook op een behoorlijke manier worden verzameld en verwerkt. Vandaar dat het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties een commissie heeft ingesteld die zich richt op behoorlijk datagebruik in de openbare ruimte. In opdracht van deze commissie is een serie essays opgesteld over behoorlijk datagebruik in de openbare ruimte. In dit essay, dat deel uitmaakt van die serie, wordt met name ingegaan op juridische en ethische kaders voor het hergebruik van gegevens in de openbare ruimte.

Voor veel steden is het niet duidelijk welke (persoons)gegevens⁷ over burgers mogen worden verzameld en verwerkt, en onder welke voorwaarden. De digitalisering in steden roept daardoor juridische en ethische vragen op over wat toegestaan en wat wenselijk is. Juridische vragen hebben onder meer betrekking op privacy, eigendom van gegevens, het verwerken van gevoelige gegevens, inzagerechten en informatiebeveiliging. Ethische vragen hebben onder meer betrekking op vertrouwen, transparantie, commercialisering van gegevens, maatschappelijk draagvlak en nudging. Bij nudging (zie paragraaf 2.4) wordt informatie ingezet om mensen subtiel, soms zonder dat ze het zelf doorhebben, te stimuleren zich op een gewenste wijze te gedragen. Een ethische vraag hierbij is hoe ver de overheid (bij smart cities is dit doorgaans de gemeente) mag gaan in het sturen van gedrag van burgers, bijvoorbeeld om te zorgen dat zij gezondere en duurzamere keuzes maken. Goede bedoelingen kunnen in dit proces mogelijk omslaan in paternalisme en het inperken van de autonomie en vrijheden van burgers. Hoewel er bij nudging sprake is van keuzevrijheid, is er tegelijkertijd sprake van manipulatie van gedrag en/of keuzes.⁸

De juridische en ethische kaders voor deze vragen zijn soms lastig toepasbaar, soms ingewikkeld en soms zelfs geheel afwezig. In dit essay wordt nagegaan welke juridische en ethische kaders kunnen worden gebruikt bij het maken van deze afwegingen. De nadruk ligt hierbij op kansen: onderzocht wordt hoe de juridische en ethische kaders ingezet kunnen worden om kansen van big data in smart cities goed te kunnen benutten, er tegelijkertijd voor zorgend dat publieke waarden en grondrechten (zoals privacy, transparantie en autonomie van burgers) niet onder druk komen te staan. Dat het gebruik van big data in smart cities volop kansen biedt, is helder. Maar er zijn vanuit juridisch en ethisch perspectief goede en minder goede manieren om deze kansen te benutten. Door goed rekening te houden met wensen van burgers zal het maatschappelijk draagvlak voor en het vertrouwen in het verzamelen en verwerken van gegevens toenemen. Wanneer burgers zich echter gemanipuleerd voelen in hun gedrag of keuzemogelijkheden, kan dat vertrouwen juist afnemen.

In dit essay zal de nadruk liggen op het hergebruiken van gegevens in slimme en duurzame steden. Bedrijven, of het nu technologiebedrijven zijn of niet, bouwen vaak hun eigen dataverzamelingen op waarmee ze bedrijfsprocessen kunnen optimaliseren en hun marktpositie kunnen verstevigen. Het speelveld van steden is echter wezenlijk anders, omdat er een veelheid aan actoren is die allemaal op hun eigen terrein gegevens verzamelen en verwerken. Dit complexe speelveld brengt met zich mee dat veel gegevens verspreid zijn over verschillende actoren. Om slim gebruik te kunnen maken van deze gegevens is het doorgaans nodig de gegevens met elkaar te delen of aan elkaar te koppelen. Zulke secundaire toepassingen van reeds verzamelde gegevens worden hergebruik van gegevens genoemd (data reuse). Het hergebruik van gegevens brengt vragen met zich mee over wanneer en onder welke voorwaarden doorgifte en zelfs verkoop van gegevens is toegestaan.

Wanneer al het bovenstaande bij elkaar wordt genomen, leidt dat tot de volgende vraag die centraal staat in dit essay: in hoeverre mogen gegevens van burgers worden hergebruikt in de context van slimme steden (smart cities)?

Het antwoord op deze vraag verschilt per situatie. Omdat het niet mogelijk is om in dit essay talloze situaties op te sommen, is ervoor gekozen om enerzijds de juridische kaders die van toepassing zijn te verduidelijken als het gaat om het hergebruik van gegevens in smart cities en anderzijds ethische afwegingskaders te

bieden waarmee actoren in smart cities zelf hun casus of situatie kunnen beoordelen.⁹ De doelgroep van dit essay bestaat uit bestuurders en beleidsmedewerkers van landelijke overheidsorganisaties en decentrale overheden, in het bijzonder gemeenten, en ondernemers. Zij kunnen de informatie en de afwegingskaders die in dit essay worden aangeboden gebruiken voor het zelfstandig beoordelen van wat toegestaan en wenselijk is als het gaat om hergebruik van gegevens. Niettemin is het advies om in alle gevallen waarin nieuwe vormen van gebruik van persoonsgegevens wordt overwogen binnen de eigen organisatie juridisch advies in te winnen bij de functionaris gegevensbescherming en keuzes te plannen en te toetsen bij alle relevante stakeholders.

Dit essay is als volgt opgebouwd. In paragraaf twee wordt verdere achtergrondinformatie geboden over de belangrijkste begrippen in dit essay. Aan bod komt onder meer wat wordt verstaan onder smart cities, Internet of Things, big data en nudging. In paragraaf drie wordt ingegaan op verschillende vormen van hergebruik van gegevens. Deze typologie draagt bij aan de verdere duiding van hergebruik en helpt bij het beoordelen van de wenselijkheid van hergebruik van gegevens. In paragraaf vier komen de meest relevante juridische kaders aan bod. Dit zijn respectievelijk privacy en mensenrechten, bescherming van persoonsgegevens en bestuursrecht (in het bijzonder de algemene beginselen van behoorlijk bestuur). In paragraaf vijf komt het ethisch perspectief aan bod. De drie hoofdstromen in de (westerse) ethiek bieden verschillende afwegingskaders die kunnen worden gehanteerd binnen de juridische kaders, bijvoorbeeld voor het opstellen van gedragscodes en andere vormen van zelfregulering. De ethische kaders die worden besproken zijn deontologie, utilitarisme en deugdenethiek. In paragraaf zes wordt de centrale vraagstelling van dit essay beantwoord en worden aanbevelingen gedaan voor vervolgstappen die kunnen worden genomen om hergebruik van gegevens op een juridisch en ethisch verantwoorde wijze in te zetten in de context van smart cities.





2. Context en terminologie

In deze paragraaf wordt verdere achtergrondinformatie geboden over de belangrijkste begrippen die worden gehanteerd in dit essay. Achtereenvolgens komen aan bod de termen smart cities (paragraaf 2.1), Internet of Things (paragraaf 2.2), big data (paragraaf 2.3) en nudging (paragraaf 2.4)

2.1 Smart cities

Smart cities zijn stedelijke gebieden waar met gebruik van sensoren grote hoeveelheden gegevens worden verzameld om uiteenlopende processen in deze steden soepel te laten verlopen.¹⁰ Het gaat dan bijvoorbeeld om het verzamelen van gegevens over verkeersstromen, energievoorziening, watervoorziening, afvalverwerkingssystemen, openbare orde en veiligheid en publieke voorzieningen zoals scholen, bibliotheken en ziekenhuizen. Door de beschikbare informatie realtime te analyseren en terug te koppelen in de informatievoorziening richting burgers en besturingssystemen kunnen beschikbare middelen en voorzieningen zo efficiënt en effectief mogelijk worden ingezet. Hieronder worden enkele voorbeelden beschreven om dit te verduidelijken.

Bij verkeersstromen is het doorgaans de bedoeling om weggebruikers zo snel mogelijk naar hun plaats van bestemming te geleiden. Meestal is de snelste route ook de kortste route. Echter, als er files staan op een bepaalde route, kan het zijn dat een andere route, waarbinnen je een klein beetje moet omrijden, sneller is. Op basis van camera's boven de wegen kan realtime worden gemonitord waar in de stad het verkeer vaststaat. Maar als bij een file op de snelweg de verkeersborden alle automobilisten adviseren een B-weg te nemen, ontstaat al snel ook een file op die B-weg. In dat geval kan het beter zijn om toch de snelweg te nemen of een ander alternatief te kiezen. Anders krijg je een soort selffulfilling prophecy:¹¹ als iedereen om de oorspronkelijke file heen wordt geleid, staan alsnog veel mensen in de file, maar dan op een andere plek. Het zou handiger zijn om bijvoorbeeld de helft van de weggebruikers om te leiden en de andere helft niet, zodat de beschikbare wegcapaciteit optimaal wordt gebruikt en de doorstroming van het verkeer maximaal is. Dat lukt niet goed met aanwijzingen die voor elke weggebruiker hetzelfde zijn. Maar via persoonlijke input op de navigatiesystemen van elke weggebruiker kan een variabel advies worden verstrekt, waardoor niet iedereen hetzelfde doet.

Een ander voorbeeld is het ophalen van afval in grote steden. In een klassiek systeem komt de vuilniswagen elke week langs elke straat om vuilcontainers op te halen. Door dit gestandaardiseerde proces komt het regelmatig voor dat burgers het ene moment een halfvolle vuilcontainer aan de straat zetten en het andere moment hun afval niet kwijt kunnen. Een vuilniswagen moet bovendien langs alle straten rijden terwijl dat soms overbodig is, hetgeen een verspilling van middelen is en een belasting voor het milieu. In een slim systeem wordt elke vuilcontainer uitgerust met een chip die aangeeft of de vuilcontainer vol is. Het systeem kan dan berekenen waar de vuilniswagen die dag langs moet rijden en hiervoor een optimale route berekenen. Bezorgdiensten als Deliveroo gebruiken reeds algoritmen om de optimale routing voor hun fiets- of scooterkoeriers te berekenen. Bedrijven als IBM hebben vergelijkbare algoritmen ontwikkeld voor het routeren van afvalophaaldiensten. Via optimale routing kan aanzienlijk op transportkosten worden bespaard, hetgeen ook goed is voor het milieu, en kunnen producten en diensten nog beter op het juiste moment worden aangeleverd bij klanten.¹² De routing wordt berekend door algoritmen, niet alleen op basis van geografische gegevens en tijden, maar ook op basis van mogelijke opstoppingen, beschikbaarheid van alternatieve routes en plotselinge veranderingen in wensen van klanten of burgers.¹³

Bezorgdiensten als Deliveroo gebruiken reeds algoritmen om de optimale routing voor hun fiets- of scooterkoeriers te berekenen

2.2 Internet of Things (IoT)

Het verzamelen van gegevens voor toepassingen zoals hierboven beschreven, vindt plaats op verschillende manieren. In de eerste plaats genereren mensen zelf heel veel gegevens, bijvoorbeeld bij het invullen van formulieren, het gebruik van sociale media of wanneer ze informatie over zichzelf op het internet zetten. Daarnaast genereert technologie veel gegevens, bijvoorbeeld via sensoren als camera's en microfoons die in veel apparaten ingebouwd zitten. Ook via *trackers* worden veel gegevens verzameld, zoals offline via RFID-tags (bijvoorbeeld in toegangspasjes voor gebouwen, OV-chipkaarten of bankpassen voor contactloos

betalen) of online via cookies die online surfgedrag monitoren. Tal van apparaten, zoals mobiele telefoons en polsbanden die beweging en gezondheid bijhouden, genereren zeer gedetailleerde en persoonlijke informatie.

Zodra apparaten op het internet worden aangesloten, wordt gesproken van het Internet of Things (IoT).¹⁴ Voor de hand liggende voorbeelden van deze apparaten zijn computers, tablets, telefoons en navigatiesystemen. Maar ook minder voor de hand liggende apparaten kunnen aan het internet worden gekoppeld.

Denk bijvoorbeeld aan auto's (nu voor navigatie en monitoring van onderhoud, in de toekomst voor zelfsturende auto's), thermostaten en wasmachines (om op afstand aan te zetten),¹⁵ elektrische tandenborstels¹⁶ (om te adviseren over het juiste moment om tanden te poetsen) en speelgoed (om met poppen te kunnen communiceren, al werden die na ophef over privacy weer van de markt gehaald).¹⁷ In de context van smart cities worden sensoren bijvoorbeeld ingebouwd in de riolering om waterstanden te reguleren,¹⁸ in vuilcontainers om ze pas op te halen als ze vol zijn,¹⁹ en in lantaarnpalen om luchtkwaliteit, geluidsoverlast en verkeersdruk te meten.²⁰ Al die apparaten verbonden aan het internet creëren enorme datastromen, vaak realtime, die op slimme manieren aan elkaar kunnen worden gekoppeld om bepaalde processen in smart cities beter te laten verlopen.²¹



Een belangrijke eigenschap van het IoT is het feit dat apparaten vaak ook onderling communiceren (en dus niet alleen met mensen). Bijgevolg kunnen de apparaten onderling gegevens uitwisselen, zaken afstemmen en bepaalde beslissingen nemen, zoals keuzes in de hierboven beschreven voorbeelden: keuzes over waar en wanneer informatie wordt aangeboden of de keuze wat de optimale route is.

2.3 Big data

Hoewel er geen vaststaande, algemeen geaccepteerde definitie bestaat van big data, is de belangrijkste eigenschap van big data de enorme omvang van de hoeveelheden gegevens.²² Dat is uiteraard de reden waarom het big data wordt genoemd. Maar ook voor de omvang bestaat geen definitie. Het is niet zo dat er sprake moet zijn van een bepaalde (minimale) hoeveelheid gegevens om te kunnen spreken van big data. Wel is duidelijk dat het gaat om vele terabytes en petabytes met gegevens.²³

Big data is echter meer dan grote datasets. Doorgaans wordt de term big data niet alleen gebruikt om grote hoeveelheden gegevens aan te duiden, maar ook om aan te duiden dat er sprake is van een technologische ontwikkeling of zelfs versnelling. Daarmee zit het concept big data in dezelfde categorie als de industriële revolutie, halfgeleiders, nanotechnologie, robotica en het internet. Big data wordt vaak gezien als het resultaat van een gestage groei in de hoeveelheden data die beschikbaar zijn. Maar deze groei is in veel gevallen exponentieel. Neem bijvoorbeeld het hierboven beschreven Internet of Things. Wanneer een gebruiker met drie apparaten communiceert (data uitwisselt), zijn er drie datastromen. Maar wanneer de apparaten ook onderling met elkaar communiceren, zijn er opeens zes datastromen.

Daarnaast is er sprake van zogeheten tipping points.²⁴ Kleine hoeveelheden data kunnen wellicht al bepaalde verbanden blootleggen, maar grote hoeveelheden data kunnen exponentieel meer verbanden laten zien (mits aanwezig uiteraard). Waar kleine aantallen een beperking vormen, bijvoorbeeld bij zeldzame fenomenen als ziektebeelden met een lage prevalentie in afzonderlijke regio's of landen, kunnen grote getallen geaggregeerd over veel landen toch nieuwe inzichten geven. Big data betreft niet alleen grote datasets, maar ook nieuwe concepten - veelal wordt pas gesproken van big data als aan meerdere van de zogeheten 3 V's is voldaan:²⁵

- **Volume:** Big data gaat om grote hoeveelheden data, er worden geen steekproeven genomen).
- **Velocity (snelheid):** Big data worden doorgaans realtime verwerkt en betreffen soms streaming data die niet worden vastgelegd of opgeslagen.
- **Variety (verscheidenheid):** Big data kan ongestructureerde data betreffen in verschillende vormen, zoals tekst, cijfers, beelden (fotobeelden en camerabeelden) en geluid.

De eerste twee eigenschappen van big data zijn hierboven al aan bod gekomen. De derde eigenschap, variety/verscheidenheid, brengt nog andere uitdagingen met zich mee. Waar gewone (relationele) databanken vaak nog doorzoekbaar zijn met *queries* (zoekwoorden en eenvoudige bevragingen) en *syntaxes* (gestructureerde doorzoekingen), is big data vaak afkomstig uit diverse bronnen die niet altijd eenvoudig koppelbaar zijn. Dat maakt dat bigdata-analyses vaak lastiger zijn dan data-analyses in relationele databanken.

In veel gevallen moeten dus eerst oplossingen worden gevonden om gegevens te combineren.²⁶ Daarnaast zijn steeds meer tools beschikbaar om andersoortige databronnen te analyseren. Denk bijvoorbeeld aan *text mining* voor ongestructureerde tekstdocumenten en gezichtsherkenning voor foto's en videobeelden.²⁷

2.4 Nudging

Met de term *nudging* wordt bedoeld op vormen van subtiele gedragsbeïnvloeding, waarbij het doel is mensen zich te laten gedragen op een bepaalde, gewenste manier. Door keuzes op een bepaalde manier te framen en door de wijze van informatie aanbieden, kan gedrag worden beïnvloed.²⁸ Nudging is daarmee in feite een duwtje geven in de goede richting door het gewenste gedrag aantrekkelijk te maken, zonder mensen daarbij in hun vrijheden te beperken. Het doel is gedrag (licht) bij te sturen via een keuzearchitectuur. Typische kenmerken hiervan zijn dat mensen altijd keuzevrijheid wordt geboden (afwijken van het gewenste gedrag is mogelijk) en dat het gewenste gedrag aantrekkelijk wordt gemaakt. Nudging is vaak mogelijk omdat mensen zich vaak gedragen volgens bepaalde automatismen (ze volgen bijvoorbeeld de massa) of bepaalde vooronderstellingen hebben (ze denken bijvoorbeeld: 'duurdere producten zullen wel beter zijn') en/of over gebrekkige informatie beschikken ('kennelijk zijn dit mijn keuzemogelijkheden').

Typische voorbeelden van nudging zijn te vinden in de supermarkt. Groente en fruit liggen daar doorgaans vlakbij de ingang, omdat dit er uitnodigend (want vers, gezond en kleurrijk) uitziet. Daarnaast liggen A-merken doorgaans op ooghoogte, terwijl goedkopere producten vaak helemaal onderin of bovenin de schappen liggen.

Nudging is mogelijk omdat mensen zich vaak gedragen volgens bepaalde automatismen

Ook online wordt nudging vaak toegepast, bijvoorbeeld door meerdere opties aan te bieden maar een van de opties extra aan te prijzen als 'meest gekozen', 'nieuw', of 'nu in de aanbieding'. Veel mensen kiezen uit een reeks opties niet de goedkoopste of de duurste variant, maar bijvoorbeeld de op-een-na goedkoopste optie. In alle gevallen blijft er ruimte voor het maken van keuzes, maar wordt de gewenste gedraging aantrekkelijk gemaakt, soms visueel en soms fysiek (door niet te hoeven bukken of reiken voor producten in de schappen).²⁹ Uiteraard kunnen al deze vormen van nudging ook worden toegepast in slimme steden om burgers aan te zetten tot bepaald, wenselijk gedrag.³⁰ Wenselijk gedrag kan bijvoorbeeld betrekking hebben op een schone omgeving (een typisch voorbeeld van zo'n toepassing is de inzet van Holle Bolle Gijs in pretpark de Efteling, waar de prullenbak aantrekkelijk mee wordt gemaakt) of gezond gedrag (zoals het aanmoedigen van wandelen door aantrekkelijke stadsparken en autoluwe zones). Het verbieden van auto's in de binnenstad of het bestraffen van afval op straat gooien zijn geen voorbeelden van nudging, omdat ze geen keuzeruimte laten en een negatieve benadering hebben.

Nudging kan aantrekkelijk zijn als vorm van gedragsbeïnvloeding, maar stuit ook op bezwaren. Soms worden mensen enigszins ongemerkt gemanipuleerd, hetgeen spanning kan opleveren met autonomie³¹ en transparantie.³² Ook kan nudging overgaan in paternalisme. Immers, zodra de overheid (bij smart cities is dit doorgaans de gemeente) gaat bepalen wat wenselijk gedrag is en daarop actief gaat sturen, heeft dat invloed op de vrijheden van burgers.

Bij nudging zijn er weliswaar altijd keuzemogelijkheden, maar degenen die nudging toepassen bepalen veelal uit welke opties kan worden gekozen. Verder is de vraag of nudging alleen aanzet tot gedragsveranderingen op korte termijn of ook tot mentaliteitsveranderingen op lange termijn.

3. Hergebruik van gegevens

Er zijn verschillende manieren waarop gegevens kunnen worden gebruikt en hergebruikt. Dit kan van belang zijn voor de wettelijke kaders en ethische afwegingen. Daarom is het relevant om deze verschillen helder te hebben. Allereerst kan slechts sprake zijn van hergebruik (secundair gebruik) van gegevens na (primair) gebruik van gegevens. Gebaseerd op eerder onderzoek zetten we hieronder een taxonomie voor het hergebruik van gegevens uiteen.³³

Gegevens verzamelen en verwerken draait altijd om een specifiek doel in een specifieke context. Deze opvatting over het gebruik van gegevens is min of meer hetzelfde voor de verschillende perspectieven waarmee je naar gegevens kunt kijken, of dat perspectief nu juridisch, organisatorisch of technologisch is. Echter, er treden wel verschillen tussen de perspectieven op wanneer wordt gekeken naar de interpretatie van gegevensgebruik. Vanuit technologisch perspectief, bijvoorbeeld, wordt het verzamelen en opslaan van gegevens gezien als iets dat plaatsvindt alvorens gegevens kunnen worden verwerkt. Het vernietigen van gegevens wordt vanuit technologisch perspectief evenmin gezien als het verwerken van gegevens – het wordt gezien als iets dat plaatsvindt nadat gegevens zijn gebruikt.

Vanuit juridisch perspectief ligt dit heel anders. In de Algemene Verordening Gegevensbescherming (AVG, zie paragraaf 4.2) worden niet expliciet de termen gebruik en hergebruik van gegevens gehanteerd, maar het concept ‘verwerken van (persoons)gegevens’. Dit is gedefinieerd in artikel 4 als een “bewerking [...] zoals het verzamelen, vastleggen, ordenen, structureren, opslaan, bijwerken of wijzigen, opvragen, raadplegen, gebruiken, verstrekken door middel van doorzending, verspreiden of op andere wijze ter beschikking stellen, aligneren of combineren, afschermen, wissen of vernietigen van gegevens”. Met andere woorden, het verwerken van gegevens wordt in juridisch opzicht veel breder gezien dan in technologisch opzicht, omdat het ook het verzamelen, opslaan, wissen en vernietigen van gegevens betreft.

Er kunnen drie verschillende vormen van hergebruik van gegevens worden onderscheiden, afhankelijk van de vraag of het doel en de context bij het hergebruik afwijken van het doel en de context bij het oorspronkelijke gebruik van de gegevens:

- **Data recycling:** gegevens worden hergebruikt voor hetzelfde doel in dezelfde context.
- **Data repurposing:** gegevens worden hergebruikt voor een ander doel dan waarvoor ze oorspronkelijk zijn verzameld, maar het hergebruik vindt wel plaats in dezelfde context.
- **Data recontextualisation:** de gegevens worden hergebruikt in een nieuwe context, anders dan de context waarin ze oorspronkelijk zijn verzameld.

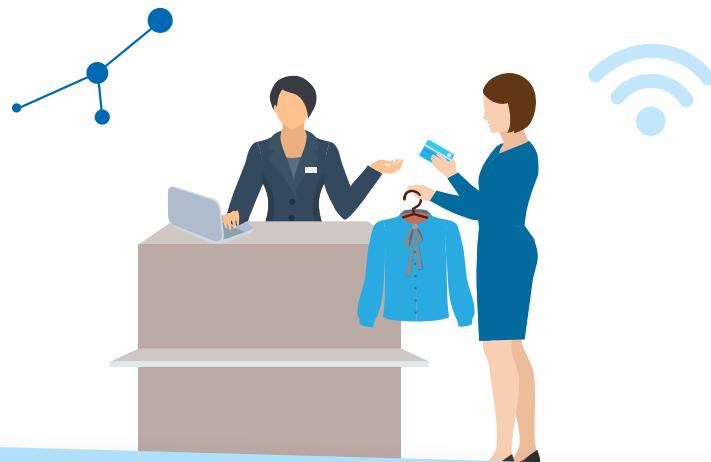
Deze verschillende vormen van hergebruik van gegevens kunnen eenvoudig worden geïllustreerd aan de hand van een voorbeeld. Veel online winkels vragen persoonsgegevens van hun klanten wanneer die klanten iets bestellen. Het gaat bijvoorbeeld om naam, adresgegevens (voor de bezorging) en financiële gegevens (voor de betaling). Wanneer een klant de volgende keer iets wil bestellen, kan hij of zij inloggen en komen de gegevens automatisch weer tevoorschijn. Dit is data recycling, omdat de gegevens worden gebruikt voor hetzelfde doel, in dezelfde context, maar niet voor dezelfde gelegenheid. Wanneer de webshop deze gegevens gaat gebruiken om klantprofielen op te bouwen, bijvoorbeeld om de kredietwaardigheid van klanten te beoordelen of om aanbiedingen te personaliseren, is er sprake van hergebruik voor doelen voorbij de oorspronkelijke doelen. Dan gaat het om *data repurposing*. Doelbinding is een sterk beginsel in de privacywetgeving (zie paragraaf 4.2) om zogeheten *function creep* te voorkomen. Function creep is in deze context het fenomeen dat gegevens in eerste instantie voor een specifiek doel worden verzameld, maar later ook voor andere, nieuwe doelen worden gebruikt. Deze vorm van hergebruik kan legaal zijn, bijvoorbeeld wanneer klanten hiervoor toestemming hebben gegeven. Wanneer dezelfde webshop de verzamelde gegevens gaat doorverkopen aan bijvoorbeeld marketingbedrijven, kunnen die gegevens worden hergebruikt in (soms totaal) andere contexten. Dan is er sprake van *data recontextualisation*.

In paragraaf 4.2 zal nader worden ingegaan op de vraag welke vormen van hergebruik van gegevens legaal zijn en onder welke voorwaarden. Maar los van deze juridische vraag is het ook van belang om vanuit ethisch perspectief naar

hergebruik van gegevens te kijken. Burgers hebben namelijk bepaalde verwachtingen omtrent wat er met hun gegevens gebeurt. Hoewel bepaalde vormen van hergebruik van gegevens vanuit juridisch perspectief weliswaar toegestaan kunnen zijn, kan het bij die vormen van hergebruik toch gebeuren dat burgers vertrouwen verliezen in bedrijven en overheden (in de context van smart cities is dit de gemeente). Dit vertrouwensverlies kan optreden wanneer zij onverwachte dingen doen. In dat kader is het van groot belang, zeker voor overheden, om transparant te zijn over het gebruik van gegevens, omdat anders het maatschappelijk draagvlak voor (her)gebruik van gegevens in smart cities kan afnemen of verdwijnen.

4. Wettelijke kaders

In deze paragraaf worden de belangrijkste wettelijke kaders aangereikt voor het (her)gebruik van gegevens in smart cities. Omdat het buiten de reikwijdte van dit essay valt om uitgebreid alle juridische kaders in detail te beschrijven, wordt hieronder vooral de nadruk gelegd op de beginselen die in deze wettelijke kaders worden gehanteerd. Paragraaf 4.1 zal ingaan op privacy en andere grondrechten, paragraaf 4.2 zal ingaan op de bescherming van persoonsgegevens en paragraaf 4.3 zal ingaan op de algemene beginselen van behoorlijk bestuur.³⁴



4.1 Privacy en andere grondrechten

Als het gaat om het verwerken van gegevens wordt vaak als eerste gesproken over de vraag of dat is toegestaan in verband met de privacy. Privacy is een belangrijk grondrecht in dit kader, maar zeker niet het enige grondrecht dat op het spel kan staan. Hieronder worden verschillende grondrechten die van belang kunnen zijn bij het verwerken van gegevens op een rij gezet.

Privacy

Het verwerken van grote hoeveelheden gegevens kan leiden tot situaties waarin gegevensbeheerders heel veel weten over eigenschappen, gedragingen en verblijfplaatsen van mensen.³⁵ Om die reden wordt privacy doorgaans het eerste genoemd als probleem of punt van zorg in de context van big data.³⁶ In de context van smart cities is dit niet anders.³⁷ Vanuit een grondrechtenperspectief hangen schendingen van privacy voor een groot deel af van de (redelijke) verwachtingen die mensen mogen hebben omtrent hun privacy. Tot op zekere hoogte is dit subjectief en kan dit afhangen van de situatie, de persoon en culturele omstandigheden. De een heeft niet de behoefte 's avonds de gordijnen te sluiten, terwijl de ander dat juist zorgvuldig doet. Sommige mensen zetten tot in detail hun persoonlijk leven op sociale media, terwijl andere niet eens een account aanmaken. Verwachtingen omtrent privacy kunnen wel worden geobjectiveerd, door na te gaan wat een gemiddeld persoon zou verwachten in een bepaalde situatie en context. Dit wordt aangeduid met 'redelijke verwachtingen van privacy'.³⁸

In de context van big data ligt de nadruk vaak op zogeheten informatiele privacy, een term die gericht is op de vraag welke persoonsgegevens worden verzameld en gebruikt en voor welke doelen (zie paragraaf 4.2). In wezen gaan verwachtingen omtrent informatiele privacy vooral over het delen, onthullen en gebruiken van gegevens op manieren die de personen die het betreft (de data-subjecten/betrokkenen) niet op prijs stellen. Zulk gebruik van persoonsgegevens is soms gerelateerd aan informatiebeveiligingsproblemen, bijvoorbeeld wanneer gegevens worden gehackt (zie het voorbeeld van de datingsite voor vreemdgangers Ashley Madison)³⁹ of wanneer gegevens lekken (soms opzettelijk, soms door onoplettendheid). Het ongewenst gebruik van gegevens kan onder andere het gevolg zijn van een gebrek aan transparantie en van *function creep*, ofwel wanneer gegevens worden gebruikt voor nieuwe doelen en/of in een nieuwe context.

De indeling voor verschillende typen hergebruik van gegevens in paragraaf 3 is daarbij relevant. Hergebruik van gegevens voor dezelfde doelen in dezelfde context (*data recycling*) zullen binnen de redelijke verwachtingen van privacy vallen, maar zodra het gaat om andere doelen (*data repurposing*) en/of een andere context (*data recontextualisation*), is het onwaarschijnlijk dat aan deze eis zal worden voldaan. Het feit dat gegevens reeds publiekelijk beschikbaar zijn, bijvoorbeeld omdat ze op internet staan, betekent niet dat ze vogelvrij zijn en voor willekeurig welk doel mogen worden gebruikt.⁴⁰

Ongewenste onthullingen kunnen ook ontstaan door andere oorzaken dan datalekken. Zelfs wanneer betrokkenen hun gegevens met niemand anders delen, kunnen deze gegevens in een context van big data toch worden blootgelegd. Dat is omdat big data de mogelijkheid biedt om op basis van grote hoeveelheden beschikbare gegevens ontbrekende gegevens te voorspellen. Dit geldt ook voor de voorspelling van gevoelige gegevens die mensen doorgaans (liever) niet delen met anderen, zoals gegevens over hun gezondheid, etnische afkomst, strafblad en seksuele voorkeuren. Gegevensbeheerders kunnen deze ontbrekende gegevens overigens betrekkelijk eenvoudig voorspellen.⁴¹ Ook eigenschappen die door de tijd heen variëren, zoals emoties en locaties, kunnen aan de hand van big data worden voorspeld, bijvoorbeeld aan de hand van berichten op sociale media of op grond van videobeelden.⁴² Als mensen bepaalde informatie over zichzelf niet willen delen en die informatie wordt via een omweg toch afgeleid, is duidelijk sprake van een privacyprobleem. Redelijke verwachtingen van privacy worden daarmee geschonden, hetgeen een inbreuk op het recht op privacy kan opleveren.

In bepaalde gevallen kunnen gegevensbeheerders zelfs meer weten over betrokkenen dan dat deze over zichzelf weten, waaronder levensverwachtingen, kansen op ernstige ziektes of auto-ongelukken, risico's op bepaalde verslavingen en inschattingen van welzijn en geluk. Een typisch voorbeeld is het uitdelen van zogeheten likes op Facebook. Met deze icoontjes kunnen gebruikers aangeven welke muziek, filmpjes, games, uitspraken, personen, etc. ze leuk vinden. Uit onderzoek blijkt dat op basis van enkele Facebook-likes tal van sensitieve persoonlijke attributen zeer nauwkeurig kunnen worden voorspeld.⁴³ Zo konden de onderzoekers bij Facebookgebruikers onder meer met hoge nauwkeurigheid voorspellingen doen over seksuele voorkeur, etniciteit, religie, politieke voorkeur, persoonlijkheidskenmerken, intelligentie, geluk, drugsgebruik en echtscheiding van

ouders. Zelfs als mensen zulke eigenschappen en kenmerken niet (willen) prijsgeven zijn ze dus te voorspellen op grond van andere gegevens die zijzelf (of anderen) wel prijsgeven.⁴⁴ Overigens kan op dezelfde manier bijvoorbeeld ook anonimisering ongedaan worden gemaakt.⁴⁵

Als mensen **bepaalde informatie over zichzelf niet willen delen** en die informatie **wordt via een omweg toch afgeleid**, is **duidelijk sprake van een privacyprobleem**

In de context van smart cities zijn er tal van gegevens die kunnen worden gecombineerd en zodoende anonimisering kunnen doorbreken en/of eigenschappen kunnen voorspellen. Alleen al op basis van postcode kunnen tientallen persoonsgegevens worden voorspeld, waaronder inkomen, gezins-samenstelling en opleidingsniveau. Waar mensen vroeger van een dorp waar iedereen elkaar kende konden verhuizen naar een stad, waar zij anoniemer konden leven (soms om een nieuw leven te beginnen, bijvoorbeeld na een schandaal of een strafblad), weten stadsbestuurders en bedrijven tegenwoordig misschien nog wel meer over haar burgers. Immers, de gegevens kunnen ook informatie onthullen die mensen niet eens over zichzelf weten, zoals hun levensverwachting. Omdat gegevens lang bewaard kunnen blijven, is beginnen met een schone lei in een digitale wereld lastig, al bestaat er inmiddels wel een recht op het laten wissen van gegevens.⁴⁶

Discriminatie, stigmatisatie en polarisatie

Wanneer de hierboven beschreven voorspellingen gevoelige gegevens (zoals etniciteit, politieke voorkeuren, seksuele voorkeuren, leeftijd en geslacht) blootleggen en deze vervolgens worden gebruikt om beslissingen over mensen te nemen, kan er ook sprake zijn van discriminatie. Discriminatie kan op verschillende niveaus voorkomen in bigdata-analyses. Ten eerste kunnen de gegevens al vooroordelen bevatten. Een typisch voorbeeld is wanneer de politie vooral surveilleert in bepaalde achterstandswijken waar veel migranten wonen.

Het waarschijnlijke resultaat van een dergelijk beleid is dat databanken van de politie gevuld worden met mensen met bepaalde etnische achtergronden. Dit is een typisch voorbeeld van selecte (in plaats van aselechte) steekproeven.⁴⁷

Wanneer dezelfde politieorganisatie vervolgens op zoek zou gaan naar verbanden in de verzamelde gegevens, teneinde uit te zoeken welke groepen hogere risico's op crimineel gedrag vertonen, zal het geen verbazing wekken dat dezelfde etnische minderheden geprofileerd kunnen worden als risicogroep voor crimineel gedrag. Omdat de gegevens al vooroordelen bevatten voordat de analyses werden uitgevoerd, is er sprake van een selffulfilling prophecy. Als de politie vervolgens de profielen gebruikt om te bepalen waar het meeste moet worden gesurveilleerd, is de cirkel rond. Door verder te surveilleren in deze wijken, komen nog meer gegevens over deze wijken in de databanken en worden de profielen verder bevestigd.

Het is van belang om op te merken dat discriminerende patronen niet altijd duidelijk zijn. Wanneer bijvoorbeeld gevoelige eigenschappen als etniciteit, al dan niet intentioneel, worden gebruikt voor profilering, is het duidelijk dat er een risico op discriminatie is. Echter, wanneer de profielen worden gebaseerd op bijvoorbeeld postcodes, kan dit ook (indirecte) discriminatie opleveren wanneer die postcodes sterk gerelateerd zijn aan etniciteit. In zulke gevallen zijn postcodes slechts een proxy, een benaderingsvariabele, voor etniciteit. In Nederland is overigens ook indirecte discriminatie verboden, doordat in art. 1 Algemene wet gelijke behandeling (waarin het discriminatieverbod in de Grondwet is uitgewerkt) direct en indirect onderscheid gelijk worden gesteld. Feit blijft dat indirect onderscheid veel lastiger is te bewijzen dan direct onderscheid. Er zijn wel technieken om discriminatie in big data op te sporen.⁴⁸

Uit onderzoek blijkt overigens dat het weglaten van gevoelige karakteristieken uit databanken niet verhindert dat er patronen kunnen worden gevonden die leiden tot indirecte discriminatie.⁴⁹ Los van het feit dat ontbrekende variabelen kunnen worden voorspeld (zie hierboven), blijkt dat veel patronen die worden blootgelegd ook te vinden zijn in andere variabelen dan de gevoelige variabelen waarop een discriminatieverbod rust.⁵⁰ Een voorbeeld hiervan is het gebruik van geografische gegevens voor profilering. Dit wordt aangeduid als *redlining* en is doorgaans verboden.⁵¹

Verder moet worden opgemerkt dat indirecte discriminatie ook niet-intentioneel kan plaatsvinden en dat gebruikers van profielen zich mogelijk van geen kwaad bewust zijn. Indien daarentegen de profielen juist worden gebruikt om discriminatie te verhullen, wordt ook wel van *masking* gesproken.

Als het gaat om de wetgeving die discriminatie op grond van specifieke eigenschappen verbiedt, gaat het enerzijds om lijsten met eigenschappen die niet mogen worden gebruikt voor beslissingen (zoals etniciteit, politieke voorkeuren, vakbondslidmaatschap, seksuele geaardheid, etc.) en anderzijds om bepaalde typen beslissingen die verboden zijn (zoals het aannemen en ontslaan van personeel, het aanbieden van producten en diensten, etc.). Niet alle beslissingen op grond van de opgesomde gevoelige karakteristieken zijn verboden. Het is bijvoorbeeld een persoonlijke keuze met wie iemand bevriend wil zijn. Niettemin kunnen zich ook bij de vorming van vriendschappen 'zwakkere' vormen van discriminatie voordoen, bijvoorbeeld in de vorm van stigmatisering van bepaalde bevolkingsgroepen. Op grotere schaal kan dat leiden tot sociale polarisatie en maatschappelijke segregatie.



Vrijheid van gedachten en meningsuiting, autonomie en zelfbeschikking

Wanneer iemand de eerste twee of drie letters van een zoekterm invoert in Google, worden die door de zoekmachine automatisch aangevuld tot een aantal suggesties. Deze suggesties komen enerzijds voort uit de grote aantallen zoekvragen die andere gebruikers eerder stelden en anderzijds uit iemands persoonlijke profiel, waarbij op basis van eerdere zoekvragen wordt ingeschat waarnaar iemand op zoek is. Waar iemand naar zoekt, wat iemand leuk vindt en waarover iemand berichten stuurt, is allemaal informatie die prijsgeeft waar iemand (in elk geval op dat moment) aan denkt. Een zoekmachine die op deze manier suggesties doet staat op gespannen voet met de vrijheid van gedachten, zoals beschermd in art. 9 EVRM en art. 10 van het Handvest van de grondrechten van de EU. Zodra gedachten niet meer privé zijn, ontstaat ook de mogelijkheid om beslissingen te nemen op basis van de gedachten die iemand (kennelijk) heeft. Daarmee kan worden ingespeeld op momenten van de dag of in de week of maand dat het waarschijnlijker is dat iemand bepaalde dure aankopen doet. Ook kunnen mensen ermee worden gevoed in bepaalde gedachten, zodat die worden versterkt, waar die gedachten anders mogelijk van voorbijgaande aard waren.

Het gevolg kan eveneens zijn dat mensen hun gedrag (preventief) aanpassen zodra ze doorhebben dat er meegekeken of meegeluisterd wordt. Dit wordt in de surveillance-wetenschap ook wel aangeduid als *chilling effects*. Hierbij komt niet alleen de vrijheid van gedachten, maar ook de vrijheid van meningsuiting (art. 7 Grondwet, art. 10 EVRM en art. 11 van het Handvest van de grondrechten van de EU) onder druk te staan. Burgers in smart cities die niet willen dat gegevens over hun worden verzameld, kunnen er in verband met de aanwezigheid van camera's bijvoorbeeld voor kiezen andere (mogelijk langere) routes te nemen door de stad, waar minder camera's staan. Een ander voorbeeld is dat burgers wellicht bewust minder gesticuleren bij camera's die dat mogelijk voor agressief gedrag aanzien. Wanneer burgers hun gedrag aanpassen op de monitoring, kan dat de gegevens die worden verzameld ook beïnvloeden.⁵²

Tegelijkertijd kan ook de autonomie van mensen onder druk komen te staan door het gebruik van big data. Dit is doorgaans het gevolg van zogeheten informatie-asymmetrie, waarbij de gegevensverwerker veel meer inzicht en overzicht heeft dan de betrokkene. In beginsel zouden gebruikers zeggenschap en controle moeten hebben over welke informatie ze verstrekken aan wie en voor welke

doeleinden. Dit wordt ook wel aangeduid als informatiele zelfbeschikking.⁵³ Gewoonlijk wordt dit geïmplementeerd op een praktische manier door het opstellen van privacybeleid (of algemene voorwaarden) waarin wordt uitgelegd welke gegevens worden verzameld, hoe deze gegevens worden verwerkt en voor welke doeleinden. Vervolgens kunnen gebruikers zelf beslissen of ze hiermee akkoord gaan of niet (geïnformeerde toestemming).

Burgers in smart cities die niet willen dat gegevens over hun worden verzameld, kunnen ervoor kiezen andere (mogelijk langere) routes te nemen door de stad waar minder camera's staan

Echter, er zijn aanzienlijk problemen met dit model, dat gebaseerd is op een beeld dat mensen autonome, rationele actoren zijn. Uit onderzoek blijkt namelijk dat als mensen alle privacyvoorwaarden op websites die ze bezoeken daadwerkelijk zouden lezen, dit hen gemiddeld 244 uur per jaar zou kosten.⁵⁴ Uit ander onderzoek blijkt echter dat gebruikers bereid zijn er slechts tussen de 1 en 5 minuten aan te besteden.⁵⁵ Daar zit een groot gat tussen. Privacyvoorwaarden zijn doorgaans lange teksten. Het privacybeleid van Facebook telt bijvoorbeeld ca. 9.500 woorden (hetgeen meer dan een uur kost om te lezen) en het privacybeleid van LinkedIn telt ca. 7.500 woorden (hetgeen ongeveer een uur kost om te lezen). Het gevolg is dat slechts zeer weinig mensen deze teksten lezen. Degenen die de moeite nemen om deze teksten te lezen, zullen mogelijk niet de volledige inhoud begrijpen. Veel privacyvoorwaarden zijn namelijk erg juridisch of technisch geformuleerd. Sommige privacyvoorwaarden zijn ingekort en vereenvoudigd (zoals het privacybeleid van Twitter), maar de vraag is of dit een goed alternatief is.

Vereenvoudigde privacyvoorwaarden bieden niet altijd voldoende details over wat er met de gegevens gebeurt, waardoor mensen alsnog onvoldoende geïnformeerd zijn om goede beslissingen te kunnen nemen.⁵⁶

De onduidelijkheden bestaan ook voor het verwerken van de gegevens. Aangenomen dat gebruikers de privacyvoorwaarden daadwerkelijk lezen en ook echt begrijpen, kan het zijn dat ze de gevolgen van de gegevensverwerkingen alsnog niet kunnen overzien. Sterker nog, in sommige gevallen zullen zelfs de gegevensbeheerders de consequenties niet kunnen zien (en dus ook niet vooraf kunnen uitleggen). Dit is bijvoorbeeld het geval wanneer gebruik wordt gemaakt van geavanceerde geautomatiseerde technieken om gegevens te analyseren, zoals data mining en machine learning, die juist zijn ontworpen om nieuwe, onverwachte verbanden bloot te leggen. Dit is een van de belangrijkste mogelijkheden die big data te bieden heeft. Echter, wanneer iemand zijn schoenmaat, hobby's en huisdieren prijsgeeft, verwacht hij doorgaans niet dat op basis daarvan de kans op een hartaanval kan worden berekend. Deze voor burgers onverwachte mogelijkheden maken het lastig om mensen voldoende vooraf te informeren.

Zelfs wanneer gebruikers weten wat ze willen, kan het zijn dat hun voorkeursoptie mogelijk niet geboden wordt.⁵⁷ In veel gevallen is er bij privacyvoorwaarden en algemene voorwaarden sprake van een alles-of-niets-beslissing. Er kan meestal alleen een vinkje worden gezet. Ruimte om te onderhandelen over de voorwaarden en bijvoorbeeld akkoord te gaan met slechts een deel van de voorwaarden is er niet. De keuzearchitectuur bepaalt dan de mogelijkheden waaruit kan worden gekozen.⁵⁸



In sommige gevallen worden *privacy settings* aangeboden, waarmee voorkeuren kunnen worden ingesteld, maar ook deze mogelijkheden zijn doorgaans beperkt en niet onderhandelbaar vanuit het perspectief van de burger. Daar komt bij dat bepaalde privacy-instellingen ertoe kunnen leiden dat de functionaliteit van een dienst afneemt. Daarnaast kunnen privacyvoorwaarden soms afwezig, onvindbaar of onbegrijpelijk zijn (zie volgende paragraaf).

Transparantie, eerlijk proces en het onschuldbeginzel

Hoewel de Algemene Verordening Gegevensbescherming (AVG, zie paragraaf 4.2) verschillende rechten en plichten schept voor het verzamelen en verwerken van persoonsgegevens, kan gesteld worden dat het lastig is deze rechten uit te oefenen en naleving van de verplichtingen te handhaven. Dit is met name het gevolg van een gebrek aan transparantie ten aanzien van wie gegevens verzamelen, hoe en voor welke doeleinden deze gegevens worden verwerkt, en met welke consequenties. Onderzoek laat zien dat mensen in Nederland met hun gegevens in honderden databanken staan.⁵⁹ Daarvan zijn mensen doorgaans niet op de hoogte.⁶⁰ Bovendien weten ze niet welke databanken dat zijn, welke gegevens daarin staan en wat voor rechten ze hebben ten aanzien van deze gegevens.

Een gebrek aan transparantie kan leiden tot situaties waarin mensen worden geconfronteerd met beslissingen over henzelf zonder dat ze weten hoe deze beslissingen tot stand zijn gekomen. Bijvoorbeeld wanneer mensen een offerte krijgen voor een hypotheek, is soms wel duidelijk welke informatie als input heeft gediend, maar onduidelijk hoe de prijsstelling vervolgens tot stand is gekomen. In een ander voorbeeld bleek dat mensen met een toevoeging aan hun huisnummer (dus huisnummer 186a in plaats van nummer 186) hogere verzekeringspremies kregen.⁶¹ Volgens de Amerikaanse hoogleraar Daniel Solove lijkt dit allemaal nogal op situaties die zijn beschreven in het werk van de schrijver Franz Kafka, in het bijzonder in diens roman *Het Proces*.⁶² In dit boek beschrijft Kafka een dystopische wereld waarin de protagonist, Joseph K., op een ochtend wakker wordt en een aantal overheidsfunctionarissen in zijn woning aantreft die hem arresteren, zonder erbij te vermelden waarom. Aangezien K. zich niet het minste vergriep kan herinneren, is hij nogal van slag. Gedurende de rest van het boek onderneemt hij vergeefse pogingen om uit te vinden waarom hij is gearresteerd en hoe zijn zaak kan worden opgehelderd en opgelost.



Het blijft echter onduidelijk welke gegevens er over K. zijn verzameld, hoe de beslissing hem te arresteren tot stand kwam, wie verantwoordelijk is voor deze beslissing en hoe de beslissing ter discussie kan worden gesteld.

In het boek van Kafka staat centraal het recht op een eerlijk proces: er wordt niet voldaan aan de eisen van art. 6 EVRM. In de context van big data is niet zonder meer sprake van schendingen van dit recht, omdat elke burger gewoon en beroep kan doen op een rechter, met een eerlijke en openbare behandeling van zijn zaak, binnen een redelijke termijn, door een onafhankelijk en onpartijdig gerecht. Toch wringt het wanneer burgers via de AVG rechten krijgen toegekend ten aanzien van hun persoonsgegevens die ze vervolgens niet goed kunnen uitoefenen.

De AVG kent verschillende rechten toe aan burgers, waaronder het recht om geïnformeerd te worden (bijvoorbeeld over welke informatie er over hen wordt verzameld, door wie, voor welke doeleinden, etc. – art. 12 AVG), het recht tot toegang tot de gegevens (art. 13 en 14 AVG), het recht van inzage (art. 15 AVG), het recht op rectificatie (art. 16 AVG), het recht om gegevens te laten verwijderen ('het recht om vergeten te worden', art. 17 AVG), het recht op beperking van de verwerking (art. 18 AVG), het recht op overdraagbaarheid van gegevens (art. 20 AVG) en het recht van bezwaar (art. 21 AVG) en het recht niet onderworpen te worden aan geautomatiseerde besluitvorming (art. 21 AVG). Betrokkenen kunnen echter alleen goed gebruik maken van deze rechten als ze weten dat ze deze rechten hebben en hoe ze die kunnen uitoefenen. Veel burgers zijn slecht op de hoogte van hun rechten op dit terrein. Het gevolg is dat de EU-wetgeving voor de bescherming van persoonsgegevens onvoldoende werkt.⁶³

Een laatste noemenswaardig punt op dit vlak is dat ook het onschuldbeginsel onder druk kan komen te staan in de context van big data. Zoals in paragraaf 4.1 werd aangegeven, wordt bij big data veel gebruik gemaakt van voorspellingen en profilering. De voorspellingen zijn uiteraard zeer bruikbaar voor bedrijven om mensen te categoriseren en bijvoorbeeld gepersonaliseerde aanbiedingen te doen. De voorspellingen kunnen ook worden gebruikt (en worden ook gebruikt) om kiezers te beïnvloeden bij hun stemgedrag.⁶⁴ De voorspellingen leveren daarbij een soort 'omgekeerde bewijslast' op voor betrokkenen: zodra een voorspelling onjuist is (het blijft statistiek), is het aan de betrokkene om aan te tonen dat de gegevens onjuist zijn.⁶⁵ Dit kan doorgaans alleen maar door de juiste gegevens te laten zien,

waarmee verder wordt ingeleverd op de privacy.⁶⁶ Tegelijkertijd zit hierin een procedurele onrechtvaardigheid, omdat de betrokkene wordt gedwongen zich te verdedigen tegen de gevolgen van onjuiste gegevens of onjuiste conclusies.

4.2 Bescherming van persoonsgegevens

Het grondrecht op privacy is in secundaire wetgeving vooral uitgewerkt voor één aspect, namelijk dat van informationele privacy. Dat is op EU-niveau gedaan in de Algemene Verordening Gegevensbescherming (AVG).⁶⁷ Het valt buiten het bestek van dit essay om een compleet overzicht te geven van alle bepalingen in de AVG en hoe die van toepassing zijn in smart cities. Hieronder wordt daarom vooral ingegaan op de beginselen voor behoorlijke verwerking van persoonsgegevens. Deze beginselen, ontwikkeld door de OESO in 1980,⁶⁸ werden opgenomen in het zogeheten Verdrag van Straatsburg van de Raad van Europa in 1981.⁶⁹ In 1995 nam de EU deze beginselen over in Richtlijn 95/46/EC, de zogeheten databeschermingsrichtlijn. Deze richtlijn werd in mei 2018 vervangen door de AVG, die directe werking heeft in de gehele EU.



De beginselen zijn:

- **Beperkt verzamelen:** er zijn grenzen aan het verzamelen van persoonsgegevens en zulke gegevens mogen alleen worden verwerkt wanneer ze rechtmatig en eerlijk zijn verkregen en, waar mogelijk, met kennis en instemming van betrokkene.
- **Datakwaliteit:** de persoonsgegevens moeten relevant zijn voor de doelen waarvoor ze worden verwerkt en, waar noodzakelijk voor die doelen, accuraat, compleet en niet verouderd zijn.
- **Doelspecificatie:** de doelen waarvoor persoonsgegevens worden verzameld moeten vooraf duidelijk worden gespecificeerd en de persoonsgegevens mogen slechts voor deze doelen worden gebruikt.
- **Beperkt gebruik:** persoonsgegevens mogen niet worden gebruikt voor andere doelen dan die doelen die vooraf zijn gespecificeerd, tenzij daarvoor toestemming door betrokkene is gegeven of daarvoor een wettelijke basis bestaat.
- **Beveiliging:** er moeten redelijke waarborgen zijn getroffen tegen risico's van verlies, ongeautoriseerde toegang, vernietiging, etc. van persoonsgegevens.
- **Transparantie:** betrokkene moet in staat zijn kennis te nemen van het bestaan en van de aard van de persoonsgegevens, de doelen waarvoor die worden verwerkt en de identiteit van de gegevensbeheerder.
- **Individuele participatie:** betrokkenen hebben het recht hun persoonsgegevens te laten verwijderen, rectificeren, aanvullen of wijzigen.
- **Verantwoording:** gegevensbeheerders zijn verantwoordelijk voor de naleving van deze beginselen.

De eerste vier beginselen richten zich vooral op de persoonsgegevens en de voorwaarden waaronder de gegevens mogen worden verzameld en verwerkt. De andere vier beginselen richten zich vooral op de plichten en verantwoordelijkheden van de gegevensbeheerders en de rechten van betrokkenen. Het is belangrijk om hier op te merken dat deze beginselen zich vooral op procedurele rechtvaardigheid richten en niet zozeer op inhoudelijke rechtvaardigheid.

Het kan dus zijn dat gegevensbeheerders zich netjes houden aan deze beginselen (duidelijke doelen en doelbinding), maar toch verder gaan dan betrokkenen verwachten. Een typisch voorbeeld is wanneer gegevensbeheerders in hun

privacyvoorwaarden zetten persoonsgegevens te zullen doorverkopen voor reclaimedoeinden of gepersonaliseerde prijsstellingen. Wanneer een betrokkene hiervoor toestemming geeft, is er juridisch geen probleem, maar wanneer een betrokkene vervolgens zijn of haar foto in een reclame terugziet of geconfronteerd wordt met hogere prijzen, kan dat voorbij de verwachtingen van de betrokkene liggen. Juridisch gezien is daar niet veel aan te doen.

Als gegevens voor andere doelen in andere contexten worden gebruikt, kan dat een schending van de privacyverwachtingen en van de doelbinding zijn

Bij hergebruik van gegevens in de context van smart cities speelt doelbinding en de redelijke verwachtingen van betrokkenen over het gebruik van hun persoonsgegevens een belangrijke rol. Als gegevens voor andere doelen in andere contexten worden gebruikt, kan dat een schending van de privacyverwachtingen en van de doelbinding zijn. Het is in zulke gevallen het beste om burgers om toestemming te vragen en om transparant te zijn over welke gegevens op welke manieren worden gebruikt.

4.3 Algemene beginselen van behoorlijk bestuur

Afgezien van het privacy- en gegevensbeschermingsrecht kent ook het bestuursrecht een aantal bepalingen die relevant zijn voor het verzamelen en verwerken van gegevens in smart cities. Omdat een uitgebreide analyse van het bestuursrecht buiten de reikwijdte van dit essay valt, ligt hieronder de nadruk op de algemene beginselen van behoorlijk bestuur (abbb).⁷⁰ Dit zijn gedragsregels voor de overheid (in de context van smart cities is dit de gemeente) ten opzichte van burgers. Deze beginselen zijn in eerste instantie uit de jurisprudentie ontstaan, maar in 1994 vastgelegd in de Algemene wet bestuursrecht (Awb).⁷¹ De beginselen kunnen worden onderscheiden in formele en materiële beginselen.

De formele beginselen zijn:

- **Legaliteitsbeginsel:** er is geen bevoegdheid zonder grondslag in wet of Grondwet.
- **Zorgvuldigheidsbeginsel:** de overheid moet een besluit zorgvuldig voorbereiden en nemen; hierbij horen een correcte behandeling van de burger, zorgvuldig onderzoek naar de feiten en belangen, het goed volgen van de procedure en deugdelijke besluitvorming (art. 3:2 Awb).
- **Motiveringsbeginsel:** de overheid moet haar besluiten goed motiveren; de feiten moeten kloppen en de motivering moet logisch en begrijpelijk zijn (art. 3:46 Awb).
- **Rechtszekerheidsbeginsel:** de overheid moet haar besluiten zo formuleren dat de burger precies weet waar hij aan toe is of wat de overheid van hem verlangt. Bovendien moet de overheid de geldende rechtsregels juist en consequent toepassen.
- **Fair-play-beginsel:** de overheid moet de publieke taken zonder vooringenomenheid vervullen (art. 2:4 Awb).
- **Verbod op détournement de procédure:** er mag geen lichtere procedure worden gevolgd om tot een besluit te komen, wanneer voor dat besluit ook een met meer waarborgen omklede procedure openstaat.
- **Vertrouwensbeginsel:** wie op goede gronden - bijvoorbeeld na een duidelijke toezegging - erop kan vertrouwen dat de overheid een bepaald besluit neemt, heeft daar ook recht op.

Deze beginselen zijn eenvoudig toepasbaar op het hergebruik van gegevens van burgers in smart cities. In feite lijken deze beginselen voor een groot deel op de beginselen in het gegevensbeschermingsrecht (zie paragraaf 4.2). De essentie is dat de overheid beslissingen op een eerlijke en zorgvuldige manier voorbereidt, neemt en onderbouwt. Bij geautomatiseerde besluitvormingsprocessen moeten gegevens dus rechtmatig zijn verzameld, accuraat en up-to-date zijn. De besluitvorming zelf moet transparant en uitlegbaar zijn. Hergebruik van gegevens moet dus passen bij de verwachtingen die burgers (mogen) hebben. Dit geldt niet alleen voor overheden die zelf gegevens verwerken, maar ook voor overheden die het verwerken van gegevens uitbesteden aan private bedrijven. Uitbesteding van bepaalde vormen van gegevensverwerking, zoals het verzamelen en opslaan van grote hoeveelheden gegevens of het uitvoeren van complexe gegevensanalyses, is heel gewoon, omdat het nu eenmaal gespecialiseerde taken betreft, maar kan

allerlei praktische problemen met zich meebrengen. Zo kan het bijvoorbeeld lastig te controleren zijn of ingehuurd organisaties zich goed aan de regels houden en kan lastig traceerbaar zijn wie allemaal toegang heeft tot de gegevens.

Hergebruik van gegevens, zeker als het voor nieuwe doelen of in een nieuwe context is, kan dan problematisch zijn, waardoor extra nadruk op toestemming en transparantie komt te liggen.

De materiële beginselen van behoorlijk bestuur zijn:

- **Specialiteitsbeginsel:** de overheid mag alleen die belangen behartigen waarvoor de betrokken wet of regeling een grondslag biedt (art. 3:4 lid 1 Awb).
- **Evenredigheidsbeginsel:** de overheid moet ervoor zorgen dat de lasten of nadelige gevolgen van een overheidsbesluit voor een burger niet zwaarder zijn dan het algemeen belang van het besluit (art. 3:4 lid 2 Awb).
- **Vertrouwensbeginsel** (materiële rechtszekerheid): een burger moet, onder bepaalde voorwaarden, kunnen vertrouwen op uitlatingen van een bestuursorgaan waarin dingen worden toegezegd maar die later niet nagekomen (kunnen) worden door het bestuursorgaan.
- **Gelijkheidsbeginsel:** de overheid moet gelijke gevallen op gelijke wijze behandelen (art. 1 Grondwet).
- **Verbod van détournement de pouvoir:** de overheid mag een bevoegdheid alleen gebruiken voor het doel waarvoor die bevoegdheid is gegeven (art. 3:3 Awb).

Deze materiële beginselen lijken op het eerste gezicht nog steeds redelijk procedureel van aard, omdat ze betrekking hebben op de wijze waarop beslissingen worden genomen. Echter, bij nadere beschouwing blijkt dat veel van deze materiële beginselen vragen om het maken van een substantiële inhoudelijke afweging per casus. Dat houdt in dat van geval tot geval de omstandigheden moeten worden afgewogen. Voor het hergebruik van gegevens moet dan steeds het perspectief van de betrokkene (de burger) worden nagegaan danwel worden afgewogen tegen andere perspectieven, zoals het algemeen belang. Uiteraard staat of valt een dergelijke afweging met de mate waarin het overheden lukt zich in te leven in het perspectief van de betrokkene. Inspraak en burgerparticipatie kunnen hieraan bijdragen (zie paragraaf 6).

Vanuit een juridisch perspectief luidt de tussenconclusie dat er weliswaar verschillende kaders zijn (grondrechten, gegevensbeschermingsrecht, bestuursrecht) die bescherming beogen te bieden voor burgers, maar dat kaders veel ruimte laten voor het maken van afwegingen. De bestaande kaders zijn breed en algemeen opgezet, met een nadruk op procedurele rechtvaardigheid, meer dan materiële rechtvaardigheid. Hoewel de juridische kaders best duidelijk afbakenen wat niet is toegestaan, zijn ze niet altijd bruikbaar voor het vaststellen van behoorlijk datagebruik en -hergebruik. Met andere woorden, wanneer de juridische regels voor het verwerken van gegevens niet worden nageleefd is er sprake van onbehoorlijk datagebruik, maar andersom is niet elke vorm van gegevensverwerking die binnen de juridische regels valt daadwerkelijk behoorlijk. Burgers kunnen immers geheel andere verwachtingen hebben over wat er met hun gegevens gebeurt of zou mogen gebeuren.

5. Ethisch perspectief - afwegingskaders

Naast beschouwing van juridische kaders, kan het ook relevant zijn bij (her)gebruik van gegevens te kijken vanuit ethisch perspectief. De juridische kaders bieden randvoorwaarden waaraan moet worden voldaan. Maar binnen die randvoorwaarden kan er ruimte zijn voor verschillende soorten afwegingen. Juridische kaders kunnen niet alles reguleren, er zitten grenzen aan de reikwijdte en hoeveelheid details van wet- en regelgeving. De ethische kaders kunnen worden gebruikt voor een invulling van de juridische kaders en een verdere basis leggen voor zelfregulering, bijvoorbeeld door het opstellen van gedragscodes.⁷² Juridische kaders geven aan hoe we ons moeten gedragen, ethische kaders geven aan hoe we ons zouden moeten gedragen. Wat dan goede afwegingen zijn, hangt af van het perspectief. Het morele perspectief houdt in dat rekening wordt gehouden met de kwetsbaarheden van anderen. Soms kan het lastig zijn verschillende belangen tegen elkaar af te wegen, zoals privacy versus veiligheid of individuele belangen versus collectieve belangen. Hieronder worden verschillende afwegingskaders beschreven die vanuit ethisch perspectief kunnen worden ingezet. Het gaat om de drie hoofdstromen in de (westerse) ethiek. Respectievelijk komen aan bod deontologie, utilitarisme en deugdenethiek.

5.1 Deontologie

De deontologie, ook wel de plichtethiek of regel-gebaseerde ethiek genoemd, gaat uit van morele normen en waarden, morele beginselen en morele regels. De morele normen en waarden beschrijven de wenselijke situatie, het doel waarnaar gestreefd wordt, bijvoorbeeld een rechtvaardige samenleving. De morele beginselen beschrijven algemene gedragsregels die bijdragen aan het bereiken van de morele waarden (bijvoorbeeld dat welvaart eerlijk verdeeld moet zijn) en de morele regels zijn prescriptieve, concrete gedragsregels (bijvoorbeeld dat stelen verboden is). De morele regels moeten worden nageleefd, vandaar de term regel-gebaseerde ethiek. Morele regels kunnen juridische regels zijn (zoals stelen is verboden), maar dat hoeft niet het geval te zijn (zoals vreemdgaan is moreel laakbaar, maar niet strafbaar). Of bepaald gedrag of bepaalde beslissingen moreel zijn, wordt beoordeeld aan de hand van morele regels.

De morele normen en waarden beschrijven de wenselijke situatie, het doel waarnaar gestreefd wordt, bijvoorbeeld een rechtvaardige samenleving

Morele regels komen uit de samenleving, bijvoorbeeld via het publieke debat. In een democratie kunnen morele regels worden omgezet in juridische regels door vertegenwoordigende organen, zoals het nationale parlement of vergelijkbare decentrale organen, hetgeen terugleidt naar de juridische kaders van de vorige paragraaf. Maar op terreinen waarvoor geen juridische kaders bestaan of waar die onvoldoende bruikbaar zijn, kunnen morele regels verder gaan. Bijvoorbeeld als het gaat om het doorverkopen van gegevens, kan het sterk van de context afhangen of dit moreel laakbaar is.⁷³ Ook op het terrein van privacy, dataficatie (al het gedrag van burgers vastleggen in gegevens) en dataveillance (surveillance op basis van gegevens) liggen ethische uitdagingen.⁷⁴ Deze morele regels kunnen ook eenvoudig op het hergebruik van gegevens worden toegepast. Wanneer bijvoorbeeld een overheid voornemens is om gegevens in smart cities te gaan doorverkopen aan verzekeraars om financiële winsten te maken, is het goed voor

de betreffende overheidsmedewerkers om zich te realiseren dat zijzelf ook burgers zijn. Vanuit hun rol als burger kunnen ze dan nagaan of dit een verstandig plan is. Ook kan het nuttig zijn om in deze situaties actief gebruik te maken van inspraak van burgers: via burgerparticipatie kan worden nagegaan wat goede (morele) regels zijn voor hergebruik van gegevens.

5.2 Utilitarisme

Een andere moreel perspectief is dat van het utilitarisme. Vanuit dit ethisch perspectief, onder meer ontwikkeld door de Engelse filosoof Jeremy Bentham, wordt getracht het welzijn ('utiliteit') te maximaliseren. Het utilitarisme, ook wel consequentialisme genoemd, beoordeelt of bepaald gedrag of bepaalde beslissingen moreel zijn aan de hand van de uitkomsten van dat gedrag. In feite moet hierbij dus onderzocht worden wat het beste resultaat oplevert.

De twee grote vragen in deze benadering zijn uiteraard wat er nou precies moet worden gemaximaliseerd en hoe dat vervolgens kan worden berekend. Bentham zelf definieerde utiliteit als de optelling van al het plezier minus al het lijden dat voortkomt uit bepaalde gedragingen of beslissingen. De Britse filosoof John Stuart Mill bedacht het 'greatest happiness principle'. Vanuit een economisch perspectief maken financiële berekeningen het berekenen van maximale welvaart soms best mogelijk, maar welvaart en welzijn zijn verschillende zaken. Het berekenen van het maximale welzijn is veel ingewikkelder. Wat belangrijk is, in dit opzicht, is dat het welzijn van elke persoon even zwaar meetelt in de berekeningen.

Naast deze praktische problemen wordt regelmatig een ander bezwaar geopperd tegen het utilitarisme, namelijk dat het resultaat zo bepalend is. Met andere woorden, het doel heiligt de middelen. Het blind toepassen van utilitarisme zou kunnen leiden tot enorme herverdelingen van de (soms zeer ongelijk verdeelde) welvaart. Dit lijkt op het eerste gezicht wellicht positief, maar zulke ingrepen zouden erg disruptief kunnen zijn voor de economie.

Ondanks de hiervoor genoemde bezwaren kan een utilistische benadering nuttig zijn, bijvoorbeeld om belangen op grotere schaal af te wegen en om belangen van het individu af te wegen tegen die van de samenleving. Een typisch voorbeeld betreft het gebruik van zelfrijdende auto's in een stad. Deze technologie is niet

zonder fouten, waardoor er slachtoffers kunnen vallen. Maar als er ondanks dit gegeven door het gebruik van zelfrijdende auto's minder slachtoffers vallen dan door menselijke bestuurders, kan dit een sterk argument zijn om toch in te zetten op zelfrijdende auto's.

Ook als het gaat om het hergebruik van gegevens kan het zinvol zijn om af te wegen welke strategieën uiteindelijk de meeste tevredenheid opleveren. Wanneer bijvoorbeeld de vraag is of beelden van verkeerscamera's ook gebruikt mogen worden om een ontvoerd kind te traceren, kan een utiliteitsargument worden ingezet. De verminderde privacy van mensen op de camerabeelden zal dan niet snel opwegen tegen het belang van het terugvinden van het kind. In juridische termen, overigens, wordt deze afweging doorgaans gemaakt door toepassing van het beginsel van proportionaliteit, waarbij wordt afgewogen of het in te zetten middel proportioneel is aan het beoogde doel.

5.3 Deugdenethiek

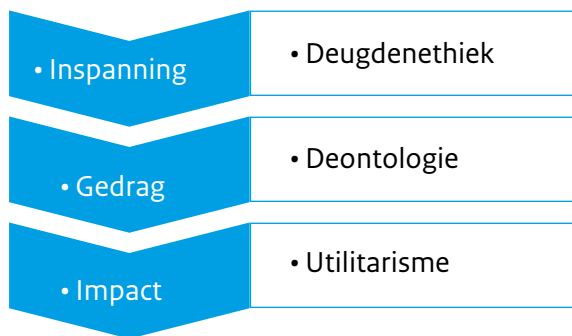
Een derde ethisch kader is dat van de deugdenethiek, waarbij wordt onderzocht hoe iemand zou moeten leven. De nadruk ligt daarbij op het handelen van een persoon en de overwegingen die iemand heeft bij zijn handelingen en beslissingen. Aristoteles wordt gezien als de grondlegger van deze benadering en bekende filosofen zoals Thomas van Aquino, David Hume en Friedrich Nietzsche hebben hierop voortgebouwd. In tegenstelling tot de deontologie (waarin gedrag en regels centraal staan) en utilitarisme (waarin het resultaat en de impact centraal staan), gaat het in de deugdenethiek om het karakter van de persoon en diens inspanningen (zie Figuur 1). Volgens Aristoteles moeten mensen onder meer moed, gematigdheid, wijsheid en rechtvaardigheid tonen. Uiteindelijk is de motivering van handelen bepalend bij het vaststellen of bepaald gedrag of bepaalde beslissingen moreel zijn.

De belangrijkste vraag bij deze benadering is uiteraard wat dan de deugden zijn waarnaar moet worden geleefd of die ten grondslag moeten liggen aan iemands intenties. Aristoteles stelde in zijn tijd al een lijst op met elf deugden, waaronder moed en rechtvaardigheid, om voortreffelijk karakter te ontwikkelen.

Latere filosofen hebben weer andere lijsten met deugden ontwikkeld. In dit opzicht is het werk van de Amerikaanse technologiefilosofe Shannon Vallor interessant.

Zij heeft op basis van deugdenethiek een specifieke set van technologisch-morele deugden ontwikkeld om deugden aan te passen aan onze moderne technologische tijd. Zij komt tot een lijst van twaalf waarden: eerlijkheid, zelfbeheersing, nederigheid, rechtvaardigheid, moed, empathie, zorg, beschaafdheid, flexibiliteit, perspectief, grootmoedigheid en technologisch-morele wijsheid.⁷⁵

Wat in de context van smart cities vooral van belang is als het gaat om deugdenethiek is wat oprecht en goed handelen is. Dat is relevanter dan welke set van deugden of morele waarden uiteindelijk wordt gekozen.⁷⁶ Het gaat er vooral om na te gaan of bepaalde intenties van handelingen of beslissingen juist zijn. Wanneer bijvoorbeeld gegevens van burgers worden verkocht om winst te maken kan dat onderliggende winstbejag een moreel ondeugdelijke intentie zijn. Wanneer de gegevens worden verkocht om vervolgens met de opbrengsten de stad te verbeteren, kan de beoordeling anders zijn. Als beslissingen worden genomen met het oog op het verbeteren van de leefbaarheid in een stad is dat een heel andere intentie dan wanneer stadsbestuurders hun stempel willen drukken op hun ambtstermijn.



Figuur 1 Algemene ethische theorieën voor nadere uitwerking van specifieke uitdagingen van hergebruik van gegevens in smart cities.

De drie bovenstaande ethische perspectieven zijn klassiek ethische benaderingen. Voor de praktijk zijn er allerlei instrumenten ontwikkeld om deze benaderingen handen en voeten te geven. Een mooi voorbeeld is De Ethische Data Assistent (DEDA), een instrument om ethische problemen te herkennen in dataprojecten,

ontwikkeld door de Universiteit Utrecht en de Gemeente Utrecht.⁷⁷ Een ander handig instrument is de Ethical Matrix.⁷⁸ Dit instrument gaat uit van drie kernwaarden (welzijn, menselijke waardigheid en rechtvaardigheid) en helpt om stapsgewijs na te gaan welke waarden in het geding kunnen zijn en wat daaraan gedaan kan worden.

6. Conclusies

In dit essay is geprobeerd antwoord te geven op de vraag: *in hoeverre mogen gegevens van burgers worden hergebruikt in de context van smart cities?* Het antwoord op deze vraag verschilt per situatie en is afhankelijk van de omstandigheden. Toch zijn er duidelijke juridische en ethische aanknopingspunten om deze vraag van geval tot geval te beantwoorden. De belangrijkste juridische kaders voor het gebruik van big data in de openbare ruimte worden gegeven door (1) privacy en andere grondrechten, (2) de algemene beginselen van behoorlijke gegevensverwerking in het gegevensbeschermingsrecht en (3) de algemene beginselen van behoorlijk bestuur in het bestuursrecht. Het grondrechtenkader biedt vooral een inhoudelijk perspectief, het gegevensbeschermingsrecht (in het bijzonder de AVG) biedt vooral een procedureel perspectief en het bestuursrecht biedt een combinatie van beide. De grote gemene deler in deze regimes is het centraal zetten van mensen (soms aangeduid als betrokkenen of burgers) in plaats van hun gegevens. Door niet te vergeten dat het verzamelen en verwerken van gegevens een middel is, dus niet een doel op zichzelf, kan steeds worden afgewogen of het gebruik en hergebruik van gegevens gerechtvaardigd is.

Daarbij is van belang dat het ook geen goede strategie is om de (juridische/procedurele) regels centraal te stellen. Hoewel het belangrijk is om de juridische regels goed na te leven, is dat niet voldoende. Met name de juridische kaders van de AVG bieden voldoende ruimte om binnen de bestaande regels toch gegevens te verwerken op manieren dat burgers teleurgesteld kunnen raken over de wijze waarop hun gegevens in smart cities worden verwerkt. Dit kan leiden tot afkalving van het vertrouwen en van het maatschappelijk draagvlak voor plannen. Om deze reden zijn de aanvullende ethische kaders noodzakelijk. Ze bieden mogelijkheden om naast meer procedurele afwegingen ook meer inhoudelijke afwegingen te maken.

De beschreven ethische perspectieven van deontologie, utilitarisme en deugdenethiek kunnen daarbij behulpzaam zijn. Zij leggen de focus op respectievelijk het gedrag, de uitkomsten en de intenties van bepaald gedrag en bepaalde beslissingen. Elk perspectief heeft zijn eigen sterke en zwakke kanten, en door complexe vraagstukken vanuit meerdere perspectieven te benaderen kan een beter beeld ontstaan over wat goede handelwijzen en beslissingen zijn. Uiteindelijk is het doel van het verder ontwikkelen van smart cities het welzijn van burgers te vergroten. Het mag triviaal klinken, maar daarbij is van belang die burgers niet uit het oog te verliezen. Op basis van bovenstaande conclusies kunnen de volgende aanbevelingen worden gedaan:

1. Gemeentelijke overheden moeten bij smart cities inzetten op behoorlijk hergebruik van gegevens via publiek-private samenwerkingen.

Het speelveld van steden is complex, met een veelheid aan actoren die allemaal op hun eigen terrein gegevens verzamelen en verwerken. In plaats van dit zelf beter te willen doen door eigen datasets te verzamelen, kunnen overheden beter inzetten op hergebruik van gegevens, waarbij datasets van stakeholders worden gedeeld of gekoppeld.



Via publiek-private samenwerking kunnen niet alleen gegevens worden gedeeld, maar ook gezamenlijke projecten voor bijvoorbeeld betere leefbaarheid en meer duurzaamheid worden opgezet en uitgevoerd.⁷⁹

De eerste stap die een gemeente kan nemen is het in kaart brengen van beschikbare datasets en bestaande datastromen in het publiek-private speelveld. Vervolgens kan voor bestaande of nieuwe vormen van hergebruik aan de hand van de kaders in dit essay worden getoetst of dit hergebruik behoorlijk is. Als de AVG bijvoorbeeld geen juridische basis biedt, is er geen sprake van behoorlijk hergebruik. Bij het toetsen van het hergebruik moeten alle juridische en ethische afwegingen in dit essay worden nagelopen – wanneer slechts een selectie wordt gemaakt (bijvoorbeeld alleen het perspectief van de AVG of van de deugdenethiek) kan niet worden gegarandeerd dat er sprake is van behoorlijk hergebruik.

2. Gemeentelijke overheden moeten bij smart cities burgers actief betrekken bij afwegingen.

Het welzijn van burgers moet centraal staan in het verder ontwikkelen van smart cities. Juridische kaders bieden daarvoor randvoorwaarden en een minimaal beschermingsniveau voor burgers, maar daarenboven is het belangrijk het morele perspectief te houden, waarin rekening wordt gehouden met de kwetsbaarheden van anderen. Door burgerparticipatie en -consultatie kan worden geborgd dat voldoende rekening wordt gehouden met de perspectieven van (groepen) burgers. Smart cities zijn geen doel op zichzelf, het gaat om de achterliggende doelen zoals welzijn, leefbaarheid en duurzaamheid.

Het heeft doorgaans weinig zin om burgers te vragen wat ze vinden van projecten waarin gegevens worden gedeeld. In de praktijk is het dan vaak dezelfde kleine groep personen die reageert, een groep die waarschijnlijk niet representatief is voor alle burgers. Een beter alternatief kan daarom zijn burgers vooral te vragen naar de achterliggende doelen: wat zijn hun ideeën over welzijn, leefbaarheid en duurzaamheid? Dit kan via de gebruikelijke wegen: enquêtes, interviews, discussiebijeenkomsten, burgerpanels, internetconsultaties, etc. Het is vervolgens aan de professionals om de vertaalslag te maken naar concrete projecten en informatiestromen, met inachtneming van de kaders in dit essay en transparantie richting burgers (zie hieronder).

3. Gemeentelijke overheden moeten bij smart cities zorgen voor transparantie met betrekking tot gegevens en afwegingen.

Transparantie is van groot belang voor het vertrouwen dat burgers hebben in overheden en de ontwikkelingen rondom smart cities. Die transparantie heeft enerzijds betrekking op gegevens (bijvoorbeeld welke gegevens over burgers worden verzameld en hoe deze gegevens worden gebruikt) en anderzijds op de wijze waarop bepaalde afwegingen tot stand komen (bijvoorbeeld hoe de noodzakelijkheid van gegevens wordt afgewogen tegen dataminimalisatie of hoe nudging en paternalisme tegen elkaar worden afgewogen). Transparantie draagt bij aan het vertrouwen van burgers en vertrouwen draagt vervolgens bij aan maatschappelijk draagvlak.

Vaak wordt getracht transparantie te realiseren via portals waarop burgers kunnen inloggen en hun gegevens kunnen inzien en beheren. In de praktijk maken burgers hier niet altijd veel gebruik van. Dergelijke vormen van transparantie kunnen weliswaar verplicht zijn onder de AVG, veel betrokkenheid creëren ze doorgaans niet. Een beter alternatief kan zijn het inzetten op de democratische instrumenten voor transparantie, bijvoorbeeld door zaken omtrent smart cities en dataverzameling en hergebruik een plek te geven op de agenda van de gemeenteraad. Uiteraard moeten debatten dan niet gaan over gegevensstromen, maar over het bereiken van doelen als welzijn, leefbaarheid en duurzaamheid. Dat creëert enerzijds bredere betrokkenheid (direct van burgers en indirect via hun vertegenwoordigers in de gemeenteraad) en anderzijds democratische transparantie en instrumenten om behoorlijk hergebruik van gegevens te toetsen.

4. Gemeentelijke overheden moeten via pilots smart cities verder ontwikkelen.

In plaats van grote plannen is het veiliger en realistischer voor overheden om via kleinschalige pilots te experimenteren met het doorontwikkelen van smart cities.⁸⁰ Het concept smart cities klinkt groots, maar bestaat doorgaans uit een verzameling van deelprojecten in verschillende maatschappelijke sectoren, zoals transport, veiligheid, milieu, etc. Door op deelgebieden nieuwe initiatieven te ontwikkelen kunnen geleidelijk informatiestromen worden geoptimaliseerd.

Ook kan via de weg van geleidelijkheid meer tijd worden vrijgemaakt om burgers te betrekken bij nieuwe ontwikkelingen en kan hun vertrouwen worden gewonnen.⁸¹

Pilots dienen altijd binnen de juridische en ethische kaders te blijven zoals die in dit essay zijn beschreven. Het kan echter, wanneer geëxperimenteerd wordt, onduidelijk zijn of een pilot geheel binnen deze kaders valt. In geval van twijfel is het aan te raden in elk geval een gegevensbeschermingseffectbeoordeling (ook wel Privacy Impact Assessment, PIA, genoemd)⁸² uit te voeren, om de belangrijkste risico's voor burgers in kaart te brengen. Wanneer uit een dergelijke analyse blijkt dat er sprake is van hoge risico's, is het verplicht de toezichthouder (de Autoriteit Persoonsgegevens) te raadplegen alvorens te starten met de pilot (art. 36 AVG). Deze kan dan adviseren. Als bij een pilot reeds sprake is van hoge risico's, is maar zeer de vraag of de plannen die getest worden überhaupt in de praktijk mogen of zouden moeten worden geïmplementeerd.⁸³

De afwegingskaders in dit essay bieden de mogelijkheid om te toetsen of er sprake is van behoorlijk (her)gebruik van gegevens. Door bestaande of voorgenomen gegevensstromen vanuit deze juridische en ethische perspectieven te beoordelen kan worden voorkomen dat het gegevensgebruik onbehoorlijk is. Daarbij kan niet worden geselecteerd uit de perspectieven (bijvoorbeeld alleen een AVG-toetsing uitvoeren), maar moeten zoveel mogelijk (lieft alle) perspectieven worden nagelopen. Bovendien is het bij voorgenomen plannen altijd raadzaam om de kaders verder aan te vullen via zelfregulering, bijvoorbeeld door gebruik van gedragscodes en convenanten.⁸⁴ Dat alles is zeer bewerkelijk, maar die bewerkelijkheid is te verwachten in een complexe omgeving als die van smart cities. Toch kunnen deze inspanningen zeer de moeite waard zijn, omdat concrete projecten en goede voornemens volledig kunnen stuklopen als blijkt dat ze onvoldoende doordacht en voorbereid zijn en er onvoldoende rekening is gehouden met verschillende perspectieven en met maatschappelijk draagvlak.

Over de auteur



Prof. mr. dr. ir. Bart Custers is hoogleraar Law and Data Science aan Universiteit Leiden.



Referenties

- 1 Iainone C. (2016) The CO-City: Sharing, Collaborating, Cooperating, and Commoning in the City. *American Journal of Economics and Sociology* 75 (2): 415–455. doi:10.1111/ajes.12145.
- 2 Ranchordás, S. and Klop, A., Data-Driven Regulation and Governance in Smart Cities, in *Research Handbook on Data Science and Law*, Edward Elgar Publishing, 2018, p. 245–273.
- 3 Page, S., Phillips, B. Siembab, W. (2003) The Millenium City: Making Sprawl Smart through Network-Oriented Development, 10(3) *Journal of Urban Technology* 63.
- 4 <https://www.denhaag.nl/nl/in-de-stad/wonen-en-bouwen/bouwprojecten/gebiedsontwikkeling-scheveningen-kust/slimme-lantaarnpalen-in-scheveningen.htm>
- 5 <https://cities-today.com/industry/eindhoven-to-create-worlds-first-crowdsourced-smart-city/>
- 6 <https://www.rtvnoord.nl/nieuws/201369/Stad-roept-hulp-bedrijven-in-bij-oplossen-drukte-fietsers-en-voetgangers>
- 7 Gegevens over burgers kunnen juridisch geduid worden als persoonsgegevens (zie paragraaf 4). Voor de leesbaarheid wordt in dit essay de term gegevens gebruikt, maar worden persoonsgegevens bedoeld.
- 8 Hansen, P., Jespersen, A. M. (2013) Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy. *European Journal of Risk Regulation* 4 (1): 3–28.
- 9 Kitchin, R. (2013) The ethics of smart cities and urban science, *Phil. Trans. R. Soc. A*, vol. 374(2083).
- 10 Albino, V., U. Berardi, and R. M. Dangelico (2015) Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology* 22 (1): 3–21; Hollands, R. (2008) Will the real smart city please stand up? *City*, vol. 12(3), p. 303–320; Kitchin, R. (2015) Making sense of smart cities: addressing present shortcomings, *Cambridge Journal of Regions, Economy and Society*, vol. 8(1), p. 131–136.
- 11 Merton, R.K. (1968) *Social Theory and Social Structure*, New York, NY: Free Press.
- 12 Eksioglou, B., Volkan Vural, A., and Reisman, A. (2009) The vehicle routing problem: a taxonomic review, *Computers & Industrial Engineering*, 57, p. 1472–1483.
- 13 Golden, B., Raghavan, S., and Wasil, E. (2008) *The Vehicle Routing Problem: Latest Advances and New Challenges*, Heidelberg: Springer.
- 14 Berkel, J.J. van, Pool, R.L.D., Harbers, M., Oerlemans, J.J., Bargh, M.S., Braak, S.W. van den (2017) *Verkeerd verbonden in een slimme samenleving: het Internet of Things, kansen, bedreigingen en maatregelen*. Den Haag: WODC.
- 15 <https://www.techware.com.au/it-services-blog/the-internet-of-things-does-anyone-really-need-a-washing-machine-that-can-send-sms>
- 16 <https://informationmatters.net/smart-toothbrushes-iot/>
- 17 <https://www.law.kuleuven.be/citip/blog/smart-dolls-a-triple-threat-to-children-and-their-rights/>
- 18 <https://www.smartcitiesworld.net/special-reports/special-reports/smart-sewers-smart-cities-start-eight-feet-below-the-ground>
- 19 <https://www.smartupcities.com/smart-waste-containers/>
- 20 <https://www.denhaag.nl/nl/in-de-stad/wonen-en-bouwen/bouwprojecten/gebiedsontwikkeling-scheveningen-kust/slimme-lantaarnpalen-in-scheveningen.htm>
- 21 Kitchin, R. (2004) The real-time city? Big data and smart urbanism, *GeoJournal*, vol. 79(1), p. 1–14; Hiller, J. and Blanck, J. (2016) Smart Cities, Big Data, and the Resilience of Privacy, *Hastings Law Journal*, vol. 68, p. 309–356.
- 22 Voor meer over het gebrek aan definitie en de complexiteit van big data, zie: Boyd, D., Crawford, K. (2011) Six Provocations for Big Data. A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, September 2011. SSRN: <https://ssrn.com/abstract=1926431> or <http://dx.doi.org/10.2139/ssrn.1926431>
- 23 Gandomi, A., & Haider, M. (2015) Beyond the hype: Big data concepts, methods and analytics, *International Journal of Information Management*, 35, p. 137–144.
- 24 Ball, P. (2004) *Critical Mass; How One Thing Leads to Another*, New York: Farrar, Straus and Giroux.
- 25 Laney, D. (2001) *3D Data Management: Controlling Data Volume, Velocity and Variety*. Gartner. Stamford CT: META Group Inc.
- 26 Calders T. & Custers B.H.M. (2013), What is data mining and how does it work?. In: Custers B.H.M., Calders T., Schermer B., Zarsky T. (Eds.) *Discrimination and Privacy in the Information Society*. no. 3 Heidelberg: Springer.
- 27 Gandomi and Haider (2015) schatten dat 95 % van de big data ongestructureerd is.
- 28 Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT, USA: Yale University Press.
- 29 Voor meer voorbeelden, zie de special issue van de *University of Chicago Law Review*, Volume 86, Issue 2, April 2019, p. 217–609. <http://lawreview.uchicago.edu/volume-86-issue-2-april-2019-217-609>.
- 30 Ranchordás, S. (2019) Nudging Citizens through Technology in the Smart City, *International Review of Law, Computers & Technology*, <https://doi-org.proxy-ub.rug.nl/10.1080/13600869.2019.1590928>.
- 31 Rebonato, R. (2012) *Taking Liberties: A Critical Examination of Libertarian Paternalism*. London: Palgrave MacMillan.
- 32 Bovens, L. (2009) The Ethics of 'Nudge'. In *Preference Change: Theory and Decision Library*. Vol. 42., edited by T. Grüne-Yanoff and S. O. Hansson, 207–219. Dordrecht: Springer.
- 33 Custers, B.H.M., and Ursic, H. (2016) Big Data and Data Reuse; A Taxonomy of Data Reuse for Balancing Big Data Benefits and Personal Data Protection, *International Data Privacy Law*, 1–12. doi: 10.1093/idpl/ipv028.
- 34 Intellectueel Eigendomsrecht wordt buiten beschouwing gelaten, omdat dit niet van toepassing is op persoonsgegevens. In Europa, in tegenstelling tot bijvoorbeeld de Verenigde Staten en China, bestaan geen eigendomsrechten op persoonsgegevens.
- 35 Custers, B.H.M., Calders, T., Schermer, B., and Zarsky, T. (eds.) (2013) *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, Heidelberg: Springer.
- 36 Boyd, D., and Crawford, K. (2012) Critical questions for big data: provocations for a cultural, technological and scholarly phenomenon, *Information, communication & society*, 15(5), p. 662–679; Tene, O., and Polonetsky, J. (2012) Privacy in the age of big data: a time for big decisions, *Stanford Law Review Online*, 64, p. 63.
- 37 Ranchordás, S. and Klop, A. (2018) Data-Driven Regulation and Governance in Smart Cities, in *Research Handbook on Data Science and Law*, Edward Elgar Publishing; Zoonen, L. van (2016) Privacy Concerns in Smart Cities. *Government Information Quarterly* 33 (3): 472–480. doi:10.1016/j.giq.2016.06.004; NESTA (2018) *Reclaiming the Smart City: Personal Data, Trust and the New Commons*. https://media.nesta.org.uk/documents/DECODE-2018_report-smart-cities.pdf; Privacy International (2017) *Smart cities: Utopian vision, Dystopian Reality*. <https://privacyinternational.org/sites/default/files/2017-12/Smart%20Cities-Utopian%20Vision%2C%20Dystopian%20Reality.pdf>.
- 38 McArthur, R.L. (2001) Reasonable expectations of privacy, *Ethics and Information Technology*, 3:123.
- 39 <https://www.theguardian.com/technology/2015/aug/19/ashley-madison-hackers-release-10gb-database-of-33m-infidelity-site-accounts>
- 40 Rechtbank Noord-Nederland 3 mei 2017, nr. C/18/170676/HA RK 16/269.

- 41 Custers B.H.M. (2012), Predicting Data that People Refuse to Disclose; How Data Mining Predictions Challenge Informational Self-Determination, *Privacy Observatory Magazine* 2012(3).
- 42 Baveye, Y., Bettinelli, J.N., Dellandrea, E., Chen, L. and Chamaret, C. (2013) A large video data base for computational models of induced emotion. In *Proceedings of Humane Association Conference on Affective Computing and Intelligent Interaction*; Borth, D., Chen, T., Ji, R.R., and Chang, S.F. (2013) Sentibank: Large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In *Proceedings of ACM Multimedia*.
- 43 Kosinski, M., Stillwell, D. & Graepel, T. (2012) Private traits and attributes are predictable from digital records of human behaviour, *Proceedings of the National Academy of Sciences (PNAS)*, www.pnas.org/content/early/2013/03/06/1218772110.
- 44 Custers, B.H.M. (2012) Predicting Data that People Refuse to Disclose; How Data Mining Predictions Challenge Informational Self-Determination, *Privacy Observatory Magazine*, Issue 3. See <http://www.privacyobservatory.org/>
- 45 Ohm, P. (2010) Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review* 57: 1701.
- 46 Voor meer over dit 'recht op vergetelheid', zie: J Ausloos, 'The "Right to Be Forgotten"—Worth Remembering?' (2012) 28 *CLSR* 143, 152; Graux, J Ausloos, and P Valcke, 'The Right to Be Forgotten in the Internet Era' (2012) 11 *ICRI Research Paper*; G Finocchiaro and A Ricci, 'Quality of Information, the Right to Oblivion and Digital Reputation' in B Custers, T Calders, B Schermer, and T Zarsky (eds), *Discrimination and Privacy in the Information Society, Data Mining and Profiling in Large Databases* (Springer Verlag, Berlin, Heidelberg 2013); Werro, 'The Right to Inform v the Right to be Forgotten: A transatlantic Clash' in AC Ciacchi, C Godt, P Rott, and LJ Smith (eds), *Haftungsbereich im dritten Millennium/Liability in the Third Millennium (Nomos, Baden-Baden 2009)* 291; BJ Koops, 'Forgetting Footprints, Shunning Shadows' (2011) 8 *SCRIPed* 3, 5; M. L. Rustad, S. Kulevska, *Reconceptualizing the right to be forgotten to enable transatlantic data flow*, 28 *Harv. J.L. & Tech.* 349, Spring, 2015; E. Bougiakiotis, *The enforcement of the Google Spain ruling*, *International Journal of Law and Information Technology*, 2016;
- 47 Voor meer over predictive policing, zie Richardson, R., et al. (2019) *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice* (February 13, 2019). *New York University Law Review Online*, Forthcoming. Available at SSRN: ssrn.com/abstract=3333423; Ferguson, A.G. (2017) *Policing Predictive Policing*. *Washington University Law Review*, Vol. 94, No. 5, 2017. SSRN: <https://ssrn.com/abstract=2765525>
- 48 Pedreschi, D., Ruggieri, S., Turini, F. (2013) The discovery of discrimination, in: Custers, B.H.M., Calders, T., Schermer, B., and Zarsky, T. (eds.) (2013) *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, Heidelberg: Springer.
- 49 Zliobaite I. & Custers B. (2016), Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models, *Artificial Intelligence and Law* (24): 183-201.
- 50 Kamiran, F., and T. Calders. 2009. Classification without discrimination. In *IEEE international conference on computer, control & communication (IEEE-IC4)*, 17-19 February 2009, Karachi, Pakistan.
- 51 Squires G (2003) Racial profiling, insurance style: insurance redlining and the uneven development of metropolitan areas. *Journal of Urban Affairs*, 25(4):391-410; Hillier A (2003) Spatial analysis of historical redlining: a methodological explanation. *Journal of Housing Research*, 14(1):137-168.
- 52 Elkin-Koren, N., Gal, M.S. (2019) *The Chilling Effect of Governance-by-Data on Data Markets*, *University of Chicago Law Review*, 86, p. 402-431.
- 53 Zie ook Westin, A. (1967) *Privacy and Freedom*, London: Bodley Head.
- 54 McDonald, A.M. and Cranor, L.F. (2008) The cost of reading privacy policies. *I/S Journal for Law and Policy for the Information Society*. <http://www.aleecia.com/authors-drafts/readingPolicyCost-AV.pdf>.
- 55 Van der Hof, S. and Van den Berg, B. (2011) Endorse Deliverable D2.2: Social Requirements and Implications. *ENDORSE Legal Technical Framework for Privacy Preserving Data Management*, p. 38.
- 56 Toubiana, V., and Nissenbaum, H. (2001) An Analysis of Google Logs Retention Policies, *Journal of Privacy and Confidentiality*, 3 (1): Article 2.
- 57 Solove, D.J. (2013) Privacy self-management and the consent dilemma. *Harvard Law Review*, 126, p. 1880-1903. Zie ook: Hintze, M. (2017) In Defense of the Long Privacy Statement (February 8, 2017). *Maryland Law Review*. Beschikbaar via SSRN: <https://ssrn.com/abstract=2910583>.
- 58 Lessig, L. (2006) *Code Version 2.0*, New York: Basic Books.
- 59 Schermer, B.W. and Wagemans, T. (2009) *Onze digitale schaduw*, Amsterdam: Considerati.
- 60 Custers B.H.M., Hof S. van der & Schermer B. (2014), *Privacy Expectations of Social Media Users: The Role of Informed Consent in Privacy Policies*, *Policy and Internet* 6(3): 268-295.
- 61 <https://www.rtlnieuws.nl/economie/home/letter-achter-je-huisnummer-verzekering-100-euro-duurder>.
- 62 Solove, D. (2004) *The Digital Person; Technology and Privacy in the Information Age*, New York: New York University Press.
- 63 Koops, B.J. (2014) The trouble with European data protection law, *International Data Privacy Law*, doi: 10.1093/idpl/ippu023
- 64 Grassegger, H. & Krogerys, M. (2017) *The Data That Turned the World Upside Down*, Motherboard, 28 januari 2017.
- 65 Custers B.H.M. (2003), *Effects of Unreliable Group Profiling by Means of Data Mining*. In: Grieser G, Tanaka Y, Yamamoto A (red.) *Lecture Notes in Artificial Intelligence*. Heidelberg, New York: Springer Verlag. 290-295.
- 66 Dit wordt ook wel aangeduid als de privacy paradox of profiling paradox. Zie Custers B.H.M. (2004), *The Power of Knowledge*. Tilburg: Wolf Legal Publishers.
- 67 In het Engels: *General Data Protection Regulation (GDPR)*.
- 68 See <http://www.oecd.org/dsti/sti/it/secur/prod/PRIV-EN.HTM>
- 69 Council of Europe, Convention no. 108, January 28th 1981. *Convention for the protection of individuals with regard to automatic processing of personal data*.
- 70 Voor meer informatie, zie Besselink, H.J.M, Dun, P.L.J.M. van, Waard, B.W.N. de (2011) *Bestuursprocesrecht*, Meppel: Boom Juridische Uitgevers. Joris in 't Veld en N.S.J. Koeman (1979), *Beginselen van behoorlijk bestuur*, Zwolle. ISBN 9027126615
- 71 <https://wetten.overheid.nl/BWBR0005537/2019-04-02>
- 72 Zie ook art. 40 AVG.
- 73 Leetaru, K. (2018) When Will Cities Begin to Monetize Their Resident's Data? *Forbes*, 19th July 2018. <https://www.forbes.com/sites/kalevleetaru/2018/07/19/when-will-cities-begin-to-monetize-their-residents-data/#4a6a63c04661>
- 74 Bianchini, D., Avila, I. (2014) Smart Cities and Their Smart Decisions: Ethical Considerations, in: *IEEE Technology and Society Magazine*, vol. 33, no. 1, pp. 34-40, Spring 2014. doi: 10.1109/MTS.2014.2301854. Zie ook Kitchin, R. *The ethics of smart cities and urban science*, *Phil. Trans. R. Soc. A*, vol. 374(2083).
- 75 Vallor, S. (2016). *Technology and the virtues: A philosophical guide for a future worth wanting*. New York: Oxford University Press.

- 76 La Fors, K., Custers, B.H.M., Keymolen, E.L.O. (2019) Reassessing values for emerging big data technologies: integrating design based and application based approaches. *Ethics and Information Technology*, <https://doi.org/10.1007/s10676-019-09503-4>
- 77 <https://dataschool.nl/deda/>
- 78 Mepham, B. (2000) A framework for the ethical analysis of novel foods: The Ethical Matrix <<https://link.springer.com/article/10.1023/A:1009542714497>>
- 79 Kort, M. and Klijn, E-H. (2013) Public–Private Partnerships in Urban Regeneration: Democratic Legitimacy and its Relation with Performance and Trust, *Local Government Studies*, vol. 39(1), p. 89–106.
- 80 Een beperkt aantal kleinschalige pilots vermindert ook het risico op gebrekkige samenhang tussen pilots.
- 81 Het opschalen van pilots naar praktische implementatie kan soms lastig zijn, maar de haalbaarheid wordt vergroot wanneer de weg van geleidelijkheid wordt gekozen. Hindernissen betreffende maatschappelijk draagvlak en juridische onderbouwing kunnen dan beter worden ondervangen.
- 82 Zie art. 35 AVG. In het Engels wordt ook wel gesproken van een Data Protection Impact Assessment (DPIA).
- 83 Met andere woorden, de (juridische en ethische) experimenteerruimte in het kader van pilots is niet groter dan de ruimte voor de praktijk.
- 84 Zie art. 40 AVG.

Digitalisering, gedragsbeïnvloeding en de overheid

Het gladde ijs tussen nudging en überveillance

Prof. dr. Wijnand IJsselsteijn
Hoogleraar Cognition and Affect in Human-Technology Interaction
Eindhoven University of Technology (TU/e)



Voorwoord

Door prof. dr. Maurits Clemens Kaptein

Het essay *Digitalisering, gedragsbeïnvloeding en de overheid* behandelt een van de - in mijn ogen - belangrijkste consequenties van datagebruik in de openbare ruimte: het gebruik van data om actief het gedrag van burgers te sturen. Zelf ben ik, als hoogleraar Data Science & Health aan de Universiteit van Tilburg en als principal investigator van het Computational Personalization lab aan de Jheronimus Academy of Data Science (JADS) te 's-Hertogenbosch, al jaren bezig met onderzoek naar het gebruik van data om menselijk gedrag te begrijpen en, uiteindelijk, te beïnvloeden. We kunnen de ontwikkelde kennis gebruiken voor positieve maatschappelijke doeleinden; we hebben bijvoorbeeld laten zien dat middels data gepersonaliseerde feedback ertoe leidt dat we effectievere eHealth-applicaties kunnen ontwikkelen. Met data kunnen we mensen dus beter ondersteunen bij het leiden van een gezonder leven. Echter, dezelfde kennis wordt ook gebruikt voor dubieuzere doelen; het gebruik van persuasion profiling in de online marketing - een manier van gedragsbeïnvloeding waar ik ook zelf veel over heb geschreven - wordt terecht kritisch bediscussieerd.

In zijn essay bespreekt prof. dr. Wijnand IJsselsteijn van de Technische Universiteit Eindhoven hoe de verzameling van data in de openbare ruimte ertoe leidt dat we, potentieel, steeds effectiever menselijk gedrag in de openbare ruimte kunnen sturen. Simpele voorbeelden van gedragsbeïnvloeding in de openbare ruimte zijn bij iedereen bekend: tijdens grote evenementen zoals de Nijmeegse Vierdaagse wordt het publiek gemonitord en door middel van grote schermen aangemoedigd bepaalde locaties te vermijden. Echter, er zijn ook veel subtielere vormen van beïnvloeding, bijvoorbeeld door middel van de straatverlichting, die niet bekend zijn bij het grote publiek.

Daarbij is het niet altijd duidelijk wie er aan de spreekwoordelijke knoppen draait bij pogingen tot gedragsbeïnvloeding in de openbare ruimte, noch is het altijd duidelijk welke data, van wie, en door wie, wordt verzameld om de technologie mogelijk te maken. Juist als het gaat over de openbare ruimte zijn dit punten waar we gezamenlijk het debat over aan moeten gaan.

Op basis van zijn ruime ervaring als onderzoeker naar de manieren waarop technologie mensen beïnvloedt, en naar manieren om gebruikers van technologie deze technologie te kunnen laten begrijpen, geeft prof. dr. IJsselsteijn nieuwe inzichten in het gebruik van data in de openbare ruimte. Er is wat mij betreft niemand die meer ervaring en kennis heeft op dit gebied dan Wijnand.

De deelname aan de commissie Behoorlijk datagebruik in de openbare ruimte was voor mijzelf erg leerzaam. Het was bemoedigend om te zien hoe overheid, wetenschap en bedrijfsleven, vanuit verschillende achtergronden, bij elkaar zijn gekomen om na te denken over dit maatschappelijk belangrijke en beladen onderwerp. Ik heb persoonlijk veel geleerd van de discussies en kan mij volledig vinden in onze aanbevelingen (waarvan ik een groot deel voorafgaand aan het proces nooit had kunnen voorzien). Ik hoop dat voor u als lezer hetzelfde geldt.

*Prof. dr. Maurits Clemens Kaptein
Hoogleraar Data Science & Health, Jheronimus Academy of Data Science (JADS)
Lid begeleidingscommissie*

Samenvatting

Digitalisering gaat snel en heeft een grote impact op onze leefwereld. Ook de overheid wordt geconfronteerd met de digitale transformatie, die zowel ongekende mogelijkheden als nieuwe uitdagingen met zich mee brengt. Een aantal pionierende steden experimenteert met digitale technologie in de publieke ruimte, in smartcity-projecten en living labs, om zodoende de efficiency van overheidsdiensten te vergroten, maar ook om de leefbaarheid en duurzaamheid van de stad te verbeteren. Op gebieden waar de overheid een actief preventiebeleid voert – veiligheid, volksgezondheid, milieu - wordt ook nagedacht over de inzet van digitale middelen in de publieke ruimte - sensoren, big data en kunstmatige intelligentie - om gedrag van burgers te meten, te voorspellen en te beïnvloeden.

We kunnen hierbij denken aan slimme straatverlichting om agressie te voorkomen of een interactief wandelpark waar gezond bewegen wordt gestimuleerd. Deze beïnvloeding vindt veelal plaats op basis van kennis uit de psychologie en staat bekend als persuasive technology of nudging. De mogelijkheden van gedragsbeïnvloeding door gebruik van digitale technologie in de publieke ruimte gaan gepaard met soms lastige technische, sociale, juridische en ethische vraagstukken. Hoe, en in hoeverre, mag de overheid sturen op gewenst gedrag van haar burgers, gebruik makend van publieke datastromen, kunstmatige intelligentie en psychologische beïnvloedingstechnieken? En hoe dient de overheid zichzelf te reguleren wanneer ze geconfronteerd wordt met de verleidingen van macht en controle die digitalisering met zich meebrengt?

In dit essay ga ik nader in op de mogelijkheden en beperkingen van digitale technologie die wordt gebruikt voor gedragsbeïnvloeding, en sta ik stil bij de ethische vragen die worden opgeworpen wanneer een overheid zich begeeft op het gladde ijs tussen nudging en überveillance. Naast een kritische analyse komen er ook oplossingsrichtingen aan de orde, waarbij participatoir ontwerp en digitale geletterdheid centraal staan.



1. Inleiding

In September 2002 kwam *Minority Report* uit, een sciencefiction-film gebaseerd op het gelijknamige boek van Philip K. Dick. De hoofdpersoon uit de film, agent John Anderton (gespeeld door Tom Cruise), is deel van PreCrime, een gespecialiseerde eenheid van de politie die poogt om criminelen aan te houden, nog voordat ze hun misdaad hebben begaan, op basis van voorkennis van drie paranormaal begaafde ‘precogs’. *Minority Report* is een *film noir* over een toekomst waarin mensen voortdurend door biometrische sensoren in de publieke ruimte worden gescand, en hun gedrag in de gaten wordt gehouden. Privacy behoort in 2054 tot het verleden. Mensen worden door een lawine van media en gepersonaliseerde *product placement* overspoeld, terwijl een technisch geavanceerde overheid tracht de criminaliteit naar nul terug te brengen, met mobiele eenheden uitgerust met jetpacks, miniatuur-robots, en heel veel interactieve beeldschermen.

Minority Report speelt zich af over 35 jaar vanaf het moment van dit schrijven. Kijken we 35 jaar terug, dan belanden we in 1984, een jaar dat geassocieerd is met een andere, nog bekendere dystopie: George Orwell’s totalitaire regime van Big Brother, waar constante surveillance en gedragssturing door de overheid bijdragen aan de totale onderwerping van de burger. Het mag duidelijk zijn dat niemand deze scenario’s als wenselijk ervaart, maar ze helpen ons om kritisch te reflecteren op de waarden die in het gedrang komen wanneer een overheid zich vergrijpt aan macht, en toegeeft aan de verleidingen van technologische controle over haar burgers. Het zijn ook de verhalen die ons confronteren met de risico’s van te eenzijdige of te ver doorgevoerde keuzes binnen bekende spanningsvelden – spanningsvelden die ook in onze huidige maatschappij voelbaar zijn: het belang van het individu versus het belang van het collectief, privacy versus veiligheid, het gereguleerde en rigide ‘systeem’ versus het onvoorspelbare individu in een immer veranderende en complexe realiteit.

Terug naar onze tijd. In het kielzog van de wet van Moore¹, worden computers steeds krachtiger, compacter, goedkoper, slimmer, meer verbonden met elkaar – het internet – én met de fysieke wereld – het *Internet of Things*. Bovendien worden computers in toenemende mate persoonlijk en intiem – we zijn van desktop naar laptop naar smartphone, smartwatch en *personal tracker* gegaan. We zijn onafscheidelijk van onze mobiele telefoons, die overal met ons meereizen, en waarmee we een breed scala aan persoonlijke informatie delen – onze locaties,

sociale interacties en netwerken, persoonlijke foto’s en video’s, gezondheid en fitness data, en onze meest persoonlijke voorkeuren en gewoontes in aankoopgedrag, voeding, entertainment, seksualiteit, enzovoort². Al deze informatie wordt verzameld, overal en continu, en gedeeld met een, letterlijk, ongekend aantal partijen – veelal met een commercieel belang om het individu aan te sporen nog meer informatie te delen, om zo met *microtargetting* en *hypernudging* mensen te overtuigen om te kopen, te consumeren, te ervaren, te *liken*, en te delen. In de nabije toekomst dient zich nog intiemere technologie aan, in letterlijke zin: nog dichter op en onder de huid. Denk aan *augmented-reality*-brillen, sensoren in kleding en sieraden, onderhuidse chips en nanotechnologie in het lichaam zelf.

Als we uitzoomen, van micro naar macro, zien we een vergelijkbare trend. Ook onze directe leefomgeving en de publieke ruimte waarin we ons bewegen worden in toenemende mate uitgerust met sensoren³, waarmee grootscheeps data worden verzameld, veelal door commerciële marktpartijen, maar ook, en steeds frequenter, door de overheid. De overheid heeft goede redenen om informatie te willen verzamelen. Meer gedetailleerde informatie – zowel over situaties als over doelgroepen – heeft belangrijke meerwaarde bij het doelmatig en efficiënt uitvoeren van kerntaken van de overheid.

De overheid heeft goede redenen om informatie te willen verzamelen

Wat de overheid beschouwt als haar pakket aan kerntaken is aan verandering onderhevig. Waar in de eerste helft van de twintigste eeuw de overheid zich vooral richtte op het handhaven van wetsorde en veiligheid, en het op orde houden van infrastructuur – de overheid als ‘nachtwakersstaat’ – heeft de Nederlandse overheid in de tweede helft van de twintigste eeuw een veel actievere rol genomen in het stimuleren van gezondheid en goed onderwijs, en het garanderen van volkshuisvesting en sociale zekerheid – de verzorgingsstaat. In de 21^e eeuw is ook deze set aan taken aan verandering onderhevig, door wisselend politiek draagvlak, uitdagingen in kostenbeheersing en efficiëntie, en de daarmee verband houdende voortschrijdende privatisering en deregulering.

In lijn met de veranderende taakopvatting, zien we de overheid steeds vaker sturen via preventiebeleid – in het bijzonder op het gebied van veiligheid en gezondheidszorg, maar ook op het gebied van sociale rechtvaardigheid, is de rol van de overheid zichtbaar aan het wijzigen. Waar sinds eeuwen wetgeving en strafrecht-toepassing, belasting en subsidie, en voorlichting en communicatie de geëigende beïnvloedingsstrategieën zijn van de overheid, zien we in recente jaren een toenemende focus op gedragsbeïnvloeding op basis van psychologische gedrags-interventies, waarbij ook gebruik wordt gemaakt van digitale technologie. In een samenleving waarin burgers geconfronteerd worden met veel verleidingen, en waar controle afbrokkelt of ontbreekt, lijkt het opportuun om burgers te helpen de juiste keuzes te maken – een duwtje in de goede richting te geven⁴.

Een overvloed aan voedsel, vooral ongezond voedsel, de alomtegenwoordigheid van tabak en alcohol, en de prevalentie van 'zittend' werk in onze kennis- en diensteneconomie, zorgen voor een hausse aan welvaartsziektes. Dit zijn ernstige chronische aandoeningen, zoals bepaalde vormen van kanker, diabetes, of hart- en vaatziektes, die sterk verbonden zijn met keuzes in levensstijl, zoals roken, te veel drinken, te weinig beweging en te veel stress. Deze ziektes komen pas na verloop van tijd tot uiting, en zijn dan veelal lastig te genezen. Bij deze ziektes is preventie de meest effectieve vorm van interventie, maar die vereist een verandering in attitude en gedrag bij de individuele burger. Het vereist tevens een heroriëntatie van de gezondheidszorg, van een focus op ziektebestrijding naar een focus op gezondheidsbevordering en preventie van ziekte. Deze benadering staat ook centraal in het Nationaal Preventieakkoord zoals recentelijk geformuleerd door de staatssecretaris van VWS, Paul Blokhuis⁵.

Naast gezondheidszorg is ook veiligheid een belangrijk thema waar preventie een steeds centralere rol speelt in het overheidsbeleid. We zijn sinds de jaren 80 van de vorige eeuw getuige van een omwenteling in het denken rondom veiligheid – van een nadruk op repressie en een justitiële benadering van criminaliteit, naar de opkomst van veiligheidsdenken, waarbij niet de rechtsorde, maar de publieke orde centraal staat. De toename van individuele vrijheid en mobiliteit, de toename in culturele diversiteit, en het afbrokkelen van sociale controle, maakt dat burgers elkaar minder goed kennen, elkaars gedrag als minder voorspelbaar ervaren, en er een minder duidelijke gemeenschappelijke gedragsnorm heerst. De overheid, en met name de politie, komt dientengevolge steeds vaker in een normgevende rol

terecht. Net zoals gezondheid meer is dan de afwezigheid van ziekte, is veiligheid is meer dan de afwezigheid van criminaliteit. Dit wordt goed gekenschetst in de brief van de minister van Binnenlandse Zaken aan de Tweede Kamer inzake het *Integraal Veiligheidsprogramma*⁶ (1999):

“Veiligheid is te omschrijven als het aanwezig zijn van een zekere mate van ordening en rust in het publieke domein en van bescherming van leven, gezondheid en goederen tegen acute of dreigende aantastingen. Onveiligheid is te omschrijven als alles wat hierop inbreuk maakt. Die inbreuken op de veiligheid kunnen feitelijke aantastingen van de veiligheidssituatie betreffen, maar kunnen ook veiligheidsrisico's en gevoelens van onveiligheid betreffen.” (p.9)

Het is begrijpelijk dat westerse overheden op terreinen van gezondheid en veiligheid nadenken over subtiële gedragssturing, op basis van psychologische inzichten, en gebruik makend van de mogelijkheden die geboden worden door digitale middelen en de inrichting van de publieke ruimte.

Het thema dat centraal staat in dit essay is **gedragsbeïnvloeding in de openbare ruimte. In het bijzonder zullen we aandacht besteden aan het gebruik**, behoorlijk of onbehoorlijk, van digitale middelen, en in het bijzonder data, door de overheid, met het doel om kennis te vergaren en invloed te kunnen uitoefenen over het gedrag van haar burgers.



In dit essay zal ik nader ingaan op de belangrijkste vormen van digitale gedragsbeïnvloeding, te weten *persuasive technology* (hoofdstuk 2) en *nudging* (hoofdstuk 3). Ofschoon beide gebieden een wat andere academische achtergrond kennen, delen ze de karakteristiek dat men kennis vanuit de psychologie, de wetenschap van menselijk gedrag en denken, inzet om tot gedragsbeïnvloeding te komen. Vervolgens zal ik aandacht besteden aan de mogelijkheden, beperkingen en risico's die geassocieerd zijn met het gebruik van kunstmatige intelligentie, en in het bijzonder kunstmatig lerende netwerken (zgn. *deep learning*), om gedragsdata te analyseren, om zo tot valide classificaties en voorspellingen van gedrag te komen (hoofdstuk 4). Dit leidt ons in hoofdstuk 5 tot een beschouwing van de potentie tot misbruik van dergelijke digitale middelen van gedragsanalyse en beïnvloeding door overheid (*überveillance*) of bedrijfsleven (*surveillance kapitalisme*). Hoofdstuk 6 en 7 zijn gericht op constructieve oplossingsrichtingen, waarbij participatoir ontwerp (hoofdstuk 6) en digitale geletterdheid (hoofdstuk 7) centraal staan. Het essay sluit af met een slotbeschouwing en enkele aanbevelingen voor beleid (hoofdstuk 8).

2. Digitale gedragsbeïnvloeding: persuasive technology

Elkaar beïnvloeden en overtuigen is een wezenlijk onderdeel van de menselijke interactie. Van de slang in het Hof van Eden tot onze huidige sociale media – we worden volcontinu blootgesteld aan pogingen om ons te overtuigen en ons gedrag te beïnvloeden. Koop ons product, stem op onze partij, stop met roken, beweeg meer, drink minder, vrij veilig, bescherm het milieu. En eet een appel.

Sinds lange tijd speelt mediatechnologie een belangrijke rol bij het beïnvloeden van gedrag – van megafoons tot billboards tot televisie. De overheid maakt hier regelmatig gebruik van, en met aanzienlijk succes. Denk aan de mediacampagnes waarmee de overheid een sociale norm probeert te beïnvloeden, zoals veilig omgaan met vuurwerk (“Je bent een rund als je met vuurwerk stunt”), niet met drank op autorijden (de Bob-campagne), de anti-roken-campagne (“Roken is zoou Gordon’s kapsel uit 1991”), of de meer recente ‘Rij Mono’-campagne om het gebruik van sociale media achter het stuur te ontmoedigen.

Technologie wordt extra krachtig als communicatiemiddel wanneer er sprake is van een interactie in plaats van eenrichtingsverkeer, dat wil zeggen, wanneer de beïnvloedende partij een boodschap qua inhoud, vorm, timing of locatie kan aanpassen aan kenmerken of reacties van de te beïnvloeden ontvanger. Dit besef heeft geleid tot de ontwikkeling van zgn. *persuasive technology* – technologie die is ontworpen met het expliciete doel om attitudes te beïnvloeden en gedrag te veranderen. Een voorbeeld van persuasive technology is een programma op de computer dat iemand helpt herinneren om af en toe op te staan, uit te rekken, of even een wandelingetje te maken, om zo RSI en hart- en vaatziekten tegen te gaan. Een ander voorbeeld is een slimme thermostaat die feedback geeft over het dagelijks energieverbruik en helpt om energiezuiniger instellingen te gebruiken. Een derde voorbeeld is het gebruik van een smartwatch of *fitness tracker* om informatie en motiverende feedback te geven die kan helpen om meer te bewegen.

Persuasive technology - een programma op de computer dat iemand helpt herinneren om af en toe een wandelingetje te maken

Deze technologie-voorbeelden hebben gemeen dat er iets gemeten wordt aan de gebruiker en/of haar context – het aantal uren onafgebroken gebruik van toetsenbord en muis, de temperatuurfluctuaties in huis, of het aantal stappen dat iemand loopt per dag – en op basis van die informatie feedback wordt gegeven en suggesties voor gedragsverandering worden gedaan. Belangrijk is dat het hier gaat om *vrijwillig* en *geïnformeerd* gebruik van deze technologie. Wanneer er sprake is van dwang of misleiding dan valt dit strikt genomen buiten de definitie-kaders van persuasive technology, hoewel we hier later toch op terug zullen komen. In de bovenstaande voorbeelden kiezen mensen zelf voor het gebruik van de persuasive technology, en zijn de doelen van de gebruikte technologie opgelijnd met doelen die de gebruiker zelf nastreeft.



Als inherent sociaal diersoort zijn mensen zelf de sterkste beïnvloeders van elkaars gedrag en mening. Interpersoonlijke interactie is dan ook de sterkste vorm van beïnvloeding. Mensen hebben een sterke sociale aanwezigheid en impact, hebben een intuïtief besef van de juiste timing en context voor effectieve beïnvloeding, kunnen de ontvankelijkheid van de te beïnvloeden persoon inschatten, en hebben, vaak onbewust, een arsenaal aan beïnvloedingsgedrag in huis. Echter, zoals BJ Fogg aangeeft in zijn boek *Persuasive Technology*⁸, computers hebben ook een aantal karakteristieken die hen uniek geschikt maken voor gedragsbeïnvloeding. Een computer kan heel geduldig of vasthoudend zijn – hij zal nooit vermoeid of geïrriteerd raken. Een algoritme velt ook geen oordeel over gedrag, anders dan de respons-patronen die deel zijn van de geprogrammeerde beïnvloedingsstrategie. In die zin kan een computer ook de anonimiteit van een gebruiker waarborgen zoals een mens dat niet kan.

Een computer kan heel geduldig of vasthoudend zijn – hij zal nooit vermoeid of geïrriteerd raken

Daarnaast is een computer in staat, en wordt die ook in staat gesteld, om metingen te verrichten en feedback te geven op tijden en plaatsen wanneer en waar andere mensen niet zonder meer welkom zijn – in de privésfeer thuis, 's nachts, tijdens het eten, in de slaapkamer, in de badkamer, etc. Zoals al eerder opgemerkt, een computer – in de vorm van een mobiele telefoon of smartwatch, maar ook als slimme weegschaal, slimme babyfles, of als *smart speaker* – is intieme technologie. Mensen verwelkomen het in hun privésfeer, delen er persoonlijke informatie mee, en laten zich erdoor beïnvloeden.

Technologische ontwikkelingen in sensor-technologie, algoritmen, communicatie-infrastructuur en mens-machine- interfaces, bekend als *ubiquitous computing*, *pervasive computing*, *context-aware computing* of *ambient intelligence*⁹, creëren nieuwe mogelijkheden voor de toepassing van *persuasive technology*. Het *persuasive* systeem zal zo over steeds meer contextuele kennis kunnen beschikken, waardoor het mogelijk wordt om gevolgtrekkingen te maken ten aanzien van het gewoonte-

gedrag of de huidige activiteit van de individuele gebruiker, en de sociale en fysieke omgeving waarin hij of zij zich bevindt. Zo kan een systeem vaststellen of iemand aan het werk is, thuis televisie kijkt, aan het sporten is, of slaapt. En of iemand alleen is of met anderen. Het is dankzij fysiologische metingen aan of in de buurt van het lichaam ook steeds beter denkbaar de emotionele staat van een gebruiker te kunnen meten¹⁰.

Dit alles kan enorme voordelen met zich mee brengen voor de doelmatigheid van de *persuasive* communicatie. Zo kunnen aan gezondheid gerelateerde berichten worden verstuurd op het juiste tijdstip en op de juiste locatie (zgn. *just-in-time*, *just-in-place messaging*), bijvoorbeeld om iemand bij binnenkomst van een gebouw te helpen herinneren de trap te nemen in plaats van de lift, of bij het winkelen te wijzen op opties voor een gezonde maaltijd¹¹. Ook kan men de berichten zo timen dat ze niet worden afgeleverd tijdens een drukke vergadering of terwijl men achter het stuur zit, maar op een moment dat de gebruiker de mogelijkheid heeft er aandacht aan te besteden. Tenslotte kan men de beïnvloedingsstrategie, zowel wat betreft vorm als inhoud, aanpassen aan persoonlijke kenmerken van de gebruiker. Dit kan op basis van persoonlijkheidskenmerken, gewoontegedrag, gevoeligheid voor (of antipathie tegen) bepaalde beïnvloedingsstrategieën, of andere relevante persoonskenmerken. Deze combinatie van *ambient intelligence* en *persuasive technology* wordt ook wel aangeduid met de term '*ambient persuasion*'¹².

3. Nudging: een duwtje in de goede richting?

In 2017 werd de Nobelprijs voor de economie toegekend aan Richard Thaler, van de Universiteit van Chicago, voor zijn werk gericht op het toepassen van psychologische kennis binnen de economie. Het werk is schatplichtig aan een uitgebreide set aan psychologische studies, in het bijzonder bevindingen uit de cognitieve psychologie rondom beslissingen en keuzegedrag van onder andere collega-Nobelprijswinnaar Daniel Kahneman, waarin wordt aangetoond dat het beeld van de mens als rationele beslisser, de *homo economicus*, onjuist is. We nemen veel beslissingen op de automatische piloot, zonder uitgebreide denkprocessen of redeneringen. We beslissen in een oogwenk, op basis van signalen uit de omgeving en een beperkt aantal heuristieken of vuistregels.

Deze manier van denken heeft het voordeel van snelheid, niet onbelangrijk gegeven ons evolutionaire *do-or-die*-verleden, maar heeft als nadeel dat het gevoelig is voor een aantal systematische denkfouten¹³. Deze denkfouten hebben te maken met het feit dat de mens beperkt is in de aandacht die men heeft voor relevante aspecten uit de omgeving, selectief is in het herinneren van informatie, kansen en risico's slecht kan inschatten, en een sterke focus heeft op onmiddellijke, korte-termijn-veranderingen en -effecten, terwijl substantiële lange-termijn-effecten worden genegeerd of onderschat.

Thaler en collega Cass Sunstein introduceerden in 2008, met hun boek *Nudge*¹⁴, het idee van *nudging*: het aanpassen van de keuzecontext waardoor mensen vrijwel automatisch betere beslissingen nemen die uiteindelijk tot een gezonder, welvarender en gelukkiger leven leiden. Nudges zijn als een vriendelijk elleboogstootje, een duwtje in de goede richting, om daarmee een keuze te beïnvloeden zonder dat het keuzevrijheid inperkt of dwingt tot bepaald gedrag. Een bekend voorbeeld is het plaatsen van gezonde producten op ooghoogte in de schappen van een winkel, waardoor mensen substantieel vaker voor gezonde producten kiezen. Een ander voorbeeld is het gebruik van kleinere borden in een schoolkantine, om zodoende de hoeveelheid eten dat wordt geconsumeerd terug te brengen. Een derde voorbeeld is het gebruik van verkeersdrempels om de snelheid waarmee een auto door een woonwijk rijdt te beperken. Met het ontwerp van een dergelijke zgn. keuze-architectuur tracht men negatieve effecten van onze inadequate psychologische beslisprocessen te voorkomen of te compenseren. Door de keuzecontext goed te structureren kan men meer wenselijke gedragsroutines en automatismen ondersteunen en activeren.

Gedragsbeïnvloeding door middel van keuze-architectuur impliceert overigens niet dat we volledig zijn overgeleverd aan de grillen van de keuze-architect, en we niet bewust en reflectief kunnen kiezen voor een alternatieve keuze. Voor Thaler en Sunstein is het van belang dat de keuzecontext onderhandelbaar blijft, en we ons vrij voelen in het overwegen van alternatieven. Deze benadering wordt door Thaler en Sunstein 'libertair paternalisme' genoemd¹⁵. Het is *paternalistisch* omdat de keuze-architecten het keuzegedrag van mensen trachten te beïnvloeden uitgaande van een positieve gedragsnorm (bijv. gezond leven), vergelijkbaar met hoe een ouder of verzorger de keuzes van zijn of haar kinderen zou willen beïnvloeden. Het is *libertair* omdat een centraal uitgangspunt van een nudge is dat deze de

vrijheden niet inperkt maar mensen in staat stelt de nudge te negeren en hun eigen gang te (blijven) gaan. We kunnen nog steeds het ongezonde product pakken in de winkel, meerdere borden met eten volscheppen in de kantine, of met de auto op hoge snelheid langs een andere route rijden. Als er geen mogelijkheid is om aan de positieve gedragsnorm te ontsnappen dan is er geen sprake van nudging, maar van manipulatie of dwang.

Het idee van nudging, en meer in zijn algemeenheid gedragsbeïnvloeding door gebruik te maken van psychologische deskundigheid, is niet onopgemerkt gebleven bij de overheid. Het libertair of zacht paternalisme biedt een extra set aan krachtige beïnvloedingsstrategieën die naast meer traditionele benaderingen kunnen worden ingezet. Thaler en Sunstein (2008) beargumenteren dat nudging over het algemeen kosten-efficiënt is, en het libertair paternalisme zich ideëel in het midden van het sociale versus liberale politieke spectrum bevindt – een overheid die sociale verantwoordelijkheid stimuleert, zonder individuele vrijheden in te perken. Daarmee lijkt libertair paternalisme een heel aantrekkelijk beleidsalternatief voor een overheid die zich geplaatst ziet voor grote maatschappelijke uitdagingen die zijn terug te voeren op individueel keuzegedrag, en waarbij de overheid in toenemende mate een rol speelt van opvoeder.



Waar de oorspronkelijke voorbeelden van nudging zich vooral richten op het herontwerp van een statische keuzecontext in de publieke ruimte, en persuasive technology zich meer heeft gericht op de dynamische interactie tussen het individu en haar ICT-tools, is door de introductie van sensoren, big data⁶, en kunstmatige intelligentie in de publieke ruimte het onderscheid tussen nudging en persuasive technology aan het vervagen. De analyse van bigdata-patronen kan leiden tot het precies en doelmatig richten van een interventie op de eigenschappen, gewoontes, voorkeuren en interesses van een individu, een groep, een wijk of een stad. Men heeft het in deze context ook wel over *hypernudging*¹⁷ of *big nudging*¹⁸. Het verschil met 'traditionele' nudges, als we daarvan mogen spreken, is dat het systeem niet enkel passieve interventies aanbiedt, zoals bijvoorbeeld snelheidsdrempels of een meer opvallende plaatsing van gezond voedsel. In plaats daarvan is er sprake van een controle-loop, waarbij vanuit metingen (data) een model kan worden opgesteld, waarna een interventie specifiek kan worden toegesneden op de situatie en de doelgroep. Vervolgens kan men door metingen ook de effectiviteit van die interventie in kaart brengen, en deze data kunnen weer gebruikt worden om het model verder te optimaliseren. Enzovoorts.

De analyse van bigdata-patronen kan leiden tot het precies en doelmatig richten van een interventie op de eigenschappen, gewoontes, voorkeuren en interesses van een individu, een groep, een wijk of een stad

Er is dus sprake van een cybernetische¹⁹ controle-lus of feedback-loop. Het achterliggende model kan variëren van een aantal eenvoudige 'als-dan'-regels tot een vrij uitgebreid kennissysteem of een neurale netwerk dat op basis van de data in combinatie met leerregels het model dynamisch kan aanpassen. In het volgende hoofdstuk (De kunstmatig intelligente overheid?) zullen we nader ingaan op de mogelijkheden en beperkingen van een dergelijke benadering.

Zowel persuasive technology als nudging gebruiken psychologische kennis rondom menselijke cognitieve vermogens en beperkingen, en sociale wetmatigheden, om gedrag zo effectief mogelijk te beïnvloeden. Beiden gebruiken technologie om te meten aan menselijk gedrag, het gedrag te modelleren en te voorspellen, en dat gedrag op haar beurt weer te beïnvloeden. Zowel persuasive technology als nudging hebben tendensen van een zachte vorm van paternalisme. Paternalisme, omdat een autoriteit buiten het individu besluit wat gezond of verantwoord gedrag is, en er op basis van deze norm expliciet wordt ingegrepen in het ontwerp van de omgeving (inclusief de technologie) van een individu, en daarmee in zijn of haar gedrag. Zacht, omdat zowel persuasive technology als nudging de claim maken dat dit alles gebeurt op basis van vrijwilligheid en openheid.

Er is in de literatuur uitgebreide discussie over de wenselijkheid en ethische implicaties van het libertair paternalisme van Thaler en Sunstein en de paternalistische uitgangspunten van persuasive technology. Paternalisme kan worden gezien als een inbreuk op de autonomie van de burger die niet als een mondige en soevereine volwassene wordt behandeld. Wanneer we technologie de macht geven om ons te beïnvloeden, worden we dan niet gestuurd door apparaten, in plaats van zelf na te denken? Paternalisme impliceert ook vertrouwen in de goede bedoelingen en superieure expertise van hen die ons beïnvloeden: zij weten wat goed voor ons is, beter dan dat we dit zelf doen.

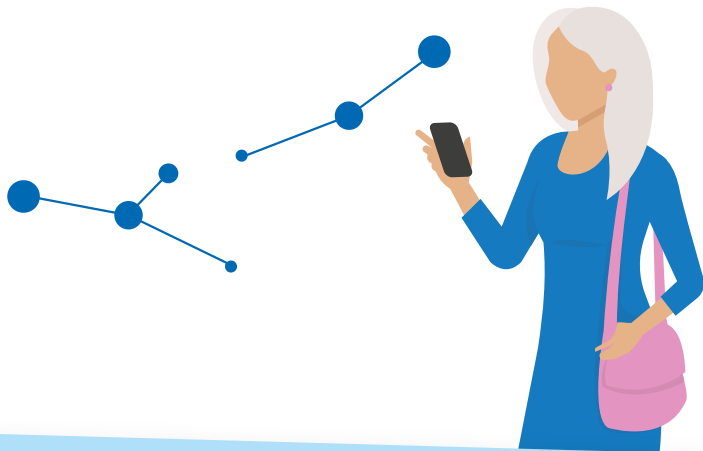
Zacht paternalisme is verdedigbaar wanneer de doelen die worden nagestreefd overeenkomen met de doelen van het individu zelf. Dat wil zeggen dat een interventie ervoor zorgt dat ofwel de persoon doet wat hij of zij zelf wil, ofwel de persoon wordt beschermd tegen krachten van buiten die het nastreven van de persoonlijke doelen in de weg staan (zoals bijv. verleiding). Dit staat in contrast met 'hard' paternalisme, waarbij een overheid of andere partij zich bedient van methodes om volledige vrijheid van handelen in te perken. Filosofen Lily Frank en Philip Nickel²⁰ beargumenteren dat het uitdagend is om precies te weten op welke manier doelen van het individu het beste kunnen worden nagestreefd door een persuasief systeem. Ofschoon op een abstract niveau de doelen van het systeem kunnen resoneren met de doelen van een persoon (bijvoorbeeld 10 kilo afvallen binnen een jaar), hoeft dit voor de diverse implementatiekeuzes niet op te gaan.

Zo kan het persuasieve systeem op enig moment voorstellen doen die niet passen bij persoonlijke voorkeur of mogelijkheden, of niet werken binnen een specifieke context of situatie. In dergelijke gevallen vragen Frank en Nickel zich af of er nog wel sprake kan zijn van zacht paternalisme.

Het werk van Thaler en Sunstein lijkt sterker gericht op de partij die de nudges uitdeelt, dan op degene die ze ontvangt. Een perspectiefwisseling is in dit opzicht een nuttige exercitie, zoals blijkt uit een review van het boek Nudge door Robert Sugden²¹, een collega van Thaler en Sunstein:

“If you are attracted by the agenda of libertarian paternalism, I urge you to think carefully about whom you are authorising to nudge you, what criteria they will use to decide when and how your own decisions are capable of being improved on, and what incentives there are to induce them to follow those criteria.” (p.373)

Deze overwegingen leiden ook tot een ander punt van kritiek: het idee dat het te beïnvloeden individu toegang heeft tot volledige informatie over de doelen en methodes van de gedragsbeïnvloeding. Dit is een essentiële aanname bij zowel Fogg's persuasive technology als het nudgen van Thaler en Sunstein. Onvolledige, onjuiste of afwezige informatie kan worden beschouwd als misleiding. Echter, nudging maakt willens en wetens gebruik van onze cognitieve 'zwakke plekken' om ons gedrag te beïnvloeden.



Wordt hiermee bewuste reflectie niet omzeilt? We kunnen ons afvragen in hoeverre de gebruikte persuasieve principes en doelen voldoende expliciet worden gemaakt. Hier valt tegenin te brengen dat ook een keuzecontext die niet expliciet paternalistisch is ontworpen, of niet expliciet is ingericht met het doel om wenselijk of gezond gedrag te beïnvloeden, nog steeds een keuzecontext is die, bewust of onbewust, invloed zal hebben op ons gedrag. Neem het voorbeeld van de verkeersdrempel. Een criticus van nudging zou kunnen beweren dat dit een paternalistische inperking is van de keuzevrijheid van de automobilist.

Tegelijkertijd is het ontwerp van de auto, en het feit dat deze met groot gemak en comfort hoge snelheden kan bereiken, ook een reflectie van een hele serie aan ontwerpkeuzes die een bepaald type gedrag (bijvoorbeeld te hard rijden) mogelijk maken, en daartoe zelfs uitlokken. Zo beschouwd heeft onze technologische omgeving *altijd* invloed op ons gedrag. Dit punt wordt het meest expliciet gemaakt door de techniekfilosoof Peter-Paul Verbeek²². Met verwijzing naar Hans Achterhuis en Bruno Latour, maakt Verbeek duidelijk dat technologie eigenlijk altijd een bemiddelende rol speelt in het handelen van mensen. Verbeek²³:

“De rol van technologie is zó fundamenteel dat we niet moeten pretenderen ons geheel autonoom te kunnen opstellen. Sterker nog, wie weigert mee te denken over wenselijke vormen van gedragsbeïnvloedende technologie, onttrekt zich aan de verantwoordelijkheid om die bemiddeling een wenselijke vorm te geven.” (p.26)

Hiermee heeft technologie inherent een morele of gedrag beïnvloedende lading, ook wanneer deze niet expliciet wordt gecommuniceerd. Het erkennen dat we niet volledig autonoom zijn in relatie tot onze technologische wereld wil overigens niet zeggen dat we geen vrijheid kunnen nastreven in onze keuzes. Verbeek nodigt vooral uit om ons op een bewuste manier te *verhouden* tot technologie, en tot de rol die we technologie laten spelen in ons leven. Wanneer het gaat om door de overheid ingezette gedragsbeïnvloeding zullen we onze invloed vooral moeten laten gelden in het vormgevingsproces van technologie – op basis van participatoir ontwerp -, en in de democratische inspraak en kritische debatten waarin de rol van de overheid en haar gebruik van technologie in relatie tot de burger wordt doorgelicht.



Het frequent, langdurig en alomtegenwoordig gebruik van gedragsbeïnvloedende technologie door de overheid of door andere partijen werpt ook nog een andere vraag op. Hoe zal dit het waardesysteem van mensen op langere termijn beïnvloeden? Wanneer we dankzij technologie voortdurend duwtjes in de 'goede' richting krijgen, leren we dan ook zelf wat goed of slecht is? Zal er sprake zijn van een morele ontwikkeling in de gewenste richting dankzij een internalisatie van het gewenste gedrag en de impliciete norm die daarmee wordt gecommuniceerd, of wordt het gedrag eerder een oppervlakkig automatisme – de weg van de minste weerstand? Hoe kunnen we trots zijn op wat we bereiken, bewust van onze inspanning en de beloning die daarop volgt, als we niets anders hebben gedaan dan luisteren naar een technologisch superego?

*Als we niet willen dat onze geest een **willoos verlengstuk** wordt van de **technologie** waarmee we zo intiem verkeren, dan is een **kritisch reflectieve houding** ten aanzien van **gedragsbeïnvloeding** onontbeerlijk*

Zijn we ook goed als we goed *doen* dankzij technologie die duwt en trekt, stuurt en stimuleert, beloont en bestraft? Onze morele groei en cultivatie hangt ook samen met *bewuste* keuzes voor datgene dat ons moeite kost, waarvoor weerstanden overwonnen moeten worden. Het is deze eigen verantwoordelijkheid voor gedrag, dit eigendom over de eigen keuzes, deze rekenschap, waarmee we onze eigen morele persoonlijkheid ontwikkelen. Verbeek heeft gelijk – we kunnen kiezen om ons bewust te verhouden tot technologie. Maar gemakzucht en geestelijke luiheid is een deel van de verleiding waarop gedragsbeïnvloeding kapitaliseert. Als we niet willen dat onze geest een willoos verlengstuk wordt van de technologie waarmee we zo intiem verkeren, dan is een kritisch reflectieve houding ten aanzien van gedragsbeïnvloeding onontbeerlijk.

4. De kunstmatig intelligente overheid?

Kunstmatige intelligentie (*Artificial Intelligence*, of AI) is het kennisgebied dat zich bezighoudt met het creëren van technologische artefacten die in staat zijn om intelligent gedrag te vertonen. Dit kennisgebied is in zichzelf multidisciplinair en bestaat uit informatica, cognitieve psychologie, neurowetenschap, logica, statistiek, en linguïstiek. Daarnaast raakt AI ook diverse vragen op juridisch, ethisch, filosofisch, en sociaal-maatschappelijk gebied, waardoor ook vakgebieden als techniekfilosofie, geesteswetenschap (bijv. *philosophy of mind*), organisatiekunde, ethiek en recht deel uit maken van AI als kennisgebied in brede zin. AI is een snel voortschrijdend vakgebied, met een aantal duidelijk zichtbare en bruikbare toepassingen, o.a. in de automatische beeldherkenning, spraakherkenning (bijv. Siri, Amazon Echo), verwerking van natuurlijke taal (bijv. Google Translate, chatbots), kennissystemen (bijv. IBM Watson), robotica (bijv. industriële robots, drones, zelfrijdende auto's) en digitale spellen.

Big data wordt de olie van de 21ste eeuw genoemd. Als data de olie is, dan is AI de verbrandingsmotor waarin deze olie haar effect heeft. De laatste jaren wordt er vooral aanzienlijke voortgang geboekt in het kunstmatig leren (*deep learning*), waarbij snelle computers een brein-achtig netwerk simuleren die patronen kan leren herkennen in grote hoeveelheden getallen, plaatjes, teksten, spraak en andere vormen van data. Ook overheden experimenteren in toenemende mate met het gebruik van AI ten behoeve van overheidsservices, bijvoorbeeld bij het opsporen van uitkeringsfraude, het voorspellen van recidive na een misdrijf of het implementeren van slimme verkeerssystemen. De aanzienlijke technische mogelijkheden van AI in combinatie met de hoeveelheid aan data waar de overheid in potentie over beschikt, en de omvang en complexiteit van de uitdagingen waar de overheid zich voor geplaatst weet, maakt AI van groeiend belang voor de publieke sector. Er liggen grote kansen voor innovatie, die kunnen leiden tot een verhoogde efficiency en kwaliteitsverbetering, maar die ook potentiële risico's met zich meebrengen.

De kwaliteit van AI-classificaties is sterk afhankelijk van de kwaliteit van de data waarmee wordt getraind; als deze eenzijdig is of niet-representatief voor de doelgroep, zal dit ook de kwaliteit van de AI-adviezen negatief beïnvloeden. Wanneer bijvoorbeeld gezichtsherkenningsoftware vooral wordt getraind met een

set van blanke, mannelijke gezichten, dan zal het meer fouten maken in het correct classificeren van mensen met een ander ras of geslacht. Zo laat onderzoek van Joy Buolamwini²⁴, oprichtster van de *Algorithmic Justice League*, zien dat gezichtsherkenningsoftware van tech-bedrijven als Amazon, Microsoft en IBM nog duidelijk te wensen over laat. Waar voor mannen met een lichte huidskleur het aantal classificatiefouten onder de 1 procent bleef, gaat het aantal fouten voor vrouwen met een donkere huidskleur naar 35 procent - inclusief de gezichten van bekende Amerikaanse vrouwen als Oprah Winfrey, Michelle Obama en Serena Williams.

Een ander controversieel voorbeeld in dit kader is het recente werk van AI-onderzoekers Wu en Zhang²⁵ die AI-algoritmen of *classifiers* trainden op een set van 1856 foto's van bestaande gezichten van Chinese mensen, waarvan ongeveer de helft het gezicht van een reeds veroordeelde crimineel betrof. Na training bleken de algoritmen betrekkelijk nauwkeurig onderscheid te kunnen maken tussen een crimineel en niet-crimineel gezicht. Een probleem met dit type werk is dat het goed denkbaar is dat mensen met een, voor andere mensen, prototypisch 'crimineel' gezicht ook een groter risico lopen op arrestatie dan mensen die rondlopen met een prototypisch 'onschuldig' gezicht.



Er bestaat dus een gerede kans dat de set van niet-criminele gezichten ook een aantal gezichten bevat van mensen die ten onrechte *niet* zijn opgepakt. Op deze manier zou een AI-algoritme een bestaand vooroordeel in de gemeenschap kunnen laten voortbestaan, en zelfs versterken. Wanneer een dergelijk algoritme zou worden ingezet door de Chinese overheid om voorspellingen te doen over potentieel crimineel gedrag – zgn. *predictive policing*, we zien hier een parallel met de preCogs uit onze introductie – dan is er sprake van een zichzelf versterkend effect, een *selffulfilling prophecy*.

Een vergelijkbare casus betreft de inzet van algoritmen om een risicoschatting te maken van de mate waarin binnen een bepaalde wijk probleemgedrag frequenter zal voorkomen dan elders. Zo kan een voorspelling van marginaal meer probleemgedrag in een wijk ten opzichte van een andere wijk leiden tot meer politie-aanwezigheid in die wijk, waardoor de kans op detectie van probleemgedrag toeneemt. Wanneer die arrestatiedata dan weer als actuele update aan het model worden toegevoegd ontstaat een recursief, zichzelf versterkend proces, waarbij een wijk, goeddeels ten onrechte, een toenemend probleem lijkt te vormen ten opzichte van andere wijken. Bovenstaande zelfversterkende processen hebben *de facto* ook zonder AI-profilering al plaats – denk bijvoorbeeld aan de kans om als zwarte Amerikaan vast te komen zitten in het justitiële systeem van de VS, versus de kans dat dit je als blanke Amerikaan overkomt voor een vergelijkbaar vergrijp.

Vooroordelen die verborgen zitten in de data kunnen door feedback loops en machinaal leren worden uitvergroet. Op deze manier worden we door de inzet van AI geconfronteerd met uitkomsten die we als maatschappij als ongewenst of oneerlijk beschouwen, maar die diep zijn ingesleten in ons maatschappelijk verkeer en ons sociale oordeel. AI is in die zin een morele spiegel van ons maatschappelijk handelen. Mogelijk dat dit probleem zichzelf kan oplossen, wanneer we als maatschappij bewuster omgaan met vooroordelen, en hiermee rekening houden in de datasets die we aanbieden. We kunnen bijvoorbeeld een algoritme hertrainen met recentere gegevens zodat zowel de mens als het algoritme stukje bij beetje hun vooroordelen kwijtraken. Er zijn inmiddels ook toolkits en richtlijnen beschikbaar om AI eerlijker te maken. Zo heeft IBM recentelijk een open source toolkit beschikbaar gesteld, AI Fairness 360²⁶, waarmee vooroordelen in kunstmatig leren kunnen worden gedetecteerd en beperkt.

Een punt van aandacht bij de inzet van AI is dat de hoeveelheid geschikte trainingsdata uit de 'echte wereld' veelal te wensen overlaat voor het creëren van een robuust model. Zo is bijvoorbeeld het aantal agressieve excessen op straat gelukkig redelijk beperkt, maar is diens gevolg de dataset van camerabeelden (of andere relevante sensoren) die kan worden gebruikt om een predictiemodel van agressieve escalatie te construeren ook beperkt. Wanneer een AI-model over te weinig data beschikt zullen de uitkomsten zeer onbetrouwbaar zijn – niet veel beter dan de voorspelkansen bij het opwerpen van een muntje. In zijn algemeenheid zijn het de omvangrijke en toegankelijke datasets waarmee algoritmen getraind kunnen worden die bepalend zijn voor de uitkomst van de classificatie of beslissing. Ondervertegenwoordigde handelingen, gebeurtenissen, eigenschappen, of groepen (bijv. minderheden) kunnen hierdoor minder goed, of met minder reliëf, worden herkend.

Wanneer een AI-model over te weinig data beschikt zullen de uitkomsten zeer onbetrouwbaar zijn – niet veel beter dan de voorspelkansen bij het opwerpen van een muntje

Kunstmatige intelligentie is vooral sterk in zeer goed afgebakende probleemgebieden met duidelijke patronen en/of regels, maar nog slechts zeer beperkt in staat tot generalisatie naar andere gebieden, zelfs al zou dit voor mensen erg voor de hand liggen. Algemeen intelligent gedrag, wat we als mensen als *common sense* beschouwen – voor de hand liggend en vanzelfsprekend – ligt nog ver voorbij de horizon van AI. AI maakt bijvoorbeeld soms aannames of fouten die voor mensen compleet wezensvreemd zijn. Daarnaast kan een zelflerend systeem in zijn algemeenheid (nog) niet redeneren over hoe een beslissing of classificatie tot stand is gekomen op een manier die inzichtelijk is voor mensen. Het kunstmatig leersysteem kan dus nog niet goed uitleggen hoe of waarom het tot een bepaald advies of classificatie is gekomen.

Dit is in iedere setting waar adviezen worden gegeven of beslissingen worden genomen moeilijk te verteren, maar wanneer het de overheid en haar beslissingen aangaat is het volkomen onacceptabel. De overheid moet altijd in staat zijn om te onderbouwen hoe haar beslissingen tot stand zijn gekomen, en moet hierover transparant zijn naar haar burgers. ‘Computer Says No’ is geen optie.

De kracht van AI-systemen lijkt vooral te liggen in de combinatie met de menselijke professional, en niet in de vervanging van deze professional. AI is het krachtigst als hulpmiddel om onze eigen cognitieve vermogens mee te ondersteunen en uit te breiden. Eerder in dit essay spraken we over de tekortkomingen in het menselijk beslisvermogen – zoals bij het inschatten van risico’s op basis van data, of het meewegen van langere-termijn-effecten. Juist op deze punten kan een computer, en kan AI, complementair zijn aan de mens. De snelheid waarmee patronen in grote hoeveelheden data kunnen worden geanalyseerd kan van grote toegevoegde waarde zijn op het langzamere, meer diepgaande en context-gevoelige denkwerk van mensen.

AI is het krachtigst als hulpmiddel om onze eigen cognitieve vermogens mee te ondersteunen en uit te breiden

Wat uit het bovenstaande duidelijk mag zijn geworden is dat AI, net als de mens, niet onfeilbaar is. Kennis over de kracht én beperkingen van AI is onontbeerlijk. Het is in deze context ook van groot belang om na te denken over de meest optimale mens-machine-rolverdeling, over het ontwerp van de mens-machine-interactie en interface, en de mate van controle en overzicht die de mens zal willen en moeten behouden in het licht van voortschrijdende automatisering. Hierop moeten professionals, bij de overheid en elders, zich voorbereiden binnen het huidige werkveld, maar ook moeten we opleidingen *future proof* maken, zodat de professionals van de toekomst de relevante kennis, vaardigheden en flexibiliteit hebben.

5. Big Brother is Watching You: het spook van überveillance

De astrofysicus Vincent Icke heeft ooit naar aanleiding van de discussie rondom het Elektronisch Patiënten Dossier (EPD) een treffende vergelijking getrokken met het boek *Gulliver’s Reizen*. Gulliver, na schipbreuk aangespoeld op het eiland Lilliput, werd daar, terwijl hij nog buiten westen was, door de Lilliputters ingesponnen in een web van honderden draadjes. Elk draadje, aldus Icke, staat symbool voor een stukje persoonlijke informatie – of het nu het EPD betreft, een socialmedia-account, een GPS-navigatiesysteem of de Bonuskaart van Albert Heijn. Elk draadje op zichzelf lijkt ongevaarlijk, maar het is de combinatie die Gulliver, en ook ons, muurvast zet. Wanneer deze draadjes tezamen zich als een web van data rond een individu sluiten, en worden gebruikt om de overheid meer mogelijkheden tot surveillance en controle te bieden, dan is er sprake van *überveillance*, aldus Michael Fusco en Michael²⁷.

In hun artikel richten Michael en collega’s zich in het bijzonder op *location based services* (LBS), dat wil zeggen services die voor hun functionaliteit gebruik maken van locatie-informatie van het individu, door middel van, bijvoorbeeld, GPS- of RFID-technologie. De applicaties lopen uiteen van GPS-systemen op trucks waarmee een werkgever de locatie van een truck kan vaststellen tot *friendfinder*-applicaties in mobiele telefonie tot ePassports met RFID-tags voor identificatie- en security-doeleinden. Hiernaast beschikt de overheid natuurlijk over een rijk geschakeerd pallet aan mogelijkheden tot individuele identificatie, inclusief camerasystemen in de publieke ruimte, biometrische data en een veelheid aan overheidsdatabases met persoonsinformatie.

Überveillance is geen enkele bron op zichzelf, maar alle bronnen tezamen. Het is sterk gerelateerd aan Roger Clarke’s concept ‘dataveillance’²⁸, gedefinieerd als “the systematic use of personal data systems in the investigation or monitoring of the actions of one or more persons”. Überveillance is dataveillance op steroiden. Het is 24/7 alomtegenwoordige elektronische surveillance. Echter, alomtegenwoordig en ‘always on’ betekent niet dat het überveillance systeem alwetend is, of zelfs ook maar kan zijn.



In principe tracht een surveillance-systeem te meten aan het 'wie' (identificatie), 'waar' (locatie), en 'wanneer' (tijd), om te komen tot een interpretatie of voorspelling van 'waarom' (intentie/motivatie), 'wat' (beoogde resultaat) en 'hoe' (de methode). Er bestaat een groot risico op misinformatie (bijv. meetfouten, ontbrekende informatiebronnen, toevallige correlaties), misinterpretatie (bijv. onjuist omgaan met onzekerheid en ambiguïteit, onjuiste inferentie van intentie) en manipulatie (bijv. selectieve rapportage van data, *confirmation bias*, hacking, politieke manipulatie) van data afkomstig uit een *überveillance*-systeem. Ondertussen is de menselijke neiging om sensor-data en algoritmische interpretaties daarvan voor 'waar' en 'objectief' en 'wetenschappelijk' aan te nemen zorgwekkend groot. In de woorden van Michael et al: "Überveillance can be a predictive mechanism for one's expected behavior, traits, characteristics, likes or dislikes; or it can be based on historical fact, or something in between. The inherent problem with *überveillance* is that facts do not always add up to truth (...), and predictions based on intelligence are not always correct." (p.1198).

Naast zorgen over de feitelijke zeggingskracht van surveillance data, is *überveillance* ethisch evident problematisch. Allereerst is er sprake van een schending van het recht op privacy, waarbij elk individu in beginsel het recht heeft op een persoonlijke levenssfeer, ongeobserveerd door derden, waarin we kunnen gaan waar we willen, omgaan met wie we willen, vrijelijk informatie tot ons kunnen nemen, en het leven kunnen leiden dat we willen, binnen de grenzen van de wet. De Nederlandse burger is in het algemeen zeer gehecht aan haar privacy²⁹, en een waargenomen schending van privacy resulteert dan ook in een vertrouwensbreuk.



De psychologe Judy Burgoon en haar collega's³⁰ maken onderscheid tussen vier soorten privacy:

- 1. Fysieke privacy** is de controle die iemand heeft over de mate waarin die persoon fysiek toegankelijk is voor anderen (denk aan zichtbaarheid, persoonlijke ruimte, fysiek contact).
- 2. Sociale/communicatieve privacy** is de mogelijkheid en moeite waarmee een individu sociale contacten kan reguleren (bijvoorbeeld de lengte en inhoud van een sociale interactie, non-verbale communicatie).
- 3. Psychologische privacy** refereert aan de mogelijkheid om cognitieve en emotionele signalen te reguleren (denk aan het voorkomen van pest-gedrag, maar ook het blootstaan aan gedragsbeïnvloeding).
- 4. Informatieele privacy** refereert aan controle over de mate en manier waarop informatie wordt gedeeld met anderen - hoe, wanneer, met wie en in hoeverre (denk aan roddelen, ongevraagd iemands mail lezen, of online foto's van iemand delen).

Wanneer er sprake is van *überveillance*, vindt de inbreuk op privacy op een combinatie van de fysieke, psychologische en informatiele niveaus plaats. Met een dergelijk dramatisch verlies aan privacy komt ook het verlies aan controle en persoonlijke autonomie. Iemand die zich voortdurend in de gaten gehouden weet, gedraagt zich anders, censureert zichzelf, vermijdt sommige ogenschijnlijk gevoelige of controversiële situaties en interacties, en zal in veel opzichten zichzelf niet meer kunnen zijn. Het verlies aan persoonlijke controle en autonomie leidt op termijn tot gevoelens van frustratie, machteloosheid en depressie, en kan leiden tot grote maatschappelijke onrust en verzet. Wat echter potentieel nog zorgelijker is, is niet de frustratie en het verzet, maar het ontbreken ervan. Wanneer een nieuwe generatie opgroeit zonder privacy, kan een extern opgedrongen gedrags-norm transformeren naar een geïnternaliseerde standaard van gedrag. Het aangepaste gedrag wordt dan 'het nieuwe normaal', zeker wanneer kritische tegengeluiden gemarginaliseerd worden.

Het Chinese 'Social Credit'-systeem³¹ is een recent voorbeeld van hoe *überveillance* er in de praktijk uit kan zien. Dit systeem is een persoonlijk reputatiesysteem geïnitieerd door de Chinese overheid, waarbij een combinatie van bigdata-analyse en geavanceerde surveillance-technieken leidt tot een persoonlijk profiel van iedere Chinese burger op het Chinese vasteland.

Online gedrag, gedrag in de publieke ruimte, politiek afwijkende uitingen – het wordt allemaal vastgelegd en leidt tot een plus of een min op een unidimensioneel scoringsstelsel, de Social Credit Score. Deze Score is vervolgens bepalend voor de toegang die men heeft tot publieke diensten, zoals reisdocumenten, toegang tot transport, scholing of een goede baan. Niet enkel het persoonlijke gedrag weegt mee; ook het gedrag van de sociale groep waartoe men behoort, maakt deel uit van de Social Credit Score. Iedereen wordt ogenschijnlijk vaandelhouder van een collectieve norm, met als gevolg grote sociale controle en onderling wantrouwen, en het isoleren van mensen die afwijkend denken. Er ontstaat zodoende een staatsgeleid model waarin uniformiteit in gedrag en denken wordt nagestreefd, en culturele, intellectuele en politieke veelvormigheid wordt onderdrukt. Dit staat haaks op individuele behoeftes aan privacy en autonomie, aan de integriteit van de persoon en haar sociale leefomgeving, maar gaat bovendien ten koste van de flexibiliteit die vereist is om een complexe maatschappij adaptief te laten reageren en adequaat te anticiperen op een veranderende wereld.

Er ontstaat zodoende een staatsgeleid model waarin uniformiteit in gedrag en denken wordt nagestreefd, en culturele, intellectuele en politieke veelvormigheid wordt onderdrukt

Naast de overheid speelt ook het grote internationale bedrijfsleven (Google, Amazon, Microsoft, Facebook, etc.) een belangrijke rol in het digitale landschap. Wereldwijd worden in hoge mate de diensten van de grote digitale bedrijven omarmd – bedrijven die het niet heel nauw nemen met onze privacy en die onze persoonlijke profiel- en gedragsdata verhandelen op een schaal en met een snelheid die ongekend is. Er is nu reeds sprake van een enorme informatie-ongelijkheid, waarbij de gemiddelde gebruiker van een online service – of het nu een Google search, een Facebook-pagina, een ‘Nest Home’-thermostaat, of een ‘Amazon Alexa’-voice assistant betreft – geen idee heeft wat er met haar informatie gebeurt: welke andere partijen dan degene die de service aanbiedt toegang krijgen tot deze data, hoe deze data gebruikt worden, wat haar data waard zijn voor

derden, welke inferenties men maakt op basis van deze persoonlijke data, en hoe dit uiteindelijk weer leidt tot persoonlijke beïnvloeding van kijk-, klik-, en koopgedrag³².

Niemand leest de honderden pagina’s online contracten, *licence agreements*, *privacy policies*, of *terms of service*, en dat ze gelezen worden zou ook een onrealistische, ja zelfs onethische verwachting zijn. Een recente studie naar een fictief sociaal netwerk, NameDrop, wees uit dat van de honderden participanten, driekwart de privacy policy en terms of service geheel oversloeg, en het overgebleven kwart deze documentatie slechts zeer vluchtig las, terwijl 97 procent van de mensen instemden met de privacy policy en 93 procent met de terms of service. Had men deze documenten beter gelezen dan was duidelijk geworden dat ze ridicule clausules bevatten over het afstaan van het eerstgeboren kind en het delen van alle persoonlijke informatie met de eigen werkgever en de Amerikaanse veiligheidsdienst NSA. Kwalitatieve data laten zien dat mensen deze documenten als hinderlijk ervaren, en dat men direct toegang wil tot de functionaliteit van de service zonder te hoeven nadenken over de manier waarop, of de condities waaronder, deze service tot stand komt.

Deze realiteit, door de auteurs bestempeld als de ‘grootse leugen op het Internet’³³, maakt misbruik door grote bedrijven kinderlijk eenvoudig, terwijl men zich ogenschijnlijk aan de letter van de wet houdt. Deze fundamentele informatie-ongelijkheid vormt de basis van wat Shoshana Zuboff ‘surveillance kapitalisme’ noemt³⁴. In haar woorden:

“Surveillance capitalism operates through unprecedented asymmetries in knowledge and the power that accrues to knowledge. Surveillance capitalists know everything about us, whereas their operations are designed to be unknowable to us. They accumulate vast domains of new knowledge from us, but not for us. They predict our futures for the sake of others’ gain, not ours.” (p.11).

Waar überveillance een Orwelliaanse dystopie oproept, van onderdrukking en manipulatie, roept Zuboff’s surveillance-kapitalisme eerder associaties op met Huxley’s *Brave New World*, een dystopie van zelf-opgelegd hedonistisch nihilisme. Wat beide systemen echter delen is het ondergeschikt en eenvormig maken van het individu – zij die zich afwijkend gedragen worden genegeerd, gemarginaliseerd, geëxcommuniceerd of anderszins bestraft.



Individuele variatie is echter een centraal ingrediënt van culturele vooruitgang en maatschappelijke flexibiliteit. In een maatschappij waar de complexiteit van uitdagingen in het digitale tijdperk enkel toeneemt, is een top-down, centralistisch, *one-size-fits-all* beleid een misvatting. Bekende spanningsvelden als die tussen privacy en veiligheid, tussen economie en ecologie, of tussen het individuele belang en dat van het collectief, maken onmiddellijk helder dat we niet zonder meer kunnen optimaliseren voor een singuliere doelfunctie. Zo'n functie bestaat niet.

Digitale middelen van meten en beïnvloeden kunnen krachtige tools zijn voor een overheid die op een data-gedreven, *evidence based* manier te werk wil gaan. De paradox is dat de middelen van surveillance die we zo kritisch onder de loep hebben genomen ook kunnen leiden tot een beter geïnformeerde overheid die in staat is om slagvaardiger te opereren en beter voor haar burger te zorgen. De metingen die nodig zijn om de cybernetische lus te sluiten, het proces te optimaliseren, zijn dezelfde metingen die onze privacy kunnen aantasten. We hebben geen Orwelliaanse repressie nodig om het gevoel te krijgen dat onze privacy op het spel staat. Onbedoelde neveneffecten van valide pogingen tot informatiegaring kunnen al tot een inbreuk op privacy leiden³⁵. De overheid moet hierin terughoudend en zelfkritisch optreden: haar pogingen tot informatiegaring of beïnvloeding mogen nooit de burger in een keurslijf vangen, het privé domein van de burger aantasten, of tot grote intransparantie en informatie-asymmetrie leiden. Participatie van burgers en andere belanghebbenden is hier van cruciaal belang.

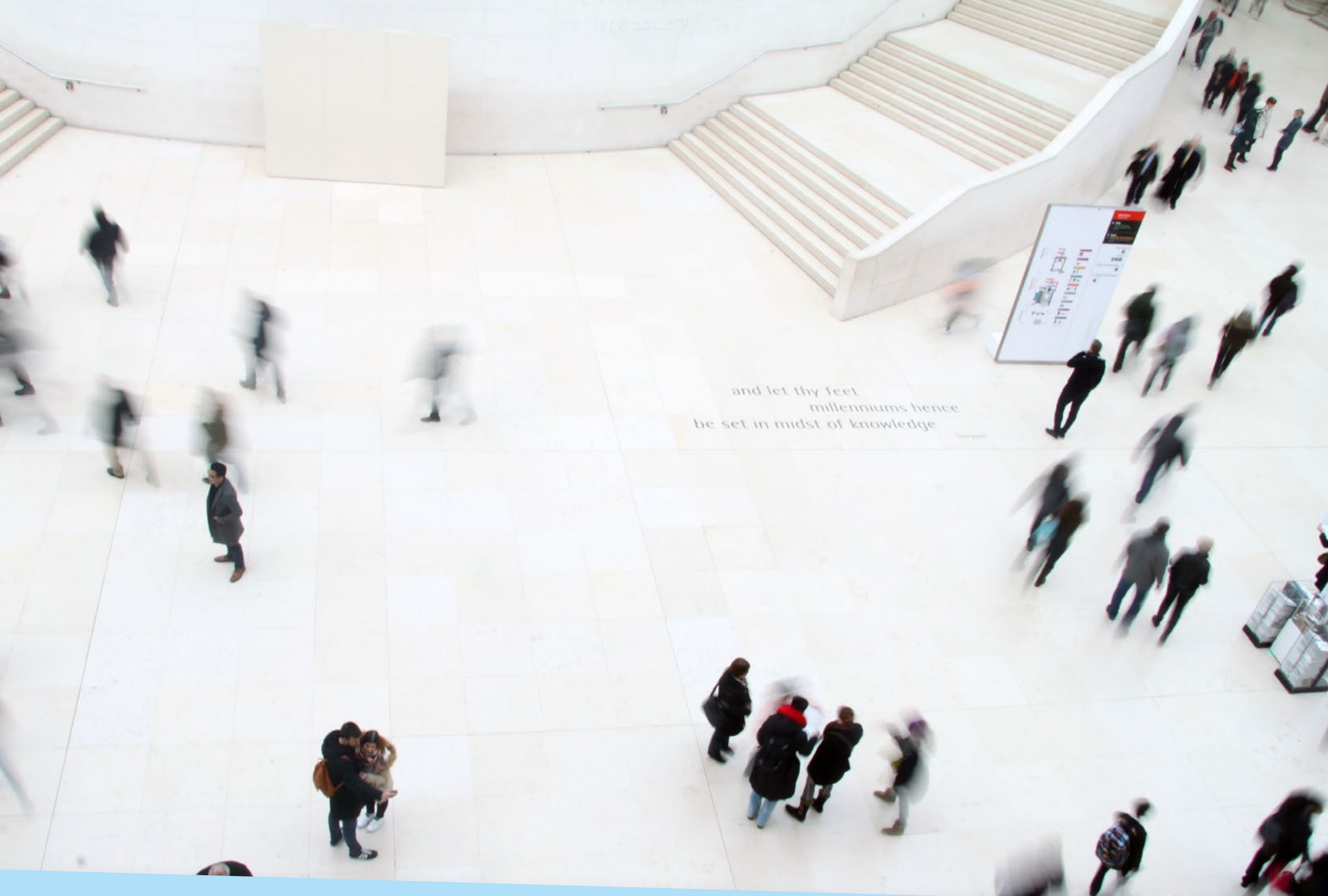


6. Naar participatoir ontwerp van de digitale openbare ruimte

Systemen zijn niet belangrijker dan mensen. Digitalisering is geen doel op zich, maar zou het belang van de burger moeten dienen, en moeten bijdragen aan het oplossen van urgente maatschappelijke problemen die burgers raken. Te vaak wordt er pas ná het introduceren van een beleidsmaatregel nagedacht over de effecten op burgers. Centraal uitgangspunt in het ontwerp, de implementatie, het bestuur en de regelgeving van de digitale samenleving – de smart society en de smart city – moet zijn dat het mensgericht is. Het is van belang om de principes van mensgericht ontwerp (*user-centred design*) en waardegericht ontwerp (*value-sensitive design*) te hanteren. Dat betekent dat betrokken partijen niet enkel worden geïnformeerd over reeds in gang gezette ontwikkelingen, of als deelnemer in proeftuinen (*living labs*) mogen meedenken, maar dat ze actief en continu in het voortraject en tijdens de ontwikkeling, introductie en opschaling van innovaties worden betrokken. Hierbij moet in het voortraject ook expliciet een behoefte-analyse worden meegenomen, om niet enkel vanuit een *tech-push*- maar ook vanuit een *user-pull*-perspectief te werken. De vraag moet niet zijn: wat kunnen we allemaal met digitale technologie? De vraag moet zijn: waar hebben mensen concreet behoefte aan? Welke problemen gaan we oplossen? Welke mensen worden hierdoor geraakt? Is iedereen voldoende betrokken en kan iedereen meedoen? Is er ruimte voor maatwerk; voor persoonlijke aandacht en hulp?

Zijn er voldoende mogelijkheden tot het bijsturen, opschorten of zelfs afblazen van een innovatie als die onvoldoende blijkt te werken? Hierbij verdienen kwetsbare groepen extra aandacht – denk bijvoorbeeld aan laaggeletterden, allochtone minderheden, mensen met een beperking en ouderen die de digitale innovaties niet helemaal kunnen bijbenen. Zij zijn, in zekere zin, de 'kanarie in de kolenmijn'. Zij zijn extra gevoelig voor ongunstige omstandigheden en daarmee ook een belangrijke indicator voor de rest van de maatschappij.

In dit kader is het goed om de meerwaarde te benoemen van proeftuinen en experimenten om ethische dilemma's invoelbaar en bespreekbaar te maken voor een breed publiek. In hun huidige incarnatie zijn *living labs* vooral van groot nut om technische innovaties in de praktijk te kunnen uitproberen.



and let thy feet
millenniums hence
be set in midst of knowledge

Echter, wat tot nu toe onderbelicht is gebleven, is het invoelbaar maken van ethische dilemma's die toenemende digitalisering van onze leefomgeving met zich meebrengt. Deze discussie lijkt momenteel vooral plaats te hebben via speculatieve designs, culturele evenementen, kritische journalistiek, wetenschap en science fiction. Dit is buitengewoon waardevol voor het maatschappelijk debat, maar in deze vormen van discussie mist veelal nog aansluiting met de 'gewone' inwoners van de stad, inclusief de hierboven genoemde kwetsbare groepen. Deze afstand kan ook leiden tot onrealistische verwachtingen, zowel in positieve als in negatieve zin, en beelden creëren die op gespannen voet staan met de 'lived experience' – de concrete, beleefde werkelijkheid van alledag. Burgers moeten in brede zin kennis kunnen maken met nieuwe technologie, in alledaagse situaties, en moeten actief betrokken worden in de dialoog rondom de dilemma's van nieuwe technologie. Dit moet structureel worden verankerd in het ontwikkelingstraject en politieke besluitvorming rondom de introductie van nieuwe technologie in de leefomgeving. Hiernaast is het van groot belang om ook waarden en belangen die veelal geen expliciete stem hebben op een principiële manier mee te nemen in de belangenafweging.

Burgers moeten in brede zin kennis kunnen maken met nieuwe technologie, in alledaagse situaties, en moeten actief betrokken worden in de dialoog rondom de dilemma's van nieuwe technologie

Denk hierbij aan de belangen van toekomstige (ook nu nog ongebooren) generaties, van bredere en complexe systemen (zoals het klimaat of de leefomgeving), van onvoorziene en ongewenste effecten over tijd, of van effecten die optreden wanneer een technologie opschaaft, of voor andere doelen wordt ingezet dan oorspronkelijk de gedachte was. Hiervoor is een breder ethisch toetsingskader nodig, bijvoorbeeld gestoeld op deugdenethiek.

Waar mensgericht ontwerp geschikt is om belangen van mondige en actieve stakeholders mee te nemen, is een meer principiële ethische overweging nodig om een bredere afweging te kunnen maken, en ook collectieve waarden te verankeren in het innovatieproces die nu nog geen duidelijke stem hebben.

7. Digitale geletterdheid, mondige burgers en een eerlijke overheid

De overheid heeft de verantwoordelijkheid om de mogelijkheden van digitalisering zo goed mogelijk te benutten ten behoeve van haar burgers, maar ook om de maatschappij te beschermen tegen uitwassen en onbedoelde effecten van digitalisering. Het belang van digitale geletterdheid is hierbij groot. Digitaal geletterde burgers zijn in staat om effectief en efficiënt gebruik te maken van digitale technologie, met alle potentiële voordelen die dit biedt in scholing en werk, ondernemerschap en politieke en maatschappelijke participatie. Een digitaal geletterde burger die gebruik kan maken van participatoire digitale netwerken en platformen, die meewerkt aan 'citizen science'-projecten, en de digitalisering van zijn leefomgeving zelf mede vorm kan geven, is een waardevolle partner voor de overheid. Een digitaal geletterde burger is ook digitaal weerbaar; beter in staat om *hacking* en *phishing* te herkennen, om identiteitsdiefstal te voorkomen en om nepnieuws van echt nieuws te kunnen onderscheiden. Maar bovendien is een digitaal geletterde burger ook een kritische burger, die beter in staat is om overheid of marktpartijen kritisch te volgen en te bevragen wanneer zaken als persoonlijke privacy of autonomie onder druk komen te staan.

In het onderwijs wordt al enige tijd fors ingezet op digitale geletterdheid, maar slechts nog binnen een beperkt spectrum aan onderwerpen, gericht op het gebruik en de risico's van sociale media, mobiele telefonie en cyberpesten. Het is wenselijk om deze inzet zowel kwantitatief als kwalitatief uit te breiden, ook richting *cyber-physical*-systemen (IoT), persoonlijke tracking (sensoren, smartwatches, etc.), surveillance-technologie (bijv. biometrie, face-reading) en de mechanismes en risico's van digitale hypertexting en microtargeting (*pointcasting*, personalisatie), zowel online en mobiel, maar ook in digitale games en in de fysieke wereld.

Hierbij moet bewustwording gecreëerd worden van zowel de rol als de verantwoordelijkheid van de overheid in een digitale maatschappij, maar ook van de actieve rol die de burger zelf kan en moet spelen. In een dergelijke bewustwording moet ook een kritische blik worden geworpen op de rol, de strategieën en de belangen van grote internationale bedrijven die van de handel in persoonlijke data en de strijd om de aandacht van de gebruiker hun businessmodel hebben gemaakt. Om een dergelijk programma vorm te kunnen geven is het van belang om een coherent curriculum te ontwikkelen, gericht op diverse niveaus van educatie en geletterdheid. Ook hier is sociale inclusiviteit en diversiteit – zoals eerder benoemd – van groot belang. Tevens ligt een koppeling met een breder cybersecurity-initiatief voor de hand. Zonder digitale hygiëne in zowel privé-, zakelijk- als overheidsdomein blijft cybersecurity immers dweilen met de kraan open.

Er is hierbij behoefte aan een coördinerende partij - mogelijk vanuit een combinatie van departementen met verschillende aandachtsgebieden - die het initiatief neemt en de curriculum-ontwikkeling coördineert en mede vormgeeft, maar dit oppakt in samenwerking met het maatschappelijk middenveld. De kennis en kunde die bestaat bij digitale ondernemers, adviesorganen en burgerinitiatieven is buitengewoon waardevol en moet worden gemobiliseerd. Hierbij verdient ook het 'right to challenge' vermelding – de overheid die in het regeerakkoord in 2017 heeft afgesproken dat het zal zoeken naar de mogelijkheid om “via een Right to Challenge-regeling burgers en lokale verenigingen de mogelijkheid te geven om een alternatief voorstel in te dienen voor de uitvoering van collectieve voorzieningen in hun directe omgeving”³⁶. Dit bevordert de democratische inspraak.

Er wordt veel gesproken en geschreven over transparantie. Transparantie in beslisprocessen van de overheid, en transparantie in AI-algoritmen, om zo het zo voor burgers inzichtelijk te maken op basis waarvan bepaalde beslissingen worden genomen. Nog belangrijker wellicht, maar vaak onbenoemd omdat het zo vanzelfsprekend lijkt, is eerlijkheid. Transparantie is geen gedragsnorm, maar een procesvoorwaarde om controle mogelijk te maken. Eerlijkheid is een morele waarde, een deugd, die bovenin het vaandel van de overheid moet wapperen. Zoals we gezien hebben in het voorgaande leidt transparantie niet automatisch tot eerlijkheid – denk aan het voorbeeld van de honderden pagina's aan kleine

lettertjes in service-contracten waarmee gebruikers van online services worden geconfronteerd. Dit is een overvloed aan informatie, maar geen transparantie. Hetzelfde geldt voor het rücksichtloos op straat gooien van vertrouwelijke informatie - denk bijvoorbeeld aan gevoelige onderhandelingen tussen meerdere partijen die tot een compromis moeten komen. Volledige transparantie zorgt hierbij voor een tekort aan privacy bij de onderhandelende partijen, waardoor het gedrag voor de bühne wordt aangepast, en de kans op polarisatie groter wordt – zeker wanneer we de echokamer-effecten van sociale media meenemen. Alleen de meest gelouterde bestuurders en politici zijn ongevoelig voor dat effect. Dit is geen pleidooi voor achterkamertjespolitiek, maar voor vertrouwen. Aan de basis van transparantie staat (een wellicht terecht) wantrouwen. Eerlijkheid, wanneer fundamenteel gerespecteerd en consequent tot in de haarvaten van de organisatie in de praktijk gebracht, brengt vertrouwen met zich mee.



8. Conclusies en aanbevelingen

Digitalisering heeft een ongekennde impact op onze wereld. De snelle technologische vooruitgang in onze digitale maatschappij gaat ook aan de overheid niet voorbij. Big data wordt op diverse plaatsen ingezet om efficiëntie te vergroten, bijvoorbeeld bij het onderhoud van de stad, of het doelmatig inzetten van politie. Ook kan men door technologie de stad duurzamer en leefbaarder maken – denk aan energiebesparende lantaarnpalen die alleen aangaan als er iemand langkomt, of aan sensoren die geluidsoverlast in een omgeving automatisch detecteren. Tegelijk worstelt de overheid ook met de snelheid en de schaal waarmee de digitale transformatie zich voltrekt. Een groot deel van de huidige infrastructuur – gebouwen, wegen, energiesystemen en openbaar vervoer – moet de digitale transformatie nog doorlopen. Een aantal pionierende steden experimenteren met smartcity-proefprojecten en living labs, om een gevoel te krijgen voor de potentiële meerwaarde van digitale technologie in de publieke ruimte, maar ook om in te kunnen schatten waar knelpunten zitten, en waar wet- en regelgeving en ethische kaders nodig zijn om alle innovaties in goede banen te leiden.

Met de transitie naar een intelligente digitale publieke infrastructuur dienen zich ook nieuwe mogelijkheden aan om het gedrag van burgers te beïnvloeden teneinde positieve maatschappelijke doeleinden te bewerkstelligen.

Om bijvoorbeeld agressieve incidenten te voorkomen, asociaal rijgedrag tegen te gaan, of gezond bewegen te bevorderen. Juist op die gebieden waar de overheid actief preventiebeleid voert – veiligheid, gezondheid, sociaal gedrag, milieu – kan de combinatie van digitale technologie met kennis van de menselijke psychologie een positief resultaat opleveren. Maar deze combinatie roept ook direct lastige vragen op. Hoe kan de overheid de privélevenssfeer van burgers beschermen, maar tevens toezicht houden op veiligheid en recht, op sociaal en gezond gedrag, op het beschermen van de leefomgeving? Hoe, en in hoeverre, mag de overheid sturen op gewenst gedrag van haar burgers, gebruik makend van publieke datastromen, kunstmatige intelligentie en digitale beïnvloedingstechnieken? En hoe dient de overheid *zichzelf* te reguleren wanneer het geconfronteerd wordt met de verleidingen van macht en controle die digitalisering met zich mee brengt?

Er zijn op dit punt geen eenvoudige oplossingen aan te dragen. De complexiteit van de maatschappelijke en ethische uitdagingen is daarvoor te groot. Het gladde ijs tussen nudging en *überveillance* is gelegen in de verleiding om als overheid te veel zaken in één hand te willen nemen, en hierbij de grenzen van het publieke en het privédoel te overschrijden. Er is niet 'one algorithm to rule them all', geen singuliere doelfunctie die alle verschillende belangen op een juiste manier kan meewegen. En zelfs als deze er wel zou zijn, dan blijft, ondanks alle digitale middelen die ons ter beschikking staan, onze kennis, en daarmee onze mogelijkheid om gedrag juist te voorspellen en te beïnvloeden, onvolmaakt. De wijsheid van onze beslissingen kennen we vaak slechts achteraf, en ons *track record* als maatschappij tot nu toe (denk aan de ecologische crisis, migratieproblematiek, het groeiende verschil tussen rijk en arm, cyberterrorisme, de polarisatie in politiek en maatschappij) noopt ons tot bescheidenheid.



Met al het voorgaande in het achterhoofd, wil ik dit essay besluiten met een aantal aanbevelingen voor beleid.

1. Wanneer de overheid gebruik wil maken van technologie ten behoeve van gedragsverandering in het publieke domein, dan dient dit transparant te gebeuren, op basis van democratische besluitvorming, ethische toetsing en wetenschappelijke evidentie, en met expliciete betrokkenheid en instemming van partijen die deze technologie direct of indirect zal gaan raken. De overheid dient hierbij te allen tijde proactief, open, begrijpelijk en eerlijk te communiceren over digitalisering in het publiek domein, of plannen daartoe, op alle niveaus – van landelijk en provinciaal tot gemeentelijk en wijkniveau.
2. Betrek vroegtijdig en continu alle partijen die, direct of indirect, in aanraking komen met digitale gedragsbeïnvloeding in de openbare ruimte, door middel van een proces van participatoir ontwerp. Neem extra moeite om ook de meest kwetsbare groepen uit de maatschappij een actieve stem te geven in dit proces. Neem ook ethische waarden mee in het ontwerp. De overheid heeft de primaire verantwoordelijkheid om dit proces te organiseren, samen met haar maatschappelijke partners.
3. De overheid moet haar eigen digitale competenties vergroten en dient zich hierbij gedegen op de hoogte te stellen van de mogelijkheden en beperkingen van de meest recente vormen van digitale technologie. Een technologisch naïeve overheid is een makkelijke prooi voor kwaadwillende partijen, een slechte partner voor de industrie en een tekortschietende hoeder van het publiek belang.
4. Data-eigendom en privacy van burgers in de publieke ruimte dienen te worden gerespecteerd en beschermd. De overheid dient geen gebruik te maken van onbewezen, onveilige of te weinig robuuste technologie. Onbedoeld, onvoorzien of illegaal gebruik van de technologie en de data moet in kaart worden gebracht en gereguleerd.
5. De overheid moet een constructief-kritisch maatschappelijk (tegen)geluid organiseren. Hierbij moet fors geïnvesteerd worden in de digitale geletterdheid van burgers, met daarbij meer aandacht voor de nieuwste technologische ontwikkelingen en hoe deze de fundamentele ethische waarden binnen onze digitale maatschappij raken. De versterking van technologisch burgerschap kan vele vormen aannemen, bijvoorbeeld debatten, cursussen, livinglab-demonstraties, kritische culturele projecten of speculatieve designs.

6. De overheid moet voldoende en onafhankelijke kritische controle en toezicht organiseren en verankeren in de wet. Hierbij moeten we denken aan continu aanzwengelen en organiseren van het waarden-debat, bespreking van relevante ethische en juridische casuïstiek, en toetsing van governance, juridische en ethische kaders via volksvertegenwoordiging op alle niveaus (gemeentelijke en provinciale raden, Eerste en Tweede kamer). Daarnaast is het van belang het maatschappelijk middenveld en relevante adviesorganen continu te betrekken en de rol en positie van toezichthouders te versterken. Hierbij moet ook onderlinge afstemming over handhavingsdomeinen heen worden georganiseerd.

Samenvattend is het belangrijkste advies dat we de veelvormigheid in onze maatschappij moeten mobiliseren, omarmen en betrekken. Niet vanuit een noodzakelijke democratische concessie naar de burger, maar omdat dit uiteindelijk tot een bredere en meer gevarieerde culturele en intellectuele basis van collectieve besluitvorming leidt. Zoals de socioloog Dirk Helbing en collega's overtuigend hebben beargumenteerd ligt er een grote kracht in collectieve intelligentie bij het besturen van onze huidige complexe digitale maatschappij. Dit vereist het organiseren van inspraak en participatie, het waarderen van variatie boven een-vormigheid, en het stimuleren van onafhankelijke informatiegaring en lokale beslissingsbevoegdheden. Socio-diversiteit is net zo belangrijk als biodiversiteit, schrijft Helbing³⁷.

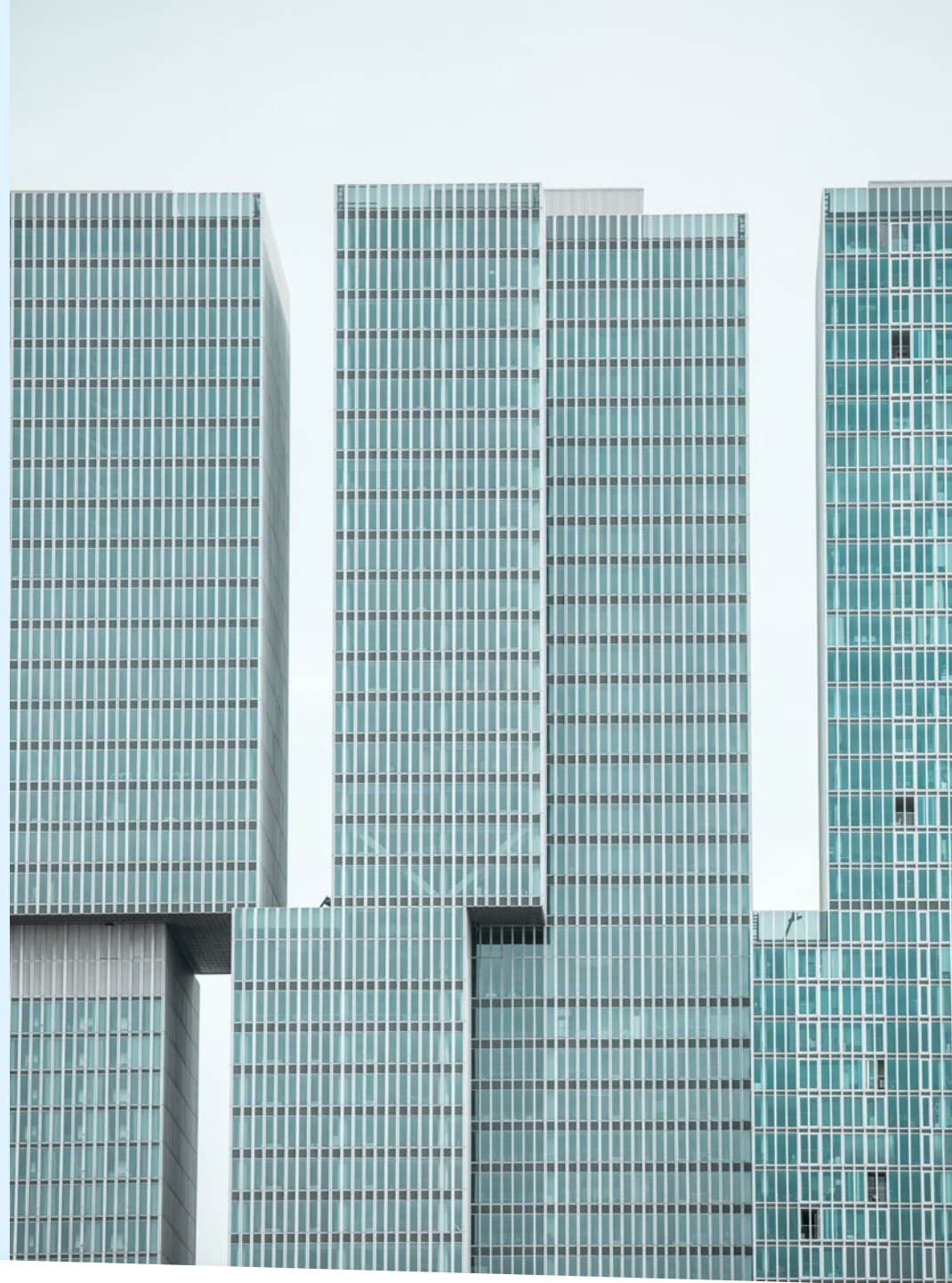
We moeten de veelvormigheid in onze maatschappij mobiliseren, omarmen en betrekken

Een digitaal verantwoordelijke overheid, ook eentje die worstelt op het gladde ijs tussen nudging en überveillance, is gebaat bij een mondige, digitaal vaardige en kritische burger, een levendig democratisch debat, participatie vanuit een breed scala aan maatschappelijk belanghebbenden, een solide wetenschappelijke basis, onafhankelijke en transparante wettelijke controle, en voortdurende ethische toetsing. De kracht zit in flexibiliteit door veelvormigheid, hetgeen leidt tot robuust en responsief bestuur van onze digitale maatschappij.

Over de auteur



Prof. dr. Wijnand IJsselstein is hoogleraar Cognition and Affect in Human-Technology Interaction en wetenschappelijk directeur van het Center for Humans & Technology aan Eindhoven University of Technology (TU/e).



Referenties

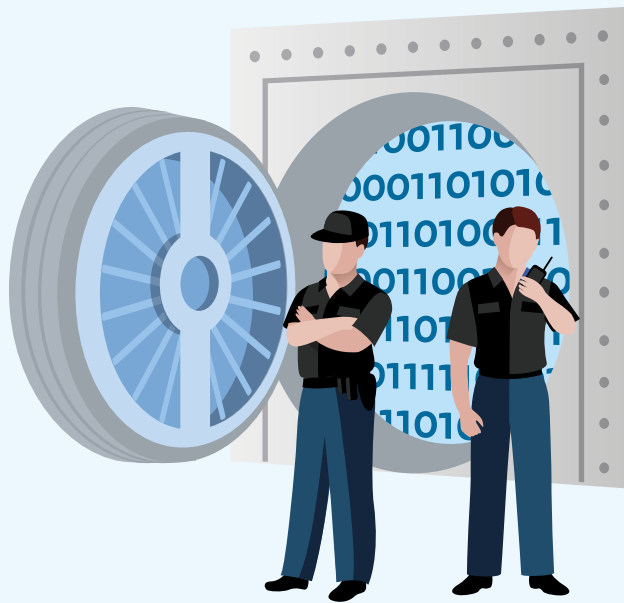
- 1 De wet van Moore, of Moore's law, werd in 1965 beschreven door Gordon Moore, één van de oprichters van Intel, en refereert aan de trend dat het aantal transistoren op een geïntegreerd circuit, of chip, ongeveer iedere 2 jaar verdubbelt. Naast de wet van Moore zien we in aanpalende gebieden van micro-elektronica vergelijkbare exponentiële trends – bijvoorbeeld in digitale opslag, optische resolutie en communicatiesnelheid. Deze decennialange exponentiële functie is de drijvende kracht achter ons gebruik van digitale rekenkracht en opslag in vrijwel alle gebieden van het menselijk bestaan. Zo diepgaand is de impact, dat in 2016 het World Economic Forum zelfs spreekt van een vierde industriële revolutie, met een convergentie van diverse technologietrends in robotica, kunstmatige intelligentie, 3D-printen, smart cities, nanotechnologie en biotechnologie. Moore's 'wet' is natuurlijk meer een marketingstrategie dan een fysieke wetmatigheid, en werd min of meer succesvol tot op heden geïmplementeerd. Er zijn redenen om aan te nemen dat verdere miniaturisatie van ICs in de nabije toekomst steeds moeizamer zal verlopen, aangezien we tegen de grenzen van wat fysiek mogelijk is beginnen te lopen, en verdere miniaturisatie in toenemende mate zal leiden tot instabiele systemen.
- 2 van Est, Q. C., Rerimassie, V., van Keulen, I., & Dorren, G. (2014). Intieme technologie: de slag om ons lichaam en gedrag. Rathenau Instituut.
- 3 Een sensor is in essentie een kunstmatig zintuig dat waarnemingen doet, en deze waarnemingsdata opslaat en/of verder doorstuurt. De waarnemingen lopen uiteen van geluid, licht, nabijheid, of beweging, tot luchtkwaliteit, luchtdruk, neerslag, of temperatuur. Een sensor is permanent in de ruimte aanwezig (in tegenstelling tot tijdelijke opstellingen, zoals bijvoorbeeld een tijdelijke snelheidscamera langs een weg) en doet de waarnemingen continu en zonder menselijke tussenkomst (zo valt bijvoorbeeld het handmatig tellen van verkeersdeelnemers hier dus buiten). Sensoren in de publieke ruimte maken doorgaans gebruik van bestaande infrastructuur voor zowel fysieke plaatsing (bijv. op lantaarnpalen, buitenmuren van openbare gebouwen, langs wegen) als datacommunicatie (bijv. bestaande bekabeling, mobiele netwerken, etc.).
- 4 Peeters, R., & Schuilenburg, M. B. (2014). De geest als aangrijpingspunt. De betekenis van gedragssturing in veiligheid, gezondheid en onderwijs. *Justitiële Verkenningen*, 40 (4), 9-23.
- 5 Kamerstuk 32793-339, brief van de staatssecretaris VWS aan de voorzitter van de Tweede Kamer, 2018, Preventief gezondheidsbeleid, https://www.tweedekamer.nl/kamerstukken/brieven_regering/detail?id=2018Z22083&did=2018D56403
- 6 Kamerstuk 26604-1, brief van de minister BZK aan de voorzitter van de Tweede Kamer, 1999, Integraal Veiligheidsprogramma, <https://zoek.officielebekendmakingen.nl/kst-26604-1.html>
- 7 Fogg, B. J. (1999). Persuasive technologies. *Communications of the ACM*, 42(5), 27-29.
- 8 IJsselstein, W., de Kort, Y., Midden, C., Eggen, B., & Van Den Hoven, E. (2006, May). Persuasive technology for human well-being: setting the scene. In *International conference on persuasive technology* (pp. 1-5). Springer.
- 9 Fogg, B. J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann.
- 9 Deze termen, ubiquitous computing, pervasive computing, context-aware computing en ambient intelligence, verwijzen op enkele subtiele verschillen na in essentie naar dezelfde technologieontwikkeling. Computers maken in toenemende mate deel uit van onze omgeving, verwerkt in alledaagse objecten – lamp, koelkast, fototoestel, televisie, thermostaat. Door gebruik te maken van sensoren en interconnectiviteit leren ze hun omgeving en hun gebruiker steeds beter kennen. Het idee is dat ze hun services aanpassen op basis van het gebruikersprofiel dat wordt opgebouwd, zodat ze de gebruiker steeds beter van dienst kunnen zijn. Tegenwoordig wordt hier meestal aan gerefereerd als het internet der dingen – Internet of Things (IoT) -, waarbij connectiviteit centraal staat. Het IoT bestaat uit semi-intelligente apparaten die via een eigen IP-adres verbonden zijn met het Internet, en daardoor o.a. op afstand (bijv. via een mobiele app) bedienbaar en programmeerbaar zijn, en onderling informatie kunnen uitwisselen.
- 10 Dit is geen science fiction. Veel smartwatches doen al aan metingen van hartslag en ademhaling, waaruit bijvoorbeeld informatie over iemands stressniveau kan worden afgeleid. In 2012 heeft Microsoft een patent ingediend, "Targeting advertisements based on emotion", waarbij men, gebruik makend van de Kinect camera (die wordt gebruikt bij Microsoft's Xbox platform), emoties meet en deze metingen koppelt aan persoonlijk "geoptimaliseerde" advertenties.
- 11 Zie bijvoorbeeld:
Aipperspach, R., Cohen, E., & Canny, J. (2006). Modeling human behavior from simple sensors in the home. In *International Conference on Pervasive Computing* (pp. 337-348). Springer, Berlin, Heidelberg.
Atienza, A. A., & Patrick, K. (2011). Mobile health: the killer app for cyberinfrastructure and consumer health. *American journal of preventive medicine*, 40(5), S151-S153.
Patrick, K., Griswold, W. G., Raab, F., & Intille, S. S. (2008). Health and the mobile phone. *American journal of preventive medicine*, 35(2), 177-181.
Rivera-Illingworth, F., Callaghan, V., & Hagaras, H. (2006). Towards the detection of temporal behavioural patterns in intelligent environments. In *International Conference on Intelligent Environments* (Vol. 1, p. 119). IET.
- 12 Voor een goed overzicht en model van verschillende beïnvloedingsstrategieën die kunnen worden ingezet bij ambient persuasion verwijst ik de lezer naar Kaptein, M. C., Markopoulos, P., De Ruyter, B., & Aarts, E. (2010). Persuasion in ambient intelligence. *Journal of Ambient Intelligence and Humanized Computing*, 1(1), 43-56.
- 13 Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185 (4157)1124-1131. Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan. Ariely, D. (2008). *Predictably irrational*. Harper.
- 14 Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- 15 Sunstein, Cass R. and Thaler, Richard H. (2003a) Libertarian paternalism, *American Economic Review*, Papers and Proceedings, 93(2), pp. 175-79. Sunstein, Cass R. and Thaler, Richard H. (2003b) Libertarian paternalism is not an oxymoron, *University of Chicago Law Review*, 70, pp. 1159-202.
- 16 Big data refereert aan het toepassen van analysetechnieken op grote, vaak ongestructureerde datasets (denk aan data uit sensoren, maar ook aan analyse van online gedrag), om zo patronen en correlaties in de data te kunnen identificeren, die kunnen leiden tot kennis en inzicht.
- 17 Deze term is in 2017 geïntroduceerd door Karen Yeung: "Unlike the static Nudges popularised by Thaler and Sunstein such as placing the salad in front of the lasagne to encourage healthy eating, Big Data analytic nudges are extremely powerful and potent due to their networked, continuously updated, dynamic and pervasive nature (hence 'hypernudge')". Zie: Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118-136.
- 18 Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., ... & Zwitter, A. (2019). Will democracy survive big data and artificial intelligence?. In *Towards Digital Enlightenment* (pp. 73-98). Springer, Cham.
- 19 Cybernetica werd door Norbert Wiener in 1948 gedefinieerd als "de wetenschappelijke studie van

- controle en communicatie in het dier en de machine." Cybernetica is een breed gebied, verwant aan AI, waarbij de doelstelling is om functies en processen van doelgerichte complexe systemen te begrijpen, waarbij waarnemingen (data) worden vergeleken met de gewenste toestand (het doel) van het systeem waarna er actie wordt ondernomen om eventuele discrepantie tussen de huidige staat en de gewenste doeltoestand te verkleinen. Een eenvoudig voorbeeld is de thermostaat – deze heeft een doeltoestand (de gewenste temperatuur in huis) en een set waarnemingen (de huidige temperatuur). Wanneer de gemeten temperatuur lager is dan de gewenste temperatuur, dan zal de verwarming aangaan (actie), tot het moment dat het verschil tussen de gemeten en gewenste temperatuur nul is. Deze circulaire of iteratieve lus van waarneming–doelfunctie–actie vinden we terug in bijna alle geautomatiseerde systemen.
- 20 Frank, L. E. & Nickel, P. J., (2017). E-coaching in de gezondheidszorg: is het zacht paternalisme? In: Wobbes, T., & van den Muijsenbergh, M. (eds.) *Verleiding tot gezond gedrag*, 89-103.
- 21 Sugden, R. (2009). On nudging: A review of nudge: Improving decisions about health, wealth and happiness by Richard H. Thaler and Cass R. Sunstein.
- 22 Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press. Verbeek, P. P. (2006). Materializing morality: Design ethics and technological mediation. *Science, Technology, & Human Values*, 31(3), 361-380.
- 23 Verbeek, P.P. (2014). *Op de vleugels van Icarus. Hoe techniek en moraal met elkaar meebewegen*. Lemniscaat.
- 24 Joy Buolamwini (2019). Artificial Intelligence Has a Problem With Gender and Racial Bias. Here's How to Solve It. <https://time.com/5520558/artificial-intelligence-racial-gender-bias/>
- 25 Wu, X., & Zhang, X. (2016). Automated inference on criminality using face images. arXiv preprint arXiv:1611.04135, 4038-4052.
- 26 Zie: <https://aif360.mybluemix.net/> en K.R. Varshney (2018). Introducing AI Fairness 360. <https://www.ibm.com/blogs/research/2018/09/ai-fairness-360/>
- 27 Michael, M. G., Fusco, S. J., & Michael, K. (2008). A research note on ethics in the emerging age of überveillance. *Computer communications*, 31(6), 1192-1199.
- 28 Clarke, R. A. (1988). Information technology and dataveillance. *Commun. ACM*, 31(5), 498-512.
- 29 Koffijberg, J., Dekkers, S., Homburg, G., & Van den Berg, B. (2009). Niets te verbergen en toch bang. Nederlandse burgers over het gebruik van hun gegevens in de glazen samenleving. Den Haag/ Amsterdam, College Bescherming Persoonsgegevens/Regioplan.
- 30 Burgoon, J. K., Parrott, R., LePoire, B. A., Kelley, D. L., Walther, J. B. and Perry, D. (1989). Maintaining and restoring privacy through communication in different types of relationship. *Journal of Social and Personal Relationships* 6, 131-158
- 31 The Economist, 17 December 2016, Big data, meet Big Brother: China invents the digital totalitarian state. The worrying implications of its social-credit project. <https://www.economist.com/briefing/2016/12/17/china-invents-the-digital-totalitarian-state>
- 32 Een saillant voorbeeld wordt gegeven door Maurits Martijn en Dimitri Tokmetzis in hun boek "Je hebt wél iets te verbergen" (2016, De Correspondent). Zij installeren een stukje sniffer-software, genaamd Charles, om het netwerkverkeer tussen een Samsung smartphone en externe partijen vast te leggen. Een online bestelling van een leren tas bij de Bijenkorf leidt onmiddellijk tot een hectisch online biedingsproces op de persoonlijke data. "In één enkele seconde is contact gelegd met servers in de Verenigde Staten, Zweden, Duitsland, Ierland en Nederland, die behoren tot bedrijven die luisteren naar namen als Improve Digital, Admeta, Adtech, Metrigo, Burst Media, Yieldlab, Switch Concepts, AppNexus, Sociomantic, Adscale, Rubicon Project, OpenX, Smart Adserver en Casale Media" (p.28).
- 33 Obar, J. A., & Oeldorf-Hirsch, A. (2018). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, 1-20.
- 34 Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.
- 35 In een artikel over de onbedoelde bijeffecten van sensoren getiteld "Maak van sensornetwerk geen buurtwacht", geeft Ardy Siegert een aantal gefingeerde, maar realistische voorbeelden. Zie: <https://ibestuur.nl/podium/maak-van-sensornetwerk-geen-nieuwsgierige-buurtwacht>
- 36 den Ouden, W., Boogaard, G. & Driessen, E.M.M.A. (2018). Right to Challenge. Een voorstudie naar de mogelijkheden voor een algemene regeling voor het 'Right to Challenge' en andere burgerinitiatieven in Nederland. Universiteit Leiden, Instituut voor Publieksrecht.
- 37 Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., ... & Zwitter, A. (2019). Will democracy survive big data and artificial intelligence?. In *Towards Digital Enlightenment* (pp. 73-98). Springer, Cham.

De bestuurlijke regie over experimentele data en algoritmen

Aanbevelingen voor een nationaal algoritmeregister en rapportagekader

Prof. dr. Gerd Kortuem
Hoogleraar Internet of Things
TU Delft



Voorwoord

Door Theo Veltman

Prof. dr. Gerd Kortuem en ik hebben elkaar enkele keren ontmoet in de context van de gemeente Amsterdam. Ik heb hem leren kennen als een aimabele, bescheiden deskundige en een gezaghebbende wetenschapper. Hij is één van de mensen die een breed veld van technologie overziet. Hij doet dat vanuit zijn grote kennis van en brede ervaring met de technologie en weet die kennis en ervaring te verbinden met vraagstukken rond ons menszijn en ons verblijf in een gedigitaliseerde wereld. Ook zijn essay illustreert dat.

Het onderzoek van Gerd richt zich op het Internet of Things (internet der dingen); het ontwerp en gebruik van data, algoritmen en slimme apparaten als belangrijke bouwstenen voor intelligente producten en diensten. Hij is lid van het AITech-initiatief voor betekenisvolle menselijke controle van autonome intelligente technologie. Gerd is verbonden aan de TU Delft als hoogleraar. Hij is daar tevens voorzitter 'Internet of Things' binnen de faculteit Industrieel Ontwerpen. Ook is hij verbonden aan het AMS (Amsterdam Institute for Advanced Metropolitan Solutions).

Het onderwerp 'onbehoorlijk datagebruik in de openbare ruimte' vind ik belangrijk omdat vrijheid een groot goed is. Het wordt in zekere zin bedreigd door de toenemende digitalisering. Zeker als het gaat om de vrijheid van bewegen, zonder gemonitord of geregistreerd te worden; te kunnen doen wat je wilt zonder (digitaal) toezicht. Ook de vrijheid van oordeelsvorming is belangrijk, dat je niet ongewild en onmerkbaar wordt beïnvloed in jouw mening, jouw keuzes. Die vrijheden staan onder druk. Digitalisering biedt voordelen. Al is het alleen maar dat elke persoon, ongeacht zijn of haar locatie, toegang kan hebben tot alle beschikbare informatie en kennis om zichzelf te ontwikkelen. Dat voordeel moet er zijn, voor iedereen. 'Onbehoorlijk datagebruik' is een ongewenst bijeffect. Het moet verminderd worden, dus worden tegengegaan.

In de begeleidingscommissie Behoorlijk datagebruik in de openbare ruimte heb ik weer geleerd hoeveel we van elkaar kunnen leren. De deelname aan deze groep van gedreven, deskundige mensen gaf mij veel energie en bracht mij meerdere nieuwe inzichten.

Theo Veltman
Lid begeleidingscommissie

Samenvatting

Data en algoritmen hebben in de loop der tijd een bepalende rol in de samenleving gekregen. Ze beïnvloeden welk nieuws we online lezen, hoe verkeer door de stad wordt geleid en hoe we contact met elkaar hebben. Steeds vaker gebruiken bedrijven en overheden machinelearning-technieken om beslissingen met significante implicaties voor mens en maatschappij te automatiseren, bijvoorbeeld in de ordehandhaving, bij het opleggen van straffen en in de sociale dienstverlening.

Algoritmische besluitvorming wordt echter ook bekritiseerd omdat de mogelijkheid bestaat dat vooroordelen en discriminatie erdoor worden versterkt. Bovendien ontbreken transparantie en verantwoording (*accountability*). Een van de belangrijkste vraagstukken voor overheden is dus hoe ze het ontwerp, de ontwikkeling en het gebruik van data en algoritmen in de samenleving zo kunnen inrichten en beheersen dat de voordelen voor het publiek worden gemaximaliseerd en maatschappelijke waarden worden gerespecteerd.

Het is om verschillende redenen moeilijk om op een doeltreffende manier de bestuurlijke regie (*governance*) te houden over data en algoritmen. Ten eerste ontwikkelen algoritmen zich razendsnel tot een niveau van complexiteit en geavanceerdheid waar zelfs computerspecialisten nog maar amper begrijpen hoe algoritmen werken en beslissingen nemen. Ten tweede leidt de gestaag vorderende toepassing van geautomatiseerde machinelearning-tools en zelflerende algoritmen tot algoritmische systemen die snel en op onverwachte en vaak onvoorspelbare manieren evolueren. Ten derde zijn de effecten van datagedreven algoritmen op de samenleving dermate groot en vergaand, dat ze onmogelijk op voorhand kunnen worden voorspeld voordat dergelijke systemen daadwerkelijk worden ingezet in de samenleving. Om deze redenen spelen *living labs* (Schuurman, 2015) een sleutelrol in de experimentele ontwikkeling van data en algoritmen.

In dit artikel onderzoek ik de morele vraagstukken die zijn verbonden aan het experimenteren met data en algoritmen en ontwikkel ik aanbevelingen voor een nationaal bestuurskader voor data en algoritmen, dat uit vijf onderdelen bestaat: 1) een **Algoritme Rapportage Initiatief**, dat tot doel heeft de maatschappelijke waarde en risico's van algoritmeprojecten in Nederland te documenteren en te volgen; 2) een **Nationaal Algoritme Register** dat een effectieve vergelijking en beoordeling van initiatieven op het gebied van data en algoritmen binnen Nederland mogelijk moet maken; 3) een **Openbaar Algoritme Forum** waarop publieke belanghebbenden de werking, het gebruik en de uitkomsten van initiatieven op het gebied van data en algoritmen kunnen bespreken, bekritisieren en betwisten; 4) een **Data en Algoritme Instituut** ter bevordering van de ontwikkeling van kennis, benaderingen, tools, infrastructuren en normen voor ethisch en verantwoord gebruik van algoritmen en 5) een **Nationale Skills Agenda voor Data en Algoritmen** om ervoor te zorgen dat alle belanghebbenden (van burgers tot organisaties) over de juiste vaardigheden beschikken om deel te nemen aan de nieuwe data- en algoritme-economie.

Deze aanbevelingen zijn gestoeld op een aantal belangrijke bronnen, waaronder *Experimental Ethics* van Van de Poel (van de Poel, 2016), de FAT/ML 'Principles for Accountable Algorithms and Social Impact Statement for Algorithms' (Diakopoulos et al., 2018), geïntegreerde rapportage (Banerjee, 2019; Eccles & Krzus, 2015), een veelgebruikt model voor het volgen van en rapporteren over financiële en niet-financiële waardecreatie binnen publieke en private organisaties, en ten slotte het werk van NESTA over AI-gerelateerde ethiek en bestuur (Mulgan, 2016).

Dit essay is vertaald vanuit het Engels.

1. Inleiding

In 2001 deed Marc Andreessen de befaamde uitspraak dat ‘software de wereld opeet’. Hiermee onderstreepte hij de drastische technologische en economische verschuiving waarin softwarebedrijven steeds grotere delen van de economie overnemen (Andreessen, 2011). Op dit moment is eenzelfde verschuiving voelbaar in de hele samenleving, veroorzaakt door de almaar groeiende rol van datagedreven algoritmen in beslissingen die grote gevolgen voor mens en maatschappij hebben, bijvoorbeeld op het gebied van ordehandhaving, het opleggen van straffen, kredietverstrekking en het aannemen van personeel. Data is op ongekende schaal beschikbaar. In combinatie met steeds krachtiger wordende machinelearning-methoden kunnen bedrijven, overheden en de publieke sector algoritmen trainen om complexe problemen aan te pakken (Willson, 2017). Datagedreven algoritmen kunnen de openbare dienstverlening verbeteren en de overheid efficiënter maken door processen te optimaliseren, feedback te analyseren en uitkomsten te voorspellen. Algoritmische besluitvorming wordt echter ook bekritiseerd omdat de mogelijkheid bestaat dat vooroordelen (bias) en discriminatie erdoor worden versterkt. Ook ontbreekt het aan transparantie en verantwoording (*accountability*). De meest fundamentele vraag met betrekking tot data en algoritmen is dan ook:

Hoe kunnen we het ontwerp, de ontwikkeling en het gebruik van data en algoritmen in de samenleving zo inrichten en aansturen dat de voordelen voor het publiek worden gemaximaliseerd en maatschappelijke waarden worden gerespecteerd?

In deze paper probeer ik die vraag te beantwoorden aan de hand van vijf concrete aanbevelingen voor een ethisch verantwoorde en doeltreffende regie over datagedreven algoritmen. In de aanloop naar deze aanbevelingen bespreek ik in paragraaf 2 en 3 zes kernargumenten:

- Datagedreven algoritmen spelen een steeds grotere rol in besluitvormingsprocessen met verreikende maatschappelijke consequenties.

- Algoritmen worden steeds complexer en geavanceerder, zodat zelfs computerspecialisten maar moeilijk kunnen begrijpen hoe datagedreven algoritmen werken.
- Organisaties werken steeds vaker met geautomatiseerde machinelearning-processen. Dit versnelt het ontwerp, de ontwikkeling en de inzet van datagedreven systemen.
- Opkomende machinelearning-systemen zijn zelflerend. Ze veranderen en ontwikkelen zich voortdurend.
- De impact van datagedreven algoritmen op de samenleving kunnen onmogelijk op voorhand worden voorspeld, voordat die algoritmen daadwerkelijk worden geïmplementeerd.
- Living labs bieden een effectieve omgeving voor de geleidelijke en experimentele introductie van datagedreven algoritmen in de samenleving.
- Er zijn geen bestuurlijke kaders om in de samenleving op een veilige en ethisch verantwoorde wijze te experimenteren met data en algoritmen.

Deze argumenten leiden tot de volgende vijf aanbevelingen, die worden beschreven in paragraaf 5:

- Aanbeveling 1: Richt een **Algoritme Rapportage Initiatief** op, dat tot doel heeft de maatschappelijke waarde en risico's van algoritmeprojecten in Nederland te documenteren en te volgen.
- Aanbeveling 2: Ontwikkel een **Nationaal Algoritme Register** dat een effectieve vergelijking en beoordeling van initiatieven op het gebied van data en algoritmen in Nederland mogelijk moet maken.
- Aanbeveling 3: Ontwikkel een **Publiek Algoritme Forum** om publieke partijen in staat te stellen de werking, het gebruik en de uitkomsten van initiatieven op het gebied van data en algoritmen te bespreken, te bekritisieren en te betwisten.
- Aanbeveling 4: Richt een **Data en Algoritme Instituut** op ter bevordering van de ontwikkeling van kennis, benaderingen, tools, infrastructuren en normen voor ethisch en verantwoord gebruik van algoritmen.
- Aanbeveling 5: Ontwikkel een **Nationale Skills Agenda voor Data en Algoritmen** om ervoor te zorgen dat alle belanghebbenden (van burgers tot organisaties) over de juiste vaardigheden beschikken om deel te nemen aan de nieuwe data- en algoritme-economie.

Deze aanbevelingen zijn gebaseerd op een literatuuronderzoek, mijn eigen ervaringen met het ontwikkelen van datagedreven algoritmen en systemen in verscheidene smart city-projecten en inzichten uit drie bronnen, namelijk *Experimental Ethics* van Van de Poel (van de Poel, 2016), de FAT/ML 'Principles for Accountable Algorithms and Social Impact Statement for Algorithms' (Diakopoulos et al., 2018) en *Integrated Reporting* (geïntegreerde rapportage) (Banerjee, 2019; Eccles & Krzus, 2015), een veelgebruikt model voor het volgen van en rapporteren over financiële en niet financiële waardecreatie van publieke en private organisaties.

2. Data en algoritmen in de samenleving

Algoritmen worden steeds belangrijker in ons privé- en publieke leven. Ze hebben een enorme invloed op de samenleving. Ze beïnvloeden welk nieuws we online lezen, hoe verkeer door de stad wordt geleid en hoe openbare en particuliere organisaties werken en diensten verlenen.

2.1 Data, algoritmen en machine learning

Traditioneel gezien zijn algoritmen altijd gedefinieerd in de zin van procedures, als een serie instructies om een bepaalde output of een bepaald resultaat te produceren. Zo definieert Kitchin algoritmen als:

“Een serie afgebakende, gestructureerde stappen voor het verwerken van instructies/data, met als doel een bepaalde output te produceren.” (Kitchin, 2017a)



En Kraemer et al. stelden de volgende definitie voor (Kraemer, van Overveld, & Peterson, 2011):

“Een algoritme is ruwweg een eindige opeenvolging van duidelijk gedefinieerde instructies die in voldoende detail omschrijven hoe een probleem moet worden opgelost.”

Deze definities onderstrepen het procedurele aspect van algoritmen en geven impliciet aan dat een algoritme volledig kan worden gespecificeerd (door een menselijke ontwikkelaar) door de stappen om een uitkomst te berekenen of een probleem op te lossen, in hun geheel op te noemen. Deze procedurele en verklarende definitie is echter niet meer afdoende in de context van *big data* en *machine learning*.

Machine learning is een categorie algoritmen waarmee softwaresystemen automatisch uitkomsten kunnen berekenen door patronen uit data aan te leren zonder expliciet te zijn geprogrammeerd door een menselijke ontwikkelaar. In plaats van een algoritme te specificeren houdt een machine learning engineer slechts toezicht op het machinelearning-proces door trainingsdata te verschaffen, het geschikte type machinelearning-algoritme te selecteren en de machine learning output te evalueren. De uitkomst van dit trainings- en leerproces is een machinelearning-model dat, op basis van de data die erin wordt gestopt, zelf beslissingen kan nemen en zijn prestaties kan verbeteren. Daar is verder weinig betrokkenheid van een menselijke trainer bij nodig.

Het grote verschil tussen een traditioneel algoritme en een machinelearning-algoritme ligt in het gebruik van data. Een traditioneel algoritme werkt met data, terwijl een machinelearning-algoritme eerst wordt getraind op data en pas daarna met die data gaat werken. Het grote voordeel van machine learning boven conventionele algoritmen is dat machine learning kan worden gebruikt in gevallen waarin er geen bekende oplossing of benadering voor handen is om een probleem op te lossen, of waarin het heel moeilijk is om 'met de hand' een algoritme te vinden en te specificeren (bijvoorbeeld beeldanalyse, tekstbegrip en het besturen en optimaliseren van complexe systemen).

Alhoewel een machinelearning-model in technische zin nog steeds een algoritme is, bestaat het uitsluitend in een interne representatie (bijvoorbeeld een neuraal net) die zelfs voor machinelearning-specialisten ondoorgroendelijk kan zijn. Met andere woorden, de ontwikkelaar begrijpt niet altijd hoe en waarom een machinelearning-systeem een bepaalde output genereert. Dit is het bekende 'black box'-probleem van machinelearning-systemen (Gasser & Almeida, 2017; Pasquale, 2015)

Het is onmogelijk om hoogwaardige algoritmen te ontwikkelen met behulp van slechte data

De directe relatie tussen data en algoritmen impliceert dat datakwaliteit de kwaliteit van (datagedreven) algoritmen bepaalt. Het is onmogelijk om hoogwaardige algoritmen te ontwikkelen met behulp van slechte data. Data en algoritmen moeten beide worden gecontroleerd op kwaliteitsattributen als nauwkeurigheid, volledigheid, betrouwbaarheid, relevantie, tijdigheid, onvooringenomenheid enz. Daarom moeten data en algoritmen in bestuur en beleid als een onlosmakelijke eenheid worden gezien.

2.2 Algoritmen vs. algoritmische systemen

Een algoritme is meer dan een technische constructie. Kitchin stelt dat algoritmen 'performatief van aard zijn en ingebed in bredere sociaal-technische assemblages' (Kitchin, 2017b). Dit betekent dat we om algoritmen in de samenleving te kunnen begrijpen, moeten analyseren waarom algoritmen worden gebruikt, hoe ze functioneren in de echte wereld, hoe gebruikers erop reageren en wat de consequenties zijn van de taken die ze uitvoeren (Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016). Eenzelfde algoritme kan in verschillende situaties tot uiteenlopende uitkomsten leiden. Volgens Kitchin 'presteren algoritmen in context (in samenwerking met data, technologieën, mensen enz. onder wisselende omstandigheden). Daarom worden hun effecten op onderling afhankelijke en relationele wijze zichtbaar, en leiden ze tot gelokaliseerde en situationele uitkomsten' (Kitchin, 2017b). Dit wil zeggen dat algoritmen meer zijn dan slechts objectieve, waardenneutrale artefacten.

De menselijke keuzes die hun gebruik, werking en effecten bepalen, worden beïnvloed door subjectieve en vaak impliciete waardeoordelen. Met andere woorden, een ethische verkenning van algoritmen moet ook rekening houden met hoe algoritmen worden gemaakt, met de algoritmen zelf, de context waarin ze worden gebruikt en hun impact op mens en maatschappij.

2.3 Het groeiende belang van algoritmen in de samenleving

Overheidsinstellingen, publieke organisaties en gemeenten in Nederland maken steeds meer en vaker gebruik van digitale systemen om grote maatschappelijke vraagstukken aan te pakken en essentiële dienstverlening te bieden. Steeds vaker bestaat de kern van zulke systemen uit datagedreven algoritmen die de voorziening van openbare diensten stroomlijnen of zelfs automatiseren (Kitchin, 2014). Zo gebruiken steden satellietbeelden en computervisie om automatisch vanuit de lucht te controleren of bouwprojecten zich aan de regels houden. Transportautoriteiten maken gebruik van ticketgegevens om forensenverkeer te volgen en de toekomstige mobiliteitsvraag te voorspellen. Energiebedrijven gebruiken data uit elektriciteitsmeters om energieverbruik en CO₂-emissie te volgen en te voorspellen. En de politie gebruikt data van camera's en microfoons om misdrijven in realtime te detecteren en criminaliteitshotspots in kaart te brengen. Daarnaast gebruiken bedrijven als Uber en Airbnb data en algoritmen om nieuwe transport- en accommodatiediensten te leveren.

2.4 Inzicht in hoe algoritmen worden gebruikt

Wanneer we kijken naar de rol van algoritmen in de samenleving en wat hierbij de morele vraagstukken zijn, moeten we drie belangrijke vragen stellen:

- Wie controleert het gebruik en de inzet van algoritmen en in wiens belang?
- Hoe wordt het gebruik van algoritmen gerechtvaardigd?
- Voor welk doel worden algoritmen gebruikt?

Wie? Publieke en particuliere algoritmen en algoritmen gericht op het algemeen belang
Algoritmen worden in de hele samenleving gebruikt en hebben invloed op alle aspecten van het dagelijks leven. Om te begrijpen wie de controle heeft over het gebruik en de inzet van algoritmen, stel ik voor algoritmen in te delen in drie categorieën: publieke en particuliere algoritmen en algoritmen gericht op het algemeen belang.

Particuliere algoritmen: de grootste invloed van algoritmen in de samenleving komt van algoritmen die worden gebruikt door particuliere bedrijven als Google, Facebook, Airbnb en Uber. De algoritmen die door zulke bedrijven worden gebruikt noem ik **particuliere algoritmen**. Ze zijn particulier omdat a) ze private (d.w.z. commerciële) belangen dienen, b) ze worden aangestuurd door kleine groepen mensen die niet noodzakelijk bekend zijn buiten het bedrijf en c) hun functie en werking doorgaans niet worden geopenbaard aan externe partijen. Het meest fundamentele voorbeeld is het PageRank-algoritme van Google, dat het relatieve belang van webbronnen berekent en zo beïnvloedt wat gebruikers wel en niet zien op het web. De enige manier waarop de samenleving inzicht kan krijgen in de werking van particuliere algoritmen, is door hun van buitenaf zichtbare gedrag en hun directe en indirecte impact op personen en de samenleving te observeren.

Publieke algoritmen: algoritmen worden steeds vaker door overheden gebruikt om beleid te formuleren en openbare dienstverlening aan te bieden. Boetes die geautomatiseerd worden opgelegd, voorspellend profileren door de politie en geautomatiseerde beslissingen over uitkeringen zijn relevante voorbeelden van wat ik **publieke algoritmen** noem. Dergelijke publieke algoritmen helpen mensen in overheidsinstellingen niet alleen meer procedurele regels toe te passen. Ze zijn uitgegroeid tot primaire besluitvormers in het overheidsbeleid (Eubanks, 2015). Eubanks beschrijft deze situatie als volgt: ‘De algoritmen die beleidsvorming domineren, vooral in openbare dienstverlening zoals ordehandhaving, uitkeringen en kindbescherming, werken niet zozeer als ‘gegevenssorteerders’ maar meer als poortwachters. Ze bepalen wie toegang krijgt tot openbare middelen, beoordelen risico’s en delen groepen mensen in in categorieën, zoals ‘zij die wel of niet steun verdienen’ en ‘verdacht’ of ‘niet verdacht’ (Eubanks, 2015). Publieke algoritmen zijn publiek omdat ze door publieke organisaties worden gebruikt in het algemeen belang. In theorie hebben burgers het democratische recht om inzicht te krijgen in hoe en waarom dergelijke algoritmen worden gebruikt, maar in de praktijk zijn publieke organisaties en overheidsinstellingen net zomin bereid om hun algoritmen te openbaren als commerciële bedrijven.

Algoritmen gericht op het algemeen belang: tussen publieke en particuliere algoritmen in zitten wat ik algoritmen gericht op het algemeen belang noem.

Dit zijn algoritmen die worden gebruikt als onderdeel van openbare dienstverlening die wordt verschaft door een particulier bedrijf of een publiek-privaat samenwerkingsverband. Hieronder vallen bijvoorbeeld algoritmen die worden gebruikt door elektriciteits- en waterbedrijven en verkeersorganisaties. Een concreet voorbeeld is het gebruik van slimme EV-laadalgoritmen in het elektriciteitsnet, die realtime beslissen hoeveel stroom individuele elektrische voertuigen mogen gebruiken wanneer ze zijn aangesloten op een laadpunt (Liu, McNamara, & McLoone, 2013; Turel, Joskin, Geerts, Kaathoven, & Schouwenaar, 2017). Zulke algoritmen zijn niet volledig publiek of particulier. Ze worden gebruikt om een essentiële openbare dienst te verlenen (en zijn dus in het algemeen belang) maar worden mogelijk toegepast door particuliere organisaties met commerciële belangen. Organisaties die deze algoritmen toepassen en controleren zijn doorgaans niet verplicht of bereid hun algoritmen te openbaren en er bestaan potentiële belangenconflicten tussen publieke belanghebbenden (de samenleving in zijn geheel) en particuliere belanghebbenden (managementteam, eigenaren enz.).

Publieke algoritmen	Algoritmen gericht op het algemeen belang	Particuliere algoritmen
<ol style="list-style-type: none"> 1. Gericht op het algemeen belang 2. Onderworpen aan publiek toezicht 3. Bepaalde mate van transparantie en openbaarheid (in theorie) 	<ol style="list-style-type: none"> 1. Gericht op algemene en particuliere belangen 2. Onder controle van particuliere bedrijven of publiek-private samenwerkingsverbanden 3. Beperkte transparantie en openbaarheid 	<ol style="list-style-type: none"> 1. Gericht op commerciële belangen 2. Streng gecontroleerd 3. Geen transparantie en openbaarheid

Figuur 1 Classificatie van algoritmen

Waarom? Het argument voor het gebruik van algoritmen in het overheidsbeleid

Particuliere algoritmen worden door commerciële organisaties gebruikt om een eenvoudige reden: in de allereerste plaats om hun commerciële en zakelijke belangen te behartigen. Gebruikers kunnen er mogelijk van profiteren maar dit is van secundair belang.

Het gebruik van publieke algoritmen in beleid en openbaar bestuur kan worden gemotiveerd door hun bijdrage aan het algemeen belang. Volgens de theorie van het utilitarisme van het maatschappelijk welzijn (Goodin, 1995) zijn beleidsmakers verplicht zo te handelen dat ze het maatschappelijk welzijn maximaliseren. Het gebruik van algoritmen kan leiden tot enorme verbeteringen in hoe openbare instellingen hun taken uitvoeren (Kroeskop, 2018) door nieuwe beleidskansen te creëren, de overheids- en openbare dienstverlening efficiënter te maken, de consistentie en objectiviteit van besluitvorming te verbeteren en menselijke fouten te verminderen (Achrekar, Gandhe, Lazarus, Yu, & Liu, 2011; Chen & Hsieh, 2014; Joseph & Johnson, 2013; Naik & Bhide, 2014; van der Voort, Klievink, Arnaboldi, & Meijer, 2019).

Wat? Meten, besluitvorming vs. gedragsverandering

Om meer inzicht te krijgen in het gebruik van data en toepassingen in de openbare ruimte, introduceer ik drie gangbare gebruikspatronen (Figuur 2):

Meten en toezicht houden: dit betreft gevallen waarin data en algoritmen voornamelijk worden gebruikt voor dataverzameling en -aggregatie maar interpretatie en besluitvorming nog steeds grotendeels mensenwerk is. Voorbeelden zijn verkeerstoezicht op afstand, het gebruik van data in stadsplanning en eenvoudige vormen van voorspellende ordehandhaving.

Geautomatiseerde besluitvorming: geautomatiseerde besluitvorming betreft gevallen waarin data en algoritmen worden gebruikt om kritieke beslissingen te automatiseren. Voorbeelden zijn geautomatiseerde verkeerssystemen die stoplichten bijstellen om opstoppingen te verminderen, de automatische aanmaak van verkeersboetes op basis van camerabeelden en meer geavanceerde vormen van ordehandhaving.

Beïnvloeding en gedragsverandering: gedragsverandering betreft het gebruik van data en algoritmen om mensen te beïnvloeden of ertoe te brengen zich op een bepaalde manier te gedragen en openbare diensten, bijvoorbeeld vervoer, op een duurzamere manier te gebruiken (Gandy & Nemorin, 2018; Ranchordás, 2019). Zo combineert Enschede realtime verkeersdata met een slimme telefoonapp (bijv. SMART-app: Self-Motivated and Rewarded Traveling) om mensen aan te moedigen alternatieve, minder drukke autoroutes te nemen of over te stappen op het openbaar vervoer of de fiets en ze te belonen voor 'goed gedrag'. Dit en andere voorbeelden zijn gebaseerd op het belangrijke werk van Sunstein en Thaler over 'nudging' (Thaler Richard H., 2008), dat de theoretische grondslag legde voor het combineren van gedrags- en voorspellende modellen om mensen zo te beïnvloeden dat ze 'verstandigere' beslissingen nemen.

Metten en toezicht	Geautomatiseerde besluitvorming	Gedragsverandering
1. Data verzamelen en analyseren	1. Data verzamelen en analyseren 2. Besluitvorming	1. Data verzamelen en analyseren 2. Besluitvorming 3. Leren, aanpassen en optimaliseren

Figuur 2 Drie patronen van data- en algoritmegebruik

2.5 De morele problematiek rond algoritmen

Technologie speelt een belangrijke rol in het definiëren van het sociaal contract van de toekomst. Data en algoritmen zijn verborgen instrumenten geworden ten behoeve van meer intensieve surveillance, bedrijven die winst willen maken en, op zijn slechtst, maatschappelijke controle. In een review van het ethische debat rond algoritmische systemen identificeren Mittelstadt et al. de volgende problemen (Mittelstadt et al., 2016):

- Omdat algoritmen gebruikmaken van inductieve correlaties, is het mogelijk dat de hieruit voortvloeiende beslissingen niet gerechtvaardigd zijn.
- Algoritmen zijn vaak niet goed toegankelijk voor inspectie. Daarnaast zijn ze vaak moeilijk te begrijpen, met name wanneer machine learning wordt gebruikt.

Het gebrek aan toegankelijkheid en begrijpelijkheid leidt tot algoritmische ondoorzichtigheid, wat het vertrouwen van mensen in algoritmische systemen en de organisaties die ze gebruiken, kan schaden.

- Elk stuk technologie ondersteunt bepaalde waarden en ondermijnt andere waarden. Algoritmen zijn hierin geen uitzondering. Deze vooringenomenheid ('bias') kan voortkomen uit de opzettelijke of onopzettelijke overdracht van de waarden van de mensen die algoritmische systemen bouwen. Daarnaast kunnen technische tekortkomingen en slechte data leiden tot algoritmen met een ingebouwde vooringenomenheid. En ten slotte kunnen onvoorziene interacties tussen het algoritme en de wereld tijdens het gebruik tot vooringenomenheid leiden.
- Alhoewel vooringenomenheid een eigenschap is van de besluitvorming zelf, is discriminatie een eigenschap van de uitkomst van een algoritmisch systeem dat een disproportionele negatieve impact heeft op een bepaalde groep mensen.
- De praktijken van profielering, nudging en personalisering doen afbreuk aan vrij menselijk handelen.
- Profielering, data-analyse en het delen van data met derden is schadelijk voor de informatieve privacy.
- De groeiende onvoorspelbaarheid van algoritmen en de verschuiving van controle van programmeurs naar andere delen van het systeem door het gebruik van machine learning leidt tot vraagstukken op het gebied van morele verantwoordelijkheid. Daarnaast kan het feit dat algoritmen elkaars output consumeren, leiden tot hiaten in de verantwoording (accountability).



Deze problemen bestrijken het hele spectrum van technische, organisatorische en publieke aspecten (Figuur 3).

Op **technologisch vlak** doen morele vraagstukken zich voor vanwege de complexiteit en ondoorzichtigheid van AI-algoritmen, wat zich vertaalt naar een gebrek aan transparantie en verklaarbaarheid, zelfs voor informaticaspecialisten.

Op **het niveau van organisaties en instellingen** bestaan vooral morele zorgen over de verantwoording van de besluitvorming en de bestuurlijke regie over algoritmische systemen. De World Wide Web Foundation definieert verantwoording (*accountability*) als 'de plicht om algoritmische besluitvorming te rapporteren, verklaren of rechtvaardigen en eventuele negatieve maatschappelijke effecten of berokkende schade te beperken en te verzachten' (World Wide Web Foundation, 2017).

Burgers en het publiek	Legitimiteit en betwistbaarheid: Hoe kunnen burgers datagedreven beslissingen bekritisieren en aanvechten?
Organisaties en instellingen	Bestuurbaarheid: Hoe kunnen openbare instellingen datagedreven systemen en hun uitkomsten beïnvloeden, bewaken en er toezicht op houden? Verantwoording: Hoe kunnen direct betrokkenen verantwoording afleggen voor de beslissing die is genomen door data en algoritmen?
Data verzamelen en analyseren	Verklaarbaarheid: Hoe kunnen algoritmen en besluitvormingsprocessen verklaarbaar worden gemaakt voor specialisten en andere geïnteresseerden? Transparantie: Welke beslissingen worden genomen door een datagedreven algoritme?

Figuur 3 Van technische tot publieke vraagstukken

De definitie van algoritmische verantwoording maakt gebruik van dezelfde mechanismen als verantwoording in conventioneel overheidsbeleid. Wanneer een algoritme wordt geïmplementeerd in een informatiesysteem voor overheidsbeleid, wordt dat systeem de 'handelende partij' aan wie een overheidsinstelling zijn autoriteit delegeert (Introna, 2016).

Op **maatschappelijk vlak** is het voornaamste vraagstuk de legitimiteit van het gebruik van technologie en het vermogen van burgers om geautomatiseerde besluitvormingssystemen op betekenisvolle wijze te kunnen tegenspreken. Veelbesproken en herhaaldelijke fouten van AI-systemen die in de openbare dienstverlening worden ingezet (mensen die onterecht worden beschuldigd misdrijven die ze niet hebben begaan of uitkeringen die onterecht worden geweigerd) ondermijnen het vertrouwen in deze technologie. Bovendien kunnen ze ook de legitimiteit van de instelling die de technologie gebruikt negatief beïnvloeden in de ogen van het publiek. Als gevolg van het groeiende gat tussen de snelheid waarmee technologie zich ontwikkelt en de mate waarin beleidsmakers, technologiespecialisten, professionals en het algemene publiek daar inzicht in hebben, spreken onderzoekers wel van een AI 'legitimiteitsgat' (Brown, 2018). Zij willen betekenisvolle gelegenheden creëren waar het publiek en ambtenaren de mogelijke toepassingen van AI in een maatschappelijke context kunnen bespreken en aan een kritische blik kunnen onderwerpen.

2.6 De risico's van algoritmen in de publieke sector

Binnen de overheid bestaat een groeiende wens om data en algoritmen in te zetten om de werking van openbare dienstverlening en de publieke infrastructuur te rationaliseren en te automatiseren (Janssen & Kuk, 2016). Zoals al eerder besproken vergroot besluitvorming op basis van algoritmen de efficiëntie en eerlijkheid van overheidsdiensten. Aan deze voordelen kleven echter ook een heleboel risico's, zoals de mogelijkheid dat er ongewenste vooroordelen/voorkeuren worden ontwikkeld in algoritmen (Angwin, Larson, Mattu, & Kirchner, 2016) of er systematische ongelijkheden ontstaan (Woodson, 2018). Wanneer algoritmen niet op de juiste manier worden ontwikkeld of geïmplementeerd, kunnen ze minder nauwkeurig zijn dan het oordeel van overheidsambtenaren. Ook kunnen ze vooroordelen/voorkeuren in de data waarmee ze worden getraind, formaliseren en verhullen. Bovendien kunnen algoritmen beleidsoordelen uitspreken, die afwijken van de voorkeuren van burgers of hun verkozen

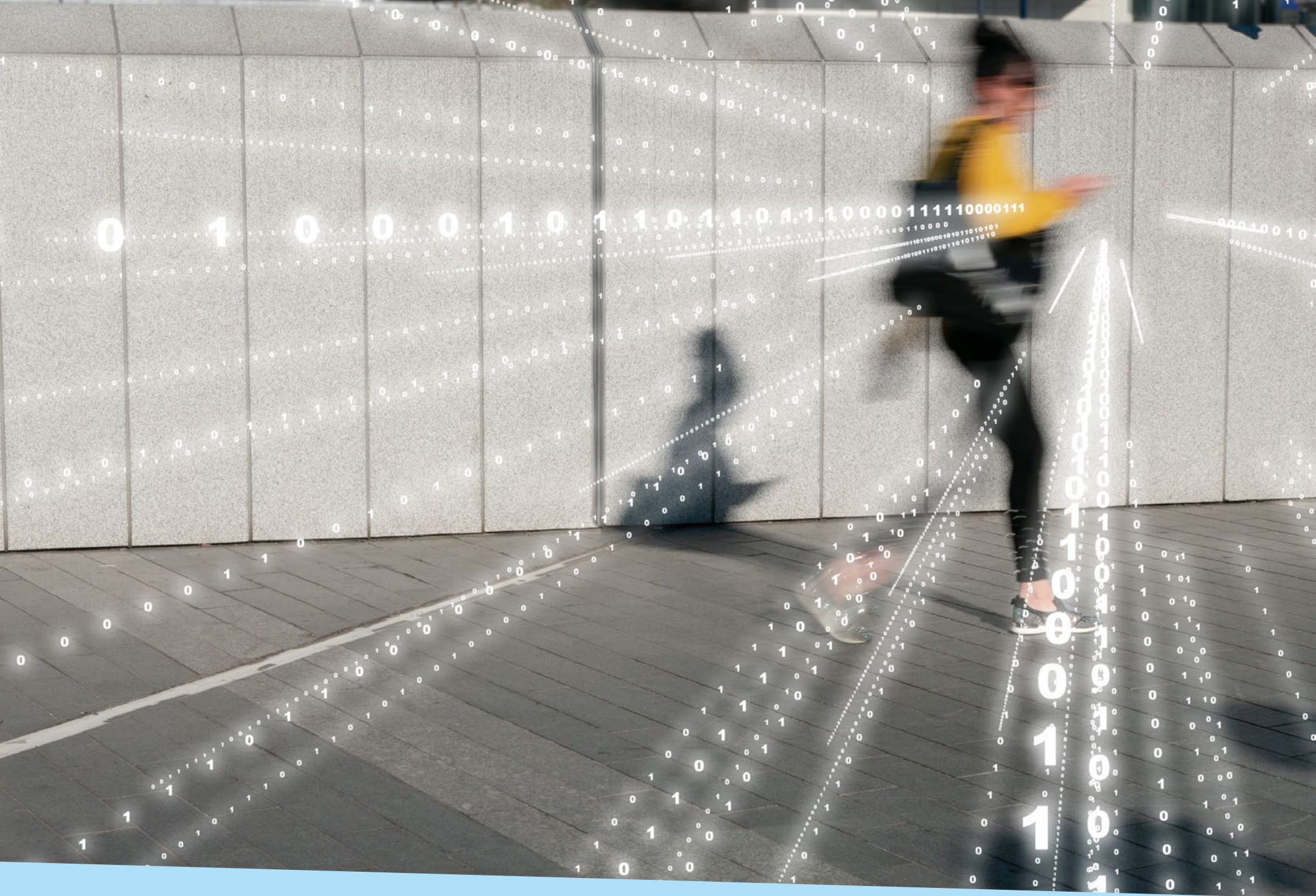
volksvertegenwoordigers (Brauneis & Goodman, 2017) Daarnaast gaat het gebruik van algoritmen mogelijk in tegen de noodzaak om beleid flexibel toe te passen en het recht van burgers om bezwaar in te dienen tegen beleidsbeslissingen of deze aan te vechten (Saward, 2010). Algoritmen kunnen dus schade toebrengen die moeilijk te herstellen is onder de bestaande juridische beschermingsinstrumenten. Brauneis en Goodman formuleerden deze situatie als volgt: 'In de publieke sector is de ondoorzichtigheid van algoritmische besluitvorming bijzonder problematisch, zowel omdat overheidsbeslissingen een enorme impact kunnen hebben als omdat democratisch verkozen overheden een diepgewortelde plicht hebben om verantwoording af te leggen' (Brauneis & Goodman, 2017) Kortom, alhoewel de kracht van data en algoritmen zeker ten goede kan worden ingezet, zijn er evenveel argumenten tegen het gebruik van algoritmen in beleids- en besluitvorming binnen overheden.

Er is een dringende noodzaak om het gebruik van data en algoritmen in de samenleving te democratiseren.

3. Algoritmen als experimentele technologie

Data en algoritmen zijn **zich ontwikkelende technologieën**; ze zijn nog niet uitontwikkeld en volledig ingebed in de samenleving (Brey, 2017). Nieuwe technologieën raken in een geleidelijk proces van co-evolutie ingebed in de samenleving. In de loop der tijd, naarmate een technologie zich ontwikkelt, wordt geïmplementeerd een aangepast, evolueert ook de samenleving. Op een bepaald moment vertraagt dit proces en worden de volledige impact en gevolgen van een technologie zichtbaar.

Dit betekent dat beleidsmakers, ontwikkelaars en de samenleving voor een fundamentele uitdaging staan: hoe zorgen we ervoor dat het evenwicht tussen gewenste en ongewenste uitkomsten naar de goede kant doorslaat voordat zaken onherroepelijk worden (Rip & Schot, 2002; van de Poel, 2016) In de vroege stadia van een opkomende technologie, wanneer de technologie en de maatschappelijke inbedding nog 'kneedbaar' zijn, is er nog maar weinig kennis over de eventuele effecten en hoe deze zich zullen manifesteren.

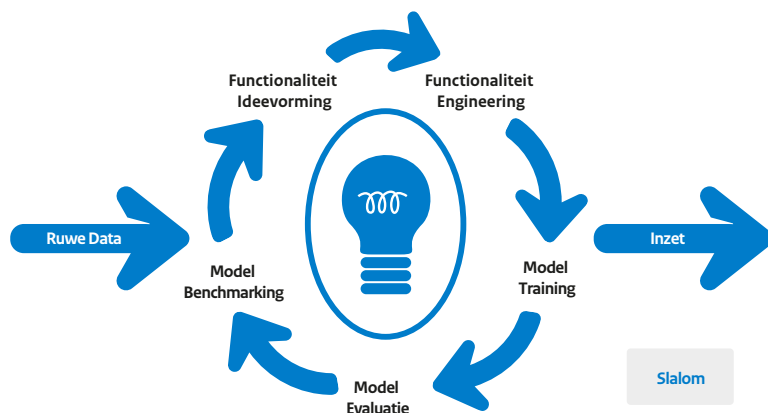


In latere stadia worden de maatschappelijke effecten misschien wel duidelijk maar is de technologie vaak zo verankerd in de samenleving, dat negatieve effecten niet meer kunnen worden verholpen. Het fundamentele onvermogen van de samenleving om technologie effectief te beheersen en te sturen werd door David Collingridge (Collingridge, 1982) het ‘beheersbaarheidsdilemma’ (control dilemma) genoemd.

3.1 Uitdagingen bij het beheersen van algoritmische technologieën in de samenleving

Data en algoritmetechnologieën zijn om vier redenen bijzonder gevoelig voor het beheersbaarheidsdilemma:

1. De technologische ontwikkeling van data en algoritmen verloopt bijzonder snel.
2. De ontwikkelings- en toepassingscyclus van data- en algoritmesystemen wordt steeds korter. Teams kunnen binnen enkele dagen of uren nieuwe functionaliteit leveren in plaats binnen van weken of maanden.
3. Zelfs kleine veranderingen in data en algoritmen (bijvoorbeeld keuze van de machinelearning-benadering) kunnen grote gevolgen hebben voor de prestaties en functionaliteit en het gedrag van zulke systemen.
4. De impact en effecten van data en algoritmen zijn ingrijpend en moeilijk te voorspellen.



Figuur 4 Agile machinelearning-cyclus (Bron: Follow (2019))

Data en algoritmen zijn kneedbare technologieën die gemakkelijk kunnen worden veranderd of aangepast. De toenemende beschikbaarheid van krachtige cloud-based data en machinelearningomgevingen in combinatie met continue software engineering (Fitzgerald & Stol, 2014, 2017). (EV hebben de ontwikkelings- en toepassingscyclus voor data- en algoritmesystemen aanzienlijk verkort (Figuur 4). Waar er een paar jaar geleden een groter team en meerdere maanden voor nodig zouden zijn geweest om een krachtig datagedreven systeem te ontwikkelen en implementeren, kunnen ontwikkelaars nu dankzij agile machine learning en machinelearning-pijplijnen elke paar uur een nieuwe versie ontwikkelen en uitbrengen (Dougherty, 2019; Kobielus, 2014; Lwakatare, Raj, Bosch, Olsson, & Crnkovic, 2019).

Technologische voortgang in de vorm van nieuwe data-analytische methoden en nieuwe machinelearning-benaderingen kan daardoor gemakkelijk in bestaande systemen worden geïntegreerd. Als een dataset eenmaal is gecompileerd, is het niet moeilijk om er nieuwe machinelearning-algoritmen op toe te passen die tot aanzienlijke prestatie- en functionaliteitsverbeteringen kunnen leiden. De snel stijgende beschikbaarheid van nieuwe en verbeterde datasets betekent omgekeerd ook dat bestaande systemen eenvoudig kunnen worden bijgewerkt door algoritmen simpelweg opnieuw te trainen met nieuwe data. Algemeen gezien betekent dit dat technologische ontwikkeling in data- en algoritmesystemen extreem snel verloopt vergeleken met de tijd die nodig is om mensen te trainen, de maatschappelijke implicaties te doorgronden en beleid en regelgeving te formuleren. Daarnaast zal de ontwikkelings- en toepassingsnelheid van data- en algoritmesystemen naar verwachting aanzienlijk toenemen wanneer organisaties machinelearning-processen gaan automatiseren om honderden machinelearning-modellen tegelijk te kunnen testen. Hierdoor zullen organisaties zonder noemenswaardig menselijk ingrijpen data en algoritmen kunnen gaan ontwikkelen, optimaliseren en toepassen.

Dit alles maakt duidelijk dat data en algoritmen passen binnen de door Van de Poel geformuleerde definitie van een **experimentele technologie**, d.w.z. een technologie waarmee slechts beperkte operationele ervaring is opgedaan en waarvan de maatschappelijke voordelen en risico's dus niet (zonder meer) kunnen worden geëvalueerd op basis van ervaring (van de Poel, 2016)

Data en algoritmen zijn bijzonder kneedbare, continu evoluerende technologieën. De efficiëntie van engineeringprocessen voor data en algoritmen neemt snel toe, wat leidt tot een steeds grotere innovatiesnelheid.

Data en algoritmen zijn experimentele technologieën waarvan de maatschappelijke voordelen en risico's pas goed kunnen worden beoordeeld en doorgrond nadat ze ook werkelijk in de samenleving zijn geïntroduceerd.

3.2 Experimenteren met algoritmen

De erkenning van het beheersbaarheidsdilemma van Collingridge heeft geleid tot een toenemend gebruik van experimentele methodes voor het ontwerpen, ontwikkelen en evalueren van technologieën. Dit heeft de vorm aangenomen van living labs en, in de stedelijke context, *urban living labs* (Baccarne, Mechant, Schuurman, De Marez, & Colpaert, 2014; Cosgrave, Arbuthnot, & Tryfonas, 2013; Gascó, 2017; Kronsell & Mukhtar-Landgren, 2018; Steen & Bueren, 2017; Steen & van Bueren, 2018). Volgens het European Network of Living Labs (Schuurman, 2015) zijn urban living labs 'op de gebruiker gerichte, open-innovation-ecosystemen gebaseerd op een systematische benadering van co-creatie door gebruikers in publiek-private-menselijke samenwerkingsverbanden, waarbij onderzoeks- en innovatieprocessen worden geïntegreerd in echte gemeenschappen en contexten'. Schaer stelt dat living labs zich verplaatsen van *in-vitro* naar *in-vivo* onderzoeksscenario's waarbij data wordt verzameld in een levende omgeving, bijvoorbeeld een gebouw of openbare ruimte (Schaer, 2017).

Living labs spelen een sleutelrol in de ontwikkeling van openbare algoritmische systemen en de co-creatie van openbare dienstverlening op het gebied van mobiliteit, energie en openbare veiligheid door publieke (gemeenten, overheden) en particuliere (technologiebedrijven) partners en het maatschappelijk middenveld (gemeenschappen). Bij veel living labs wordt data gebruikt, die actief of passief wordt verstrekt door de burgers en gebruikers. Dit ligt dus heel gevoelig voor wat betreft de rechten van burgers en in relatie tot het openbaar bestuur.

Urban Living Labs

Inmiddels worden living labs beschouwd als een van de belangrijkste instrumenten voor innovatie in steden. Een van de meest besproken voorbeelden van een living lab is de Quayside-ontwikkeling door Sidewalk Labs van Google in Toronto (Canada), waar een grootschalig aan het water gelegen gebied wordt ontwikkeld tot een hightech digitale wijk. Het onlangs uitgebrachte ontwikkelingsplan (*Sidewalk Labs Master Innovation and Development Plan*, 2019) omvat grootschalige experimenten met innovaties, waaronder zelfrijdende auto's, openbare wifi en nieuwe oplossingen voor de gezondheidszorg. Alhoewel de plannen voor Sidewalk ongebruikelijk zijn qua schaal en ambitie, bestaan er gelijksoortige experimentele projecten in de meeste grote steden over de hele wereld.

Living Labs in Nederland

In 2017 stelden de Gemeente Den Haag en de Hague Security Delta (HSD), een cluster van meer dan 200 nationale cyber- en stedelijke beveiligingsorganisaties, een programma op voor beveiligingsinnovatie in de Internationale Zone in Scheveningen. Het doel is het gebied steeds aantrekkelijker en veiliger te maken voor zowel internationale organisaties als inwoners ('International Zone,' g.d.). In de Internationale Zone zijn internationale organisaties als Europol, de OPCW en Eurojust gevestigd.



Deze wordt ook gebruikt als conferentielocatie voor regeringsleiders van over de hele wereld. Het veiligheidsinitiatief heeft ten doel een geïntegreerde benadering te ontwikkelen om de veiligheid in de Internationale Zone te vergroten en de overlast voor werknemers, inwoners en bezoekers te verlagen. De partners in dit initiatief zijn de Gemeente Den Haag, Politie Den Haag, het Ministerie van Buitenlandse Zaken, het Rijksvastgoedbedrijf, Europol, Eurojust, de OPCW, het Vredespaleis, het Catshuis, TNO, Thales, Siemens en Crowd Sense. Begin 2019 werd gestart met de transformatie van het veiligheidsinitiatief tot een Security Living Lab, waar bedrijven, kennisinstellingen en beveiligingsdiensten testen hoe data van sensoren (zoals camera's en vrij toegankelijke wifipunten) samen met AI-algoritmen kunnen worden ingezet ten behoeve van beveiliging en andere doeleinden in een openbare ruimte. Particuliere bedrijven die zijn verbonden aan het initiatief hebben toegang tot de data uit het project voor productontwikkeling en in het kader van een doorlopend kennisdelingsproject met de gemeente.

In Stratumseind, een uitgaanswijk in Eindhoven, is ook een living lab gestart als testomgeving voor nudging-strategieën om criminaliteit en verstoringen van de openbare orde terug te dringen.



Het Stratumseind living lab is een multidisciplinair samenwerkingsverband tussen de gemeente, de politie en partners uit het bedrijfsleven. Met behulp van camera's en geluidssensoren worden bezoekersstromen en gedrag op straat bijgehouden. Straatverlichting wordt op basis van live data en algoritmen dynamisch bijgesteld om het gedrag van bezoekers te beïnvloeden. Net als bij veel andere living labs wordt data gedeeld door publieke en private partners met als doel samen systemen te ontwikkelen die maatschappelijke en commerciële waarde hebben.

Straatverlichting wordt op basis van live data en algoritmen dynamisch bijgesteld om het gedrag van bezoekers te beïnvloeden

Het living lab in Helmond richt zich op innovatie op het gebied van duurzame energie, hergebruik van water, autodelen en geautomatiseerd parkeren, slimme straatverlichting en meer.

Hier is een kleine wijk met woningen en openbare infrastructuur voorzien van sensoren om energie- en waterverbruik, alsmede het gebruik van gedeelde auto's, te meten. Daarnaast wordt het gedrag van huishoudens en individuele personen gemeten, bijvoorbeeld de slaaptijd, het gebruik van huishoudelijke apparaten en de tijd die wordt doorgebracht op sociale media. In ruil voor hun deelname ontvangen inwoners een financiële beloning en betalen zij minder huur (Kuyper, 2019).

Andere living labs zijn ingericht om toekomstige intelligente energiesystemen te ontwerpen en te testen, waarbij vooral veel aandacht wordt besteed aan elektrische mobiliteit zoals in het Vehicle 2 Grid-project ('Vehicle2Grid,' 2016). Centraal in deze living labs staat het idee dat er in de toekomst AI-algoritmen nodig zullen zijn om te kunnen voldoen aan de elektriciteitsbehoefte van een snelgroeiend aantal elektrische voertuigen op de weg. Wanneer de vraag groter is dan het aanbod, moeten datagedreven algoritmen gaan beslissen wie er hoeveel stroom krijgt. Daarbij zou gebruik kunnen worden gemaakt van de status van accu's en andere informatie over de bestuurder.

The Hague Security Living Lab in de Internationale Zone in Scheveningen	Een geïntegreerde benadering ontwikkelen om de veiligheid in de Internationale Zone te vergroten en de overlast voor werknemers, inwoners en bezoekers te beperken	Data over burgers afkomstig van openbare camera's en wifi-netwerken
Stratumseind, uitgaanswijk in Eindhoven	Toezicht houden op en beïnvloeden van gedrag van het publiek om criminaliteit en verstoringen van de openbare orde tegen te gaan	Data over burgers afkomstig van openbare camera's
Helmond Living Lab	Ontwikkelen van innovatieve wijkdienstverlening voor duurzame energie, hergebruik van water, autodelen en geautomatiseerd parkeren	Data over het huishoudelijk gedrag van inwoners en hun gebruik van water, energie en mobiliteitsdiensten
Vehicle 2 Grid	Ontwerpen en testen van toekomstige intelligente energiesystemen om te kunnen beantwoorden aan de toekomstige vraag vanuit elektrische mobiliteit	Gegevens over mobiliteit en laadgedrag van elektrische voertuigen

Tabel 1 Urban living labs in Nederland

Living Labs spelen een sleutelrol in de experimentele ontwikkeling van data- en algoritmesystemen.

3.3 Ethische aspecten van algoritmische experimenten

Living labs worden beschouwd als alternatief voor een institutionele benadering van smart cities, waar burgers slechts worden gezien als passieve dataleveranciers, wier data is gemaakt om de doelen van stedelijke systemen en organisaties te dienen (Ranchordás, 2019). Men wil dat burgers een actieve rol spelen als

co-innovatoren om 'producten, services, systemen en technologieën in een levensechte situatie te creëren, er prototypes van te maken, te valideren en te testen' (Westerlund & Leminen, 2018) (Steen & van Bueren, 2018). Het toenemende gebruik van experimentele technologieën in de openbare ruimte en van data en machinelearning-algoritmen leidt echter tot fundamentele vragen over de bestuurlijke en ethische kant van deze initiatieven.

Schaer (Schaer, 2017) noemt vier belangrijke vraagstukken: 1) Keuzevrijheid: kunnen mensen ervoor kiezen om wel of niet mee te doen aan het experiment? 2) Geïnformeerde deelname: begrijpen mensen de aard van het experiment? 3) Bestuurlijke regie en beheersbaarheid: wie bepaalt hoe data en algoritmen worden gebruikt en wie is verantwoordelijk voor de resultaten en uitkomsten? 4) Geen negatieve gevolgen: zijn de data en algoritmen vrij van vooringenomenheid en discriminatie?

Keuzevrijheid en geïnformeerde deelname: door de snelle instroom van sensoren, data en algoritmen in de stedelijke ruimte beschikken we binnenkort over hetzelfde vermogen om mensen te beïnvloeden (manipuleren) en experimenten in de openbare ruimte uit te voeren dat bedrijven als Facebook al online hebben. In veel gevallen is het voor deze experimenten nodig dat mensen onbewust deelnemen, omdat hun gedrag kan worden beïnvloed door te veel informatie over het experiment. Dat zou de experimentele opzet het experiment tenietdoen. Aan de andere kant zou te weinig informatie verschaffen betekenen dat mensen mogelijk meedoen aan een experiment waaraan ze, als ze volledig geïnformeerd waren, nooit zouden hebben meegedaan. Dit dilemma werd benadrukt in een rapport uit 2007 van de European Expert Group on Science and Governance: 'Als burgers routinematig en zonder overleg zouden worden onderworpen aan experimenten die niet bij naam worden genoemd, dan zou dat ernstige morele en maatschappelijke vraagstukken opleveren' (Felt et al., 2007).

Bestuurlijke regie en beheersbaarheid: een tweede vraagstuk heeft te maken met de overmatige invloed van particuliere bedrijven op het ontwerp en gebruik van datagedreven technologieën. Publieke instellingen gebruiken steeds vaker data en AI-technologieën die worden gemaakt, verkocht en beheerd door particuliere bedrijven, wat mogelijk ernstige gevolgen kan hebben voor de burgerlijke vrijheden. Deze overmatige invloed kan vele vormen aannemen (Joh, 2017).

Zo is het mogelijk dat publieke instellingen onder een contract geen informatie aan het publiek mogen geven, zelfs al zijn ze dit wettelijk of beleidsmatig verplicht. Publieke organisaties kunnen worden gedwongen om een contract te tekenen met bedrijven die min of meer een monopolie hebben op bepaalde technologische gebieden, wat kan leiden tot stijgende kosten of ontwerpkeuzes die niet in het algemeen belang zijn. Dit wordt door (Joh, 2017) omschreven als 'productontwerp als beleid'. Ten slotte kan een gebrek aan expertise binnen publieke organisaties betekenen dat ze weinig invloed hebben op ontwerpkeuzes of de implicaties van technologische keuzes niet kunnen overzien. Uiteindelijk betekenen deze factoren dat het publiek beperkte informatie krijgt over de technologie die ten grondslag ligt aan belangrijke beslissingen over publieke kwesties en dat de logica en ratio van geautomatiseerde beslissingen zelfs voor de organisatie die van deze technologieën gebruikmaakt ondoorzichtig blijft. Hierdoor wordt de bestaande 'informatie-asymmetrie' tussen technologiegebruikers en de mensen die de gevolgen van deze geautomatiseerde beslissingen ondervinden, d.w.z. de burgers, nog verder versterkt. (Joh, 2017).

Vooringenomenheid en eerlijkheid: alhoewel machinelearning-technieken zo kunnen worden geoptimaliseerd dat er in een laboratorium fantastische resultaten mee worden behaald, is de juistheid en betrouwbaarheid van machinelearning-toepassingen in de echte wereld vaak ronduit slecht. Recent onderzoek liet de gevaren zien van verborgen vooringenomenheid in AI-systemen. (Campolo, Sanfilippo, Whittaker, & Crawford, 2017). Ook van commerciële systemen is herhaaldelijk aangetoond dat ze onjuist en vooringenomen zijn (Angwin et al., 2016; Lebovits, 2019; Misra, 2018; Richer, 2018). Zo werd ontdekt dat een tool om de neiging tot recidive van veroordeelde criminelen te voorspellen niet beter presteerde dan een willekeurige online poll onder mensen die geen enkele strafrechtelijke achtergrond hadden (Dressel & Farid, 2018). Dit ondanks het feit dat de London Metropolitan Police een aantal interne tests had uitgevoerd om dergelijke problemen te voorkomen. Ondanks dat het bewijs dit tegenspreekt, blijft de London Metropolitan Police volhouden dat deze technologie een foutpercentage van minder dan 1 op de 1000 heeft. De organisatie heeft de methodologie tot dusver nog niet openbaar willen maken ('London police's face recognition system gets it wrong 81% of the time,' 2019).



Privacy: data die is verzameld in de openbare ruimte bevat vaak privé-informatie (bijv. seksuele oriëntatie, lidmaatschap van een vakbond of politieke partij, economische en uiterlijke kenmerken) die per ongeluk of met opzet kan worden geëxtraheerd door externe dataconsumenten (bijv. data-annotatoren, partnerorganisaties, burgers). Zo kan meer informatie worden geopenbaard dan oorspronkelijk de bedoeling was en kan deze voor een ander doel worden ingezet. De huidige oplossingen voor automatische privacybescherming, zoals bewerking van de inhoud of het gebruik van edge computing, kunnen de privacy vergroten maar uiteindelijk niet garanderen dat alle privé- en gevoelige informatie is verwijderd. Dit heeft belangrijke gevolgen in het kader van de Algemene Verordening Gegevensbescherming (AVG) (Bourgeois, Kortuem, & Kawsar, 2018). Daarom zouden living labs vooral baat hebben bij 'openbaarmaking op basis van begrensde toegang' (Fan & Jackson, 2015), een model waarin alleen gekwalificeerde professionals toegang hebben tot informatie. Deze professionals zijn in staat om data effectief te gebruiken en moeten zich houden aan de regelgeving op het gebied van gegevensbescherming.

Zo kan meer informatie worden geopenbaard dan oorspronkelijk de bedoeling was en kan deze voor een ander doel worden ingezet

3.4 De behoefte aan bestuurlijke regie over experimenten met algoritmen

Living labs in Nederland en elders maken pas nu een begin met deze ethische vraagstukken. Bij Security Living Lab in Den Haag wordt iedereen die de Internationale Zone bezoekt potentieel opgenomen door een van de vele camera's in het gebied. Toch is er nergens een publieke kennisgeving dat er videobeelden worden opgenomen en mogelijk gedeeld met partnerorganisaties binnen het living lab. Bezoekers aan de Internationale Zone die het gratis openbare wifi-netwerk in de zone willen gebruiken, moeten de algemene gebruiksvoorwaarden ondertekenen voordat ze toegang krijgen.

Het is echter onduidelijk of bezoekers zich bewust zijn van hun deelname aan het project en de implicaties hiervan voor hun privacy op de korte en lange termijn. Onder de huidige regelgeving is het volgen van individuele personen niet toegestaan maar de techniek om dit te kunnen doen bestaat al. Projectpartners hebben ook sterke aanleiding om te experimenteren met deze mogelijkheden. In een poging om smart energie- en mobiliteitssystemen te ontwerpen in samenwerking met burgers, organiseerde het Vehicle-to-grid-project workshops. Hier werd potentiële gebruikers gevraagd om aan te geven hoe de algoritmen volgens hen zouden moeten worden geprogrammeerd. Bedrijven geven echter niet graag geheime prijs en willen het ontwerp van algoritmen onder eigen beheer houden om de commerciële waarde te maximaliseren. Als het aankomt op het technische ontwerp, nemen commerciële partners maar zelden maatschappelijke zaken als eerlijkheid en non-discriminatie in overweging.

Wat de voordelen van data en algoritmische systemen ook zijn, er zijn potentiële kosten mee gemoeid voor wat betreft het vertrouwen van het publiek en de legitimiteit van organisaties. Dat roept de vraag op of er toepassingen zijn van data- en algoritmesystemen die helemaal niet zouden moeten worden gebruikt (Calo, 2017). Alhoewel de open en participatieve benadering van living labs potentieel meer rekening houdt met maatschappelijke waarden, heeft ook experimenteren in de echte wereld zijn nadelen. Living lab-zones in steden zullen no go areas worden voor mensen die niet willen deelnemen aan het onderzoek, tenzij er nieuwe ethische en ontwerp oplossingen komen.

Er is al een hoop werk verricht om goede praktijken aan te moedigen of te verplichten bij de ontwikkeling en het gebruik van data- en algoritmesystemen. Zo worden participatieve ontwerpbenaderingen en value-sensitive design (VSD)-methoden regelmatig gebruikt om burgers die direct of indirect worden geraakt door zulke systemen, bij het ontwerp te betrekken. Een andere benadering is de transparantie van algoritmische systemen te verbeteren. De gedachte is dat, door systemen transparant te maken, mensen begrijpen hoe ze werken. Dat moet weer leiden tot een betere verantwoording. Wetgeving ten aanzien van gegevensbescherming, zoals de AVG/GDPR, schrijft voor hoe persoonlijke gegevens gebruikt mogen worden. Dit leidt effectief tot een 'recht op verklaring', waarbij gebruikers kunnen vragen waarom een bepaalde algoritmische beslissing over ze is genomen (Malgieri & Comandé, 2017).



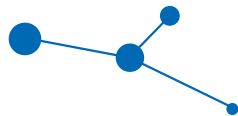
Deze visie is echter betwist door Wachter et al., die zich afvragen 'Waarom er in de GDPR geen recht op een verklaring van geautomatiseerde besluitvorming is opgenomen' (Wachter et al., 2017)

Bestaande benaderingen hebben echter ernstige tekortkomingen: participatieve benaderingen vereisen een welwillende organisatie die wil luisteren naar de zorgen van de burger en lenen zich niet goed voor het oplossen van problemen die optreden nadat het systeem al is ingezet. Uit inspanningen om algoritmische transparantie te vergroten, zijn nog geen manieren voortgekomen die burgers en niet-specialisten inzicht geven in de complex en onvoorspelbare aard van algoritmische systemen. Tot nu toe is het maar zeer beperkt mogelijk om de maatschappelijke impact van algoritmische systemen op het gebied van problemen als discriminatie, vooroordelen, inclusiviteit enz. te voorspellen.

Kortom, de huidige situatie wordt gekenmerkt door drie fundamentele problemen:

- Ontwerpers en technische specialisten kunnen maar zeer beperkt de maatschappelijke impact van data en algoritmische systemen voorspellen
- Alle systemen die tegen de maatschappelijke waarden ingaan, ondermijnen het vertrouwen in en de legitimiteit van organisaties die van deze technologie gebruikmaken
- Gemeenten zijn maar zeer beperkt in staat data en algoritmische systemen aan te sturen
- Burgers zijn niet in staat betekenisvolle kritiek te leveren en zich te verzetten tegen algoritmische systemen en machinevoorspellingen.

Er is behoefte aan richtlijnen voor het openbaar bestuur ten aanzien van openbare experimenten met data en algoritmen.



4. Bestuurlijke modellen voor experimenten met algoritmen

Over het algemeen is men het erover eens dat er voor het gebruik van data en algoritmen een nieuw controlesysteem moet komen om het publiek te beschermen (Kitchin, 2016, 2019). In de vorige paragrafen gaf ik aan dat we algoritmen als experimentele technologie moeten beschouwen en ons ervan bewust moeten zijn dat veel innovatie-initiatieven nu een experimentele benadering volgen. De fundamentele vraag is dan ook: *Hoe kunnen we op ethisch verantwoorde wijze experimenteren met data en algoritmen in de samenleving als data en algoritmen an sich al fundamentele ethische problemen met zich meebrengen?*

In deze paragraaf zet ik eerst uiteen waarom huidige pogingen om ethische richtlijnen voor AI minder zinnig zijn dan ze lijken. Daarna geef ik een beknopte beschrijving van twee benaderingen die als basis kunnen worden gebruikt voor het aansturen van experimenten met algoritmen in de samenleving.

4.1 De beperkingen van ethische richtlijnen

Er is een groot aantal initiatieven, die als doel hebben om richtlijnen, normen en praktijken te formuleren voor verantwoord en ethisch gebruik van digitale technologieën in de samenleving. Sommige initiatieven richten zich op steden, zoals Barcelona Digital City, Amsterdam Tada!, Detroit Digital Justice Coalition, Seattle Community Technology Advisory Board, Smart London Board, Smart Dubai AI Ethics Advisory Board en de Cities Coalition for Digital Rights.

Andere concentreren zich op *Artificial Intelligence (AI)*, bijvoorbeeld initiatieven van de overheid (bijv. de EU High-Level Expert Group on AI), universiteiten (Stanford Institute for Human-Centred Artificial Intelligence), bedrijven (zoals Google en Microsoft) en maatschappelijke organisaties (bijv. Data for Black Lives en Partnership on AI). De meeste van deze initiatieven hebben ethische richtlijnen voor AI opgesteld.

Volgens het 'Principled Artificial Intelligence Project' van Harvard, dat 32 ethische richtlijnen voor AI analyseerde (Figuur 5), richten de meeste initiatieven zich op acht hoofdthema's ('The Principled Artificial Intelligence Project,' 2019): verantwoording, eerlijkheid en non-discriminatie, menselijke controle over technologie, privacy, professionele verantwoordelijkheid, bevordering van menselijke waarden, veiligheid en beveiliging en transparantie en verklaarbaarheid. AI-richtlijnen omvatten doorgaans aanbevelingen met betrekking tot **data** (bijv. publiceer en minimaliseer vooringenomenheid in de trainingsdata), **algoritmen** (bijv. zorg dat de code van een AI-systeem transparant is en toegankelijk voor inspectie, zorg dat de functie van het AI-systeem kan worden toegelicht), **uitkomsten** (bijv. zorg ervoor dat de bedoelde en werkelijke uitkomsten van een AI-systeem eerlijk, transparant, wettig en overeenkomstig menselijke waarden zijn) en **bestuur** (zorg dat er een proces van toezicht en evaluatie is en stel iemand verantwoordelijk voor de genomen beslissingen).

Ethische richtlijnen voor AI richten zich ook op rechten, principes, waarden en eisen. Volgens ('The Principled Artificial Intelligence Project,' 2019) worden de volgende **rechten** vaak genoemd: respect voor menselijke waardigheid, individuele vrijheid, respect voor democratie, rechtvaardigheid en de rechtsstaat, gelijkheid, non-discriminatie en solidariteit, waaronder ook de rechten van minderheden en burgerrechten (gebaseerd op: EU Handvest van de grondrechten).

Veelgenoemde **principes** zijn: weldoen (doe wat goed is), geen kwaadwilligheid (doe niet wat slecht is), autonomie (bescherm het beslissingsrecht van de mens, rechtvaardigheid (wees eerlijk) en verklaarbaarheid (operationele transparantie) (uit: AI4People—An Ethical Framework for a Good AI Society); de laatste, verklaarbaarheid, is relatief nieuw en specifiek voor AI.

Veelgenoemde **waarden**: de richtlijnen 'willen niet nog een lijst van kernwaarden verschaffen' omdat er al veel bruikbare lijsten beschikbaar zijn, zoals die van Asilomar, Montreal, IEEE en EGE (deze worden besproken in: AI4People—An Ethical Framework for a Good AI Society).

Veelgenoemde **eisen** zijn: verantwoording, bestuurlijke regie over data, design for all, aansturing van AI-autonomie (menselijk toezicht), non-discriminatie, respect voor (en bevordering van) menselijke autonomie, respect voor privacy, robuustheid, veiligheid en transparantie.

Hoewel ze grotendeels welkom zijn, hebben deze ethische initiatieven vier grote beperkingen: Ten eerste zijn ethische initiatieven voor AI bekritiseerd omdat ze te breed en te simplistisch zijn en weinig praktische richtlijnen bieden om met de complexiteit van AI in de echte wereld om te gaan (Copeland, 2019). Copeland noemt drie niveaus van complexiteit van AI, namelijk 1) human-experts hand-craft machinelearning-modellen 2) AI-algoritmen bepalen zelf welke factoren relevant zijn en 3) een doorlopend geautomatiseerd proces waarin machinelearning-modellen zich ontwikkelen in snelle iteraties op basis van inkomende datastromen. Copeland merkt op dat ethische richtlijnen voor AI feitelijk alleen bruikbaar zijn voor AI-systemen met de laagste complexiteit.

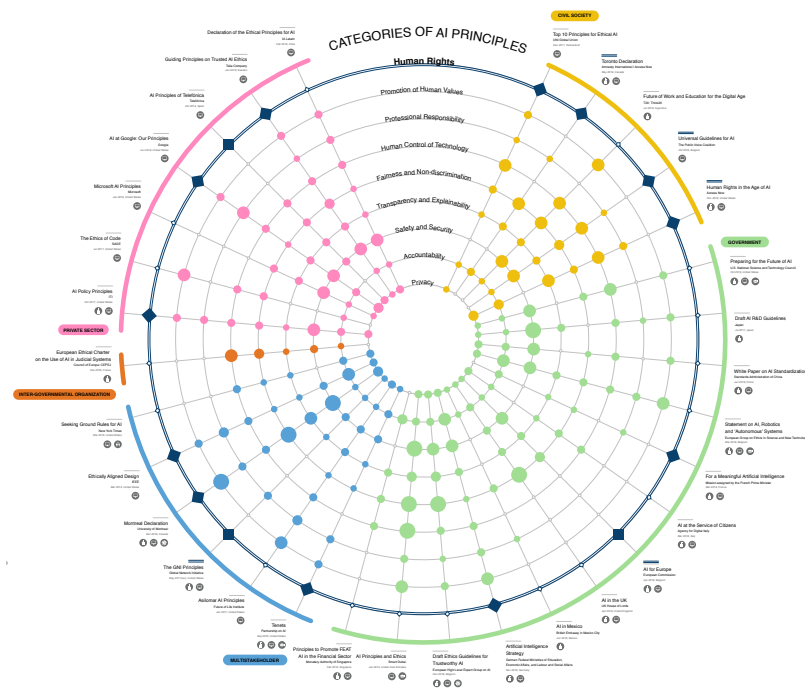
Een tweede, meer fundamentele vorm van kritiek is geuit door (D'Ignazio & Klein, 2019), die opmerken dat huidige discussies over AI vaak 'fundamentele normatieve vragen negeren over wat voor samenleving we willen en niet gaan over bredere morele filosofische concepten als rechtvaardigheid, onderdrukking, gerechtigheid, burgerschap en het algemeen belang'. In plaats daarvan, zo zegt Kitchin, richten ethische richtlijnen voor AI de aandacht in de eerste plaats op 'meer procedurele technische en juridische factoren, zoals vooringenomenheid, eerlijkheid, transparantie, consumentenrechten, verantwoording, naleving en herstel' (Kitchin, 2019). Hij stelt dat deze richtlijnen de bron van ethische problemen vooral zoeken in individuele personen en technische systemen in plaats van in structurele macht, en dus slechts symptomen bestrijden en niet de onderliggende oorzaken.

Huidige discussies over AI negeren vaak fundamentele normatieve vragen over wat voor samenleving we willen

Mittelstadt (Mittelstadt, 2019) uitte een derde vorm van kritiek. Volgens hem steekt AI scherp af tegen de geneeskunde, een 'morele gemeenschap' (Frankel, 1989) met gezamenlijke doelen, waarden en opleiding, waar ethische principes, richtlijnen en codes goed verankerd zijn. Hij stelt dat in tegenstelling tot de geneeskunde,



'In AI-ontwikkeling (1) gezamenlijke doelen en fiduciaire plichten ontbreken, evenals (2) professionele geschiedenis en normen, (3) beproefde methoden om principes te vertalen naar de praktijk en (4) robuuste wettelijke en professionele verantwoordingsmechanismen. ... Daarom moeten we niet te enthousiast zijn over de consensus rond hoogwaardige principes die een diepliggende politieke en normatieve onenigheid verhullen. Gedeelde principes zijn niet genoeg om 'betrouwbare AI' of 'ethisch verantwoorde AI' in de toekomst te waarborgen. Zonder een fundamentele verschuiving in de regelgeving zal het vertalen van principes naar de praktijk een competitief proces blijven in plaats van een coöperatief proces.' (Mittelstadt, 2019)



Figuur 5 Kaart met 32 ethische richtlijnen voor AI. Bron: ("The Principled Artificial Intelligence Project", 2019)

Hij geeft vervolgens vier aanbevelingen voor de volgende stappen in AI-gerelateerde ethiek (Mittelstadt, 2019):

1. 'De langetermijn-doelen en wegen-naar-invloed van principiële initiatieven moeten duidelijk gedefinieerd worden.' Hiermee bedoelt hij dat er bindende verantwoordingsstructuren moeten worden ingericht, evenals duidelijke implementatie- en evaluatieprocessen op sector- en organisatieniveau, en documentatie van modellen en datasets (in de laatste paragraaf over aanbevelingen kom ik hierop terug).
2. 'Steun 'bottom-up' AI-gerelateerde ethiek.' Mittelstadt pleit ervoor ethiek per geval te benaderen, waarbij lokale praktijken en lessen van onderaf ontstaan en zich door het veld en de branche verspreiden. Hiervan kunnen dan principes en precedents worden afgeleid (in de laatste paragraaf met aanbevelingen kom ik hierop terug).
3. 'Licentieer ontwikkelaars van hoogrisico-AI'. Mittelstadt suggereert dat het wellicht nodig is om AI-ontwikkeling formeel als een beroep te definiëren met eenzelfde status als andere hoogrisico-beroepen, zoals artsen en advocaten (ik zie beperkte kansen voor deze aanbeveling gezien het beperkte succes van de poging om software engineering als beroep in te stellen).
4. 'Verplaats de focus van professionele ethiek naar zakelijke ethiek': hier stelt hij voor de focus van ethiek te verschuiven van individuele personen naar de ethiek van zakelijke praktijken en businessmodellen.
5. 'Benader ethiek als een proces en niet als technologisch solutionisme': Mittelstadt stelt dat we ethiek niet slechts als een technisch ontwerpaspect moeten zien dat kan worden 'opgelost'. In plaats daarvan moeten we AI-gerelateerde ethiek zien als een doorlopend, nooit eindigend proces dat veranderingen in de samenleving weerspiegelt.

Een vierde en laatste argument dat aangeeft waarom de huidige ethische richtlijnen voor AI beperkt bruikbaar zijn, heeft te maken met de eerdere discussie over het feit dat data en algoritmen (d.w.z. de bouwblokken van AI) experimentele en snel ontwikkelende technologieën zijn. De huidige ethische richtlijnen voor AI gaan ervan uit dat systemen op basis van AI van tevoren zo kunnen worden ontwikkeld dat negatieve gevolgen worden vermeden. Het beheersbaarheidsdilemma van Collingridge geeft echter al aan dat dit misschien niet mogelijk is voor snel ontwikkelende technologieën, zoals data en algoritmen, waarbij de maatschappelijke risico's en effecten pas kunnen worden beoordeeld nadat de technologieën al zijn geïmplementeerd.

4.2 Mogelijke richtingen

Als ethische richtlijnen onvoldoende zijn, wat kunnen we dan doen om experimentele algoritmen effectief te beheersen?

Experimental Ethics van Van de Poel

Living Labs willen het beheersbaarheidsdilemma van Collingridge omzeilen door vooruitzichten en voorkomen te vervangen door de geleidelijke en experimentele introductie van een technologie in de samenleving en die te controleren op de maatschappelijke effecten ervan, om die technologie en het gebruik ervan in de samenleving vervolgens te verbeteren. Met andere woorden, de experimentele introductie van een technologie in de samenleving wordt gezien als een soort **sociaal experiment**, dat de samenleving in staat stelt om te leren door ervaring en fouten (van de Poel, 2016).

Bestuur van een sociaal experiment is niet hetzelfde als bestuur van een technologie want het verschuift de blik van de technologie naar de context waarin die technologie wordt gebruikt. Van de Poel (van de Poel, 2016) introduceerde een kader om problemen met technologische experimenten in de echte wereld aan te pakken op basis van ethische principes die voortkomen uit de bio-ethiek, namelijk niet-kwaadwilligheid (bijv. geen schade berokkenen, risico's minimaliseren), weldoen (bijv. voordelen creëren of vergroten), respect voor autonomie (bijv. autonome keuzes van personen en groepen beschermen en waarborgen) en rechtvaardigheid (bijv. kwetsbare groepen beschermen, exploitatie voorkomen). Deze ethische principes grijpen terug op het Belmont-rapport, een sleutelwerk op het gebied van ethiek en onderzoek in de gezondheidszorg (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979).



Op basis van deze vier bio-ethische morele principes noemt Van de Poel zestien specifieke voorwaarden waaronder hij experimenten met nieuwe technologie in de samenleving als moreel aanvaardbaar beschouwt (van de Poel, 2016):

1. *Ontbreken van andere redelijke manieren om kennis te vergaren over risico's en voordelen*
2. *Toezicht houden op data en risico's, rekening houdend met privacykwesties*
3. *Mogelijkheid en bereidheid om het experiment aan te passen of stop te zetten*
4. *Indammen van risico's voor zover redelijkerwijze mogelijk*
5. *Weloverwogen opschalen om schade op grote schaal te voorkomen en het leerproces te verbeteren*
6. *Flexibele inrichting van het experiment en vermijden van vergrendelen (lock-in) van de technologie*
7. *Vermijden van experimenten die veerkracht ondermijnen*
8. *Redelijk om te verwachten dat het experiment maatschappelijke voordelen oplevert*
9. *Duidelijke verdeling van verantwoordelijkheden voor het inrichten, uitvoeren, toezicht houden op, evalueren, aanpassen en stopzetten van het experiment*
10. *Personen die deelnemen aan het experiment worden hierover geïnformeerd*
11. *Het experiment is goedgekeurd door democratisch goedgekeurde organen*
12. *Personen die deelnemen aan het experiment kunnen invloed uitoefenen op het inrichten, uitvoeren, toezicht houden op, evalueren, aanpassen en stopzetten van het experiment*
13. *Personen die deelnemen aan het experiment kunnen zich terugtrekken uit het experiment*
14. *Kwetsbare personen nemen niet deel aan het experiment, worden extra beschermd of profiteren in het bijzonder van de experimentele technologie (of een combinatie hiervan)*
15. *Een eerlijke distributie van mogelijke gevaren en voordelen*
16. *Omkeerbaarheid van schade, of indien dit onmogelijk is, compensatie voor schade*

Een aantal van deze voorwaarden weerspiegelen concepten die terug te vinden zijn in ethische richtlijnen voor AI of in bestaande wetgeving, zoals de AVG/GDPR. Zo komt de voorwaarde dat deelnemende personen moeten worden geïnformeerd nauw overeen met het concept van geïnformeerde toestemming in de AVG/GDPR. Het kader van Van de Poel gaat echter verder dan de meeste bestaande ethische richtlijnen voor AI omdat het zich richt op aansturing (*governance*) van het proces van experimentatie en niet zozeer op aansturing van de technologie zelf. Terwijl ethische richtlijnen voor AI een statische visie op technologie hebben, verwijst het kader van Van de Poel naar de continu evoluerende aard van data en AI-algoritmen.

Op deze wijze komt het kader van Van de Poel overeen met de experimentele benadering van living labs. Het is echter generiek en heeft niet specifiek betrekking op aspecten van data en algoritmen. Vanwege de aard van veel algoritmische systemen, die kan worden vergeleken met een zwarte doos, en het hoge kennisniveau dat vereist is om algoritmen te begrijpen, is het moeilijk om te voldoen aan een aantal van de voorwaarden van Van de Poel. Dit geldt vooral voor voorwaarden 10 tot en met 12. Het is bijvoorbeeld niet duidelijk hoe deelnemende personen invloed kunnen uitoefenen op het inrichten, uitvoeren, toezicht houden op, evalueren, aanpassen en stopzetten van een experiment dat fundamenteel is gestoeld op de inzet van bijzonder complexe en ondoorzichtige algoritmen.

FAT/ML 'Principles for Accountable Algorithms and Social Impact Statement for Algorithms' De FAT/ML 'Principles for Accountable Algorithms and Social Impact Statement for Algorithms' (Diakopoulos et al., 2018) is een voorstel van een groep onderzoekers uit de computerwetenschap, dat aansluitend op de jaarlijkse conferentie over 'Fairness, Accountability, and Transparency in Machine Learning' (eerlijkheid, verantwoording en transparantie in machine learning) werd uitgebracht (fatml.org, 2018). De Principles for Accountable Algorithms zijn bedoeld om ontwikkelaars, ontwerpers en organisaties te ondersteunen bij de ontwikkeling van algoritmische systemen die **publieke verantwoording moeten afleggen**, waarbij publieke verantwoording wordt gedefinieerd als:

'... de plicht om algoritmische besluitvorming te rapporteren, verklaren of rechtvaardigen en eventuele negatieve maatschappelijke effecten of berokkende schade te beperken en te verzachten'. (Diakopoulos et al., 2018)

Het FAT/ML-voorstel definieert vijf **Principes voor verantwoorde algoritmen** (namelijk verantwoordelijkheid, verklaarbaarheid, nauwkeurigheid, controleerbaarheid (*auditability*) en eerlijkheid) en geeft vervolgens een **Verklaring van de maatschappelijke impact van algoritmen**, die erop is gericht inachtneming van de voornoemde principes te waarborgen.

De auteurs raden makers van algoritmen aan een Social Impact Statement te ontwikkelen en te publiceren en deze (minstens) drie keer in de loop van het ontwerp- en ontwikkelingsproces te herzien en te evalueren: in de ontwerpfase, voor de lancering en na de lancering. De structuur van de Social Impact Statement

hangt nauw samen met de vijf principes en de inhoud van de verklaring geeft informatie die betrokken publieke partijen in staat moet stellen te evalueren in welke mate de makers van het algoritme de principes in acht nemen.

De Social Impact Statement is een belangrijke stap die verdergaat dan de algemene principes en richtlijnen en onderscheidt zich in het bijzonder van de ethische richtlijnen voor AI die eerder werden besproken. Een van de sterkste punten van dit voorstel is de eis dat een Social Impact Statement wordt gepubliceerd. Deze moet een openbare discussie over het betreffende algoritmische systeem faciliteren en betrokken publieke partijen in staat stellen de werking van algoritmen aan te vechten. Aan de andere kant lijkt de FAT/ML-verklaring zich exclusief te richten op 'bouwers van algoritmen' d.w.z. ontwikkelaars en productontwerpers, waarbij de mogelijke input van andere betrokkenen, zoals contractpartijen, beleidsmakers en overheidsinstellingen wordt genegeerd (Rohde, 2018; Selbst, Boyd, Friedler, Venkatasubramanian, & Vertesi, 2019).

De Social Impact Statement gaat verder dan algemene principes en richtlijnen

Kader voor geïntegreerde rapportage

Openbare verslaglegging is een standaardmethode om ervoor te zorgen dat externe belanghebbenden de prestaties en uitkomsten van publieke en private organisaties kunnen beoordelen, bijvoorbeeld ten aanzien van financiële verslaglegging en duurzaamheidsrapportage. Een bijzonder interessante en relevante benadering is geïntegreerde rapportage (IR, *integrated reporting*) (Banerjee, 2019; Eccles & Krzus, 2015), een benadering die waardecreatie van een organisatie op holistische waarde wil omvatten en dus verdergaat dan simpelweg de financiële waarde en traditionele duurzaamheidsmaatregelen. Het Kader voor geïntegreerde rapportage Kader heeft ten doel de bestuurlijke regie, verantwoording en vertrouwen in organisaties te verbeteren met behulp van periodieke rapportage van 'geïntegreerde' waardecreatie in de loop der tijd.

Een geïntegreerd rapport is een beknopte communicatie over hoe de strategie, het bestuur, de prestaties en vooruitzichten van een organisatie leiden tot de creatie van waarde op de korte, middellange en lange termijn (IIRC, 2016). Het Kader voor geïntegreerde rapportage is aangenomen door vele organisaties in de publieke sector. Het is voor steden over de hele wereld uitgegroeid tot een belangrijk instrument om over duurzaamheid te rapporteren (Niemann & Hoppe, 2018; Oprisor, Tiron-Tudor, & Nistor, 2016; Sulkowski, 2017; Tirado-Valencia, Rodero-Cosano, Ruiz-Lozano, & Rios-Berjillos, 2016).

Alhoewel geïntegreerde rapportage zich voornamelijk richt op rapportage over financiën en duurzaamheid, verschaft het vier lessen die relevant zijn voor de verdere ontwikkeling van het FAT/ML-voorstel.

Ten eerste moet rapportage worden gezien als een vorm van communicatie tussen verschillende groepen betrokkenen. Een rapport is uitsluitend bruikbaar als het de betrokken partijen helpt een doorlopende, vruchtbare dialoog tot stand te brengen. Opstellers van een rapport moeten dus inzicht hebben in hun publiek en een taal gebruiken die hun publiek kan begrijpen. Een rapport dat alleen ‘makers van algoritmen’ kunnen begrijpen, heeft geen waarde.

Ten tweede moet rapportage toekomstig gedrag beïnvloeden en organisaties (of makers van algoritmen) helpen hun praktijken en processen te verbeteren.

Een rapport moet niet slechts worden gezien als rechtvaardiging van gedrag in het verleden. Met andere woorden, ‘makers van algoritmen’ moeten deze rapportage integreren als een vorm van zelfreflectie.

Ten derde legt geïntegreerde rapportage de nadruk op *waardecreatie* door de vraag te stellen wat voor waarde er wordt gecreëerd en voor wie. De waardedimensie is op dit moment echter niet aanwezig in het FAT/ML-voorstel. Terwijl het FAT/ML-voorstel onderwerpen noemt die kunnen worden geïnterpreteerd als ‘waardevernietiging’ (vooroordelen, oneerlijkheid enz.), omvat het geen informatie over de positieve waarde die wordt gecreëerd door algoritmen.

Ten vierde benadrukt de opkomst van de geïntegreerde rapportage hoe een langdurig, door de gemeenschap geleid proces kan leiden tot een wereldwijde beweging met aanzienlijke impact (Banerjee, 2019), dat op grote schaal wordt overgenomen in de private en publieke sector (Niemann & Hoppe, 2018; Oprisor, Tiron-Tudor, & Nistor, 2016; Sulkowski, 2017; Tirado-Valencia, Rodero-Cosano, Ruiz-Lozano, & Rios-Berjillos, 2016).

5. Aanbevelingen

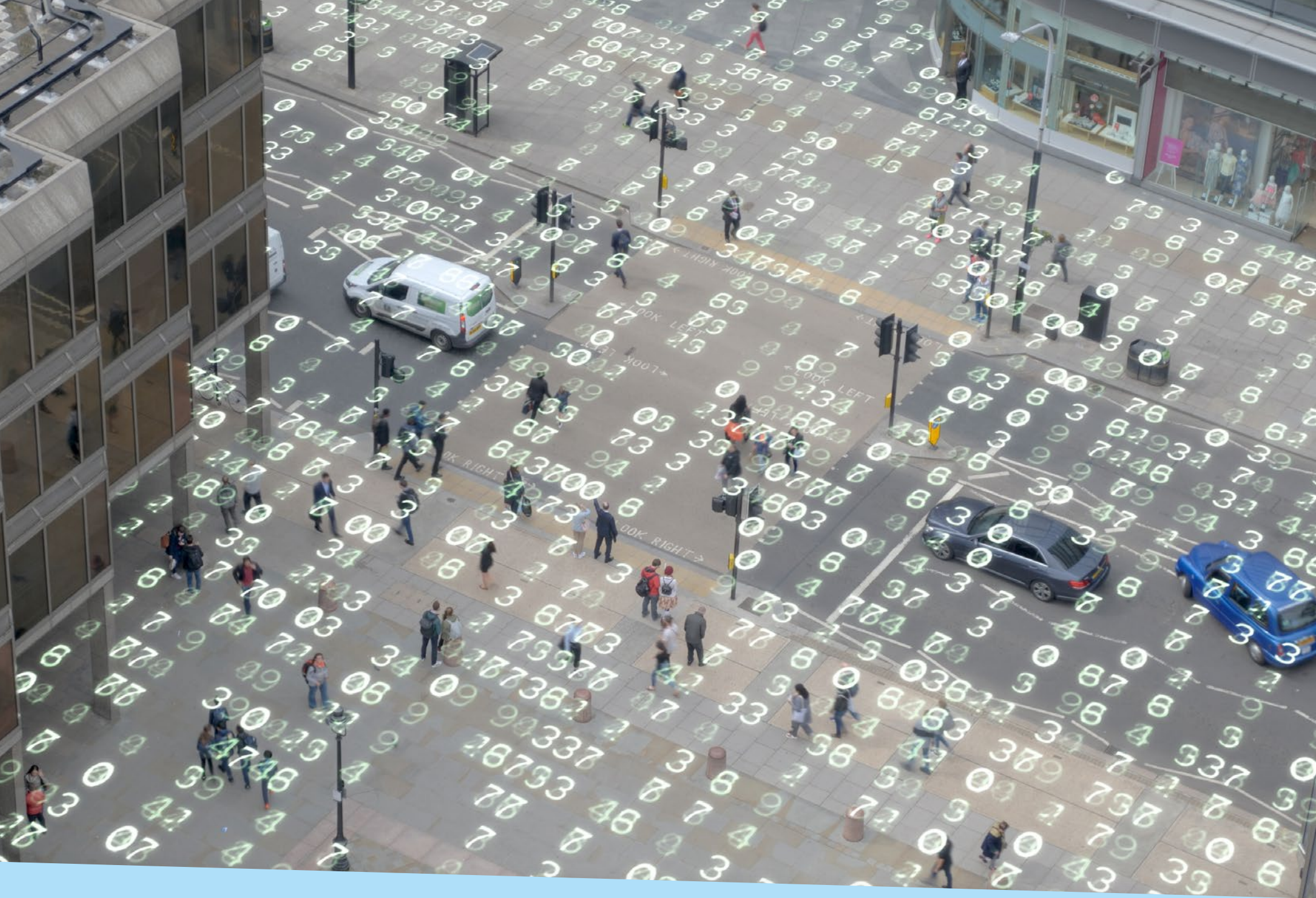
Ten aanzien van data en algoritmen luidt de voornaamste vraag als volgt:

Hoe kunnen we het ontwerp, de ontwikkeling en het gebruik van data en algoritmen in de samenleving zo inrichten en aansturen dat de voordelen voor het publiek worden gemaximaliseerd en maatschappelijke waarden worden gerespecteerd?

De discussie die in dit essay wordt gepresenteerd, belicht de brede ethische vraagstukken rond het gebruik van data en algoritmen en leidt tot de volgende belangrijke inzichten:

1. Datagedreven algoritmen spelen een steeds grotere rol in besluitvormingsprocessen met verrijkende maatschappelijke implicaties. Met andere woorden, algoritmen activeren data. Bestuur en beleid moeten data en algoritmen dus als eenheid behandelen.
2. Algoritmen worden nu snel zo complex en geavanceerd dat zelfs computerspecialisten maar moeilijk kunnen begrijpen hoe algoritmen werken en beslissingen nemen.
3. Data en algoritmen zijn bijzonder kneedbaar en onderhevig aan continue verandering en evolutie. De snelheid waarmee ze veranderen zal alleen nog maar toenemen wanneer organisaties agile data science engineering-praktijken (DevOps) ontwikkelen en toepassen en machinelearning-processen steeds verder automatiseren.
4. Algoritmen moeten worden gezien als experimentele technologie. Het is onmogelijk de maatschappelijke implicaties van algoritmen te voorspellen. Daardoor ontstaat de behoefte aan ethische aansturing van het experimentele proces omtrent data en algoritmen.
5. Living labs zijn een doeltreffende methode om verschillende belanghebbenden bij elkaar te brengen in een innovatief leerproces om te bepalen hoe data en algoritmen maatschappelijke en commerciële waarde kunnen opleveren.
6. We missen richtlijnen en regelgeving die ons vertellen hoe dit experimentele proces moet worden bestuurd. Ethische richtlijnen voor AI bieden waardevolle aanknopingspunten voor waar aandacht aan moet worden besteed maar zijn nog niet concreet genoeg.





5.1 Principes

Antwoorden op de bovenstaande kernvraag vereisen een brede insteek waarbij ethiek, beleid en regelgeving worden betrokken. Vertrouwen op regelgeving alleen is niet genoeg omdat er ook schadelijke en onethische gevolgen kunnen optreden als men zich streng aan de regelgeving houdt.

Ik stel voor dat toekomstige beleidsinspanningen moeten worden gebaseerd op vijf principes:

- **Concentreren op data en algoritmen, niet alleen data.** Met andere woorden, toekomstige beleidsinspanningen moeten invloed uitoefenen op hoe data wordt gebruikt in (geautomatiseerde) besluitvormingsprocessen en hoe besluitvorming (in de publieke en private sector) wordt versleuteld in steeds complexere en meer geavanceerde algoritmen.
- **Aanmoedigen van een veilig en ethisch experimenteel leerproces rond data en algoritmen.** Het experimentele leerproces dat bijvoorbeeld door living labs wordt geïmplementeerd, is fundamenteel voor economische en maatschappelijke ontwikkeling. Beleid moet experimenteren dus zien als een kans, niet als bedreiging.
- **Nadruk op publieke betwistbaarheid.** Publieke betwistbaarheid richt zich op de vraag hoe het maatschappelijk middenveld kritiek kan uitoefenen op algoritmen en bezwaar kan maken tegen de beslissingen die eruit voortvloeien. Dit gaat verder dan transparantie en richt zich op mechanismen en instellingen om feedback en kritiek vanuit de samenleving te verzamelen, te analyseren en ernaar te handelen. Ook moeten belanghebbenden in staat worden gesteld lopende ontwikkelingen aan te vechten.
- **Concentreren op het stimuleren van een openbaar debat** over de manier waarop data en algoritmen in de openbare ruimte worden gebruikt; een debat door en voor het publiek.
- **Ondersteunen van doorlopende inspanningen door Nederlandse gemeenten** en daarbij dubbele inspanningen vermijden. Nederlandse steden hebben een internationaal vooraanstaande positie in de ontwikkeling van ethische praktijken en beleidsmaatregelen op het gebied van data en algoritmen. Bij toekomstig beleid moeten er lessen worden getrokken uit deze ervaringen en moet men zich eerst richten op het verzamelen van bewijs en het identificeren van best practices voordat uiteindelijk regels en normen worden vastgesteld.

Naar aanleiding van de analyse die in dit hoofdstuk is gepresenteerd, stel ik beleidsaanbevelingen voor op drie gebieden: bestuur, instellingen en vaardigheden.

5.2 Bestuur

De potentiële impact van data en algoritmen is verreikend en onvoorspelbaar. Overheden en gemeenten moeten hun beslissingen over het gebruik van data en algoritmen baseren op **democratisch overleg**, waarbij het publiek inspraak heeft in hoe deze ontwikkelingen worden vormgegeven. Door heel Nederland werken publieke en private partijen in living labs samen aan de ontwikkeling van data- en algoritmesystemen. Vaak zijn ook burgers en gemeenschappen daar nauw bij betrokken. De praktijken en normen van deze inspanningen variëren echter enorm en er is een gebrek aan transparantie ten aanzien van welke praktijken en normen waar in Nederland worden toegepast. De nationale overheid kan een sleutelrol spelen in het samenbrengen van deze versnipperde lokale inspanningen om een gemeenschappelijk inzicht te verkrijgen in alle best practices en beleidsmaatregelen. Zo kan uiteindelijk een aantal gezamenlijke principes voor heel Nederland worden ontwikkeld.



Het doel van de overheid moet drievoudig zijn:

1. Transparantie creëren ten aanzien van wie met welk doel welke algoritmen ontwikkelt
2. Continue evaluatie van algoritmen en de impact ervan mogelijk maken
3. Partijen uit het maatschappelijk middenveld in staat stellen de ontwikkeling en het gebruik van algoritmen ter discussie te stellen via openbare debatten

Aanbeveling 1: Richt een Algoritme Rapportage Initiatief op, dat tot doel heeft de maatschappelijke waarde en risico's van algoritme-projecten in Nederland te documenteren en te volgen.

Op dit moment is er weinig inzicht in welke soorten datagedreven algoritmen er worden ontwikkeld en gebruikt in Nederland en wat de voordelen en risico's zijn. Het voornaamste probleem is dat er niet voldoende openbare informatie voorhanden is over zulke projecten en dat er geen standaard manier is om algoritmen te documenteren. In de vorige paragraaf besprak ik voorstellen om via rapportages verantwoording af te leggen over algoritmen en de maatschappelijke impact ervan, maar deze voorstellen zijn nog niet in de praktijk getest (Diakopoulos, 2014, 2015, 2019; Diakopoulos & Friedler, 2016). Ik heb ook besproken hoe geïntegreerde rapportage is voortgekomen uit een proces op gemeenschapsniveau (Banerjee, 2019), en deze praktijken zijn inmiddels door publieke en private organisaties wereldwijd in gebruik zijn genomen (Niemann & Hoppe, 2018; Oprisor, Tiron-Tudor, & Nistor, 2016; Sulkowski, 2017; Tirado-Valencia, Rodero-Cosano, Ruiz-Lozano, & Rios-Berjillos, 2016).

Daarom bestaat mijn eerste aanbeveling eruit een groep betrokken partijen (onderzoekers, ontwikkelaars, beleidsmakers, burgers enz.) bij elkaar te brengen met het doel een **Algoritme Rapportage Kader** te ontwikkelen. Dit is hoogstwaarschijnlijk een meerjarig initiatief waarbij bestaande algoritme-initiatieven in kaart worden gebracht, er 360 graden inzicht wordt verkregen in dergelijke initiatieven en er rapportagepraktijken worden ontwikkeld en getest.

Met het oog op de hoge innovatiesnelheid van algoritmen en de experimentele aard van het merendeel van de ontwikkelingswerkzaamheden, zouden we voorzichtig moeten opereren en moeten we niet te snel proberen normen vast te stellen of aannemen dat we de impact van algoritmen al volledig kunnen

doorgronden. Het is dan ook belangrijk dat we algoritmerapportage niet beschouwen als een soort audit of certificering. Voor audits en certificering zijn duidelijk gedefinieerde criteria en methoden nodig, die slechts beschikbaar zijn voor een klein deel van de met algoritmen geassocieerde vraagstukken. Er is nog onvoldoende inzicht in de meest diepgaande effecten van algoritmen. Net als bij het Kader voor geïntegreerde rapportage is de voornaamste waarde van algoritmerapportage dat personen en organisaties anders over deze zaken gaan denken; het gaat om het denkwerk dat aan het gepubliceerde rapport voorafgaat en niet om het rapport zelf. In navolging van de aanbevelingen van Mittelstadt om 'bottom-up AI-ethiek te steunen' (Mittelstadt, 2019), wil ik hier benadrukken dat het noodzakelijk is om verder te bouwen op bestaande initiatieven die momenteel plaatsvinden in living labs en gemeenten door heel Nederland en geleidelijk aan een Algoritme Rapportage Kader te ontwikkelen via een participatieve bottom-up benadering.

Er is nog onvoldoende inzicht in de meest diepgaande effecten van algoritmen

Het is zeker niet te vroeg om te speculeren wat een Algoritme Rapportage Kader allemaal zou kunnen omvatten. Ik heb ideeën uit verslaglegging van maatschappelijke impact, geïntegreerde rapportage en de experimental ethics van Van de Poel gecombineerd en kom daarmee tot de conclusie dat een algoritme-rapportage in grote lijnen zou kunnen bestaan uit vier informatiecategorieën, namelijk: Waarde, Mechanismen, Experimentatie en Publiek:

Waarde

- **Waardecreatie en voordelen.** Welke waarde creëren de algoritmen? Wie profiteert van het gebruik van dit algoritme en hoe? Zijn deze voordelen publiek (voor de hele samenleving) of particulier (dragen ze bij aan commerciële belangen)?
- **Waardevernietiging en schade.** Welke waarde wordt vernietigd, met name met betrekking tot aspecten van eerlijkheid, zoals mogelijke vooroordelen, discriminatie en exclusie. Welke potentiële impact en schade kunnen er voor personen en specifieke groepen voortvloeien uit het gebruik van algoritmen?

- **Menselijke verantwoordelijkheid en betrokkenheid.** Wie is er betrokken, wie heeft directe controle over het algoritme, wie houdt toezicht en legt verantwoording af? Deze informatie moet betrekking hebben op zowel individuele personen als bedrijven en organisaties die zijn betrokken bij de creatie, toepassing en het gebruik van het algoritme en de doelen, doelstelling en intenties ervan verwoorden.

Mechanisme

- **Data.** Door welke data wordt het algoritme aangedreven? Wat is er bekend over de kwaliteit van de data, zoals de nauwkeurigheid, volledigheid en onzekerheid evenals de tijdigheid? Hoe is de data verzameld, getransformeerd, gecontroleerd en bewerkt? Hoe wordt de data beheerd en opgeslagen en onder wiens controle?
- **Model.** Welke kenmerken of variabelen worden gebruikt in het algoritme? Welke trainingsdata is gebruikt en hoe kan dit worden geclassificeerd als het gaat om correctheid, volledigheid enz.? Hoe ziet het algemene proces voor de bouw van het model eruit?
- **Inferenties.** Welke inferenties maakt het algoritme, bijvoorbeeld classificaties of voorspellingen? Wat zijn de standaardmetingen, zoals nauwkeurigheid enz.? Welke stappen worden er genomen om bekende fouten te herstellen?

Deze eerste twee categorieën zijn in de eerste plaats gebaseerd op de waardefocus van het Kader voor geïntegreerde rapportage (International Integrated Reporting Council, 2013) en de FAT/ML Principles for Accountable Algorithms (Diakopoulos et al., 2018). Overeenkomstig (van de Poel, 2016) stel ik daarnaast voor dat de rapportage informatie moet bevatten over de experimentele constructie waarin algoritmen worden ontwikkeld:

Experimentatie

- **Participatie.** Hoe worden personen van wie de data wordt gebruikt of voor wie de algoritmen gevolgen hebben, geïnformeerd en betrokken bij de inrichting van het experiment? Hoe geven personen hun toestemming en hoe kunnen ze zich terugtrekken uit het experiment (inclusief hun data)? Hoe kunnen personen die deelnemen invloed uitoefenen op het inrichten, uitvoeren, toezicht houden op, evalueren, aanpassen en stopzetten van het experiment?

- **Bescherming.** Hoe worden kwetsbare personen die deelnemen beschermd? Welke stappen worden ondernomen om deze risico's te identificeren en te verhelpen? Hoe kan potentiële schade worden teruggedraaid?
- **Toezicht.** Wie geeft goedkeuring voor en stuurt de experimentele ontwikkeling en het gebruik van algoritmen aan? Zijn dit democratisch goedgekeurde organen, particuliere partijen of anderszins? Hoe worden de effecten van het algoritme geëvalueerd en publiekelijk beschikbaar gesteld?

Ten slotte wil ik een categorie voorstellen die gerelateerd is aan de noodzaak om een effectief publiek debat en kritiek mogelijk te maken, deels gebaseerd op de FAT/ML Principles for Accountable Algorithms:

Publiek

- **Controleerbaarheid (auditability) & betwistbaarheid.** Hoe kunnen maatschappelijke betrokkenen het gedrag van het algoritme onderzoeken, begrijpen en controleren en hierop feedback geven, potentiële negatieve gevolgen aangeven en verbeteringen voorstellen? Wie ontvangt en evalueert deze input? Welke feedback en suggesties zijn ontvangen en hoe zijn ze verwerkt? Welke veranderingen zijn hieruit voortgekomen?



Samen leggen deze rapportcategorieën mogelijk een zware last op de partijen die verantwoordelijk zijn voor de ontwikkeling en het gebruik van algoritmen. Het is dan ook van groot belang dat er best practices worden gedefinieerd en flexibele normen worden gecreëerd, die in een praktisch rapportagekader worden verankerd dat ook werkelijk doet wat het moet doen: het vertrouwen in algoritmen vergroten en waarborgen dat algoritmen overeenkomen met maatschappelijke waarden.

Het doel van publieke rapportage is om het bestuur en de verantwoording van en het vertrouwen in organisaties (de eenheid, het team of de groep die verantwoordelijk is voor algoritmen) te vergroten. Dit kan alleen worden gerealiseerd als rapportage de organisatie toegevoegde waarde biedt en niet alleen externe belanghebbenden. Het Kader voor geïntegreerde rapportage benadrukt het belang van het creëren van een interne cultuur waarin het rapportagekader wordt gebruikt om discussies te stimuleren over hoe een organisatie waarde creëert (of vernietigt). Datzelfde moet gelden voor een algoritmisch rapportagekader.

Daarnaast moet rapportage een iteratief, periodiek proces zijn. De waarde van publieke rapportage ligt niet in de eerste plaats in het gepubliceerde rapport maar in het volgen van de voortgang door de tijd en de bijbehorende lessen in het verbeteren van controle, betwistbaarheid enz.

Aanbeveling 2: Ontwikkel een Nationaal Algoritme Register dat een effectieve vergelijking en beoordeling van initiatieven op het gebied van data en algoritmen in Nederland mogelijk moet maken.

De tweede aanbeveling betreft het inrichten van een centraal register van algoritmen, dat informatie over algoritmen beschikbaar maakt aan zoveel mogelijk geïnteresseerden. Waar Aanbeveling 1 zich richt op hoe algoritmen moeten worden gedocumenteerd, concentreert Aanbeveling 2 zich op het verzamelen van informatie over alle relevante algoritme-initiatieven door heel Nederland.

In eerste instantie kan dit register worden samengesteld uit informatie afkomstig van overheidsinstellingen, gemeenten en vrijwillige bijdragen van anderen.

Mettertijd, wanneer er een rapportagekader is ontwikkeld, kan de overheid bepalen dat elk algoritme-initiatief moet worden opgenomen in dit register. De informatie in het Algoritme Register moet openbaar zijn, wat betekent dat het voornamelijk is gericht op publieke algoritmen en algoritmen gericht op het algemeen belang.

Het Nationaal Algoritme Register kan een belangrijke informatiebron worden voor onderzoekers, bedrijven en beleidsmakers. Onderzoekers kunnen informatie opvragen om meer inzicht te krijgen in de risico's van algoritmen. Bedrijven kunnen inzicht krijgen in hoe ze de effecten van hun algoritmen kunnen evalueren. Beleidsmakers kunnen meer doeltreffende manieren ontwikkelen om algoritmen aan te sturen en te controleren en de impact op het beleid in de loop der tijd te volgen.

Het Nationaal Algoritme Register kan een belangrijke informatiebron worden voor onderzoekers, bedrijven en beleidsmakers

Hiertoe moet het Nationaal Algoritme Register een voor iedereen toegankelijke website hebben en regelmatig worden bijgewerkt om te waarborgen dat de informatie actueel, relevant en bruikbaar is.

Aanbeveling 3: Ontwikkel een Publiek Algoritme Forum om publieke partijen in staat te stellen de werking, het gebruik en de uitkomsten van initiatieven op het gebied van data en algoritmen te bespreken, te bekritisieren en te betwisten.

Kritieke technologieën zoals algoritmen die een vergaande impact kunnen hebben op de samenleving, moeten zo breed mogelijk worden besproken. Momenteel is dit niet mogelijk omdat het publiek niet weet welke algoritmen er worden gebruikt in de samenleving, hoe ze hun mening kunnen uiten en aan wie ze hun zorgen kunnen richten. Daarom raad ik aan een Publiek Algoritme Forum op te richten, waar publieke partijen de werking, het gebruik en de uitkomsten van initiatieven op het gebied van data en algoritmen kunnen bespreken, bekritisieren en betwisten.



Het Algoritme Forum moet een openbare website hebben en daarnaast een brede spectrum van methoden aanwenden om in gesprek te gaan met het publiek, waaronder openbare debatten, educatieve workshops en participatieve ontwerpessies. De feedback van het publiek die tijdens deze evenementen wordt verzameld, moet worden opgenomen op de website van het Algoritme Forum.

Een **Publieke Rapportage Hotline** moet deel uitmaken van het Algoritmen Forum. Hier kan het publiek specifieke zorgen melden die ze hebben over door een algoritmisch systeem genomen beslissingen of negatieve ervaringen met zo'n systeem. Dit voorstel is geïnspireerd op publieke hotlines in de lucht- en zeevaartsector, zoals het FAA Hotline Reporting Form (<https://hotline.faa.gov>) en het Britse UK Confidential Reporting Programme for Aviation and Maritime (<https://www.chirp.co.uk/>) die vertrouwelijke rapporten registreren, de hieruit voortvloeiende data analyseren en belangrijke informatie doorspelen naar relevante gemeenschappen en overheidsinstanties.

5.3 Instelling

Aanbeveling 4: Richt een Data en Algoritme Instituut op ter bevordering van de ontwikkeling van kennis, benaderingen, tools, infrastructuren en normen voor ethisch en verantwoord gebruik van algoritmen.

De missie van het **Data en Algoritme Instituut** zou als volgt moeten luiden: de maatschappelijke en economische voordelen van data en algoritmen voor de Nederlandse samenleving maximaliseren en de overheid, start-ups, gemeenten, maatschappelijke organisaties enz. bijstaan in het ontwikkelen van ethische data en algoritmepraktijken waarover verantwoording wordt afgelegd. Meer specifiek moet zo'n organisatie zich richten op:

- Het stimuleren van open innovatie in technologie, ethiek, bestuur en ontwerp
- Het aanbieden van training en onderwijs
- Contact met het publiek
- Onderzoek

Een belangrijke rol van dit instituut is het **Algoritme Rapportage Kader** aan te jagen, het **Nationaal Algoritme Register** en **Algoritme Forum** te ontwikkelen en te

onderhouden en in het algemeen op te treden als clearinghouse voor informatie over publieke algoritmen en data en algoritmen en data gericht op het algemeen belang. Dit omvat:

1. Het publiceren van informatie over lopende initiatieven op het gebied van data en algoritmen van de overheid, gemeenten en andere partijen.
2. Het ontwerpen van doeltreffende mechanismen om burgers en organisaties in staat te stellen lopende initiatieven op het gebied van data en algoritmen te bekritisieren en te betwisten (bijvoorbeeld door middel van crowdsourcing-initiatieven),
3. Het ontwerpen van doeltreffende mechanismen om zulke feedback te analyseren er actie op te ondernemen en
4. Periodiek openbaar te rapporteren over hoe de feedback en kritiek zijn verwerkt.

Om innovatie te stimuleren moet het voorgestelde **Data en Algoritme Instituut** sterke juridische, sociaalwetenschappelijke en ontwerpvaardigheden in huis hebben, evenals technische kennis op het gebied van data, informatiearchitecturen en businessmodellen. Het moet beschikken over de mogelijkheid het end-to-end-proces van data- en algoritmesystemen te analyseren, evenals de ethische en juridische implicaties van dergelijke systemen.

Het **Data en Algoritme Instituut** moet samenwerken met en ondersteuning bieden aan Nederlandse gemeenten, zodat lopende activiteiten effectiever worden verzameld en geanalyseerd om het opkomende ecosysteem van data en algoritmen te versterken. Een dergelijke organisatie kan gemodelleerd worden naar het voorstel voor een Machine Intelligence Commission van Geoff Mulgan (Mulgan, 2016) om de ontwikkeling van nieuwe benaderingen ter bescherming van het algemeen belang te ondersteunen.

De huidige inspanningen van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties om een Transparency Lab (Braak, 2019) op te richten, zijn een stap in de juiste richting. Zo'n lab dient echter te worden ingericht als onafhankelijke instelling buiten overheidshierarchieën en met een sterke betrokkenheid van bredere maatschappelijke sectoren. Naar het voorbeeld van het Open Data Institute, NESTA en de Open Knowledge Foundation moet zo'n instelling een duidelijke publieke missie hebben, worden geleid door een raad van

bestuur wier diversiteit de maatschappelijke belangen in algoritmen weerspiegelt en worden voorzien van de middelen om aanzienlijke expertise in huis te ontwikkelen. Het belangrijkste is echter dat het **Data en Algoritme Instituut** zich niet moet richten op het auditen van algoritmen, wat voornamelijk een methode voor risicobeoordeling en -beperking is, maar op benaderingen op het gebied van rapportage en betwistbaarheid zoals hierboven beschreven en op het stimuleren van innovatie in technologie, ethiek en bestuur.

Ik vermeld hierbij graag dat het in Nederland, in vergelijking tot het Verenigd Koninkrijk, ontbreekt aan sterke organisaties in de publieke sector zoals het Open Data Institute, NESTA en de Open Knowledge Foundation, die daar innovatie en publieke betrokkenheid ten aanzien van data en algoritmen stimuleren. Het Rathenau Instituut verricht uitstekend werk met onderzoek naar en publieke betrokkenheid bij maatschappelijk relevante aspecten van wetenschap en technologie.

Er is echter behoefte aan instellingen met een bredere opdracht op het gebied van innovatie, technologie, design en beleid, en die beter in staat zijn technische en beleidsontwikkelingen vorm te geven en te onderbouwen

De Waag in Amsterdam is van essentieel belang in het opschalen van grassroots-initiatieven en brengt maatschappelijke zorgen en vraagstukken voor het voetlicht. Er is echter behoefte aan instellingen met een bredere opdracht op het gebied van innovatie, technologie, design en beleid, en die beter in staat zijn technische en beleidsontwikkelingen vorm te geven en te onderbouwen. Er zijn vooral nieuwe vaardigheden en capaciteiten nodig ten aanzien van de interactie van de technologische, creatieve en beleidssectoren en voor de toepassing van design thinking en innovatieve methoden voor beleidsontwikkeling.

5.4 Vaardigheden

Aanbeveling 5: Ontwikkel een Nationale Skills Agenda voor Data en Algoritmen om ervoor te zorgen dat alle belanghebbenden (van burgers tot organisaties) over de juiste vaardigheden beschikken om deel te nemen aan de nieuwe data- en algoritme-economie.

Wetgevers en beleidsmakers hebben nieuwe vaardigheden nodig om verdergaand inzicht te verkrijgen in nieuwe technologieën en businessmodellen. Computer-experts hebben nieuwe vaardigheden nodig om inzicht te krijgen in de ethische en beleidsaspecten van data en algoritmen. Ontwerpers van producten en diensten hebben nieuwe vaardigheden nodig om effectief verantwoorde datagedreven producten en diensten te kunnen ontwerpen. Overheidsmedewerkers hebben nieuwe vaardigheden nodig om de juiste beslissingen te kunnen nemen over hoe en waar data en algoritmen worden gebruikt in publieke dienstverlening en geautomatiseerde besluitvormingsprocessen. En burgers moeten nieuwe vaardigheden krijgen om op betekenisvolle wijze input te kunnen leveren voor het gebruik van data en algoritmen in het publieke domein en dit gebruik ook te kunnen betwisten. Om deze vaardigheden de komende 25 jaar te kunnen ontwikkelen, is een strategische aanpak over de hele educatieve breedte vereist, van de basisschool tot het beroepsonderwijs en wetenschappelijke vorming.

Daarom raad ik aan naar een overkoepelende **Nationale Skills Agenda voor Data en Algoritmen** toe te werken, met drie doelstellingen in gedachten:

1. Vaststellen welke behoefte er is aan vaardigheden op het gebied van data en algoritmen door alle lagen van de samenleving, om het toekomstige economische en maatschappelijke welzijn van de Nederlandse samenleving te kunnen waarborgen.
2. Vaststellen welke vaardigheden op het gebied van data en algoritmen moeten worden onderwezen in verschillende onderwijs- en opleidingspaden, van de basisschool tot professionele en wetenschappelijke vorming.
3. Nieuwe educatieve initiatieven initiëren die gericht zijn op cruciale behoeftegebieden.



Deze aanbeveling is gebaseerd op het besef dat de huidige inspanningen op het gebied van onderwijs in data en algoritmen onvoldoende tegemoetkomen aan het tekort aan vaardigheden op dit gebied in de hele samenleving. De Nederlandse overheid investeert meer in wetenschappelijk onderwijs aan Nederlandse universiteiten, deels gericht op het versterken van het wetenschappelijk onderwijs in data, algoritmen en AI. Daarnaast wordt een aantal professionele master-opleidingen ontwikkeld, waar specialisten op verschillende vakgebieden data-vaardigheden kunnen opdoen. Ook zijn er MOOC's (Massive Open Online Courses) opgezet, die het publiek meer inzicht moeten geven in data, algoritmen en AI. Een voorbeeld is de Nationale AI-Cursus. Er moet echter nog veel meer worden gedaan. In het Verenigd Koninkrijk biedt het Open Data Institute al enkele jaren een educatief programma om datageletterdheid te vergroten, waarin online, gemengde en persoonlijke cursussen worden gecombineerd.

Tot dusver hebben meer dan 25.000 professionals en andere geïnteresseerden op deze manier vaardigheden op het gebied van data kunnen ontwikkelen. Het programma is gebaseerd op een Data Skills Framework (ODI, 2018) dat een breedgedragen datageletterdheidsprogramma beschrijft en op behoeften gebaseerde opleidingsroutes biedt. Het gaat daarbij niet alleen om technische datavaardigheden maar ook om ethiek en bestuur op het gebied van data.

De Nederlandse samenleving kan zich niet veroorloven dat slechts een selecte groep mensen inzicht heeft in data en algoritmen. Dit betekent dat de strategische aanpak zich niet alleen moet richten op beroeps- en wetenschappelijk onderwijs maar juist ook op de basisscholen. In de afgelopen jaren is het besef gegroeid dat computeronderwijs op scholen verder moet gaan dan IT-basisvaardigheden en zich moet richten op een breder aanbod van onderwijs op het gebied van datageletterdheid. In Engeland heeft een aantal recente initiatieven aangetoond hoe een modern onderwijsaanbod op dit gebied er op scholen uit kan zien, zoals het BBC micro:bit-programma (Schmidt, 2016; Sentence, Waite, Hodges, Macleod, & Yeomans, 2017), het Internet of Schools-programma (Moreira, Magalhães, Ramos, & Vairinhos, 2018; Moreira, Vairinhos, & Ramos, 2018) (met de bedoeling leerlingen Internet of Things-vaardigheden bij te brengen) en de Urban Data School (Wolff, Kortuem, & Caverio, 2015) (een initiatief waaraan ik heb meegewerkt en dat kinderen van 8 tot 12 datageletterdheid moet bijbrengen).

De Urban Data School in het bijzonder toonde aan dat datageletterdheid op betekenisvolle wijze kan worden geïntegreerd in het bètaonderwijs op scholen en dat het haalbaar is om zelfs jonge leerlingen de basis van datageletterdheid bij te brengen. Daarbij kunnen schijnbaar gecompliceerde onderwerpen als acquisitie, voorbereiding, analyse, communicatie en ethiek van data aan bod komen (Wolff, Gooch, Montaner, Rashid, & Kortuem, 2016).

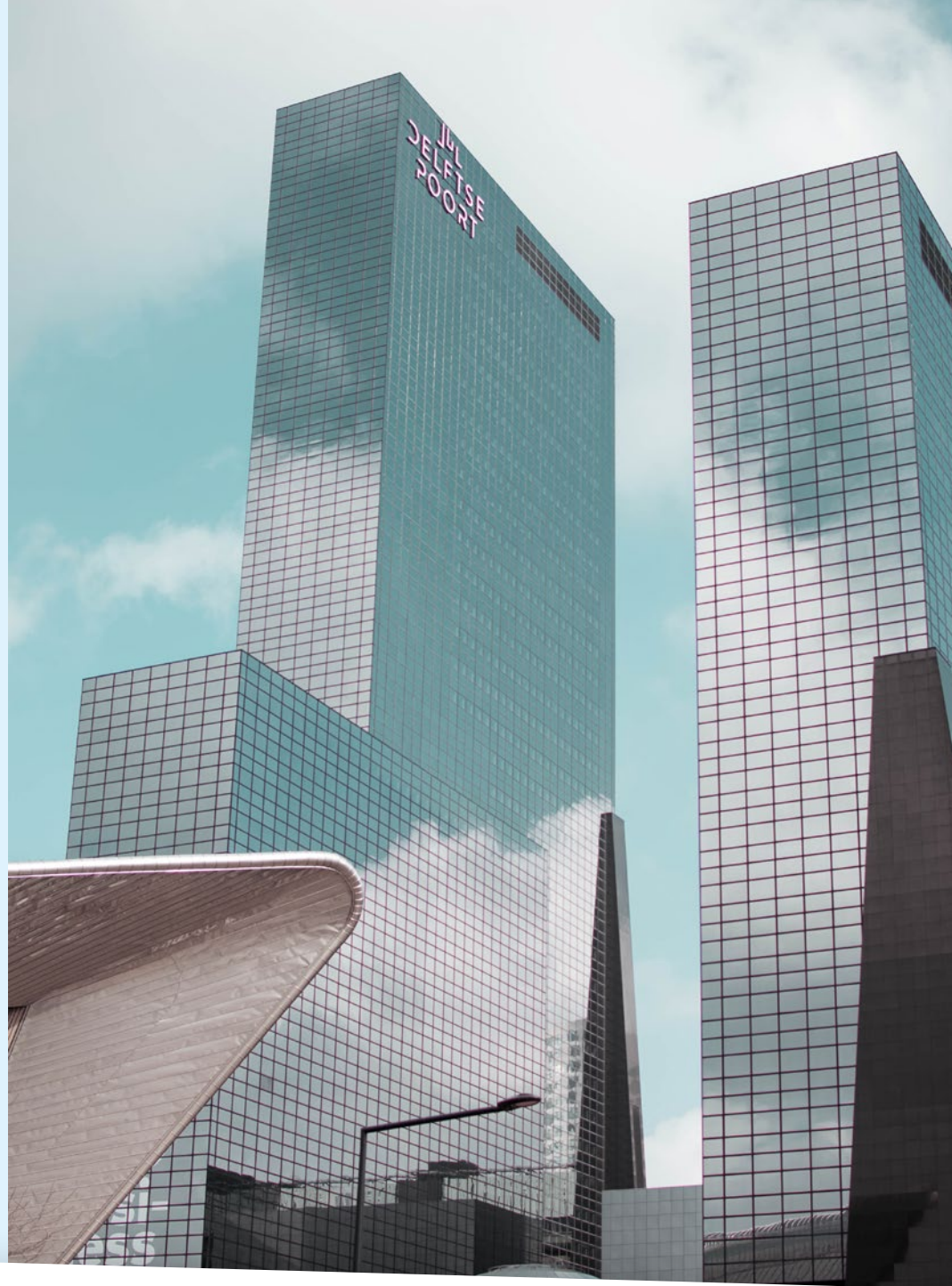
Het voornaamste doel van de voorgestelde Nationale Skills Agenda voor Data en Algoritmen is het ontwikkelen van een langetermijnvisie voor de vaardigheden op het gebied van data en algoritmen die nodig zijn in de samenleving en om tot innovatieve benaderingen te komen voor de ontwikkeling ervan.



Over de auteur



Prof. dr. Gerd Kortuem is hoogleraar Internet of Things aan TU Delft.



Referenties

- Achrekar, H., Gandhe, A., Lazarus, R., Yu, S. H., & Liu, B. (2011). Predicting flu trends using twitter data. In 2011 IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPs 2011. <https://doi.org/10.1109/INFCOMW.2011.5928903>
- Andreessen, M. (2011). Why Software Is Eating The World. *Wall Street Journal*.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias. *ProPublica*.
- Baccarne, B., Mechant, P., Schuurma, D., De Marez, L., & Colpaert, P. (2014). Urban Socio-technical Innovations with and by Citizens. *Interdisciplinary Studies Journal*, 3(4), 143–156. Retrieved from <https://biblio.ugent.be/publication/4365378>
- Banerjee, J. (2019). Integrating Multi-Capital Thinking into Business Decisions. In J. A. Allan (Ed.), *The Oxford Handbook of Food, Water and Society*. Oxford University Press. Oxford University Press.
- Bass, L. (2017). The Software Architect and DevOps. *IEEE Software*. <https://doi.org/10.1109/MS.2017.4541051>
- Bourgeois, J., Kortuem, G., & Kawsar, F. (2018). Trusted and GDPR-compliant research with the internet of things. <https://doi.org/10.1145/3277593.3277604>
- Braak, R. (2019). Transparant over algoritmen.
- Brauneis, R., & Goodman, E. P. (2017). Algorithmic Transparency for the Smart City. *Ssrn*, 103, 103–176. <https://doi.org/10.2139/ssrn.3012499>
- Brown, A. (2018). *Waving the flag for legitimate AI in government*. Retrieved from <https://www.centreforpublicimpact.org/waving-the-flag-for-legitimate-ai-in-government/>
- Calo, R. (2017). *Artificial Intelligence Policy: A Primer and Roadmap*. SSRN. <https://doi.org/10.2139/ssrn.3015350>
- Campolo, A., Sanfilippo, M., Whittaker, M., & Crawford, K. (2017). *AI Now 2017 Report*.
- Chen, Y.-C., & Hsieh, T.-C. (2014). Big Data for Digital Government: Opportunities, Challenges, and Strategies. *International Journal of Public Administration in the Digital Age*. <https://doi.org/10.4018/ijpada.2014010101>
- Collingridge, D. (1982). *The social control of technology*. New York: St. Martin's Press.
- Copeland, E. (2019). Does the public sector really need a code of AI ethics? Retrieved July 1, 2019, from <https://www.nesta.org.uk/blog/does-public-sector-really-need-code-ai-ethics/>
- Cosgrave, E., Arbutnot, K., & Tryfonas, T. (2013). Living labs, innovation districts and information marketplaces: A systems approach for smart cities. *Procedia Computer Science*, 16(Cser 13), 668–677. <https://doi.org/10.1016/j.procs.2013.01.070>
- D'Ignazio, C., & Klein, L. (2019). Data Feminism - o. Introduction. In M. P. Open (Ed.), *Data Feminism*. Retrieved from <https://bookbook.pubpub.org/data-feminism>
- Diakopoulos, N. (2014). Algorithmic accountability reporting: On the investigation of black boxes. *Columbia Journalism School*, 1–33. <https://doi.org/10.1002/ejoc.201200111>
- Diakopoulos, N. (2015). Accountability in Algorithmic Decision-making. *Queue*, 59(2), 56–62. <https://doi.org/10.1145/2857274.2886105>
- Diakopoulos, N. (2019). 6. Algorithmic Accountability Reporting. In *Automating the News* (pp. 204–239). Cambridge, MA and London, England: Harvard University Press. <https://doi.org/10.4159/9780674239302-007>
- Diakopoulos, N., & Friedler, S. (2016). How to Hold Algorithms Accountable Algorithmic systems have a way of making mistakes. *MIT Technology Review*. Retrieved from <https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/>
- Diakopoulos, N., Friedler, S., Arenas, M., Barocas, S., Hay, M., Howe, B., ... Zevenbergen, B. (2018). Principles for Accountable Algorithms and a Social Impact Statement for Algorithms. *Fatml.Org*, 1–4. <https://doi.org/10.1016/j.molliq.2013.05.029>
- Dougherty, P. (2019). How and Why to Use Agile for Machine Learning.
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*. <https://doi.org/10.1126/sciadv.aa05580>
- Ebert, C., Gallardo, G., Hernantes, J., & Serrano, N. (2016). DevOps. *IEEE Software*. <https://doi.org/10.1109/MS.2016.68>
- Eubanks, V. (2015, April). The Policy Machine. *Slate*. Retrieved from <https://slate.com/technology/2015/04/the-dangers-of-letting-algorithms-enforce-policy.html>
- Fan, M. D., & Jackson, H. M. (2015). A Bounded Access Model of Disclosure.
- fatml.org. (2018). Fairness, Accountability, and Transparency in Machine Learning.
- Felt, U., Wynne, B., Gonçalves, M. E., Jasanoff, S., Callon, M., Jepsen, M., ... Tallacchini, M. (2007). *Taking European knowledge society seriously. Office for Official Publications of the European ...*
- Fitzgerald, B., & Stol, K.-J. (2014). Continuous software engineering and beyond: trends and challenges. In *Proceedings of the 1st International Workshop on Rapid Continuous Software Engineering - RCoSE 2014*. <https://doi.org/10.1145/2593812.2593813>
- Fitzgerald, B., & Stol, K. J. (2017). Continuous software engineering: A roadmap and agenda. *Journal of Systems and Software*. <https://doi.org/10.1016/j.jss.2015.06.063>
- Follow, P. D. (2019). How and Why to Use Agile for Machine Learning. Retrieved from <https://medium.com/qash/how-and-why-to-use-agile-for-machine-learning-384b030e67b6>
- Frankel, M. S. (1989). Professional codes: Why, how, and with what impact? *Journal of Business Ethics*. <https://doi.org/10.1007/BF00382575>
- Gandy, O. H., & Nemorin, S. (2018). Toward a political economy of nudge: smart city variations. *Information Communication and Society*. <https://doi.org/10.1080/1369118X.2018.1479699>
- Gascó, M. (2017). Living labs: Implementing open innovation in the public sector. *Government Information Quarterly*, 34(1), 90–98. <https://doi.org/10.1016/j.giq.2016.09.003>
- Gasser, U., & Almeida, V. A. F. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*. <https://doi.org/10.1109/MIC.2017.4180835>
- Goodin, R. E. (1995). *Utilitarianism as a Public Philosophy*. Cambridge University Press.
- IIRC. (2016). What is integrated reporting?
- Intellipaat. (2016). Data Modelling Concepts in Data Science. Retrieved from <https://intellipaat.com/tutorial/data-science-tutorial/modeling-the-data/>
- International Integrated Reporting Council. (2013). *Value Creation background paper for <IR>. Leadership Excellence*. Retrieved from <http://integratedreporting.org/wp-content/uploads/2013/08/Background-Paper-Value-Creation.pdf>
- International Zone. (n.d.). Retrieved from <https://www.thehaguesecuritydelta.com/innovation/living-labs/lab/4-international-zone>
- Introna, L. D. (2016). Algorithms, Governance, and Governmentality: On Governing Academic Writing. *Science Technology and Human Values*. <https://doi.org/10.1177/0162243915587360>
- Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*. <https://doi.org/10.1016/j.giq.2016.08.011>
- Joh, E. E. (2017). The Undue Influence of Surveillance Technology Companies on Policing. *Ssrn*, 101–130. <https://doi.org/10.2139/ssrn.2924620>
- Joseph, R. C., & Johnson, N. A. (2013). Big data and transformational government. *IT Professional*. <https://doi.org/10.1109/ITP.2013.6542222>

doi.org/10.1109/MITP.2013.61

- Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1–14.
- Kitchin, R. (2016). The ethics of smart cities and urban science. <https://doi.org/10.1098/rsta.2016.0115>
- Kitchin, R. (2017a). Thinking critically about and researching algorithms. *Information Communication and Society*. <https://doi.org/10.1080/1369118X.2016.1154087>
- Kitchin, R. (2017b). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29. <https://doi.org/10.1080/1369118X.2016.1154087>
- Kitchin, R. (2019). The ethics of smart cities.
- Kobielus, J. (2014). The ground truth in agile machine learning. *IBM Data Management Magazine*.
- Kraemer, F., van Overveld, K., & Peterson, M. (2011). Is there an ethics of algorithms? *Ethics and Information Technology*, 13(3), 251–260. <https://doi.org/10.1007/s10676-010-9233-7>
- Kroeskop, J. R. (2018). *Algorithms and Public Policy*. Utrecht University. Retrieved from <https://dspace.library.uu.nl/handle/1874/366831>
- Kronsell, A., & Mukhtar-Landgren, D. (2018). Experimental governance: the role of municipalities in urban living labs. *European Planning Studies*, 26(5), 988–1007. <https://doi.org/10.1080/09654313.2018.1435631>
- Kuyper, T. (2019). *Living Labs: Privacy & Ethics*. Retrieved from <https://medium.com/datadriveninvestor/living-labs-privacy-ethics-87bb2f9f9ff9>
- Lebovits, H. (2019). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. *Public Integrity*. <https://doi.org/10.1080/10999922.2018.1511671>
- Liu, M., McNamara, P., & McLoone, S. (2013). Fair charging strategies for EVs connected to a low-voltage distribution network. In *IEEE PES ISGT Europe 2013* (pp. 1–5). IEEE. <https://doi.org/10.1109/ISGTEurope.2013.6695268>
- London police's face recognition system gets it wrong 81% of the time. (2019). Retrieved from <https://www.technologyreview.com/f/613922/london-polices-face-recognition-system-gets-it-wrong-81-of-the-time>
- Lwakatara, L. E., Raj, A., Bosch, J., Olsson, H. H., & Crnkovic, I. (2019). A taxonomy of software engineering challenges for machine learning systems: An empirical investigation. In *Lecture Notes in Business Information Processing*. https://doi.org/10.1007/978-3-030-19034-7_14
- Malgieri, G., & Comandé, G. (2017). Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. *International Data Privacy Law*. <https://doi.org/10.1093/idpl/ixp019>
- Misra, T. (2018, February). The Rise of "Digital Poorhouses." *CityLab*. Retrieved from <https://www.citylab.com/equity/2018/02/the-rise-of-digital-poorhouses/552161/>
- Mittelstadt, B. (2019). AI Ethics – Too Principled to Fail? *SSRN Electronic Journal*, 1–15. <https://doi.org/10.2139/ssrn.3391293>
- Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 205395171667967. <https://doi.org/10.1177/2053951716679679>
- Moore, J., Kortuem, G., Smith, A., Chowdhury, N., Cavero, J., & Gooch, D. (2016). DevOps for the urban IoT. In *ACM International Conference Proceeding Series* (Vol. 24–25–May–). <https://doi.org/10.1145/2962735.2962747>
- Moreira, F. T., Magalhães, A., Ramos, F., & Vairinhos, M. (2018). The power of the internet of things in education: An overview of current status and potential. In *Smart Innovation, Systems and Technologies*. https://doi.org/10.1007/978-3-319-61322-2_6
- Moreira, F. T., Vairinhos, M., & Ramos, F. (2018). Internet of Things in education: A tool for science learning. In *Iberian Conference on Information Systems and Technologies, CISTI*. <https://doi.org/10.23919/CISTI.2018.8399234>
- Mulgan, G. (2016). A machine intelligence commission for the UK: how to grow informed public trust and maximise the positive impact of smart machines.
- Naik, G., & Bhide, S. S. (2014). Will the future of knowledge work automation transform personalized medicine? *Applied and Translational Genomics*. <https://doi.org/10.1016/j.atg.2014.05.003>
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report* | HHS.gov.
- ODI. (2018). *ODI Data Skills Framework*.
- Pasquale, F. (2015). *The Black Box Society*. <https://doi.org/10.4159/harvard.9780674736061>
- Preimesberger, C. (2017). Predictions 2018: Random Thoughts on AI, DevOps and Other Topics. *EWeek*.
- Ranchordás, S. (2019). Nudging citizens through technology in smart cities. *International Review of Law, Computers and Technology*, 0(0), 1–23. <https://doi.org/10.1080/13600869.2019.1590928>
- Richer, A. D. (2018, February 7). Boston police's social media surveillance unfairly targeted Muslims. *Boston Globe*. Retrieved from <https://www.bostonglobe.com/metro/2018/02/07/boston-police-social-media-surveillance-unfairly-targeted-muslims-aclu-says/9JUpzPmy8Tsr5RLxvCm61M/story.html>
- Rip, A., & Schot, J. W. (2002). Identifying loci for influencing the dynamics of technological development. In *Shaping technology, guiding policy: Concepts, spaces and tools* (pp. 155–172).
- Rohde, N. (2018). *Quality Criteria for Algorithmic Processes Analysing the Strengths and Weaknesses of Selected Compendia*. <https://doi.org/10.11586/2018028>
- Saward, M. (2010). *The Representative Claim*. <https://doi.org/10.1093/acprof:oso/9780199579389.001.0001>
- Schaer, P. (2017). Living Labs – An Ethical Challenge for Researchers and Platform Operators. *Internet Research Ethics for the Social Age: New Challenges, Cases, and Contexts*, (September), 167–176.
- Schmidt, A. (2016). Increasing Computer Literacy with the BBC micro:bit. *IEEE Pervasive Computing*. <https://doi.org/10.1109/MPRV.2016.23>
- Schuurman, D. (2015). Living labs: a systematic literature review. *ENoLL OpenLivingDays*. Retrieved from <http://scholar.google.be/>
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertes, J. (2019). Fairness and Abstraction in Sociotechnical Systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19* (Vol. 1, pp. 59–68). New York, New York, USA: ACM Press. <https://doi.org/10.1145/3287560.3287598>
- Sentence, S., Waite, J., Hodges, S., Macleod, E., & Yeomans, L. (2017). Creating cool stuff" – Pupils' experience of the BBC micro:bit. In *Proceedings of the Conference on Integrating Technology into Computer Science Education, ITICSE*. <https://doi.org/10.1145/3017680.3017749>
- Sidewalk Labs Master Innovation and Development Plan. (2019). Retrieved from <https://www.sidewalktoronto.ca>
- Steen, K., & Bueren, E. van. (2017). *Urban Living Labs: A living lab way of working*. *Urban Living Labs*. <https://doi.org/10.4324/9781315230641-2>
- Steen, K., & van Bueren, E. (2018). The Defining Characteristics of Urban Living Labs. *Technology Innovation Management Review*, 7(7), 21–33. <https://doi.org/10.22215/timreview/1088>
- Thaler Richard H., C. R. (2008). *Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven. *Yale University Press*.
- The Principled Artificial Intelligence Project. (2019). Berkman Klein for Internet and Society, Harvard Law

School. Retrieved from <https://clinic.cyber.harvard.edu/2019/06/07/introducing-the-principled-artificial-intelligence-project/>

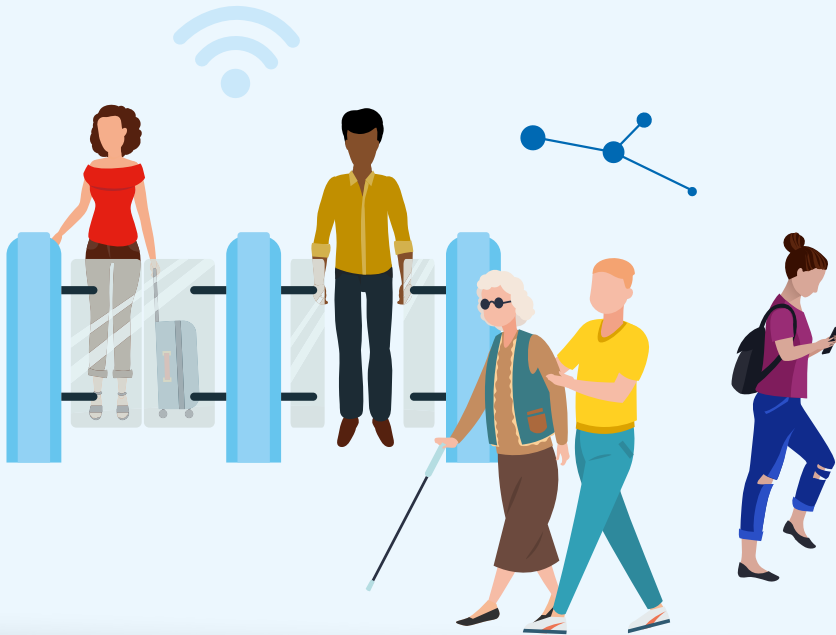
- Turel, T., Joskin, D., Geerts, F., Kaathoven, E. Van, & Schouwenaar, M. (2017). Designing a Transparent Smart Charge Point. *Evs30*.
- van de Poel, I. (2016). An Ethical Framework for Evaluating Experimental Technology. *Science and Engineering Ethics*, 22(3), 667–686. <https://doi.org/10.1007/s11948-015-9724-3>
- van der Voort, H. G., Klievink, A. J., Arnaboldi, M., & Meijer, A. J. (2019). Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly*, 36(1), 27–38. <https://doi.org/10.1016/j.giq.2018.10.011>
- Vehicle2Grid. (2016). Retrieved from <https://amsterdamsmartcity.com/projects/vehicle2grid>
- Wachter, S., Mittelstadt, B., Floridi, L., Watchter, S., Mittelstadt, B., Floridi, L., ... Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(4), 243–265. <https://doi.org/10.2139/ssrn.2903469>
- Westerland, M., & Leminen, S. (2018). Managing the Challenges of Becoming an Open Innovation Company: Experiences from Living Labs. *Technology Innovation Management Review*. <https://doi.org/10.22215/timreview489>
- Willson, M. (2017). Algorithms (and the) everyday. *Information Communication and Society*. <https://doi.org/10.1080/1369118X.2016.1200645>
- Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *Journal of Community Informatics*.
- Wolff, A., Kortuem, G., & Caverio, J. (2015). Urban Data in the primary classroom: bringing data literacy to the UK curriculum. In *Data Literacy Workshop*. Oxford. Retrieved from <http://oro.open.ac.uk/43855/>
- Woodson, T. (2018). Weapons of math destruction. *Journal of Responsible Innovation*. <https://doi.org/10.1080/23299460.2018.1495027>
- World Wide Web Foundation. (2017). *Algorithmic Accountability: Applying the concept to different country contexts*. Retrieved from http://webfoundation.org/docs/2017/07/WF_Algorithms.pdf

De noodzaak van een digitale omgevingsvisie

Van principes naar praktijk

Dr. Jiska Engelbert
Senior onderzoeker
LDE Centre for BOLD Cities

Prof. dr. Liesbet van Zoonen
Academisch directeur
LDE Centre for BOLD Cities



Voorwoord

Door Aantink Yeh

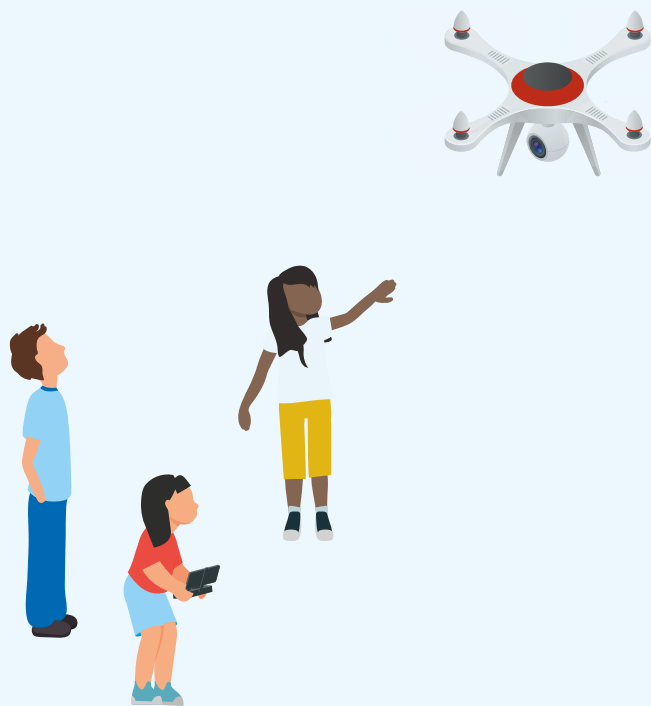
De inzet van nieuwe technologieën en het verzamelen en gebruik van data in de openbare ruimte gaat gepaard met soms lastige technische, sociale, juridische en ethische vraagstukken. Gemeenten (onder anderen) hebben de taak om te zorgen voor een toegankelijke en veilige fysieke leefomgeving. Zij moeten ook zorgen voor de toegankelijkheid, veiligheid en behoorlijk datagebruik in de digitale openbare ruimte.

Als adviseur Innovaties bij VNG Realisatie coördineer ik het traject om te komen tot een set aan gezamenlijke principes voor Nederlandse gemeenten, om hen richting te bieden en houvast te geven op dit onderwerp.

Tegelijkertijd zien we een scala aan soortgelijke initiatieven in binnen- en buitenland om publieke waarden te borgen in tijden van toenemende digitalisering. Het Centre for BOLD Cities, een onderzoeksgroep van Universiteit Leiden, TU Delft en Erasmus Universiteit Rotterdam, is gespecialiseerd in digitalisering en gebruik van data in de stad, in het bijzonder vanuit het ethisch perspectief en dat van de inwoners. Ik heb daarom dr. Jiska Engelbert en prof. dr. Liesbet van Zoonen gevraagd om onderzoek te doen naar (andere) initiatieven om publieke waarden te borgen en wat dit betekent voor de opgave van Nederlandse overheden.

De overleggen die we hebben gehad met de begeleidingscommissie om te komen tot deze essaybundel, hebben aangetoond hoe belangrijk het is dat wetenschap en beleid de dialoog blijven voeren. Maar met alleen theorie en overleg zijn we er niet. Dit essay (maar ook de andere in de bundel) gaat ons helpen om onze principes door te ontwikkelen en opvolging te geven door meer inzicht in het hoe en waarom van andere initiatieven en benodigde instrumenten en maatregelen. Principes zijn een lege huls als we er niks mee doen.

Aantink Yeh
Adviseur Smart City VNG Realisatie
Lid begeleidingscommissie



Samenvatting

In dit essay analyseren we het brede landschap aan principes voor de digitale samenleving dat zich in de afgelopen vijf jaar ontwikkeld heeft. We vinden ten eerste een onderscheid tussen principes die over ethiek, datakwaliteit en standaardisering gaan en constateren, ten tweede, dat de meeste vrijblijvend geformuleerd en vormgegeven zijn, in termen van mensenrechten waar niemand bezwaar tegen kan hebben. Een derde kenmerk van het landschap is dat de principes vrijwel nergens tot controleerbare en afdwingbare maatregelen leiden. Desalniettemin onderscheiden ze zich van een uitsluitend economisch gericht discours over standaardisering dat we ook aantreffen en waarin geen plaats is voor burgers en andere maatschappelijke belanghebbenden. De verzameling principes heeft, tenslotte, vooral betrekking op het proces van digitalisering en nauwelijks op specifieke, regionaal of lokaal geformuleerde einddoelen.

We laten aan de hand van twee voorbeelden (burgermetingen van geluidsoverlast rondom Schiphol en politiek protest tegen slimme lantaarnpalen op de Utrechtse Heuvelrug) zien dat de schijnbare harmonie waarin de ethische en kwaliteitsprincipes en de wens tot standaardisering circuleren, belangrijke maatschappelijke tegenstellingen verbloemt (bijvoorbeeld tussen bedrijfsleven, overheden en burgers) en de verantwoordelijke overheden geen duidelijk handvat geeft om een goede oplossing tussen strijdende partijen te bieden. Dat komt deels omdat burgers en andere betrokken partijen vaak pas achteraf met digitale oplossingen worden geconfronteerd, wanneer er geen andere inbreng dan accepteren of protesteren mogelijk is. Uit de principes blijkt echter de noodzaak om digitale oplossingen vanaf het begin in gezamenlijkheid te ontwerpen. Daarom werken we als laatste onderdeel van dit essay het idee uit om overheden te vragen naast een ruimtelijke omgevingsvisie ook een digitale omgevingsvisie te gaan ontwikkelen.



1. Inleiding

Gemeenten, provincies, ministeries, waterschappen en diverse andere organisaties vragen zich dringend af wat de principes voor het gebruik van digitale en data-technologieën in de samenleving en de openbare ruimte zouden moeten zijn. In toenemende mate concentreren die discussies zich op de automatische dataverwerking en –analyses die met artificiële intelligentie mogelijk worden gemaakt.¹ Ook kennisinstellingen en sociale ondernemingen hebben zich in dit debat gemengd. De gemeenten Amsterdam en Eindhoven formuleerden gezamenlijk waarden als inclusie, openheid, transparantie, privacy en data-eigendom; ze hebben veel weerklank gevonden², onder andere in de bondige TADA-principes.³ Wie dieper in de materie duikt, stuit op talloze andere pogingen om datagebruik en artificiële intelligentie ethisch te gronden en van maatschappelijke inbreng te voorzien. Zo hebben Universiteit Utrecht en de gemeente Utrecht samen een ‘ethical data assistant’ gelanceerd⁴, publiceerde het Rathenau Instituut suggesties tot het opwaarderen van publieke waarden in de digitale samenleving⁵, werkt GroenLinks aan een Handvest voor de Slimme Stad⁶, en schrijft de Vereniging Samenwerkende Nederlandse Universiteiten (VSNU) een vierdelige set van data-technische en maatschappelijke standaarden voor.⁷ Ook internationaal staan de digitale samenleving en verantwoord datagebruik hoog op de agenda van overheden. Met name de steden die zich als smart profileren, lopen voorop in de ontwikkeling van toolkits voor verantwoord datagebruik. Deze zijn vaak opgesteld door allianties van steden en regio’s, zoals de Cities Coalition for Digital Rights⁸ en, speciaal voor gemeentelijke overheden, de Sharing Cities Alliance⁹. De Britse nationale overheid laat zich leiden door zogenaamde GEMINI-principes die stellen dat datagebruik doelmatig, effectief en vertrouwenwekkend moet zijn¹⁰; voor de overheid van Dubai echter, is het belangrijkste principe om met data en digitale technologieën de burgers gelukkig te maken¹¹, terwijl de stad Barcelona juist wil dat burgers geen last hebben van digitalisering¹².

Er is duidelijk voor elk wat wils tussen de overdaad aan data-principes die er momenteel nationaal en internationaal in omloop zijn. Wat betekent dat voor de Nederlandse overheden die op zoek zijn naar richtlijnen? Moeten zij het wiel zelf uitvinden of bieden de bestaande principes behulpzame aanknopingspunten? Hoe kunnen bestaande principes ruimte laten voor specifiek provinciaal, regionaal of stedelijk beleid en cultuur?

En bij wie ligt de verantwoordelijkheid voor de uitvoering en naleving van die principes? Dat zijn de vragen die we in dit essay zullen beantwoorden.

We gaan eerst nog even verder met de analyse van het bestaande landschap van principes om uit te vinden of er misschien algemeen gedeelde waarden in te vinden zijn. We zullen zien dat de principes vooral proceswaarden beschrijven, en dat de eindwaarden die met digitale en data-technologieën bereikt zouden moeten worden, minder helder zijn. Dat Dubai zegt naar geluk van alle burgers te streven, is uitzonderlijk expliciet. Ook zullen we zien dat de principes weinig sturend zijn en zeker niet geoperationaliseerd worden naar specifieke praktijken of ambtelijke verantwoordelijkheden. We richten ons daarom vervolgens op de vraag hoe principes en richtlijnen tot een uitvoeringspraktijk kunnen leiden en welke governance daarbij hoort. We zullen uitkomen op de noodzaak voor de diverse overheden om, in analogie met de Omgevingswet, een digitale omgevingsvisie te ontwikkelen waarin verantwoord datagebruik een product wordt van de gezamenlijke belanghebbenden met hun specifieke wensen en doelen.

2. Een weids landschap aan principes

Om een beetje greep op alle vigerende principes te krijgen, moeten we eerst constateren dat ze van verschillende orde zijn: bij sommige gaat het om ethiek en maatschappelijke betekenis, bij andere vooral om betrouwbaarheid en kwaliteit van de datasystemen zelf, en weer andere zijn gericht op standaardisering. Die verschillen hebben alles te maken met de groepen en organisaties die ze propageren, zoals we zullen zien. We bespreken voor elke categorie de exemplarische voorbeelden.¹³

2.1 Ethische principes

Principes die de ethische en maatschappelijke betekenis van datagebruik proberen vorm te geven, zijn geformuleerd uit frustratie dat burgers de controle over hun eigen data kwijt zijn geraakt, en geen inzicht hebben in de manier waarop overheden, bedrijven en internationale platforms hun data gebruiken. De TADA-principes werden bijvoorbeeld ontwikkeld met het argument dat steden met behulp van data schoner, veiliger, gezonder en leefbaarder kunnen worden, maar “only as long as people maintain control over data, and not the other way around.”¹⁴



De Cities Coalition for Digital Rights, een gezamenlijk initiatief van Amsterdam, Barcelona en New York waar inmiddels zo'n 30 steden bij aangesloten zijn¹⁵, streeft naar "policies, tools and resources to promote and protect resident and visitor rights online."¹⁶ En het Handvest voor de Slimme Stad van GroenLinks stelt dat "burgers en politici weer greep moeten krijgen op de ontwikkeling van technologie."¹⁷ Ook de Sharing Cities Alliance legt de nadruk op politiek en overheid als actoren, die, net als burgers, de controle kwijt zijn geraakt en streeft daarom naar de empowerment van gemeentelijke overheden, in het bijzonder tegenover de multinationale platforms.¹⁸

De gedeelde observatie is dat burgers, hun vertegenwoordigers en hun overheden digitaal kwetsbaar zijn geworden en dat een set aan goed geformuleerde principes nodig is om die kwetsbaarheid te mitigeren en om te draaien. De aanvoerders van deze beweging zijn in eerste instantie progressieve gemeenten en hun besturen geweest, in het bijzonder Amsterdam en Barcelona. Die voorlopersrol is niet toevallig; zij beschikken over het benodigde electorale mandaat en een ambtelijk apparaat dat de principes kan uitvoeren. Barcelona heeft bijvoorbeeld per beleidsterrein geformuleerd hoe participatie, beleidsvorming, aanbesteding en vergunning volgens de stedelijke digitale principes uitgevoerd moeten worden.¹⁹

De geformuleerde principes laten echter een veel algemenere verzameling aan publieke waarden zien dan de oorsprong in progressief beleid suggereert. De principes in het Handvest voor de Slimme Stad van GroenLinks bijvoorbeeld (democratie, verbondenheid, menselijke waardigheid, privacy, duurzaamheid en gelijkheid) zijn direct te verbinden aan de breed gedragen Duurzame Ontwikkelingsdoelen van de Verenigde Naties. Ook de TADA-principes getuigen eerder van een algemeen mensenrechtendiscours dan van een scherpe progressieve ideologie. Het Rathenau Instituut sprak in 2017 in dit verband al over 'digitale mensenrechten'.²⁰ Die relatieve neutraliteit van de zich aldus ontwikkelende ethiek maakt het mogelijk dat een brede verzameling aan actoren zich eraan committeert: van overheidsinstanties en maatschappelijke groeperingen tot sociale ondernemingen en lokaal en nationaal bedrijfsleven. Ook staat die neutraliteit uiteenlopende, lokale keuzes toe: de gemeente Barcelona maakt het Uber onmogelijk in de stad te opereren²¹ en de gemeente Amsterdam gebruikt Google om de stedelijke luchtkwaliteit te meten²².

Die mobiliserende kracht wordt bovendien versterkt door de vorm van manifest of handvest waarin deze ethiek zich manifesteert: de ondertekening van dergelijke initiatieven is open en gratis voor zowel individuen als organisaties. Men sluit zich daarmee aan bij een gemeenschap van gelijkgestemden die met elkaar een ethische intentie formuleren die een gezamenlijke verantwoordelijkheid is en het directe eigenbelang overstijgt.

In de mobiliserende kracht van de relatief neutrale digitale ethiek die op deze manier allerwege vorm begint te krijgen, schuilt ook nogal wat vrijblijvendheid: wie spreekt elkaar aan op praktijken die de principes overschrijden? Mogen alle mogelijke partijen zich bij een manifest of handvest aansluiten? En wie bepaalt dat dan? Wat als er conflicten tussen ondertekenende partijen ontstaan? Daar waar een overheid zelf digitale principes heeft geformuleerd of overgenomen (bijvoorbeeld in de provincie Zuid-Holland²³ lijkt deze de voor de hand liggende hoeder van de digitale principes te zijn, ware het niet dat de operationalisatie van de principes in concrete activiteiten en maatregelen achterblijft bij het gesprek erover. Maar daar komen we later op terug.

2.2 Kwaliteitsprincipes

Een iets andere verzameling aan principes waarover een intensief, maar minder hoorbaar gesprek gaande is, gaat over de kwaliteit en betrouwbaarheid van data en de bijbehorende systemen. Immers, het ethisch gebruiken en toepassen van data in welke grote of kleine vorm dan ook, is onmogelijk als die data om te beginnen niet goed verzameld, geïndexeerd, bewaard, geschoond, geanalyseerd, gepresenteerd en toegepast worden. Het gaat hier in feite om de beginselen van de datatechniek en de correcte en betrouwbare uitvoering ervan.

In Nederland zijn het Centraal Bureau voor de Statistiek (CBS) en de kwantitatieve sociale en gedragswetenschappen van oudsher belangrijke actoren in het voorleven en handhaven van kwaliteitsprincipes die klassiek geformuleerd zijn in termen van betrouwbaarheid en validiteit. Met de data-explosie van de laatste decennia hebben zich ook andere spelers in dit veld gemengd, onder de hoede van de VSNU. Haar in 2018 gepubliceerde Onderzoeksagenda voor de Digitale Samenleving²⁴ hanteert de onder datawetenschappers gangbare concepten van FAIR data, ROBUST-systemen en FACT-algoritmen, waarbij elke afkorting specifieke elementen behelst. In tegenstelling tot de ethische principes, gaat het

hier veel meer om toetsbare criteria die zelfs tot een kwaliteitscijfer voor de datapraktijk van uiteenlopende overheidsinstanties, organisaties en bedrijven kunnen leiden.

Het principe van FAIR data houdt in dat data Findable, Accessible, Interoperable en Re-usable moeten zijn. Deze vier criteria zijn geoperationiseerd in vijftien specifieke voorschriften die onder meer betrekking hebben op de toewijzing van metadata en protocollen.²⁵ FACT is een afkorting die Nederlandse data-wetenschappers gebruiken in hun streven om algoritmen te ontwikkelen die nauwkeurig, transparent en zonder vooroordelen werken, met respect voor ieders privacy; het gaat er met andere woorden om dat ze Fair, Accurate, Confidential en Transparent zijn.²⁶ ROBUST-systemen, tenslotte, vereisen een veilig en betrouwbaar ontwerp dat inspectie van buiten mogelijk maakt, bestand is tegen systeemschade, en altijd een menselijk oog omvat: Resilient, Open, Beneficial, User-Oriented, Secure en Trustworthy. Steeds meer wetenschappers omarmen deze criteria en passen ze toe. Maar dat geldt veel minder voor de toegepaste domeinen van de overheid en het bedrijfsleven. Daar is het de vraag of deze criteria überhaupt bekend zijn, of ze als richtsnoer functioneren en of hun datapraktijken aan al die normen (kunnen) voldoen.

Een van de auteurs van dit essay publiceerde onlangs een gedetailleerde analyse van drie typen dataprojecten die populair zijn bij het gemeentelijke sociale domein: data warehouses, dashboards en predictive analytics.²⁷



Zij constateerde dat op het niveau van de data zelf de valideit en betrouwbaarheid niet gegarandeerd kan worden, dat daardoor de gebruikte voorspellende modellen noch eerlijk, noch nauwkeurig zijn en dat de mensen over wiens data het gaat, zelden geïnformeerd of betrokken worden. Haar conclusie luidt niet alleen dat gemeenten hiermee waarschijnlijk de nieuwe privacywet overtreden, maar ook dat het gebrek aan kwaliteit van de gegevens en de analytische modellen tot fouten en stigmatisering dreigen te leiden.²⁸

De kwaliteitsprincipes van de goede datawetenschap zijn veel minder prominent in het debat over verantwoord datagebruik aanwezig dan de ethische principes, al komen de twee samen in de angst voor autonome algoritmen waarvan niemand begrijpt wat ze doen. Volgens de ethische principes moeten burger, overheid en politiek de controle daarop terugwinnen; volgens de kwaliteitsprincipes zouden ze om te beginnen niet autonoom moeten kunnen opereren en eveneens maximaal transparant moeten zijn. Het is echter duidelijk dat het voor gemeenten en andere overheden veel eenvoudiger is om de algemene ethiek van verantwoord datagebruik te omarmen, dan om te zorgen dat de datasystemen robuust zijn, de algoritmen eerlijk en feitelijk, en de gegevens vindbaar en uitwisselbaar. Die drie kwaliteitsprincipes vereisen een serie organisatorische maatregelen en financiële impulsen die voor de meeste gemeenten in Nederland (en daarbuiten overigens ook) zeer hoog gegrepen zijn.²⁹

2.3 Standaardisering

Een derde set aan principes waaraan diverse actoren werken, betreft standaardisering van de data-infrastructuren. Het gaat er daarbij vooral om uitvoering te geven aan onderdelen van het FAIR-principe, namelijk *interoperability* en *re-usability*. De uitwisseling van gegevens tussen overheden, hun diverse diensten en eventuele maatschappelijke gebruikers wordt momenteel zeer bemoeilijkt door de enorme variabiliteit in methodieken van verzameling, opslag en toegankelijkheid. Er zijn inmiddels dan ook talloze initiatieven tot standaardisatie ondernomen. Het Europees onderzoeksproject Espresso inventariseerde in 2016 al 88 organisaties uit 23 verschillende landen die zich op de één of andere manier met standaardisatie van en in de *smart city* bezig hielden, en verzuchtte in zijn conclusies dat het onmogelijk is om overzicht te krijgen.³⁰

In Nederland is NEN (Nederlandse Normalisatie Instituut), in algemene zin bezig met de ontwikkeling van standaarden, onlangs een gesprek begonnen met gemeenten, kennisinstellingen en bedrijven over standaarden voor de smart city, met als reden dat alleen met standaardisatie “allerlei partijen applicaties [kunnen] ontwikkelen waarmee ze, verrijkt met data uit andere bronnen, betere en duurzamere diensten kunnen leveren.”³¹ De eerste aandacht richt zich op de standaardisering van open *urban data platforms* omdat daar, volgens NEN, grote kansen liggen voor inwoners en bedrijven in steden om met elkaar nieuwe oplossingen voor de stedelijke problemen te vinden.³² NEN koppelt de noodzaak tot standaardisatie van de infrastructuur dus direct aan de ethiek van openheid en gezamenlijkheid die we eerder in de diverse manifesten en handvesten zagen. Die koppeling tussen ethiek en standaardisering is ook te vinden in de uitkomsten van het Espresso-project over standaardisering. De onderzoeksgroep vraagt zich onder meer af of belangrijke ethische vraagstukken, in het bijzonder die over privacy, wel met standaardisering kunnen worden opgelost.

De uitwisseling van gegevens tussen overheden, hun diverse diensten en eventuele maatschappelijke gebruikers wordt momenteel zeer bemoeilijkt door de enorme variabiliteit in methodieken van verzameling, opslag en toegankelijkheid

Dit is vooral opvallend als we de argumentatie van NEN en Espresso vergelijken met de manier waarop de International Organization for Standardization (ISO) de noodzaak voor standaarden voor de smart city op haar website introduceert. Standaarden, zo lezen we daar, zijn de “*holy grail of an interoperable, plug-and-play world where cities can mix and match solutions from different vendors without fear of lock-in or obsolescence or dead-end initiatives*”³³ In de verdere uitleg gaat het om de optimale benutting van datamogelijkheden en verbetering van de dienstverlening aan burgers.



Eenzelfde marktoriëntatie is zichtbaar in de uitleg van het British Standards Institute bij zijn smartcity-principes: het gaat er allereerst om dat standaardisering de beste oplossingen biedt voor de commerciële en technische belangen van *smart city business* en dan om de hulp aan de markt om innovatie te versnellen.³⁴

Al voert het te ver om zonder nadere analyse al van een Nederlands-Europese tegenover een Brits-internationale opvatting over standaardisering te spreken, het contrast tussen de redenering van NEN-Espresso enerzijds en BSI-ISO anderzijds past in een algemenere discussie over de wenselijkheid van een Europees publiek model van de digitale samenleving tegenover de *corporate* invulling door multi-nationale grootmachten als Apple, Alfabet, Amazon, Google en Microsoft en een totalitaire staatsinvulling zoals in China vorm krijgt.³⁵ De verschillende argumentaties laten ook zien dat standaardisering geen neutrale technische afstemmingsexercitie betreft, maar ook altijd belangrijke ethische keuzes in zich draagt.

2.4 Eerste conclusies

Onze verkenning van het landschap aan principes leidt tot een aantal voorlopige en partiële conclusies. Ten eerste lijken de principes waar diverse overheids- en maatschappelijke actoren zich mee verbonden hebben vrijblijvend geformuleerd en vormgegeven.



Vrijwel nergens ontstijgen ze het niveau van algemene mensenrechten waar niemand tegen kan zijn en vrijwel nergens leiden ze tot controleerbare en afdwingbare maatregelen. Dat ze desalniettemin bijzonder zijn, blijkt pas op het moment dat ze geen onderdeel zijn van het discours zoals bij de ISO- en BSI-visies op standaardisering. Ten tweede valt op dat de verzameling aan principes vooral betrekking heeft op het proces van digitalisering: dat moet 'goed' in tweeërlei opzicht zijn: ethisch verantwoord doordat iedereen mee kan doen en datatechnisch van betrouwbare kwaliteit. Voor zover er einddoelen geformuleerd worden (waar moet dit allemaal dan toe leiden?), gaat het om abstracte waarden als leefbaarheid en efficiëntie. Wat betekent deze relatief neutrale procesethiek voor de diverse overheden die hun datapraktijk volgens ethische en kwaliteitsprincipes willen inrichten?

3. De opgave voor de verschillende overheden

De vraag naar de betekenis van de principes voor de datapraktijken van Nederlandse overheden kunnen we het beste verkennen met een paar concrete voorbeelden die laten zien waar de problemen zitten.

Kijken we om te beginnen naar de zogeheten burgermeet-beweging: met allerlei sensoren of telefoon-apps meten burgers zelf de overlast in hun omgeving. Het Rijksinstituut voor Volksgezondheid en Milieu is een belangrijke speler in deze beweging en heeft haar eigen metingen uitgebreid met een grote hoeveelheid metingen door burgers die met goedkopere sensoren werken. Omwonenden van Schiphol meten met de app Explane hoeveel decibel een passerend vliegtuig afscheidt omdat de officiële metingen niet aansluiten bij hun eigen geluidservaring.³⁶ Net zoiets gebeurt inmiddels met de milieuhinder die wordt veroorzaakt door Tata Steel³⁷ en het fijnstof in Rotterdamse woonwijken³⁸. Al deze burgerparticipatie in dataverzameling sluit perfect aan bij de ethische principes die we eerder bespraken: mensen zeggenschap geven over data en samen data ontwikkelen. Met het proces is dus niets mis, maar als we de principes van datakwaliteit erbij nemen, stuiten we op een probleem. De systemen waarmee die data verzameld worden, wijken nogal van elkaar af, de burgermetingen zijn niet officieel gevalideerd en of de initiatieven herhaald zullen worden is nog maar de vraag. Van standaardisering van burgermetingen kan geen sprake zijn.

Dat zou er op neerkomen dat je alleen zelf mag meten als je dat volgens de standaarden van anderen doet. Het resultaat van al deze burgerparticipatie zou ook zomaar kunnen zijn. Het voorspelbare resultaat zal zijn dat bij de discussie over de uitbreiding van Schiphol, de vervuiling door Tata Steel of de verkeerscirculatie in Rotterdam meerdere datasets tegenover elkaar komen te staan: die van de officiële instanties en die van de burgers. Welke keuze moet de betreffende lokale overheid dan maken? Een vleugje van die confrontatie zien we al in het commentaar van de officiële meters van het Schipholgeluid op de burgermetingen met Explane: de app maakt geen onderscheid tussen vlieggeluid en achtergrondgeluid, de verschillende telefoons zijn qua meting niet vergelijkbaar en het maakt nogal uit of je in de stad of het platteland meet.³⁹

Een tweede voorbeeld komt van de gemeente Utrechtse Heuvelrug. Daar moest in 2018 de openbare verlichting vernieuwd worden, en de gemeente wilde de gelegenheid te baat nemen om te experimenteren met lantaarnpalen waaraan allerlei nieuwe slimme toepassingen opgehangen konden worden zoals camera's, 5G-technologie, sensoren, laadpalen en lichtscenario's. De gemeente had er ook een businessmodel bij bedacht: innovatieve ondernemers konden hun eigen slimme plek op de lantaarnpaal kopen. Er werd toen al, onder andere in Hengelo, Eindhoven en op Texel, met zulke slimme lantaarnpalen geëxperimenteerd.⁴⁰ In 2017 passeerde een beleidsnota met slimme lantaarnpalen zonder ophef de Amsterdamse gemeenteraad.⁴¹ Maar in Utrechtse Heuvelrug stond een groep burgers op die zich ernstig zorgen maakte over de slimme lantaarnpaal en dan met name over het stralingsgevaar dat mogelijke 5G-zenders op de palen zouden opleveren. Onder het motto 'Stralen doen we liever zelf' organiseerden ze gesprekken met raadsleden, een campagne op sociale media en een aantal buurtbijeenkomsten. De aldus ontstane maatschappelijke onrust deed de gemeenteraad besluiten om voorlopig van slimme lantaarnpalen af te zien en alleen de verlichting te vernieuwen.⁴² Het is een prachtig voorbeeld van burgerparticipatie, alleen met een uitkomst die lijnrecht tegen het vooruitgangsdiscours van de smart city ingaat. Dat suggereert namelijk dat we gezamenlijk, met zeggenschap voor iedereen, met de juiste kwaliteitsnormen en standaarden tot goede digitale of data-oplossingen kunnen komen voor de stedelijke problematieken. Er is niet voorzien in de mogelijkheid dat er groepen zijn die dergelijke oplossingen überhaupt niet willen, zoals in Utrechtse Heuvelrug.

De twee voorbeelden laten zien dat de schijnbare harmonie waarin de ethische en kwaliteitsprincipes en de wens tot standaardisering circuleren, belangrijke maatschappelijke tegenstellingen verbloemt (bijvoorbeeld tussen bedrijfsleven, overheden en burgers) en de verantwoordelijke overheden geen duidelijk handvat geeft om een goede oplossing tussen strijdende partijen te bieden. Wiens data in de afwegingen over de uitbreiding van Schiphol of de lantaarnpalen in Utrechtse Heuvelrug het zwaarst zullen wegen, is een kwestie van economische macht en van lokale politieke verhoudingen, zelfs als alle strijdende datasets gestandaardiseerd zouden zijn en aan de gewenste ethische en kwaliteitsprincipes zouden voldoen.⁴³ De voorbeelden laten ook zien dat voor een overheid die beslissingen over digitalisering moet nemen, de principes vooral noodzakelijke voorwaarden zijn ('dit moet in elk geval geregeld zijn'), maar geen voorwaarden opleveren die voldoende zijn om keuzes op te baseren. Dat komt omdat ze proceswaarden behelzen en de eindwaarden voorstellen als operationele kwesties: stedelijke problemen worden efficiënter opgelost, markten worden innovatiever en overheids- en andere organisaties kunnen sneller en persoonlijker dienstverleners.

De schijnbare harmonie waarin ethische en kwaliteitsprincipes en de wens tot standaardisering circuleren, verbloemt maatschappelijke tegenstellingen

Technologiecriticus Evgeny Morozov heeft dat soort denken als 'solutionisme' bestempeld; de overtuiging dat elk probleem een technologische oplossing heeft.⁴⁴ Noch de technologie, noch de nagestreefde oplossingen zijn echter neutraal en voor iedereen even gunstig,⁴⁵ zoals de voorbeelden laten zien. Daarom zou het huidige landschap aan principes ook veel meer ruimte moeten bieden aan richtingaanwijzers naar scherpe, inhoudelijke eindwaarden. In grote lijnen vallen die momenteel, zoals gezegd, uiteen in de waarden van de multinationale markteconomie (VS) enerzijds en de waarden van de centraal gestuurde staat (China) anderzijds. In het ene model is de burger consument en de overheid op afstand of dienend; in het andere is de burger onderdaan en de overheid sturend

en controlerend.⁴⁶ In geen van beide gevallen hebben burgers of hun vertegenwoordigers enige zeggenschap over hoe (proceswaarden) en in welke richting (eindwaarden) hun digitale samenleving zich beweegt. Wat dat betreft zijn we in de Nederlandse en Europese steden al een stuk verder, zoals onze analyse van het landschap van principes al liet zien. Daaruit bleek dat er redelijke consensus is over hoe we de digitale samenleving moeten doorontwikkelen. Over de richting en de eindwaarden wordt, los van het operationele discours, nog geen uitgebreid debat gevoerd. Hoe kunnen we die eindwaarden toch in gesprek brengen en welke rol speelt de overheid daarin?

4. Digitale omgevingsvisie

In Nederland hebben de diverse overheidsapparaten geen zelfstandige rol in het definiëren van eindwaarden voor de samenleving. Binnen de grenzen van de wet en mensenrechtenverdragen is oneindig veel ruimte voor burgers en bewegingen om die eindwaarden te verkennen, te formuleren, te promoten en tot beleid te laten maken. Dat gebeurt simpelweg door verkiezingen, maar ook door tussentijdse inspraak en in toenemende mate door burgers direct te betrekken bij beleidsvoorbereiding en -uitvoering. We denken bijvoorbeeld aan *living labs*, proeftuinen, leer-ateliers, wijkondernemingen, enzovoort. Burgers mogen ook zelf overheidstaken overnemen als ze denken het zelf beter te kunnen, onder het motto *'The right to challenge (R2C)'*. De Omgevingswet die in 2021 ingaat, verplicht overheden zelfs om bij de ruimtelijke ontwikkelingen "rekening te houden met de verschillende belangen in een gebied. Nu beslissen de overheden vaak alleen over een deelproject."⁴⁷

Die laatste zin, dat overheden vaak alleen beslissen over projecten, is interessant als we hem tegen de besproken ethische principes aanhouden. De proceswaarden die daarin gepropageerd worden, demonstreren eenzelfde behoefte om in gezamenlijkheid de samenleving vorm te geven en dat niet over te laten aan de alleenmacht van ofwel overheden ofwel grote bedrijven of een congres van de twee. Je zou kunnen zeggen dat uit de ethische principes een behoefte aan een digitale omgevingsvisie spreekt. Het is ook niet zo vreemd om inspiratie bij het ruimtelijk beleid te vinden voor een verdere uitwerking van principes voor de digitale samenleving.



We hebben er immers lange tijd in ruimtelijke metaforen over gesproken: de elektronische snelweg, de digitale stad, proeftuin of data-pakhuis. “Als we iedereen gezamenlijk laten meedenken en meepraten over de inrichting van de fysieke ruimte, waarom doen we dat dan niet ook voor het ontwerp van de data- en digitale ruimte; voor de datapakhuisen, dashboards, analytics en algoritmen?”⁴⁸

Je kunt de kunst regelrecht afkijken van al het werk dat in de Omgevingswet is gestoken. Het Ministerie van BZK heeft immers al een uitgebreide verzameling hulpmiddelen in de aanbieding om een omgevingsvisie te maken en precies aangegeven welke stappen in het proces genomen kunnen worden. Dat begint – onvermijdelijk – met een analyse van bestaand beleid en de samenstelling van een brede projectgroep. Maar daarna moet er toch een hele serie externe partners betrokken worden die gezamenlijk de eindambitie formuleren en ‘het verhaal’ maken.⁴⁹ Volgens de logica van de omgevingsvisie is het dus niet genoeg als de proceswaarden geborgd zijn. Als iedereen heeft meegepraat, en de bouw en ruimtelijke inrichting (datastructuur) van technisch goede kwaliteit zijn, moet het eindresultaat een gedeelde ambitie en verhaal zijn. Dat kan per definitie geen keihard economisch doel voor de digitale samenleving zijn, zoals we in de BSI- en ISO-standaarden herkennen.



Het kan ook geen top-down sturingsmodel van de overheid impliceren, zoals het schrikbeeld van China ons voorschotelt⁵⁰; niet voor niets raadt het ministerie van BZK de gemeentelijke overheden aan om bij de ontwikkeling van hun omgevingsvisie over hun eigen rol na te denken, die kan variëren tussen regulerend, samenwerkend en faciliterend.⁵¹

Leggen de proceswaarden de nadruk op gezamenlijkheid, dan is het logisch dat de eindwaarden ook de nadruk leggen op een gedeeld verhaal; een *shared story*, zoals de VSNU propageert in haar Onderzoeksagenda voor de Digitale Samenleving⁵². Daarin gelden, naast de eerder al genoemde kwaliteitsprincipes, SHARED-waarden als beoordelingscriteria voor de digitale samenleving.⁵³ Ze staan voor de volgende (Engelstalig geformuleerde) eindwaarden:

- **Sustainable**, dat wil zeggen dat de digitale visie en beleid niet alleen moeten passen bij de ecologische omgevingsagenda maar ook implementeerbaar en langdurig bruikbaar moeten zijn. De samenwerking met alle betrokken partijen kan derhalve niet eenmalig zijn maar moet steeds opnieuw herhaald worden.
- **Harmonious**, hetgeen inhoudt dat de digitale visie en het beleid open en inclusief moet zijn, wettelijke, morele kaders en omgangsvormen respecteren en bestaande tegenstellingen niet verder verscherpen.
- **Affective**, omdat een digitale visie en beleid ook moeten erkennen en meewegen dat technologie niet alleen maar met puur rationele argumenten te maken heeft, maar ook een scala aan positieve en negatieve gevoelens bij mensen en groepen kan oproepen.
- **Relevant**, waarmee bedoeld wordt dat een digitale visie en beleid juist die groepen en belangen moet betrekken die het meest te maken zullen hebben met de effecten van data en digitale technologieën. In het sociaal domein betekent dat bijvoorbeeld dat uitkeringsgerechtigden veel meer dan nu het geval is mee moeten praten over de inrichting van de datatransities in het sociaal domein.⁵⁴
- **Empowering**, oftewel de digitale visie en beleid moeten het ook mogelijk maken dat alle betrokkenen de betreffende technologie kunnen begrijpen en beoordelen, en daar waar mogelijk ook zelf moeten kunnen gebruiken.
- **Diverse**, tenslotte, is de waarde waarmee niet alleen de diversiteit in de samenleving (h)erkend wordt maar ook te beoordelen valt of de betreffende technologie zelf op diverse manieren inzetbaar en veranderbaar is.

De SHARED-waarden zijn breed genoeg geformuleerd om een scala aan operationalisaties mogelijk te maken én ze laten de noodzakelijke blijvende dynamiek in digitale visies en beleid toe. Technologie en samenleving veranderen momenteel immers zo snel dat voortdurende aanpassing in het ontwerp van digitale en datatechnieken ingebouwd moet zijn. In innovatiekringen spreekt men daarom graag van ‘permanent bèta’⁵⁵, onder kritische technologie-onderzoekers gebruikt men liever de term ‘contestable by design’⁵⁶.

Van de SHARED-waarden moeten we echter niet denken dat ze tegenstellingen en conflict zullen voorkomen. Al wordt de samenwerking tussen overheden, bedrijven, kennisinstellingen en burgers (of hun vertegenwoordigers) in de zogeheten ‘quadruple helix’ als noodzakelijk gezien voor succesvolle, gedragen innovatie, er ontstaan onvermijdelijk tegenstellingen, vertragingen, ruzies, ergernissen, frustraties en mislukkingen. De overheid, op welk niveau dan ook, heeft daar een dubbele rol in; enerzijds is ze partner in de samenwerking, anderzijds heeft ze, wat ons betreft, als enige in deze complexe uitwisselingen een overkoepelende rol als hoeder van de gemeenschappelijke proces- en eindwaarden. We kunnen noch van burgers, noch van bedrijven verwachten dat ze altijd rekening houden met ieders ideeën en belangen. Het is ook niet de taak van kennisinstellingen, maar de overheid, gelegitimeerd door haar eigen burgers, kan en moet die verantwoordelijkheid wel nemen.

We kunnen noch van burgers, noch van bedrijven verwachten dat ze altijd rekening houden met ieders ideeën en belangen

Toegegeven, dat is makkelijker gezegd dan gedaan omdat er binnen de verschillende overheden ook uiteenlopende, tegengestelde en conflicterende opvattingen zijn, zeker wat betreft digitale en datavisies. Tussen ministeries en rijksinstellingen verschillen soms de opvattingen over wat het doel van de inzet van digitale en datatechnologieën zou moeten zijn⁵⁷.

Met de onder de AVG afgedwongen aanstellingen van gemeentelijke privacy-officers zijn ook tal van spanningen binnengehaald tussen afdelingen die meer met digitalisering en data willen doen en de manier waarop de betreffende privacy-bewakers de nieuwe wet interpreteren.⁵⁸

Hoe moet de wethouder van Utrechtse Heuvelrug nu verder met zijn verlichting⁵⁹ en hoe moet de overheid in algemene zin verder met de principes voor de digitale samenleving? De proceswaarden voor het ontwerp van digitale en data-technologieën zijn inmiddels goed uitgedacht, zoals bleek in het eerste deel van dit essay. Het is ook duidelijk dat de overheid zich moet buigen over de wenselijke eindwaarden van de digitale samenleving, zoals we in deel twee betoogden, én we hebben in het laatste deel een universitair gedragen aanbeveling voor die eindwaarden gegeven en gesuggereerd dat een digitale omgevingsvisie het middel is om data en digitale technologieën ‘goed’ (in ethisch en technisch opzicht) en in gezamenlijkheid te ontwerpen. Daarmee houdt onze rol voor nu even op, maar laten we afsluiten met de gedachte dat Nederland, zijn provincies, waterschappen en gemeenten met zulke digitale omgevingsvisies een uniek voorbeeld kunnen zijn voor een digitale samenleving die noch door de staat, noch door de grote platforms gestuurd wordt, maar van en voor ons allemaal is.

Over de auteurs



Dr. Jiska Engelbert is senior onderzoeker bij het LDE (Leiden-Delft-Erasmus) Centre for BOLD Cities.



Prof. dr. Liesbet van Zoonen is academisch directeur bij het LDE Centre for BOLD Cities.



Referenties

- 1 <https://www.rathenau.nl/nl/digitale-samenleving/overzicht-van-ethische-codes-en-principes-voor-ai>
- 2 <https://innovationorigins.com/nl/eindhovense-smart-society-principes-gaan-landelijk/>
- 3 <https://tada.city/>
- 4 <https://dataschool.nl/deda/>
- 5 <https://www.rathenau.nl/nl/digitale-samenleving/opwaarderen>
- 6 <https://wetenschappelijkbureaugroenlinks.nl/artikel/handvest-voor-de-slimme-stad>
- 7 <https://www.thedigitalsociety.info/about/data-principles/>
- 8 <https://citiesfordigitalrights.org/>
- 9 <https://sharingcitiesalliance.com/>
- 10 <https://www.cdbb.cam.ac.uk/system/files/documents/TheGeminiPrinciples.pdf>
- 11 <https://www.smartdubai.ae/>
- 12 <https://www.barcelona.cat/digitalstandards/en/digital-services/o.1/>
- 13 Dit is gebaseerd op een nog niet afgeronde systematische internationale verkenning en identificatie van principes die in een latere wetenschappelijke publicatie gerapporteerd zal worden.
- 14 <https://amsterdamsmartcity.com/projects/tada-data-disclosed>
- 15 <https://citiesfordigitalrights.org/cities>, dd. 21-8-2019
- 16 <https://oascities.org/amsterdam-barcelona-and-new-york-city-launch-global-coalition-to-protect-digital-rights/>
- 17 <https://wetenschappelijkbureaugroenlinks.nl/artikel/handvest-voor-de-slimme-stad>
- 18 <https://sharingcitiesalliance.com/about-us>
- 19 <https://www.barcelona.cat/digitalstandards/en/init/o.1/index.html>
- 20 Van Est, R. & J. Gerritsen (2017). Human Rights in the Robot Age. Rathenau Instituut, with the assistance of Linda Kool. <https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf>
- 21 <https://www.bbc.com/news/business-47071710>
- 22 <https://www.google.com/earth/outreach/special-projects/air-quality/>
- 23 <https://www.zuid-holland.nl/@23133/visie-digitale/>
- 24 <https://www.thedigitalsociety.info/wp-content/uploads/2018/04/VSNU-Digital-Society-Research-Agenda.pdf>
- 25 <https://www.force11.org/group/fairgroup/fairprinciples>
- 26 <https://redasci.org/>
- 27 Van Zoonen, L. (2019). Opnieuw fatale remedies: een kritische reflectie op datatransities in het sociaal domein. Sociologie, 15(1), p. 19-45.
- 28 Idem, p.32-33
- 29 VNG/Berenschot (2018) Datagedreven sturing bij gemeenten. https://vng.nl/files/vng/_nieuws_attachments/2018/datagedreven_sturing_bij_gemeenten_lr.pdf
- 30 http://espresso-project.eu/content/deliverables/:deliverable_d7.4, p.11
- 31 <https://www.nen.nl/Normontwikkeling/Doe-mee/Normcommissies-en-nieuwe-trajecten/NEN-smart-cities.htm>
- 32 idem
- 33 <https://www.iso.org/sites/worldsmartcity/>
- 34 <https://www.bsigroup.com/en-GB/smart-cities/The-Cities-Standards-Institution/>
- 35 Van Dijk, J., Poel, T. & M. de Waal (2016). De platformsamenleving. Strijd om publieke waarden in een online wereld. Amsterdam: Amsterdam University Press.
- 36 Stil, H. (2019). Omwonenden Schiphol meten geluidshinder zelf. Het Parool, 30 juni. <https://www.parool.nl/amsterdam/omwonenden-schiphol-meten-geluidshinder-zelf-b66f51c8/?referer=https%3A%2F%2Fwww.google.com%2F>
- 37 idem
- 38 <https://www.rijnmond.nl/nieuws/183759/Onderzoek-naar-stikstof-Mensen-zijn-benieuwd-wat-er-gebeurt-in-hun-tuin>
- 39 Idem, noot 34.
- 40 Van Dijk, J. (2018). Adviesrapport slimme lantaarnpaal in Enschede. Bacheloropdracht Civiele Techniek, Universiteit Enschede. <https://essay.utwente.nl/75944/1/Dijk-Jarka-van.pdf>
- 41 Van Zoonen, L. (2017). Linke Lantaarnpaal. Sociologie Magazine, maart, p. 9.
- 42 <https://www.earth-matters.nl/11/14368/verborgen-nieuws/utrechtse-heuvelrug-tegen-smart-lantaarns.html>
- 43 Zie ook Van Zoonen, L. (2014). Data delirium. Sociologie Magazine, september, p. 10.
- 44 Morozov, E. (2012). The net delusion. How not to liberate the world. London: Penguin Books.
- 45 <https://www.open-overheid.nl/interview/marleen-stikker/>
- 46 Noot 33, idem.
- 47 <https://www.rijksoverheid.nl/onderwerpen/omgevingswet/vernieuwing-omgevingsrecht>
- 48 Van Zoonen, L. (2019). Datadrang. In: Data in de stad, Platform 31, in druk.
- 49 <https://aandeslagmetdeomgevingswet.nl/wetsinstrumenten/gemeente/omgevingsvisie/starten/>
- 50 Diverse critici stellen overigens dat de manier waarop de nationale en gemeentelijke overheden in Nederland de datamogelijkheden in het sociaal domein gebruiken, niet veel anders is dan de Chinese digitale staatscontrole: <https://bijvoorbeeldverdacht.nl/syri-medio-2019-voor-de-rechter/>
- 51 <https://aandeslagmetdeomgevingswet.nl/wetsinstrumenten/samenhang/omgevingswet-alleen/kies-sturingsstijl/>
- 52 De SHARED-waarden werden geïnitieerd door het LDE Centre for BOLD Cities.
- 53 <https://www.thedigitalsociety.info/wp-content/uploads/2018/04/VSNU-Digital-Society-Research-Agenda.pdf>
- 54 Zie voor een complete uitwerking van dit argument: Van Zoonen, L. (2019). Opnieuw fatale remedies: een kritische reflectie op datatransities in het sociaal domein. Sociologie, 15(1), p. 19-45.
- 55 Welling, W. (2017). Permanent bèta. i-Bestuur online, 11 oktober. <https://ibestuur.nl/weblog/permanent-beta>
- 56 Zie bijvoorbeeld het werk van Gerd Kortuem van de TU Delft: [https://pure.tudelft.nl/portal/en/persons/gw-kortuem\(d3f86f6b-35c6-438d-957d-e82b650d4bac\)/publications.html](https://pure.tudelft.nl/portal/en/persons/gw-kortuem(d3f86f6b-35c6-438d-957d-e82b650d4bac)/publications.html)
- 57 Zie bijvoorbeeld een brief van de minister van Economische Zaken en Klimaat aan de Tweede Kamer inzake de publieke belangen rond datavoorziening. Tweede Kamer, vergaderjaar 2018–2019, 35 000 XIII, nr. 81.
- 58 Zie bijvoorbeeld Harthoff, S. (2017). Gemeenten hopen op duidelijke regelgeving voor sensoren. Binnenlands Bestuur, 31 oktober, <https://www.binnenlandsbestuur.nl/digitaal/nieuws/gemeenten-hopen-op-duidelijke-regelgeving-voor.9574094.lynkx>; en Schoemaker, R. (2018). Gemeenten moeten kiezen tussen twee kwaden. IB&P, dat gaat je niets aan, 7 maart. <https://privacy-gemeenten.nl/2018/03/gemeenten-moeten-kiezen-persoonsgegevens-2/>
- 59 Auteurs dezes zijn overigens wel in gesprek met de betreffende wethouder om mogelijke scenario's te verkennen.

AI in de digitale samenleving

Kwaliteit van algoritmen en besluitvorming

Prof. dr. Eric Postma
Hoogleraar Artificial Intelligence
Tilburg University



Voorwoord

Door Ran Haase

In 2015 dachten wij in Eindhoven na over wat de rol van de gemeente nu eigenlijk is als het gaat over verzamelen van data in de publieke ruimte. De reden hiervoor was dat we in toenemende mate met diverse partijen werden geconfronteerd die data over bezoekers van de stad wilden verzamelen.

Ons gedachtenexperiment leidde ons langs fundamentele vragen over het publiek belang waar wij als gemeente voor staan; vragen over het eigendom van data, grondrechten, de waarde die met data kan worden gecreëerd en het belang van de inwoner. Uiteindelijk kwamen we tot de Open Data Principles en het IoT Charter. Deze principes zijn daarna doorontwikkeld tot de Spelregels voor de Digitale Stad van Amsterdam en Eindhoven, welke richting geven aan verantwoord datagebruik.

Het denken over verantwoord datagebruik heeft sindsdien gelukkig niet stilgestaan. Het debat over data en – recentelijk – AI wordt binnen gemeenteland in allerlei domeinen steeds breder gevoerd. De kansen en uitdagingen worden steeds duidelijker, maar zijn voor menig politicus, ambtenaar en inwoner nog moeilijk te vatten. Hoe dan ook, het afgelopen jaar zien we in gemeenteland een groeiend pleidooi voor het maken van ethische afwegingen bij dataprojecten.

Het is goed om te zien dat het ministerie van Binnenlandse Zaken en Koninkrijksrelaties het initiatief heeft genomen voor de proeftuin behoorlijk datagebruik. En dat het ministerie prof. dr. Eric Postma bereid heeft gevonden om bijgaand essay te schrijven. Prof. dr. Eric Postma is als hoogleraar AI verbonden aan het departement Cognitive Science & AI van Tilburg University en aan de Jheronimus Academy of Data Science (JADS) te 's-Hertogenbosch, een samenwerkingsverband van Eindhoven University of Technology en Tilburg University. Zijn onderzoek richt zit op de ontwikkeling en toepassing van machine learning in de automatische herkenning van beelden en signalen.

De aanbevelingen van prof. dr. Postma onderschrijf ik van harte: AI dient zich niet in een vacuüm te ontwikkelen. Er dient geïnvesteerd te worden in hoogwaardig onderzoek. Tegelijkertijd is het noodzakelijk dat er een ethisch, juridisch en maatschappelijk debat wordt gevoerd over wat voor een samenleving wij willen zijn.



Het is daarbij van belang dat de overheid investeert in kennisontwikkeling bij politici en beleidsmakers, zodat zij op verantwoorde en betrouwbare wijze hun bijdrage aan dit debat kunnen leveren.

De aanbevelingen van prof. dr. Postma zijn belangrijke vervolgstappen in een debat dat gevoerd moet blijven worden.

Ran Haase
Juridisch adviseur gemeente Eindhoven
Lid begeleidingscommissie

Samenvatting

Dit essay verschaft een nuchtere blik op AI (Artificial Intelligence) en haar rol in de samenleving van nu en in de toekomst. De nadruk ligt op het gebruik van data en AI-algoritmen in de openbare ruimte. Er wordt een antwoord gegeven op de vraag op welke wijze kan worden voorkomen dat er zaken misgaan in het datagebruik en bij de toepassing van AI-algoritmen in onze samenleving.

Na een inleidende beschouwing over algoritmen, machine learning, en de recente 'AI-revolutie', wordt vastgesteld dat: (1) het lastig is voor beleidsmakers om AI en verantwoord datagebruik te doorgronden, (2) de AI-revolutie zowel een technische als sociale revolutie is die een interdisciplinaire aanpak vereist, (3) er een nijpend tekort is aan professionals in allerlei domeinen met voldoende kennis van de mogelijkheden en gevaren van AI-algoritmen, en (4) certificering van AI-algoritmen door een onafhankelijk orgaan noodzakelijk is.

Op basis van deze vaststellingen worden vijf beleidsvoorstellen geformuleerd, te weten: (1) school politici en beleidsmakers regelmatig bij op het gebied van datagebruik en de laatste mogelijkheden van AI in relatie tot de sociale en juridische aspecten, (2) investeer in opleidingen die interdisciplinaire professionals afleveren, om zodoende het tekort aan professionals aan te zuiveren, (3) verbeter de omstandigheden voor AI-onderzoekers om de brain drain van onderzoekers te stoppen, (4) investeer in onderzoek om de transitie naar een goede digitale samenleving te faciliteren, en (5) de overheid zou toetsing van AI-algoritmen verplicht kunnen stellen om de kans op misstanden te minimaliseren. Aan de hand van deze beleidsvoorstellen, kan worden voorkomen dat er zaken misgaan bij de toepassing van AI-algoritmen in onze samenleving.

1. Inleiding

De kunstmatige intelligentie, kortweg AI (Artificial Intelligence), staat volop in de belangstelling. Wilde voorspellingen en waarschuwingen over AI zorgen voor onrust en onzekerheid. Dit essay verschaft een nuchtere blik op AI en haar rol in de samenleving van nu en in de nabije toekomst. De focus ligt op data, die beschouwd kunnen worden als de brandstof van AI-algoritmen in de context van de openbare ruimte. Op basis van vier observaties worden vijf beleidsvoorstellen geformuleerd.

Alvorens in te gaan op de kernvraag van dit essay, wordt hieronder in grote lijnen uiteengezet wat AI is en waarom het momenteel zo'n grote rol speelt in de media en in het publieke debat.

1.1 Wat is AI en waarom is het belangrijk voor de digitale samenleving?

Het onderzoeksgebied AI heeft een lange geschiedenis. In 1956 stelde John McCarthy de term Artificial Intelligence voor. Het doel van AI is om 'intelligente systemen' te begrijpen en te maken. Het standaard tekstboek voor AI¹ geeft de volgende vier beknopte definities van AI:

- Systemen die denken zoals mensen
- Systemen die rationeel denken
- Systemen die handelen zoals mensen
- Systemen die rationeel handelen

Deze vier definities combineren de concepten 'denken' en 'handelen' met 'zoals mensen' en 'rationeel'. Bij denken en handelen zoals mensen gaat het om het nabootsen en begrijpen van mensen. Bij rationeel denken en handelen valt eerder te denken aan een niet-menselijk, puur rationeel systeem. Dit beeld vinden we terug in allerlei sciencefictionfilms, waarin een mensachtige robot volstrekt logisch (rationeel) redeneert, maar geen emoties of empathie toont. Het bestaan van verschillende definities duidt op de veelheid aan interpretaties van AI.

Het is verhelderend om te beschouwen wat de mogelijkheden van AI zijn. AI heeft bijvoorbeeld geleid tot zeer krachtige schaakcomputers die superieur zijn aan menselijke schakers. Dergelijke systemen zijn rationeel in de zin dat ze gebaseerd zijn op pure logica waarbij de consequenties van alle mogelijke zetten en

tegenzetten worden geanalyseerd en geëvalueerd. Dit is mogelijk doordat de wereld van het schaakspel zeer beperkt is en zich daarom goed laat beschrijven in termen van logica. Om te kunnen schaken, hoeft de schaakcomputer niet te begrijpen wat een schaakstuk of een schaakbord is. In andere woorden, het schaakspel is relatief eenvoudig formaliseerbaar in termen van logische regels, die op hun beurt gemakkelijk te implementeren zijn in een computerprogramma. Vanwege de eenvoudige formaliseerbaarheid van het schaakspel versloeg de IBM-computer Deep Blue al in 1997 de wereldkampioen Gary Kasparov.

De grote droom van de eerste generatie AI-onderzoekers was het creëren van een rationeel denkend AI-systeem. Een dergelijk systeem zou puur logisch redeneren. Het succes van Deep Blue werd gezien als een belangrijke stap voorwaarts in de realisatie van een kunstmatig rationeel denkend systeem. Immers, Deep Blue was in staat om beter dan mensen te 'redeneren' over het schaakspel. Een veel grotere uitdaging is een schaakcomputer of robot die interacteert met het schaakbord door middel van bijvoorbeeld een videocamera en een robotarm. Een dergelijke schaakcomputer dient de schaakstukken te herkennen, ongeacht de belichting, het aanzicht en de mogelijke gedeeltelijke oclusie door andere schaakstukken. De visuele herkenning van schaakstukken, het schaakbord, en de tegenstander, is een grote uitdaging omdat het zich niet gemakkelijk laat formaliseren. Het is om deze reden dat we al vrij lang beschikken over schaakcomputers, snelle rekenmachines, en automatische planningsystemen, maar nog niet zo lang over systemen die automatisch afbeeldingen kunnen 'herkennen' of beschrijven. Die laatste ontwikkeling is het gevolg van de recente 'AI-revolutie' die heeft geresulteerd in doorbraken op moeilijk te formaliseren taken zoals de analyse van beeld, geluid en tekst. De opkomst van datagedreven AI-algoritmen waarmee patronen in beeld, geluid en tekst geanalyseerd en herkend kunnen worden, gaat gepaard met een groter belang van de kwaliteit van data. De kwaliteit van een AI-algoritme is direct afhankelijk van de kwaliteit van de data.

1.2 Vraagstelling

De nieuwe AI-algoritmen hebben een enorme maatschappelijke impact. De veelheid aan mogelijke toepassingen leidt tot verschuivingen in de verdeling van verantwoordelijkheden tussen mens en machine. Dit noopt tot een herinrichting van onze samenleving, voor een deel door de aanpassing van regelgeving en normering. Data staan hierbij centraal.

Vrijwel alle AI-innovaties komen tot stand door het verzamelen en voorbereiden van data. Dit leidt tot de volgende vraagstelling.

Op welke wijze kan worden voorkomen dat er zaken misgaan in het datagebruik en de toepassing van AI-algoritmen in onze samenleving?

De beantwoording van deze vraag komt voort uit een beschouwing van de 'oude' en de 'nieuwe' algoritmen (hoofdstuk 2), machinelearning-algoritmen (hoofdstuk 3), en de AI-revolutie (hoofdstuk 4). Dit resulteert in een viertal observaties en de beantwoording van onze vraagstelling (hoofdstuk 5), culminerend in vijf beleidsvoorstellen (hoofdstuk 6).

2. Deterministische algoritmen versus probabilistische algoritmen

Gewone algoritmen spelen al sinds een aantal decennia een cruciale rol in onze samenleving. Iedere computer functioneert op basis van algoritmen die zorg dragen voor de correcte afhandeling van allerlei digitale handelingen variërend van laag-niveau-taken, zoals het lezen en schrijven van data naar het computer-geheugen, tot hoog-niveau taken, zoals het adequaat vertalen van toetsenbord-commando's naar databehandelingen. Dergelijke algoritmen zijn volledig deterministisch. Dit betekent dat gegeven een begintoestand het algoritme altijd dezelfde eindtoestand bereikt. Wanneer bijvoorbeeld een computergebruiker een bestand wil opslaan, kan zij gebruik maken van de toetscombinatie ctrl-S. Het commando ctrl-S activeert een algoritme dat een voorschreven reeks digitale handelingen uitvoert, uiteindelijk resulterend in de opslag van het bestand. De programmeertaal die wordt gebruikt om dergelijke handelingen te programmeren is een geformaliseerde verzameling van instructies, vaak gebaseerd op ALS-DAN-constructies. Om de functionaliteit ctrl-S te programmeren bevat een typisch algoritme de volgende constructie:

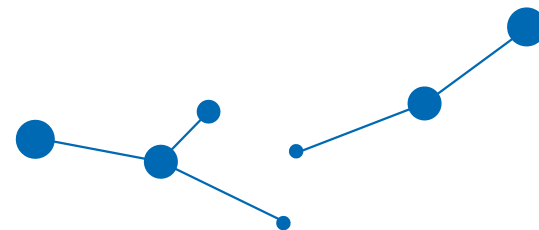
**ALS toets = 'ctrl' EN toets = 's'
DAN sla_bestand_op**

Het DAN-gedeelte specificeert de digitale handelingen om het document op te slaan. Het gebruik van ALS-DAN-regels in algoritmen is een uiting van de formele basis van computeralgoritmen. Theoretisch informatici onderzoeken algoritmen door gebruik te maken van formele logica. Op basis hiervan kan de correcte werking van algoritmen worden aangetoond. De aantoonbaarheid is relatief eenvoudig voor geformaliseerde taken zoals regelgeving. In regelgeving dient veelal aan een aantal voorwaarden voldaan te zijn om de regel van toepassing te laten zijn. In een computeralgoritme is dit eenvoudig te specificeren.

**ALS voorwaarde 1 EN voorwaarde 2 EN...
DAN de regel is van toepassing**

Zolang regelgeving ondubbelzinnig en consistent is, laat die zich gemakkelijk in een algoritme vertalen.

Probabilistische algoritmen onderscheiden zich van bovenstaande deterministische algoritmen door het hebben van een probabilistisch element. Als voorbeeld beschouwen we een algoritme voor een loterij. Om het winnend lotnummer te bepalen wordt een zogeheten *random-number-generator* gebruikt, dat een willekeurig getal genereert.² Door dit probabilistische element is niet meer met zekerheid te zeggen wat de uitvoer van het algoritme zal zijn. In plaats daarvan kan enkel een uitspraak gedaan worden in termen van waarschijnlijkheden. Voor een dobbelsteen-algoritme is bijvoorbeeld enkel te voorspellen dat de kans op een '1' als uitvoer gelijk is aan 1/6de, oftewel dat gemiddeld een op de zes keer dat het algoritme wordt uitgevoerd de uitvoer van het algoritme '1' zal zijn. Het onderscheid tussen deterministische en probabilistische algoritmen is van belang voor het begrip van moderne machinelearning-algoritmen.



3. Machinelearning-algoritmen

Machinelearning-algoritmen zijn probabilistische algoritmen die op basis van data voorspellingen trachten te doen. Er bestaan drie typen van machinelearning-algoritmen die verschillen in de wijze waarop ze leren: gesuperviseerd leren, ongesuperviseerd leren, en reinforcement leren.

Gesuperviseerd leren betreft het leren aan de hand van voorbeelden waarvan de identiteit wordt meegegeven tijdens het leren. Wanneer bijvoorbeeld een machinelearning-algoritme door middel van gesuperviseerd leren wordt getraind op het herkennen van verdacht gedrag op videobeelden, bestaan de trainingsvoorbeelden uit videosegmenten van personen met daarbij de identiteit of het label 'verdacht' of 'niet verdacht'. In feite bepaalt het label de aard van de herkenningstaak. Immers, voor dezelfde video's zouden de labels 'man' en 'vrouw' het algoritme trainen op de herkenning van het geslacht van de persoon in de video.

Ongesuperviseerd leren is gebaseerd op enkel de voorbeelden, zonder de labels. In het voornoemde voorbeeld, zal een door middel van ongesuperviseerd leren getraind machinelearning-algoritme de video's clusteren op basis van de meest prominente visuele overeenkomst. De door ongesuperviseerd leren verkregen clusters kunnen gebaseerd zijn op allerlei overeenkomsten, zoals bijvoorbeeld het weertype of de hoeveelheid beweging in de video. De afwezigheid van labels resulteert in minder controle over de uitkomst. Om die reden is ongesuperviseerd leren minder succesvol dan gesuperviseerd leren. Doorgaans wordt ongesuperviseerd leren gebruikt om inzicht in de data te krijgen, aangezien het inzicht geeft in de clustering van datapunten. Aangezien er voor veel toepassingen geen labels voor handen zijn, of moeilijk zijn te verkrijgen, bestaat er in de AI de ambitie om ongesuperviseerd leren meer in te zetten.

Reinforcement leren wordt doorgaans gebruikt voor het leren van reeksen van acties, zoals het laten navigeren van een robot. In reinforcement leren genereert het machinelearning-algoritme acties die worden beloond indien ze correct zijn. De leermethode komt overeen met de methode die beproefd is voor het aanleren van complex gedrag door dieren. Iedere keer dat het dier het gewenste gedrag vertoont, wordt het beloond met een snoepje of een aai.

De psycholoog Skinner leerde in 1950 duiven met hun snavels te tafeltennissen door middel van reinforcement leren.³ In de AI wordt reinforcement leren momenteel intensief gebruikt met successen in formaliseerbare domeinen, zoals videospellen of computervarianten van bordspellen zoals het eeuwenoude Chinese spel Go.

In het hiernavolgende beperken we ons tot gesuperviseerd leren, aangezien dit momenteel de meest dominante en succesvolle variant van machine learning is.

De probabilistische aard van machinelearning-algoritmen komt niet enkel voort uit het gebruik van een random-number-generator, maar ook uit de probabilistische aard van de data gebruikt voor het 'trainen' van deze algoritmen. Grote internet-bedrijven maken veel gebruik van machinelearning-algoritmen, bijvoorbeeld om te voorspellen of een klant geïnteresseerd is in bepaalde producten. Google verzamelt enorm veel data over het internet zoek-, navigatie- en koopgedrag van haar klanten, omdat op basis van deze data machinelearning-algoritmen kunnen worden getraind door middel van gesuperviseerd leren. Het trainen kan verschillende vormen aannemen, maar resulteert altijd in een statistisch model dat op basis van een deel van de data (de input) andere data (de output oftewel het label) voorspelt. Na training kunnen deze modellen bijvoorbeeld voor nieuwe klanten op basis van hun internetgedrag (de input) voorspellen in wat voor producten ze geïnteresseerd zijn (de output). Na training gedragen de meeste machinelearning-algoritmen zich deterministisch. Gegeven een input zal het altijd dezelfde output genereren. Zoals gezegd, komt het probabilistische karakter van machinelearning-algoritmen voornamelijk voort uit de data.

Laten we een zeer eenvoudig voorbeeld beschouwen. Stel dat als input het geslacht van een internetgebruiker bekend is. Het internetbedrijf kan dan bijhouden hoe vaak vrouwen aankopen doen uit de categorieën gereedschap en keukenspullen en hoe vaak mannen dat doen. Op basis van de waargenomen frequenties van de aankoop van gereedschap en keukenspullen, kan het machinelearning-algoritme van het internetbedrijf de waarschijnlijkheid afleiden dat een vrouw of man gereedschap of keukenspullen koopt. Na training kan het algoritme op basis van het geslacht van een nieuwe klant (de input), de output voorspellen. De output is in dit geval de voorspelling of een klant geïnteresseerd is

in gereedschap of in keukenspullen. Het machinelearning-algoritme maakt een voorspelling van de aankoop op basis van het geslacht van de klant.

Hoewel voornoemd voorbeeld erg eenvoudig is, vormt het een goede illustratie van de kern van machinelearning-algoritmen. Na het trainen, vormen deze algoritmen statistische modellen van de werkelijkheid zoals die in de data gerepresenteerd is. De kwaliteit van de algoritmen hangt daarom af van de kwaliteit van de data. Alle zaken die voor goede statistiek van belang zijn, zijn dat ook voor de goede werking van machinelearning-algoritmen. Een voorbeeld hiervan is dat de 'steekproef', oftewel de data waarop het algoritme wordt getraind, representatief moet zijn voor de populatie waar het op wordt toegepast. Wanneer bijvoorbeeld eerdergenoemd algoritme getraind wordt op data van mannen en vrouwen die afkomstig zijn uit een bevolkingsgroep waar vrouwen voornamelijk huishoudelijke functies vervullen, dan zullen de voorspellingen totaal anders zijn dan wanneer het algoritme getraind wordt op een bevolkingsgroep waar mannen voornamelijk het huishouden voor hun rekening nemen. Vanwege de probabilistische aard van de data voor het trainen van machinelearning-algoritmen, vormen de basisprincipes van de statistiek het belangrijkste fundament voor een begrip van de werking, mogelijkheden, en beperkingen van machine learning.

3.1 De statistische basisprincipes van machine learning

Het gesuperviseerd trainen van een machinelearning-algoritme vereist een groot aantal input-output-instanties. Verwijzend naar ons eenvoudige voorbeeld, wordt een input-output-instantie gevormd door de data van een klant die een aankoop heeft gedaan, bijvoorbeeld een mannelijke klant die keukenspullen heeft gekocht. De instantie bestaat dan uit het input-output-paar 'man'-'keukenspullen'. Bij het voorspellen van verkiezingsuitslagen vergroten onderzoekers de betrouwbaarheid van de voorspellingen door (1) een representatieve steekproef te nemen, en (2) zo veel mogelijk respondenten te verzamelen. Er is een directe parallel met machinelearning-algoritmen. De kwaliteit van de verkregen predicties van een getraind machinelearning-algoritme hangt af van (1) de representativiteit van de instanties en (2) de kwantiteit van de instanties.

Het eerste punt, (1) representativiteit, is hierboven al genoemd. Een recent voorbeeld betreft de objectherkenningsalgoritmen van Google, Amazon en andere

internetbedrijven. Uit recent onderzoek blijkt dat alledaagse objecten, zoals tandpasta, goed herkend worden op afbeeldingen uit rijke landen, maar bijzonder slecht op afbeeldingen uit minder rijke landen. De reden is dat de algoritmen zijn getraind op input-output-instanties (zoals een afbeelding van een tube tandpasta als input en 'tandpasta' als output) die voor het overgrote deel zijn verkregen uit rijke westerse landen. Een foto uit, bijvoorbeeld, Burundi toont tandpasta tegen een achtergrond van hout in plaats van een betegelde badkamer, en geeft daarom de foutieve output 'hout'.⁴

Het tweede punt waar de kwaliteit van predicties van een machinelearning-algoritme vanaf hangt is (2) kwantiteit, en betreft het aantal instanties voor het trainen van een machinelearning-algoritme. Voor een goede werking dienen voldoende gelabelde instanties voorhanden te zijn. Een belangrijke reden voor de grote vlucht van machine learning en AI is het bestaan van grote internet- en andere technologiebedrijven die beschikken over enorm grote databestanden.

Algoritmen zijn getraind op input-output-instanties die vooral zijn verkregen uit rijke westerse landen

3.2 Complexe machinelearning-algoritmen

Eenvoudige machinelearning-algoritmen bestaan al heel lang. Het eerste voorstel voor het zogenaamde 'nearest-neighbour'-algoritme gaan terug naar 1951.⁵ Sinds dit machinelearning-algoritme zijn lerende algoritmen in vele gedaanten onderzocht en toegepast. In plaats van machine learning werden termen gebruikt als patroonherkenning, *datamining* en *knowledge discovery*. Hoewel deze variatie in aanduidingen verwijst naar subtiele verschillen in de benadering of toepassing, is er een grote gemene deler: het creëren van een voorspellend model (of algoritme) op basis van data. Tegenwoordig is machine learning de alom gebruikte verzamelnaam voor lerende algoritmen.

Er bestaat een groot aantal verschillende machinelearning-algoritmen. Deze algoritmen verschillen van elkaar in de wijze waarop vanuit de trainingsdata

een voorspelling wordt gemaakt. Gegeven een voorspellingstaak, bijvoorbeeld het voorspellen of een persoon in de openbare ruimte crimineel gedrag gaat vertonen op basis van gedetailleerde locatiegegevens van haar mobiele telefoon, is het op voorhand niet te zeggen welk algoritme het beste werkt. Om die reden experimenteren AI-onderzoekers met verschillende algoritmen om te bepalen welk algoritme de beste predicties levert voor een gegeven taak. Een belangrijke dimensie waarop machinelearning-algoritmen verschillen is de complexiteit van de predictietaak. Stel dat crimineel gedrag eenvoudig uit locatiegegevens is af te leiden, bijvoorbeeld doordat er een vaste locatie is waar criminelen zich ophouden. In dat geval is de voorspellingstaak erg eenvoudig en kan worden volstaan met een eenvoudig machinelearning-algoritme. Immers, enkel de locatie voorspelt crimineel gedrag. In de praktijk is het voorspellen van crimineel gedrag natuurlijk veel complexer en vereist het vinden van een voorspellend element in de combinatie van locatiegegevens een complexer machinelearning-algoritme.

Het voordeel van een complexer machinelearning-algoritme is dat deze doorgaans beter in staat is om voorspellingen te maken. Deze betere prestatie in het voorspellen gaat ten koste van de uitlegbaarheid van de voorspelling. In het geval van het eenvoudige machinelearning-algoritme, is de uitlegbaarheid groot. Indien het algoritme crimineel gedrag voorspelt, is dat vanwege de 'criminele locatie' waar de persoon zich ophoudt. De reden van de voorspelling is daarom eenvoudig uit te leggen. De uitlegbaarheid van een complex machinelearning-

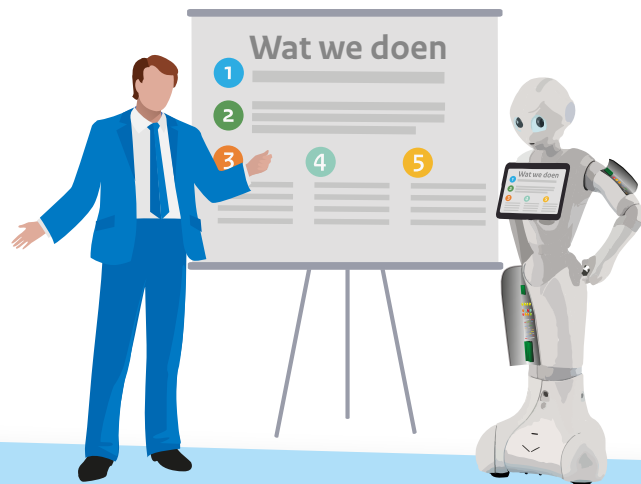
algoritme is veel lastiger, vanwege de abstracte combinatie en weging van data-elementen die zich veelal moeilijk in natuurlijke taal laten uitleggen. Zoals de European Group on Ethics in Science and New Technologies het formuleerde in haar statement in 2018: "...AI's inner workings can be extremely hard - if not impossible - to track, explain and critically evaluate."⁶

De *trade-off* tussen voorspellende prestatie en uitlegbaarheid vormt een van de kernproblemen van de AI-revolutie. Met name de krachtige *deeplearning*-algoritmen, die aan de basis staan van de AI-revolutie, bereiken enorme prestaties in hun voorspellingen, maar vragen een hoge prijs door hun gebrekkige uitlegbaarheid. Om die reden wordt er veel onderzoek verricht naar zogenaamde *explainable AI*, oftewel AI met een uitlegfaciliteit. Momenteel is het bijvoorbeeld mogelijk om door middel van een zogeheten Q&A-interactie een AI-systeem te laten uitleggen waarom het een afbeelding, video of tekst op een bepaalde manier classificeert. Wanneer het AI-systeem een afbeelding classificeert als zijnde een afbeelding van een skiër, en de gebruiker vraagt 'waarom?', dan markeert het systeem het deel van de afbeelding waar ski's te zien zijn en genereert het de zin 'omdat ik twee parallelle rechte lijnen zie tegen een witte achtergrond'. Deze vorm van uitleg komt tot stand door het systeem te trainen op een groot aantal afbeeldingen en bijbehorende verklaringen.

Peter Norvig, de gerenommeerde AI-onderzoeker en onderzoeksdirecteur bij Google, betwijfelt of explainable AI een zinvolle onderneming is. Hij stelde:

*"You can ask a human, but, you know, what cognitive psychologists have discovered is that when you ask a human you're not really getting at the decision process. They make a decision first, and then you ask, and then they generate an explanation and that may not be the true explanation. ... So we might end up being in the same place with machine learning where we train one system to get an answer and then we train another system to say – given the input of this first system, now it's your job to generate an explanation. ... Explanations alone aren't enough, we need other ways of monitoring the decision making process."*⁷

AI-systemen zouden op eenzelfde wijze een verklaring kunnen genereren, nadat een beslissing is genomen. Net zoals bij mensen het geval is zou de motivatie een mooi en overtuigend verhaal zijn, maar niet noodzakelijkerwijs op waarheid



berusten. Ondanks deze terechte twijfel, is ieder inzicht in de wijze waarop een AI-systeem tot haar beslissing komt van belang, omdat de zogeheten *narrow* intelligentie van het systeem gecompliceerd en gecontroleerd dient te worden door de brede intelligentie van de mens.

4. De AI-revolutie

De AI-revolutie betreft een klasse van nieuwe machinelearning-algoritmen die wordt aangeduid met de term *deep learning*. Weliswaar bestaan de achterliggende ideeën van deep learning al 30 jaar, maar het is pas sinds de afgelopen tien jaar dat de technologie voor allerlei doorbraken zorgt. De reden is tweeledig: (1) dankzij de opkomst van het internet en de grote technologiebedrijven is er enorm veel gelabelde data beschikbaar, en (2) de door computerbedrijven als NVIDIA ontwikkelde grafische kaarten, oorspronkelijk bedoeld voor het genereren van realistische videogames, zorgen voor een enorme versnelling bij het trainen van deep learning-algoritmen.

Deep learning heeft vooral voor doorbraken gezorgd op het gebied van beeld- en videoherkenning, spraakherkenning en tekstanalyse.⁸ Vrijwel iedere week verschijnen er berichten in de media over nieuwe doorbraken die worden bereikt op een van deze gebieden. Alvorens we een overzicht geven van toepassingen die relevantie hebben voor datagebruik in het publieke domein, is het goed om de globale grenzen aan te geven van AI op basis van deep learning. Dit doen we aan de hand van een voorbeeld.

Wanneer we een deep learning-algoritme trainen op miljoenen afbeeldingen van voorwerpen, waarbij iedere afbeelding is gelabeld met de naam van het getoonde object (bijvoorbeeld 'Fido') of de klasse waartoe het behoort ('hond'), dan is het algoritme vervolgens in staat om afbeeldingen die het nooit eerder heeft gezien van een juist label te voorzien. Oftewel, wanneer we een nieuwe foto maken van Fido en dat als input aan het deep learning-algoritme aanbieden, dan geeft het in ca. 95 tot 99 procent van de gevallen als output 'Fido' of 'hond'. Wanneer we het algoritme trainen op miljoenen afbeeldingen met bijbehorende beschrijvingen, dan kan het netwerk zelfs voor een nieuwe afbeelding een juiste beschrijving geven. Het was tien jaar geleden ondenkbaar dat een computersysteem hiertoe in

staat zou zijn. Er is daarom sprake van een enorme wetenschappelijke en praktische doorbraak. De doorbraak maakt de verleiding groot om een parallel te trekken met menselijke waarneming. Immers, mensen kunnen ook een afbeelding herkennen en classificeren. Er is echter een cruciaal verschil tussen de wijze waarop deep learning afbeeldingen 'herkent' en de wijze waarop mensen dat doen. Een deep learning-algoritme weet niets van honden. Het heeft geen kennis van dieren en weet niet hoe een hond voelt of zich gedraagt. Het enige dat het netwerk 'weet' is dat bepaalde visuele patronen geassocieerd zijn met het label 'hond'. Om die reden spreekt men in de AI over deep learning als een 'narrow AI'-methode. Het algoritme presteert excellent op een zeer nauwgedefinieerde taak. Hoewel het gebruikelijk is om uitspraken te doen als 'het deep learning-netwerk herkent de hond', is dit in feite zeer misleidend, omdat herkenning veel meer omvat dan het associëren van een afbeelding met een label.⁹

We beschouwen ter illustratie een onderzoeksproject in de openbare ruimte dat wij in het verleden in samenwerking met het Dutch Institute for Technology, Safety & Security (DITSS)¹⁰ hebben uitgevoerd op het Piusplein in Tilburg en op het Stratumseind in Eindhoven. Het doel van het project was het voorspellen van agressief gedrag op basis van videobeelden. Een van de observaties was dat onschuldig wild gedrag van een vriendengroep nauwelijks te onderscheiden was van agressief gedrag. Het bleek dat de menselijke waarnemer door haar algemene maatschappelijke kennis van hoe een groep vrienden zich gedraagt, veel beter in staat is een escalatie van agressie te voorspellen. Concreet betekent dit dat de inzet van deep learning voor de 'narrow AI'-taak van het herkennen van mogelijk agressief gedrag gecombineerd dient te worden met een verificatie door een menselijke waarnemer. Je zou kunnen zeggen dat de narrow intelligentie van een AI-systeem complementair is aan de brede intelligentie van de mens.

4.1 AI-toepassingen in de openbare ruimte

De sterke prestaties van AI en deep learning-algoritmen op het gebied van de 'herkenning' van beeld en geluid, resulteert in een breed scala aan mogelijke toepassingen in de openbare ruimte. De afgelopen jaren wordt de toepassing van deep learning en andere AI-methoden veelvuldig onderzocht in de context van *smart cities*. Het is de verwachting dat door de opkomst van het *Internet of Things (IoT)* en met name het beschikbaar komen van snelle 5G-technologie, dergelijke projecten een grote vlucht gaan nemen. In het hiernavolgende bespreken we de



drie voornaamste toepassingen in de openbare ruimte: (1) het monitoren en sturen van bezoekersgedrag, (2) het monitoren en sturen van verkeersstromen en (3) het monitoren en sturen van veiligheid.

4.2 Het monitoren en sturen van bezoekersgedrag

Het volgen van de bezoekers van een stadscentrum of een andere openbare ruimte kan op verschillende manieren plaatsvinden: met behulp van videocamera's, door middel van het tracken van mobiele telefoons, of door speciale sensoren, geplaatst in de openbare ruimte. De visuele herkenning van voetgangers en het voorspellen van voetgangsgedrag heeft de laatste jaren een enorme vlucht genomen door de opkomst van deep learning en zelfrijdende auto's. Ieder jaar verbetert de kwaliteit van de herkenning door de voortschrijdende ontwikkelingen in machine learning.

Het is in principe mogelijk om de identiteit van voetgangers vast te stellen op basis van videobeelden, al is dit afhankelijk van de kwaliteit van het beeld en van andere zaken, zoals de zichtbaarheid van gezichtskenmerken. In vele gevallen zijn de systemen nog niet zo robuust dat ze beter zijn in het herkennen van gezichten dan dat mensen dat zijn.¹¹ Daarnaast hebben veel systemen een bias in hun herkenningsprestaties. Zo is bijvoorbeeld Amazons gezichtsherkenningssysteem Rekognition omstreden door de impliciete raciale bias.¹² Het systeem presteert slechter op bijvoorbeeld de gezichtsherkenning van zwarte vrouwen, vanwege de gebrekkige representativiteit van zwarte vrouwen in de dataset waarop Rekognition is getraind, een reflectie van de statistische basisprincipes zoals beschreven in paragraaf 3.1.

Het is tot op zekere hoogte mogelijk om gezichtsexpressie te herkennen uit videobeelden, zodat er een schatting gemaakt kan worden van de gemoedstoestand van een bezoeker van de openbare ruimte. Daarnaast komen er methoden beschikbaar die uit videobeelden de lichaamshouding van een bezoeker kunnen herkennen en volgen. Hiermee ontstaat de mogelijkheid om gedragingen te herkennen en tot op zekere hoogte te voorspellen.

Ongeacht de wijze waarop de schattingen van de locaties en gedragingen van bezoekers worden verkregen, is het sturen van gedrag (*nudgen*) een logische volgende stap. Wanneer bijvoorbeeld blijkt dat de maximale bezoekerscapaciteit van een populaire winkelstraat is bereikt, kunnen allerlei verleidingsmethoden

worden ingezet om de bezoekers naar andere winkelstraten of bestemmingen te leiden. In het living lab op het Stratumseind in Eindhoven wordt druk geëxperimenteerd met nudging door het aanpassen van de intensiteit en kleur van de verlichting.¹³

In de nabije toekomst zou nudging meer gepersonaliseerd kunnen worden. In de sciencefictionfilm *Minority Report* (2002), is er een scène waarin de hoofdrolspeler Tom Cruise gefrustreerd door een winkelcentrum loopt. De winkels zijn uitgerust met camera's die zijn (film)identiteit en gemoedstoestand herkennen en daar direct op reageren door het aanbieden van een vakantie en door de bestemming vervolgens te tonen op een scherm. Op dit moment, zeventien jaar na het verschijnen van de film, is de technologie om een dergelijke gepersonaliseerde benadering te realiseren beschikbaar. Op basis van deep learning-algoritmen is de herkenning van de identiteit en gemoedstoestand van bezoekers van de openbare ruimte mogelijk. De gepersonaliseerde boodschap kan echter veel subtieler zijn dan het aanbieden van een vakantie op een beeldscherm.

In de nabije toekomst zou nudging meer gepersonaliseerd kunnen worden

Op het scherm zou een virtueel gezicht kunnen verschijnen dat door middel van subtiele verbale en non-verbale signalen de bezoeker verleidt om een gewenste handeling te verrichten. Ook de technologie voor het genereren en manipuleren van virtuele gezichten is beschikbaar. Het door IBM opgekochte bedrijf Soul Machines creëert niet van echt te onderscheiden dynamische virtuele gezichten.¹⁴ Het is een kwestie van tijd voordat dergelijke systemen breed worden ingezet. Ons onderzoek heeft laten zien dat mensen gevoelig zijn voor de non-verbale signalen (zoals bijvoorbeeld een glimlach of het subtiel fronsen van de wenkbrauwen) van virtuele gezichten en daar onbewust non-verbaal op reageren.¹⁵ Een dergelijke vorm van nudging grijpt in op ons sociale brein met alle mogelijkheden en gevaren van dien. Zoals een effectieve verkoper leert hoe hij of zij adequaat reageert op een klant om zo veel mogelijk te verkopen, zo kan een deep learning-algoritme (op basis van reinforcement leren) leren om het virtuele gezicht de optimale combinatie van sociale signalen te laten genereren,

opdat een bezoeker van de openbare ruimte het gewenste gedrag vertoont. Hoewel een dergelijk systeem nog niet bestaat, is het technologisch mogelijk en een kwestie van tijd voordat er mee geëxperimenteerd gaat worden. Dit realistische toekomstscenario is een illustratie van het gegeven dat de grote mogelijkheden en gevaren van de AI-revolutie plaatsvinden op het snijvlak van de technologie en de mens.

4.3 Het monitoren en sturen van verkeersstromen

Het volgen en regelen van verkeersstromen in een stad is een grote uitdaging. Veel steden maken al gebruik van 'Adaptive Signal Control'-technologie om de doorstroom van verkeer te optimaliseren. Dergelijke systemen tracken het verkeer en gebruiken traditionele AI-algoritmen om het gedrag van stoplichten te optimaliseren. De opkomst van 5G-technologie en zelfrijdende auto's zal voor de nodige innovaties zorgen die de doorstroom in een stadscentrum zullen verbeteren. De concrete toepassingen van deep learning hebben betrekking op het visueel tracken van auto's aan de hand van de automatische herkenning van de kentekenplaat (iets dat al op grote schaal gebeurt in parkeergarages) en het automatisch herkennen van onveilig rijgedrag of van verkeersovertredingen. Een concrete toepassing waar veel steden mee experimenteren is smart parking. Op basis van videobeelden of sensorinformatie worden vrije parkeerplaatsen gedetecteerd en vervolgens worden automobilisten die op zoek zijn naar een parkeerplaats automatisch naar deze plaatsten genavigeerd.

4.4 Het monitoren en sturen van veiligheid

Het bewaken en stimuleren van de veiligheid in de openbare ruimte is van groot maatschappelijk belang. Verschillende steden experimenteren met automatische video-analyse, geluidscamera's of Internet-of-Things-sensorsystemen, om escalerende situaties tijdig te detecteren en te volgen en om waar nodig in te grijpen. In Eindhoven wordt er in het kader van een living lab op het Stratumseind, een populair uitgaansgebied in het centrum, geëxperimenteerd met nieuwe technologie. Een belangrijk uitgangspunt voor het living lab is dat er geen privacygevoelige informatie wordt verzameld. Ook in Enschede en Utrecht worden er vele experimenten uitgevoerd in de publieke ruimte. In een artikel in de *Guardian* worden deze verschillende Nederlandse projecten kritisch belicht vanwege de privacygevoeligheid van de verzamelde data en vanwege de vermenging van publiek onderzoek en private belangen.¹⁶



Gezien het experimentele karakter van de projecten en de onduidelijke wetgeving rondom de verzameling van data zijn dergelijke kritische geluiden onvermijdelijk en nuttig. Het toont aan hoe belangrijk het is om te experimenteren om op die wijze te komen tot een goede aanpak die alle ethische toetsen kan doorstaan. Voor wat betreft de vermenging van uit publieke middelen gefinancierd onderzoek en private belangen is het van belang om op te merken dat in het wetenschappelijk onderzoek de publiek-private samenwerking gestimuleerd wordt. Initiatieven zoals het living lab op het Stratumseind danken hun succes mede aan de intensieve samenwerking van de overheid met bedrijven. Dit neemt niet weg dat voor dergelijke publiek-private samenwerkingsprojecten duidelijke richtlijnen gehanteerd dienen te worden, zoals die bijvoorbeeld zijn geformuleerd door de Vereniging van Samenwerkende Nederlandse Universiteiten.¹⁷

Initiatieven zoals het living lab op het Stratumseind danken hun succes mede aan de intensieve samenwerking van de overheid met bedrijven

In de Verenigde Staten hanteert men vooralsnog ruimere opvattingen over het verzamelen van data in de openbare ruimte. Zo detecteert en lokaliseert het bedrijf *ShotSpotter* het afvuren van een vuurwapen op basis van een netwerk van microfoons geplaatst in de openbare ruimte.¹⁸ Wanneer op een bepaalde locatie een vuurwapen wordt gebruikt, dan wordt het geluid opgepikt door verschillende naburige microfoons. Met behulp van AI-technologie wordt het specifieke geluid gedetecteerd en gelokaliseerd, ongehinderd door allerlei ruis in de vorm van reguliere stadsgeluiden veroorzaakt door verkeer, weer en mensen. De website van het bedrijf suggereert dat de inzet van *ShotSpotter* leidt tot een vermindering van de frequentie van het afvuren van wapens.

Een Nederlandse variant van deze technologie zou zich kunnen richten op andere geluidsbronnen, zoals de detectie van menselijke 'distress calls' (bijvoorbeeld gillen van angst) of van het afgaan van een autoalarm.

De introductie van een microfoonnetwerk zou uiteraard leiden tot zorgen over privacy. Microfoons pikken uiteraard ook menselijke conversaties op. Het is mogelijk om AI-systemen, met name deep learning-algoritmen, zodanig te trainen dat bepaalde aspecten van de input niet meegenomen kunnen worden.

In China wordt het gebruik van AI-technologie in de openbare ruimte ver doorgevoerd. Er wordt zelfs geëxperimenteerd met het automatisch herkennen van criminelen op videobeelden. In 2016 publiceerden twee onderzoekers uit Shanghai een controversiële studie¹⁹ waarin een deep learning-algoritme werd toegepast op het automatisch screenen van gezichten. Zij trainden het algoritme op pasfoto's van ruim 1000 niet-criminelen en ruim 700 criminelen. Het algoritme bleek in bijna 90 procent van de gevallen in staat te zijn om voor niet eerder getoonde pasfoto's te voorspellen of het een crimineel of een niet-crimineel betrof. De impliciete aanname dat criminelen visueel te onderscheiden zijn van niet-criminelen is uiteraard zeer omstreden. Belangrijker nog is dat de gebruikte trainingsmethode een statistische bias heeft. De trainingsdata van criminelen werd verkregen middels foto's gemaakt van reeds gevangene criminelen. Het is zeer waarschijnlijk dat criminelen die prototypische criminele gezichtskenmerken hebben, eerder opgepakt worden dan criminelen die die kenmerken niet hebben. Met prototypische criminele gezichten wordt bedoeld: gezichten van criminelen zoals ze in films worden afgebeeld. Je zou dus kunnen zeggen dat de meest succesvolle criminelen - dat wil zeggen: criminelen die niet opgepakt worden - er juist niet als prototypische criminelen uit zien.

Op AI gebaseerde videosurveillance biedt de mogelijkheid om criminele handelingen of intenties tijdig waar te nemen. Hierbij valt te denken aan de automatische detectie van (de intentie tot) zakkenrollen of diefstal. Het is al langer bekend dat experts en niet-experts goed in staat zijn om op basis van video-opnamen in de publieke ruimte crimineel gedrag te voorspellen.²⁰ Onderzoekers van het Massachusetts Institute of Technology (MIT) hebben in 2016 al laten zien dat het op basis van videobeelden voorspellen van intenties heel goed mogelijk is. De onderzoekers trainden een deep learning-algoritme op meer dan 600 uur videofragmenten van tv-series zoals *The Office* en *Desperate Housewives*. Na training was het systeem in staat om te voorspellen of twee personen bij aanvang van een ontmoeting elkaar zouden omhelzen, kussen, hun handen zouden schudden, of zouden 'high-fiven'.

De automatische herkenning van de intentie tot crimineel gedrag wordt al geruime tijd onderzocht, bijvoorbeeld van de intentie tot zakkenrollen.²¹ De opkomst van deeplearning-algoritmen zal de komende jaren resulteren in allerlei ‘*predictive policing*’-systemen. Het voornaamste obstakel voor het realiseren van dergelijke systemen is niet gelegen in de technologie, maar in de data.

Deeplearning-systemen vereisen omvangrijke datasets. Het is niet eenvoudig om voor een handeling als zakkenrollen voldoende videovoorbeelden te verzamelen. Ongetwijfeld zijn er nu al bedrijven en instellingen die investeren in het verzamelen van data voor dit soort toepassingen. Zodra er commerciële producten beschikbaar komen kunnen die een belangrijk ondersteuningsmiddel vormen voor het monitoren en sturen van de veiligheid in de openbare ruimte.

5. Observaties

De huidige ontwikkelingen op technologisch en maatschappelijk gebied beschouwend, komen we tot de volgende vier observaties.

1. Het is lastig voor beleidsmakers om AI en verantwoord datagebruik te doorgronden.
2. De AI-revolutie is zowel een technische als sociale revolutie en vereist daarom een interdisciplinaire aanpak.
3. Er is een nijpend tekort aan professionals in allerlei domeinen met voldoende kennis van de mogelijkheden en gevaren van AI-algoritmen.
4. Certificering van AI-algoritmen door een onafhankelijk orgaan is noodzakelijk.

Hieronder wordt elk van deze observaties nader toegelicht.

5.1 Complexiteit

Het belang van AI en verantwoord datagebruik voor de ontwikkeling en strategische positie van Nederland wordt alom erkend, maar voor politici en beleidsmakers is het moeilijk grip te krijgen op de onderliggende materie. Hiervoor zijn vier redenen aan te wijzen.

De eerste reden is dat de zuiverheid van de discussie over de ontwikkeling van AI wordt geschaad door toekomstvoorspellingen die appelleren aan Hollywood-

scenario's en niets te maken hebben met de realiteit. Een prominent voorbeeld is Ray Kurzweil die in 2005 het concept van de *Singularity* introduceerde als het punt waarop de exponentiele groei van technologische mogelijkheden resulteert in een samensmelting van mens en machine.²² Volgens Kurzweil zal de Singularity plaatsvinden in 2045. De voorspellingen van Kurzweil zijn deels verantwoordelijk voor ongefundeerde uitingen over de ontwikkeling van AI. Elon Musk, de innovatieve technologie-entrepreneur, is een prominent voorbeeld van een AI-onheilsvoorspeller die voorspelt dat AI de mensheid gaat overnemen. Dit is niet de plaats om in te gaan op de details van de voorspellingen, maar een duidelijk manco van Kurzweil-achtige toekomstscenario's is dat ze gebaseerd zijn op de extrapolatie van een exponentiele groei. In de praktijk zit er een limiet aan exponentiele groei.

De exponentiele groei van het aantal internetgebruikers wordt bijvoorbeeld beperkt door de omvang van de wereldbevolking, waardoor de exponentiele groei verzadigt en de vorm aanneemt van een S-vormige curve. Op eenzelfde wijze bereikt de exponentiele groei van de technologische mogelijkheden van een bepaalde technologie haar verzadigingspunt. De grote impact van het idee van een Singularity is ook te relateren aan de plausibiliteit vanuit het perspectief van Hollywood-films. Veel succesvolle sciencefictionfilms spelen met het idee van de machine die de mens overneemt. Om tegenwicht te bieden aan de ogenschijnlijk plausibele Kurzweil-scenario's, zou iedere politicus of beleidsmaker kennis moeten nemen van het artikel *The Seven Deadly Sins of AI Predictions* van de Amerikaanse AI- en robotonderzoeker Rodney Brooks.²³ In dit artikel rekent Brooks genadeloos en overtuigend af met de misleidende AI-voorspellingen.

De tweede reden waarom beleidsmakers moeilijk grip krijgen op AI is het ten onrechte generaliseren van prestaties van AI op formaliseerbare domeinen om die vervolgens te vertalen naar de echte wereld. Een formaliseerbaar domein verwijst hierbij naar een schaakspel, of een bordspel zoals Go, of een spel dat zich afspeelt in een virtuele wereld, zoals in videogames. Het bedrijf Deep Mind heeft de laatste jaren indrukwekkende prestaties bereikt met ‘deep reinforcement leren’, een combinatie van deep learning met reinforcement leren. Zo versloeg hun AI-systeem AlphaGo de menselijke wereldkampioen Go.

Een latere versie, AlphaGo Zero, leerde het spel door bijna 5 miljoen keer tegen zichzelf te spelen. In 100 wedstrijden van AlphaGo Zero tegen haar voorganger Alpha Go, won AlphaGo Zero alle wedstrijden.²⁴ Daarnaast ontwikkelde Deep Mind een AI-systeem dat superieur aan mensen presteert in videogames. Deze indrukwekkende resultaten zijn bereikt dankzij het gegeven dat de wereld van Go en van videogames eenvoudiger en beperkter is dan de echte wereld. Bij iedere stap in het spel is er een beperkt repertoire van acties en mogelijkheden. Ondanks deze beperking, is het aantal mogelijke sequenties van acties duizelingwekkend groot. Dat is de reden waarom het voor menselijke spelers veel tijd kost om de spellen onder de knie te krijgen. Het is verleidelijk om de AI-prestaties in games een op een te vertalen naar mogelijke AI-prestaties in de alledaagse wereld. Echter, de complexiteit van de alledaagse wereld is onvoorstelbaar veel groter.

De derde reden voor de ongrijpbaarheid van AI is de snelheid van de ontwikkeling van AI-technologie. Door deze snelheid zijn de innovaties voor niet-experts moeilijk bij te houden en te duiden. De snelheid betreft hier niet de eerder bekritiseerde exponentiële snelheid, maar de snelheid die ontstaat door relatief kleine innovaties die een enorme impact kunnen hebben op bijvoorbeeld het datagebruik. Ter illustratie beschouwen we een voorbeeld betreffende verantwoord datagebruik in relatie tot de opkomst van zogeheten generatieve modellen. De opkomst van generatieve modellen, een variant van deep learning, biedt de mogelijkheid om *fake videos* en *fake images* te genereren.



Een relatief onvoorzien effect van generatieve modellen betreft de privacybescherming voor data uit bijvoorbeeld het publieke domein. De-identificatie is een beproefde methode om aan privacybescherming te doen bij het gebruik van publieke data. Bij de toepassing van de-identificatie-methoden worden alle directe verwijzingen naar individuele gegevens verwijderd. Vele door de-identificatie geanonimiseerde datasets worden wereldwijd gedeeld voor onderzoek en analyse. De laatste jaren is meermaals gebleken dat uit ge-de-identificeerde datasets, toch individuele gegevens te verkrijgen zijn. Door het gebruik van geavanceerde machinelearning-algoritmen blijkt re-identificatie vaak mogelijk te zijn. Dit resulteerde in een publieke discussie met als belangrijke uitkomst dat re-identificatie enkel mogelijk is wanneer de volledige dataset beschikbaar is. Aangezien veelal enkel een deel van de data publiek beschikbaar is, is de kans op re-identificatie verwaarloosbaar klein. Echter, een recent artikel²⁵ laat zien dat door gebruik te maken van een generatief model, re-identificatie op basis van een deel van publieke data wel degelijk mogelijk is. Dit heeft een direct gevolg voor de Europese privacywetgeving GDPR (General Data Protection Regulation) en voor overige richtlijnen voor het delen en openbaar maken van data. Dit voorbeeld laat zien dat wat goed beleid is ten aanzien van het delen van data, plotseling kan veranderen door een relatief bescheiden innovatie in machine learning.

De vierde en laatste reden waarom beleidsmakers moeilijk grip krijgen op AI is de doorgaans niet-technische achtergrond van politici en beleidsmakers. Ter illustratie citeer ik hieronder uit het 'Profiel nieuwe Tweede Kamer 2017' zoals gepubliceerd op de website www.parlement.com.

Het opleidingsniveau van Tweede Kamerleden blijft onveranderd hoog. Bijna twee derde (94 leden) heeft een academische opleiding gevolgd en 31 leden volgden een opleiding in het HBO (of daarmee vergelijkbare opleiding). Van enkele PVV-leden is het onderwijsniveau niet bekend. Van de academisch opgeleide leden blijven juristen in de meerderheid (26 leden). Ook economen (17), politicologen (15), bestuurskundigen (11) en historici (10) zijn goed vertegenwoordigd. In de Kamer zitten ook nog een filosoof, een rechtsfilosoof, een architect en een archeoloog. Er zit geen medicus meer in de Kamer. Leden met een bèta-achtergrond zijn er weinig, namelijk slechts vijf. (https://www.parlement.com/id/vkclmgzmvqya/profiel_nieuwe_tweede_kamer_2017)

Ondanks het hoge opleidingsniveau van de Kamerleden, is hun technische affiniteit nihil. Hierdoor is het lastig voor de meeste Kamerleden, politici en beleidsmakers in het algemeen, om de AI-innovaties op een juiste wijze in te schatten. Om deze vier redenen zijn politici en beleidsmakers onvoldoende op de hoogte van de mogelijkheden en gevaren van AI-algoritmen.

5.2 Interdisciplinaire aanpak

De impact van de ontwikkeling van AI heeft alles te maken met de sociale dynamiek van mensen in interactie met technologie. Om die reden is de kern van de AI-problematiek interdisciplinair. Eerdere technische innovaties zoals de opkomst van mobiele telefonie, het internet en sociale media, werden ook gekenmerkt door de verwevenheid van technologie met sociale dynamiek. De ontwikkeling van GSM-technologie is gekoppeld aan de sociale wens om altijd bereikbaar te zijn en te kunnen communiceren. De technologische ontwikkeling van het internet en sociale media was gekoppeld aan de sociale wens om efficiënt te communiceren en zichzelf of een product te adverteren. De verwevenheid van AI-technologie met sociale dynamiek is minder evident, maar potentieel veel ingrijpender. Dit komt goed naar voren in een belangrijk opiniestuk in *Scientific American*²⁶, geschreven door Dirk Helbing, hoogleraar computational social science bij de Zwitserse universiteit ETH in Zürich, en acht internationale collega's uit verschillende disciplines. Zij stelden dat de huidige samenleving zich op een kruispunt bevindt.

Een voorbeeld van onschuldige nudging is het stimuleren van milieubewust gedrag in de openbare ruimte

Verwijzend naar Singapore als een door data gecontroleerde maatschappij, met China in haar voetsporen, constateren Helbing en collega's dat de sturing van menselijk gedrag door *persuasion* en *nudging* weliswaar onschuldig begint, maar snel ernstige vormen aan kan nemen. Een voorbeeld van onschuldige nudging is het stimuleren van milieubewust gedrag in de openbare ruimte. Daar staat tegenover het schandaal rondom Cambridge Analytica, een voorbeeld van een ernstige vorm van nudging waarbij het stemgedrag van internetgebruikers werd gemanipuleerd.²⁷

Het vinden van een juiste balans tussen nuttig datagebruik en datamisbruik was een van de aandachtspunten in de verslagen van twee *World Cafés* over behoorlijk datagebruik in de openbare ruimte in Groningen en Eindhoven.²⁸ In het kader van het verzamelen van data over voetgangers en fietsers in de binnenstad, werd tijdens de bijeenkomst in Groningen gesproken over het nut en gevaar van nudging, bijvoorbeeld wanneer dit wordt ingezet om verkeersstromen te reguleren. De deelnemers gaven aan meer te willen weten over nudging en de impact ervan op bezoekers van de binnenstad. In Eindhoven was de bijeenkomst gericht op het verzamelen van gegevens over bezoekers van Stratumseind, waar de nadruk ligt op het verhogen van de veiligheid in de openbare ruimte. De deelnemers plaatsten kanttekeningen bij, onder andere, de transparantie van dataverzameling en nudging, de legitimiteit van nudging en de beheersbaarheid van datagedreven nudging.

De bijeenkomsten in Groningen en Eindhoven, aangevuld met een soortgelijke bijeenkomst in Amsterdam,²⁹ resulteerden in de identificatie van een viertal knelpunten en vragen.³⁰ Hieronder worden de vier voornaamste discussiepunten benoemd en voorzien van kort commentaar.

1. *Er is een spanning tussen privacy en techniek. Wat kan er? Wat mag er? Wat moet je (niet) willen?*

Zoals dit essay duidelijk maakt, kan er heel veel, al moet de kracht van AI-technologie niet overschat worden aangezien het narrow intelligentie betreft. Wat er mag is niet geheel duidelijk, omdat de ontwikkeling van de wetgeving noodzakelijkerwijs achterloopt bij de ontwikkeling van de technologie.

2. *Er is (a) spanning tussen bedrijven en burgers (hierbij gaat het om de belangen van private commerciële partijen) en (b) spanning tussen het individu en het collectieve, publieke belang.*

De spanning tussen bedrijven en burgers kwam eerder naar voren in paragraaf 4.4 (waar het ging over het kritische artikel in de Guardian). Ten aanzien van de spanning tussen het individu en het collectieve belang is er een reëel gevaar dat de datagedreven maatschappij het belang van het individu overschaduwet.



3. De mate van transparantie en uitlegbaarheid van datagebruik en verzameling in de publieke ruimte.

Transparantie en uitlegbaarheid zijn (in beperkte mate) technologisch realiseerbaar, maar het betreft hier ook transparantie van de gehanteerde regels en richtlijnen.

4. Wat is de rol van de overheid binnen de smart city?

Deze vraag speelt een centrale rol in dit essay. In een algemene maatschappelijke context, bieden Helbing en collega's een goede denkrichting. Zij geven tien aanbevelingen om de technologische en sociale revolutie in goede banen te leiden. Deze aanbevelingen sluiten - in algemenere zin - goed aan bij de uitkomsten van de World Cafés en bovenstaande vragen en richten zich onder meer op: grotere transparantie ter vergroting van het vertrouwen, ruimte voor sociale en economische diversiteit, decentralisatie van AI-systemen en het promoten van verantwoord gedrag in de digitale wereld. Voor dit essay is het van belang te benadrukken dat de aanbevelingen van Helbing en zijn collega's tot stand gekomen zijn dankzij een interdisciplinaire beschouwing van de opkomst van technologie en van onze maatschappij. In samenhang met de uitkomsten van de World Cafés en de hierboven beschreven vier knelpunten, laat het artikel overduidelijk het belang van een interdisciplinaire benadering zien om de AI-revolutie het hoofd te bieden en om tot verantwoord datagebruik te komen.

5.3 Tekort aan professionals

In hun blog *AI en de Nederlandse Belangen* stellen Jurgen Opper en Aaron Arends: "Ondanks dat meerdere Nederlandse deskundigen de laatste maanden waarschuwen voor 'brain drain' op het gebied van artificial intelligence, lijkt het alsof nog steeds niet volledig wordt beseft dat onze economische belangen op het spel staan. Nederland heeft een goede startpositie, maar staat op het punt het startschot te missen."³¹ Het AI-onderzoek in Nederland staat onder druk door een zogeheten brain drain.³² Als motor van de innovatie is er meer aandacht voor AI-onderzoek noodzakelijk. Om gezonde innovatie te faciliteren is aandacht voor zowel technische als sociale innovatie onontbeerlijk. In Noord-Brabant worden in september 2019 twee AI-initiatieven gestart.

De Technische Universiteit Eindhoven lanceert het *Eindhoven Artificial Intelligence Systems Institute* dat zich voornamelijk richt op engineering, bijvoorbeeld door AI in te zetten voor sector automotive en voor hi-tech-systemen. Tilburg University start haar initiatief *AI for Society* met de eerste geaccrediteerde academische AI-opleiding in Brabant, *Cognitive Science & AI (CSAI)*, verzorgd door het gelijknamige departement. In deze opleiding komen technologische en sociale aspecten samen. In de *Jheronimus Academy of Data Science (JADS)* in 's-Hertogenbosch, een samenwerkingsverband van de universiteiten in Eindhoven en Tilburg, komen beide benaderingen samen in een gezamenlijk onderzoeks- en onderwijsprogramma. De nieuwe onderwijsprogramma's zullen de komende jaren veel AI-professionals afleveren. Deze unieke initiatieven in Noord-Brabant verdienen navolging in de rest van Nederland.

5.4 Certificering

De roep om regulering van de grote technologiebedrijven wordt steeds luider. Op 25 juli van dit jaar kondigde het Amerikaanse *Department of Justice* een breed antitrust-onderzoek aan naar de grote technologiebedrijven.³³ Afgezien van de commerciële belangen van grote bedrijven, vormen machinelearning-algoritmen een bron van problemen. Zelfs wanneer een bedrijf of overheidsinstelling op beperkte schaal machine learning toepast op data uit de publieke ruimte kan dat om allerlei redenen schadelijk zijn. Een van de redenen hiervoor is dat ieder machinelearning-algoritme een bias heeft, een soort impliciete aanname over de wijze waarop verschillende observaties met elkaar worden vergeleken. Dankzij de bias is het algoritme in staat om te generaliseren en ongeziene voorbeelden te herkennen. Gemiddeld genomen levert de bias een voordeel op (betere herkenningsprestaties), maar voor een individu kan de bias nadelig uitpakken. Een sterk vereenvoudigd voorbeeld van een dergelijke bias is de wijze waarop machinelearning-algoritmen omgaan met verschillende persoonskenmerken, zoals leeftijd en salaris. Sommige machinelearning-algoritmen, zoals zogeheten '*decision tree*'-algoritmen, beschouwen ieder kenmerk afzonderlijk, terwijl andere algoritmen, zoals eerdergenoemde '*nearest neighbour*'-algoritmen, de kenmerken impliciet met elkaar in verband brengen.

Gemiddeld genomen presteert een van de algoritmen wellicht beter dan de ander, maar voor individuele gevallen kan het zijn dat het ene algoritme meer recht doet aan de situatie dan de ander. Het onderkennen van dergelijke subtiele biases in

machinelearning-algoritmen en het verifiëren van mogelijke methodologische fouten bij het trainen of operationaliseren van machinelearning-algoritmen, vereist een gedegen screening. Een onafhankelijke door de overheid opgezette certificeringsorganisatie kan een dergelijke screening uitvoeren. Individuen die zich benadeeld voelen door een beslissing voortvloeiend uit het gebruik van een machinelearning-algoritme, zouden bij deze organisatie een verzoek tot onderzoek kunnen indienen. Meer in het algemeen, kan de organisatie het algoritme op nader te bepalen statistische en ethische criteria toetsen.

Bovenstaande vier observaties bieden handvaten voor de overheid om tot concrete beleidsmaatregelen over te gaan. Onze verkenning van de AI-technologie en observaties maken het mogelijk om een antwoord te geven op onze centrale vraag. We herhalen onze vraagstelling: *Op welke wijze kan worden voorkomen dat er zaken misgaan in het datagebruik en de toepassing van AI-algoritmen in onze samenleving?* Om te voorkomen dat er zaken misgaan in het datagebruik en de toepassing van AI-algoritmen in onze samenleving is het noodzakelijk dat...

- politici en beleidsmakers begrijpen wat de mogelijkheden en beperkingen zijn van AI. Het is vooral van belang dat de 'olie' waar AI op functioneert bestaat uit data. Basale kennis van de statistische principes die ten grondslag liggen aan een verantwoorde dataverzameling is een vereiste.
- wordt onderkend dat de AI-revolutie niet enkel een zaak is van technici. Vrijwel alle AI-systemen functioneren in interactie met mensen en daardoor ontstaat (in de woorden van Helbing *et al*) een delicaat feedback-systeem. Om die reden is er een grote behoefte aan interdisciplinaire professionals.
- de AI-onderzoeksinfrastructuur aantrekkelijk is voor talenten, zodat de digitale samenleving verzekerd blijft van de aanvoer van AI-professionals en AI-innovaties.
- onderzoeksprojecten worden gedefinieerd die de aanpak zoals geschetst in het artikel van Helbing *et al*. in de praktijk brengen.
- er een centraal certificeringsorgaan komt dat AI-algoritmen evalueert.

We besluiten het essay met vijf beleidsvoorstellen die direct voortvloeien uit de observaties.

6. Beleidsvoorstellen

De impact van machine learning en datagebruik in de openbare ruimte op de samenleving zal de komende jaren alsmaar toenemen. Om de overgang naar een door machine learning gedomineerde digitale samenleving goed te laten verlopen, formuleer ik hieronder vijf beleidsvoorstellen die voortvloeien uit de observaties in hoofdstuk 5.

1. School politici en beleidsmakers regelmatig bij op het gebied van datagebruik en de laatste mogelijkheden van AI in relatie tot sociale en juridische aspecten.

De snelle technologische ontwikkelingen leiden tot het beschikbaar komen van nieuwe databronnen en nieuwe technologie. Het is bijvoorbeeld pas sinds kort mogelijk om synthetische afbeeldingen en video's te maken die bijna niet van echt te onderscheiden zijn. Het is voor politici en beleidsmakers onmogelijk om de mogelijkheden en gevaren van de nieuwe databronnen en nieuwe technologie te overzien. De geadviseerde bijscholing betreft een update op een relatief hoog niveau en heeft twee doelen: (1) om als gelijke gesprekspartner met de markt te kunnen communiceren en (2) om regie te kunnen voeren over wat er nodig is aan technologische innovatie bij de overheid. Met een basale kennis van machine learning is met een korte bijscholing een helder beeld te verkrijgen, zonder dat daarbij de technische details aan bod hoeven te komen. Ook voor de sociale en juridische aspecten van datagebruik en AI-technologie geldt dat de snelle technologische ontwikkelingen zorgen voor voortschrijdend inzicht en veranderende opvattingen. Ook op dit vlak zou een periodieke bijscholing wetgevers, politici en beleidsmakers in staat kunnen stellen de snelle ontwikkelingen het hoofd te bieden en adequaat te handelen. Een gecombineerde bijscholing van technische, sociale, en juridisch aspecten zou bijvoorbeeld eens per jaar kunnen plaatsvinden gedurende een 'technologische bijscholingsdag'.

2. De digitale samenleving vereist professionals die thuis zijn in zowel de technische als de menselijke en sociale kant van AI. Investeer in opleidingen die dergelijke professionals afleveren, om zodoende het tekort aan professionals aan te zuiveren.

De AI-revolutie is zowel een technologische als een sociale of culturele revolutie. De technische mogelijkheden van AI hebben een impact op menselijk gedrag en beïnvloeden de wijze waarop de maatschappij functioneert. Professionals zijn nodig voor alle domeinen waar technologie en mens samenkomen en betreft dus het volledige scala inclusief het sociale domein, veiligheid, en mobiliteit.

3. Probeer de brain drain van onderzoekers te stoppen door de omstandigheden voor AI-onderzoekers te verbeteren.

Om een belangrijke rol te blijven spelen in de technologische ontwikkelingen is het werven en behouden van goede onderzoekers een vereiste. Het aantrekkelijker maken van Nederland als AI-kennisland voor onderzoekers resulteert in meer innovatie, beter onderwijs en een meer leidende rol in de sturing van de digitale samenleving.

4. Investeer in onderzoek om de transitie naar een goede digitale samenleving te faciliteren.

Investeringen in AI-onderzoek naar toepassingen van datagebruik in de openbare ruimte dragen bij aan een succesvolle transitie. Het ligt voor de hand om aansluiting te zoeken bij bestaande initiatieven, zoals het Stratumseind-project in Eindhoven.

5. Een onafhankelijk CBS-achtig orgaan dat AI-technologie screent op genoemde statistische principes voorkomt mistanden. Gezien de voortdurende upgrades van machinelearning-algoritmen, is een periodieke toetsing noodzakelijk. De overheid zou toetsing verplicht kunnen stellen om de kans op mistanden te minimaliseren.

Een dergelijk orgaan voor toezicht op en transparantie van datagebruik zou op nationaal of Europees niveau opgezet kunnen worden. Onlangs³⁴ kondigde de Australische overheid plannen aan voor het opzetten van een overheidsorganisatie voor het screenen van algoritmen zoals gebruikt door de technologiegiganten Facebook en Google. De organisatie wordt ondergebracht bij de *Australian Competition and Consumer Commission*. Het gebruik van algoritmen en data in de openbare ruimte, al dan niet door technologiegiganten, zou op eenzelfde wijze gescreend kunnen worden.

Over de auteur



Prof. dr. Eric Postma is hoogleraar Artificial Intelligence aan Tilburg University en Jheronimus Academy of Data Science (JADS) in 's-Hertogenbosch.



Referenties

- 1 Russell, S. & Norvig, P. (2009). Artificial Intelligence. A Modern Approach (3rd Edition).
- 2 Er bestaan computeralgoritmen die schijnbaar willekeurige getallen genereren, maar toch deterministisch zijn. Deze zogenaamde "pseudo random number" generatoren worden uiteraard niet gebruikt voor loterijen.
- 3 <https://www.youtube.com/watch?v=vGazyH6fQQ4>
- 4 DeVries, T., Misra, I., Wang, C., & van der Maaten, L.J.P. (2019). Does Object Recognition Work for Everyone? CVPR Workshop on Computer Vision for Global Challenges.
- 5 Fix, E. & Hodges, J.L. (1951). Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties. Technical Report 4, USAF School of Aviation Medicine, Randolph Field.
- 6 European Group on Ethics in Science and New Technologies (2018). Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems, Brussels, 9 March 2018. (page 7).
- 7 <http://www.techcentral.ie/googles-research-chief-questions-value-explainable-ai/> at 12-1-2018.
- 8 LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521, 436-444.
- 9 Postma, E.O. (2018). De angst voor AI is ongegrond. NRClive, 12 juni 2018. <https://nrclive.nl/de-angst-voor-ai-is-ongegron/>
- 10 <https://www.brainport.nl/ditss>
- 11 Rigby, S. (2019). Live facial recognition: how is it used? Science Focus, August 2, 2019. <https://www.sciencefocus.com/future-technology/live-facial-recognition-how-is-it-used/>
- 12 Raji, I.D. & Buolamwini, J. (2019). Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. Conference on Artificial Intelligence, Ethics, and Society. http://www.aies-conference.com/wp-content/uploads/2019/01/AIES-19_paper_223.pdf
- 13 <https://www.smartdatacity.org/stratumseind-living-lab/>
- 14 <https://www.ibm.com/case-studies/soul-machines-hybrid-cloud-ai-chatbot>
- 15 Mattheij, R.J.H., Postma, M., & Postma, E.O. (2015). Mirror, Mirror in the Wall: Is there mimicry in you all? Journal of Ambient Intelligence and Smart Environments, 7(2), 121-132.
- 16 <https://www.theguardian.com/cities/2018/mar/01/smart-cities-data-privacy-eindhoven-utrecht>
- 17 <https://www.vsnv.nl/publiek-private-samenwerking.html>
- 18 <https://www.shotspotter.com/>
- 19 Wu, X. & Zhang, X. (2016). Automated Inference on Criminality using Face Images. arXiv:1611.04135v1 [cs.CV] 13 Nov 2016.
- 20 Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2003). What happens next? The predictability of natural behaviour viewed through CCTV cameras. Perception, 33, 87-101.
- 21 Bouma, H., Baan, J., Burghouts, G.J., Eendebak, P.T., van Huis, J.R., Dijk, J., van Rest, J.H.C. (2014). Automatic detection of suspicious behavior of pickpockets with track-based features in a shopping mall. Proceedings SPIE 9253, Optics and Photonics for Counterterrorism, Crime Fighting, and Defence X; and Optical Materials and Biomaterials in Security and Defence Systems Technology XI, 92530F (7 October 2014).
- 22 Kurzweil, R. (2005). The Singularity Is Near: When Humans Transcend Biology. Viking Publishers. 23
- 23 Brooks, R. (2017). The Seven Deadly Sins of AI Predictions. MIT Media Review, October 6, 2017. <https://www.technologyreview.com/s/609048/the-seven-deadly-sins-of-ai-predictions/>
- 24 Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. Nature, 550, pp 354-359.
- 25 Rocher, L., Hendrickx, J.M., & de Montjoye, Y-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. Nature Communications, 3069, 23-7-2019.
- 26 Helbing, D., Frey, B., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., van den Hoven, J., Zicari, R., & Zwitter, A. (2017). Will Democracy Survive Big Data and Artificial Intelligence? Scientific American, Feb. 25.
- 27 Het artikel van Helbing et al werd gepubliceerd een jaar voor het Cambridge Analytica schandaal aan het licht kwam.
- 28 Fortunati, C., van Beusekom, H., van Enk, P., & Kerkvliet, M. (2018). Behoorlijk datagebruik in de openbare ruimte. Verslag World Café "Fietsers- en voetgangersmonitoring Binnenstad Groningen" 13 december 2018, PBLQ.
- 29 Fortunati, C., Meijer, D., van Beusekom, H., & Hartog, M. (2018). Behoorlijk datagebruik in de openbare ruimte. Verslag World Cafe Stratumseind Eindhoven, 21 november 2018.
- 30 van der Hoeven, P. (2019). Nota. Proeftuin Behoorlijk datagebruik in openbare ruimte: knelpunteninventarisatie. Directoraat-Generaal Overheidsorganisatie, 19 april 2019.
- 31 Engels-Lobrecht, R., Hooghiemstra, T., Ligthart, J., Haans, E., & Wassink, A. (2018). Inventarisatie dilemma's bij behoorlijk datagebruik in de openbare ruimte in Amsterdam. Project 006662, versie 1.2. 26 november 2018. PBLQ.
- 32 Oppel, J. & Arends, A. (2018). AI en de Nederlandse belangen. Blog Clingendael Netherlands Institute of International Relations. 26 November 2018.
- 33 Dutch Artificial Intelligence Manifesto. Special Interest Group AI. <http://ii.tudelft.nl/bnvti/wp-content/uploads/2018/09/Dutch-AI-Manifesto.pdf>
- 34 Benaich, N. & Hogarth, I. (2019). State of AI Report. June 28, 2019. <https://www.stateof.ai/>
- 35 Justice Department announces broad antitrust review of big tech. The Washington Post, July 25, 2019.
- 36 BBC News, 26-7-2019. Australia to police tech giants' algorithms. <https://www.bbc.com/news/technology-49125845>

Data makes the world go round

Voorstel voor onderzoek naar drie instrumenten ter versterking van de (digitale) autonomie

Theo Veltman en Rob van Kranenburg



Voorwoord

Door Peter van Hoesel

In het kader van het project 'Behoorlijk datagebruik in de openbare ruimte' waren oorspronkelijk vijf essays gepland. Na de bijeenkomst van de begeleidingsgroep medio juli, waar van alle gevraagde essays een hoofdlijn voorlag, is geconstateerd dat er iets miste. Kort gezegd: het vraagstuk van veiligheid en individu in de digitale wereld, nu en straks. Hoewel het oorspronkelijk niet de bedoeling was dat 'trekkers' van essays in de begeleidingscommissie zelf een artikel zouden schrijven, is in overleg met de opdrachtgevers van het ministerie van BZK anders besloten. Vlak voor de laatste vergadering van de begeleidingscommissie hebben Rob van Kranenburg en Theo Veltman een zesde essay ingezonden, dat door de commissie graag in ontvangst is genomen en beoordeeld.

De commissie vindt het een mooie bijdrage aan deze bundel essays, niet alleen vanwege de inhoud maar ook omdat je het kunt lezen als een hartenkreet vanuit het veld, of liever gezegd vanuit de frontlinie die wordt geconfronteerd met de altijd weerbarstige praktijk. Deze bijdrage is niet bedoeld als een wetenschappelijke analyse, het komt uit de praktijk. Dit essay vormt juist daarom een mooie aanvulling op de vijf andere essays, die stuk voor stuk wel zijn geschreven als wetenschappelijke analyses, met daarbij overigens ook de nodige aanknopingspunten voor de praktijk.

Theo en Rob kennen elkaar uit de innovatiewereld in en rond Amsterdam. Zij delen een gedrevenheid als het gaat om de borging van de vrijheid van het individu om zich in digitaal opzicht ongemontord en onbeïnvloed te kunnen bewegen in gebouwen en in de openbare ruimte. Tegelijkertijd vinden zij dat iedereen ook zoveel mogelijk de voordelen van digitale technologie moet kunnen benutten. Dat vraagt om een balans tussen veiligheid aan de ene kant en gemak, nut/plezier aan de andere kant. Daar gaat hun artikel over.

Rob en Theo hebben beiden een ruime achtergrond en ervaring als het gaat om mens, technologie en organisatie. Beiden hebben voldoende fouten (of bijna fouten) gemaakt in hun werk om veel geleerd te hebben. Zij kennen verschillende aspecten van digitalisering, vanuit verschillende rollen. Hun kennis, ervaring en inzichten vullen elkaar aan.

Peter van Hoesel

Emeritus hoogleraar bestuurskunde Erasmus Universiteit Rotterdam

Voorzitter begeleidingscommissie

Samenvatting

Er is de laatste tijd veel geschreven over de risico's in de digitale wereld. Het is nu en dan wat zoeken naar de samenhang en concrete, grijpbare aanbevelingen. In dit essay bevelen wij drie maatregelen aan. Vooraf gaat een samenvattende context voor die maatregelen vanuit het perspectief: data - digitaal platform - voorkeuren van mensen en maatschappelijke ontwikkelingen.

Kernvraag in dit essay is hoe en waarmee het behoud van de (digitale) autonomie' geborgd kan worden in een wereld waar data, en daarmee kennis, geld en macht, steeds meer geconcentreerd raken. Er worden steeds meer persoonsgebonden data¹ verzameld op internet en in de openbare ruimte. Daardoor kunnen velen tot in detail gekend zijn en ongemerkt beïnvloed worden in voorkeuren en keuzen. Terwijl de bescherming van persoonsgebonden data een grondrecht is, in Nederland (grondwet, artikel 10, lid 1) en in de EU.

De kernvraag naar het behoud van (digitale) autonomie kent geen eenvoudig antwoord. Met het Rathenau Instituut menen de auteurs dat er een verandering nodig is van mindset, op weg naar een wereld waarin waarden steeds meer centraal staan. Dat vergt een lange adem. De tijd dringt omdat de risico's snel toenemen. Er moet 'nu' gehandeld worden om gebruik te kunnen blijven maken van de voordelen zonder de autonomie van individuen in gevaar te brengen. Het is zaak om, nu het nog kan, de monopoliepositie van de grote organisaties, de platforms, te begrenzen en de positie van het individu te versterken. Daar is veel voor nodig, ook internationaal. In dit artikel bevelen de auteurs drie maatregelen aan voor het versterken van de regie van het individu over haar persoonsgebonden data en voor het voorkomen en tegengaan van het 'onbehoorlijk datagebruik' door organisaties die data verzamelen in de openbare ruimte van de stad, te weten:

1. Realiseer een nationale, onafhankelijke en open digitale vertrouwensinfrastructuur voor identificatie, authenticatie en autorisatie van persoonlijke data inclusief de bijbehorende governance op basis van de bestaande bouwstenen zoals Tippiq en Irma. Gebruik de whitepaper (Dunnen c.s.; 2019) als vertrekpunt.

2. Onderzoek het nut van en de mogelijkheden voor de realisatie van een hardware embedded voorziening, een 'datakill-switch' in slimme apparaten zoals sensoren, waarmee het verzamelen van persoonsgebonden data standaard op 'uit' staat. Standaard kan dan alleen data verzameld worden, waarvan democratisch besloten is dat deze anoniem zijn en nodig voor de borging van de openbare orde en de veiligheid en voor het beheer van de openbare ruimte van de stad.

3. Onderzoek het nut van en de mogelijkheden voor de realisatie van een landelijk 'datasafety-keurmerk' door een democratisch geborgd onafhankelijk instituut, op basis van een wettelijke grondslag voor het voorkomen en zo nodig aanpakken van datadiefstal en -misbruik, in aanvulling op de privacywetgeving AVG (Algemene Verordening Gegevensbescherming).

1. Data, eeuwenoude bron van kennis, macht en geld

Liza Minelli c.s. illustreren het belang van geld, welk gemak en welke invloed het biedt, in hun onnavolgbare vertolking van het lied *Money makes the worlds go round* uit 1972¹. Die situatie is nog niet veranderd. De financiële crisis die zovelen enkele jaren terug trof, werd onder meer veroorzaakt door hebzucht. Zie bijvoorbeeld de documentaires *The inside job* (2011), *Enron: the smartest guys in the room* (2005) en *Capitalism: A Love Story* (2009) van Michael Moore. De honger naar geld en macht creëert hypes, van tulpen en vastgoed tot digitale geesteskinderen. De hypes worden gevoed door kennis over wat er gaande is bij enkelen en door indrukken bij velen dat men er snel bij moet zijn. 'Kennis is macht'.

Data bieden een belangrijke mogelijkheid om kennis, en daarmee macht en geld te verwerven. Met data kun je individuen en groepen manipuleren. Het is maar welke data je verzamelt en gebruikt, op basis van welke criteria je dat doet; hoe je die data analyseert, op basis van welke veronderstellingen, regels en criteria, en hoe je de resultaten interpreteert. En ook belangrijk is: hoe vaak en via welke kanalen je de data presenteert. Als je data selectief gebruikt en de boodschap genoeg herhaalt, ontstaan er verhalen en suggesties die grote invloed kunnen hebben op voorkeuren en keuzemogelijkheden van mensen en organisaties². Dankzij de sociale media is het makkelijker dan ooit om dat op grote schaal te doen. Om de informatie en berichten goed te kunnen waarderen is dus kennis nodig over welke data gebruikt zijn, uit welke bronnen, wie die bronnen heeft aangelegd en op basis van welke criteria de analyse is gedaan³.

De aandacht voor data is geen hype. Dat de toegang tot correcte data van groot belang is, ook voor de eigen (economische) ontwikkeling, is al eeuwen het geval. De rol en het gebruik van data veranderen wel. Zo wil men tegenwoordig zo dicht mogelijk bij digitale knooppunten zitten. Hierdoor kan men milliseconden eerder inzicht hebben in de data om op basis daarvan te handelen op financiële markten. Er zijn en worden door een kleine groep miljoenen mee verdiend. Samengevat:

- Data = de bron van alle wijsheid én de bron van alle bedrog
- Data = de basis van verdienmodellen

1.1 Data, ook een bron van bedrog bij een beperkte bescherming door wet en overheid

Data worden al lang gebruikt om geld te verdienen en macht te creëren. Nieuw is dat er anno 2019 veel meer data zijn en dat er met data veel meer gedaan kan worden, terwijl de analysesoftware steeds minder eisen stelt aan de data. De ontwikkeling van internet, de smartphone, steeds meer digitale diensten (o.a. apps) en de hypes rond *big data* en *Artificial intelligence* (AI) hebben gezamenlijk geleid tot een onoverzichtelijk geheel van informatie met een vaak onduidelijke kwaliteit.

Antivaccinatiepropaganda (op Facebook) toonde aan vrouwelijke gebruikers boven de 25 van wie – dankzij het aflezen van hun surfgedrag – bekend was dat ze nadachten over zwangerschap. Deze gerichte advertenties, die miljoenen keren zijn bekeken, waren ingekocht door activisten die inenting van kinderen tegen mazelen willen ontmoedigen. Zuckerbergs platform was ze graag tot dienst, net zoals het bereid was vanuit Rusland verspreid nepnieuws te laten circuleren tijdens de Amerikaanse verkiezingen.
Uit: Thomas, Casper; (2019)⁴

Data kunnen gebruikt worden om het nieuws, de handel, een persoon en naasten of organisatie te manipuleren⁵. Beïnvloeding, misleiding en fraude komen veel vaker voor dan menigen denkt⁶.

Mensen en organisaties worden bedreigd en beschadigd met *deepfakes*⁷. Er zijn nep-datingsites, die eenzaamheid omzetten in klinkende munt⁸. Er wordt bewust *fake news* verspreid om mensen te beïnvloeden, voor economisch, maatschappelijk en/of politiek⁹ gewin. Op de sociale media zijn nauwelijks van echt te onderscheiden 'digitale personen' actief, aangestuurd door een bedrijf¹⁰. Soms hebben zij miljoenen volgers, en daardoor een grote impact. Hackers, al dan niet crimineel, spelen een heel eigen rol met de voortdurende aanvallen of bedreiging via bijvoorbeeld *ransomware* en *bots*¹¹. Dat houdt niet op¹². Men gebruikt alle middelen, soms ook die van u: "*IoT is a powerful Internet of Things (IoT) botnet comprised primarily of compromised home routers, TVs, DVRs, and IP cameras*" (Check Point Research, 2019).

Er wordt een digitale oorlog gevoerd, niet met kogels en bommen maar met gemanipuleerde data¹³. Niet alleen tussen staten is er oorlog. In zekere zin is er ook een digitale oorlog gaande tussen digitale platforms en ‘de consument’; tussen grootkapitaal, het midden- en kleinbedrijf en het individu. Allen strijden ze in meer of mindere mate om hun (digitale) autonomie¹⁴ en positie, ruimte voor ontwikkeling en bewegingsvrijheid. Wapens zijn: desinformatie, desintegratie en despotisme. De Raad voor Openbaar Bestuur (ROB) onderkent deze wapens (de ‘digitale beproevingen’) en biedt enkele borgingsmogelijkheden (zie tabel)¹⁴.

Drie digitale beproevingen

1. **Desinformatie:** een vrees is dat door digitalisering desinformatie sneller vervaardigd en (gericht) verspreid kan worden. Denk aan nepnieuws, complottheorieën en propaganda.
2. **Desintegratie:** een risico is dat door digitalisering parallelle werelden ontstaan die niet meer met elkaar kunnen worden verbonden, bijvoorbeeld door filterbubbels, echokamers en ‘digitale schandpalen’.
3. **Despotisme:** een gevaar is dat burgers door digitalisering makkelijker beïnvloed kunnen worden zonder dat dat zij dit door hebben, zoals bij heimelijke micro targeting door bedrijven of (buitenlandse) politici.

Borgen van waarheidsvinding

1. Vergroot het vertrouwen in instituties voor waarheidsvinding door zelf het goede voorbeeld te geven.
2. Maak burgers weerbaar tegen desinformatie door kritisch burgerschap te stimuleren.
3. Voorkom desintegratie door met anderen te bouwen aan platforms voor democratische gedachtewisseling.
4. Bestrijd despotisme door tegenmachten te organiseren. Doorbreek waarheidsschroom door het gesprek over waarheid te blijven voeren.

De ROB geeft in haar rapportage ‘Zoeken naar waarheid’ een inspirerende filosofische verkenning naar wat waarheid is of kan zijn en wat dat betekent voor waarheidsvinding. Er zijn vele waarheden, afhankelijk van de bril waarmee het individu naar de wereld kijkt. Juist daarom moet het individu diverse meningen kunnen zien en de onderliggende data kunnen beoordelen op aspecten als juistheid, volledigheid en tijdigheid. Op zijn minst moet bekend zijn hoe de data zijn verwerkt (bijvoorbeeld welk algoritme gebruikt is), wanneer die zijn verzameld, waarvoor en door wie. Het is jammer dat de ROB deze basisvoorwaarde om informatie, meningen en kennis te kunnen beoordelen niet expliciet heeft genoemd bij de borgingsmogelijkheden van waarheid.

Op zijn minst moet bekend zijn hoe de data zijn verwerkt (bijvoorbeeld welk algoritme gebruikt is), wanneer die zijn verzameld, waarvoor en door wie

Temeer omdat het voldoen aan deze voorwaarde de digitale weerbaarheid faciliteert van burgers die dat willen. Wij^{iv} denken dat dit effectiever is dan het blijven voeren van ‘het gesprek over de waarheid’. Hoe realistisch is het voeren van dit gesprek? Hebben burgers wel zin en gelegenheid om het gesprek met je te voeren, om kritische burgers te zijn? Volgens Kahneman hebben mensen de neiging te kiezen voor het gemak en de korte termijn; men vindt het lastig om na te denken over risico’s of het ongemak op langere termijn¹⁵. En die digitale diensten zijn eigenlijk best wel gemakkelijk... De ROB beveelt ook aan om in te zetten op het vergroten van vertrouwen in instituties, te beginnen met het geven van het goede voorbeeld. Daar moet zeker aan gewerkt worden¹⁶. Er moet echter meer gebeuren, waaronder het effectief monitoren of de instituties wel borgen in hun werkwijze wat zij hebben toegezegd. Een voorbeeld hiervan zijn de toegezegde *checks and balances* bij het dilemma veiligheid versus privacy, naar aanleiding van de discussie rond de wetsvoorstellen Computercriminaliteit III en modernisering Wet op de inlichtingen- en veiligheidsdiensten 2017 (Wiv of ‘sleepwet’).

Het rapport 'Zoeken naar waarheid' illustreert het misbruik van data (informatie) en het gevaar voor 'de waarheid' en onze vrijheid van oordeelsvorming en bewegen, onze (digitale) autonomie. De aanbevelingen lijken 'hoog over', terwijl de meer concrete wet- en regelgeving al beperkingen heeft.

Een voorbeeld is de werking van de Algemene Verordening Gegevensbescherming (AVG) of General Data Protection Regulation (GDPR)¹⁷. Deze schrijft voor: "Natuurlijke personen moeten bewust worden gemaakt van de risico's, regels, waarborgen en rechten in verband met de verwerking van persoonsgegevens, alsook van de wijze waarop zij hun rechten met betrekking tot deze verwerking kunnen uitoefenen" (excerpt grond 39-art. 5). Als de toelichting weg wordt gestopt in FAQ's en Algemene Voorwaarden, wordt mogelijk voldaan aan de regel. Maar als je er bewust naar moet zoeken, is het de vraag of er ook wordt voldaan aan de intentie van de regel, 'de geest'. Kunnen mensen hun rechten makkelijk vinden en begrijpen? Verder geeft de AVG zes rechtsgronden voor de verzameling van persoonsgebonden gegevens¹⁸. Eén van de rechtsgronden is 'het algemeen belang', een ander is 'toestemming van de persoon'. Door slimme apparaten als sensoren, camera's en trackers in de openbare ruimte, worden veel data verzameld die niet nodig zijn voor 'het algemeen belang'. Hoe te zorgen dat dit gebeurt met toestemming van het individu?



Er zijn andere voorbeelden, bijvoorbeeld het gebruik van cookies. Veel bedrijven maken het moeilijk om cookies in te stellen. Het instellen kost veel tijd, los van of je begrijpt wat die cookies doen. Het is daardoor onoverzichtelijk en ongemakkelijk, terwijl de AVG voorschrijft dat dergelijke voorzieningen overzichtelijk en in duidelijke, eenvoudige taal moeten worden aangeboden. Een ander voorbeeld is de PSD2, de nieuwe Europese wet (richtlijn) voor het betalingsverkeer van consumenten en bedrijven. Deze kent ook risico's voor de autonomie van het individu, ondanks de goede bedoelingen¹⁸. Verder zijn niet alle ondernemers even integer bij de borging van de privacy van het individu en de regelgeving die daarvoor geldt, terwijl handhaving lastig is. Al is het maar omdat er te weinig (deskundige) mensen beschikbaar zijn. De moeilijkheid om de autonomie te borgen in de openbare ruimte wordt vergroot doordat slimme apparaten door voortschrijdende technologische ontwikkeling steeds kleiner worden en dus moeilijker te herkennen. Verder opereren digitaliseringsorganisaties in een virtuele wereld waar allerlei grenzen op zijn minst doorzichtig zijn, ook de nationale grenzen. Ook dat bemoeilijkt de handhaving.

Effectieve wet- en regelgeving en de bijbehorende handhaving zijn nodig voor de bescherming en facilitering van 'het algemeen belang', waar bijvoorbeeld de veiligheid van identiteit en autonomie onder vallen van individu, groep en organisatie. Wet- en regelgeving is een belangrijk instrument van maatschappij en overheid; een overheid die zoekt naar haar rol in de digitale wereld.

<...> So what can we do to protect and preserve a local imagined community, and mobilise it as a force for good? <...>

- 1. Create a safe and secure urban environment. If we fail to ensure that everyone feels equally safe and secure, we cannot expect all our citizens to contribute to a strong civic democracy with a shared identity. A city that is strong and resilient.*
- 2. Protect the public space. The city belongs to everyone. Our neighborhoods, squares and streets are public spaces that we must protect so that all citizens are assured of a safe and peaceful life. It is their space. <...>*

Uit: speech Burgemeester Amsterdam, NYC: Columbia University d.d. 9 april 2019

Die rol van de overheid moet worden geconcretiseerd, met inbegrip van vragen waar de grens ligt bij het verzamelen en het gebruik van persoonsgebonden data door de overheid voor de openbare orde en veiligheid, bij het beheer van de openbare ruimte en bij haar digitale dienstverlening aan burgers¹⁹. Het gaat om vragen als: hoe ver de overheid mag gaan om onze digitale vrijheid te beschermen; hoe ver ze mag gaan om de openbare orde en de veiligheid te borgen en te handhaven; en hoe ver in de rationalisering van de oordeelsvorming op basis van data²⁰. Antwoorden op die vragen zijn niet altijd e nduidig te geven. Er zitten terugkerende dilemma's bij, operationeel-tactisch²¹ en bestuurlijk, bijvoorbeeld het dilemma data verzamelen voor veiligheid versus privacy. Aan het verzamelen van data zitten risico's voor ieders autonomie. Persoonsgebonden data zijn een bron van kennis over de voorkeuren van een persoon en haar naasten.

Ambtenaar: Met behulp van bigdata-technieken als machine learning kan men de kans op (dus ook het risico op) een bepaalde uitkomst berekenen. Bijvoorbeeld het risico op voortijdig schoolverlaten, of op uithuisplaatsing. De voor veel van mijn collega's aantrekkelijke gedachte is om een groep te identificeren waarbij sprake is van een verhoogd risico en daar dan beleid op te maken. Daarmee ben je eigenlijk aan het profileren, wat niet mag. Bovendien is de interventie ten aanzien van individuen die uiteindelijk op basis van deze identificatie plaatsvindt door ons als overheid niet goed uit te leggen. Er is geen (democratisch vastgestelde) regel met een vast criterium, maar een lerend algoritme dat zelf steeds een nieuw criterium op basis van steeds weer nieuwe data vaststelt.

Een bijkomend probleem is dat de interventies als zodanig niet voorafgaand aan de uitvoering van het onderzoek en de in dit kader gewenste datalevering worden beschreven, maar integendeel onder het containerbegrip 'Beleidsonderzoek' worden gevat. Op deze manier is afwijzing van het dataverzoek op basis van onrechtmatig gebruik onmogelijk. Mag dat?

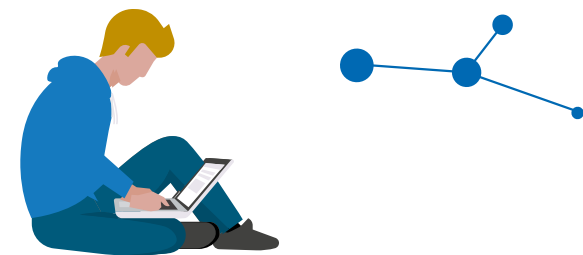
Uit: Engers-Lobrecht, R. van; Hooghiemstra, T.; Ligthart, J.; Haans, E.; Wassing, A.; (2018: 23)

1.2 Data, basis voor verdienmodellen

Persoonsgebonden data zijn geld waard. Op basis van persoonsgebonden data kan een gedetailleerd persoonlijk profiel worden opgebouwd van culturele, economische, politieke, religieuze, seksuele, sociale voorkeuren en interesses. Met dit profiel kunnen onze mening, onze keuzes en ons gedrag be nvloed worden. Om die reden heeft men de meeste interesse in de data over wat het individu digitaal doet en wil doen en waar zij interesse in toont. Men koopt en verkoopt niet voor niets elke handeling die wordt verricht op het internet²². Men laat bijna geen gelegenheid voorbij gaan om persoonsgebonden data te verzamelen. Deze worden op veel plaatsen verzameld, bijvoorbeeld door slimme apparaten in de openbare ruimte, door zoekmachines en spelletjes op het internet en door apps, ook als de apps niet zijn ingeschakeld. Elke keer als iemand een digitaal platform gebruikt, geeft hij of zij data weg via o.a. zoekopdrachten, foto's en berichten, soms ongezien. Terwijl anders is voorgeschreven (AVG; excerpt grond 39-art. 5). Digitale platforms²³ als Facebook (ook eigenaar van Whatsapp en Instagram), Google en Yahoo zijn grote handelaren in data.

Onderzoekers van het International Computer Science Institute testten onlangs 24.000 apps die gebruik maken van Google Android. Zeventig procent daarvan volgt permanent de bewegingen van gebruikers en wisselt gegevens uit met andere apps, ontdekten ze. De wetenschappers stelden Google bijna een half jaar geleden hiervan op de hoogte, maar kregen geen antwoord.

Uit: Thomas, Casper; (2019)





2. De digitale platforms rijk en dominant, riskant voor (digitale) autonomie

We maken het onszelf het liefst gemakkelijk. Hoeveel van ons klikken zonder bewust een keuze te maken op OK wanneer er een melding over cookies verschijnt, en geven zo hun persoonsgebonden data weg? En hoe vaak gebruiken we wifi in een café of op straat, zonder acht te slaan op wie de faciliteit biedt en onder welke voorwaarden, zo nodig inloggend via Facebook? In de Verenigde Staten doet kennelijk ongeveer 76% dat laatste²⁴. Ook kunnen we bijna niet meer zonder platforms als Facebook. Deze hebben inmiddels een grote 'sociale waarde': "Sinds jij van Facebook af bent weet ik niet meer wat je doet". Als bijvoorbeeld lagere scholen gebruik maken van Facebook om events en ander nieuws over de school en klassen te delen in groepen, moet je daar als ouder van een kind op die school eigenlijk ook op zitten. Als al je collega's Whatsapp gebruiken om informatie over projecten te delen, moet je daar eigenlijk ook mee werken. Anders mis je te veel, kan je een 'digital outcast' worden. De grote sociale waarde creëert een *lockin*. Ook zijn we gewend aan zoekmachines, die door het gebruik door velen steeds meer leren. Als je ook gebruik maakt van hun diensten (bijvoorbeeld Gmail van Google), is zo'n lockin al snel aan de orde.

Als al je collega's Whatsapp gebruiken om informatie te delen, moet je daar eigenlijk ook mee werken

De gegevens van jouw contacten en jouw (mail)archief zitten in zo'n dienst. Dit soort zaken maken het moeilijk voor nieuwe aanbieders om marktaandeel te verwerven ten koste van de huidige dominante platforms. Er wordt daardoor nog weinig gebruik gemaakt van alternatieve platforms die geen persoonsgebonden gegevens vastleggen, zoals 'Signal' (ipv Whatsap) en de zoekmachine 'Duck Duck Go' (in plaats van Google en Yahoo)²⁵. Hierbij is sprake van een vicieuze cirkel, waarin 'we' onszelf gevangen maken van enkele platforms die de gedigitaliseerde samenleving anno 2019 domineren.

Die dominantie is niet zonder risico's. De algoritmen van een platform kunnen er voor zorgen dat individuen in een informatietunnel^{VI} van gelijkgestemden terecht komen, waardoor zij zelden of nooit een andere mening of andere argumenten zien dan waar zij al in geloven. Dit effect past bij ons groepsgedrag. Het bevordert ook de slagingskans van beïnvloeding en van op maat gemaakte advertenties (*microtargeting*). Het stimuleert de omzet van digitale platforms. Het stimuleert eveneens de eenzijdigheid van opinie, het wij/zij-denken, het ontstaan van conflicten en radicalisering. Vergeleken met microtargeting, is traditionele reclame als een schot hagel op een mug.

American researchers have found they can use mathematical formulas to segment huge populations into thousands of subgroups according to defining characteristics like religion and political beliefs or taste in TV shows and music. Other algorithms can determine those groups' hot-button issues and identify "followers" among them, pinpointing those most susceptible to suggestion. Propagandists can then manually craft messages to influence them, deploying covert provocateurs, either humans or automated computer programs known as bots, in hopes of altering their behavior.

Uit: Calabresi (2017); Time

De risico's van deze effecten raken steeds meer bekend. Het 4th Global Consumer Trust Report (2018)²⁶ van *The Mobile Ecosystem Forum* leidt uit verschillende onderzoeken af dat 57 procent van de gebruikers van mobiele apparatuur het verzamelen van hun persoonlijke data door leveranciers van datadiensten als een risico ziet. 66 procent geeft aan dat ze zelf, of iemand in hun nabijheid data-gerelateerde schade hebben opgelopen. 63 procent geeft aan zelf hun persoonlijke data te willen beheren. Ook neemt de onzekerheid toe. Mensen weten niet welke berichten echt en welke niet echt zijn²⁷. Onduidelijk is wat er van hen wordt vastgelegd en wat daarmee gebeurt. Toch gaan we door met het gebruik van digitale diensten, waaronder platforms. Dat is de 'privacy-paradox'. Het gemak en de korte termijn prevaleren (Kahneman, 2016).



Professional from a large energy company: “I feel extremely visible: check ins on Facebook, everything you post on Twitter, Google, that knows through your phone every step you take pretty much, everything you post using Gmail. I am pretty sure everything is scanned and collected and aggregated”.

Uit: Onderzoek van Shazade c.s.(2018)

Met ons gedrag maken ‘we’ de digitale platforms rijk en versterken ‘we’ hun data-dominantie. In een negatieve vorm kan dit despotisme versterken (ROB, 2019). In een positieve vorm helpt het ons om kennis te ontwikkelen en in vrijheid met elkaar samen te leven. Om het laatste te behouden, moeten we de dominantie van enkelen voorkomen. Dat is lastig.

Digitale platforms verdienen veel geld met persoonsgebonden data en met zichtbaar en onzichtbaar verbonden diensten, waaronder advertenties^{vii}. Met dit geld kunnen ze macht en invloed kopen om hun positie te behouden en te vergroten. Daarnaast kunnen ze ermee investeren in de ontwikkeling van nieuwe, soms gevaarlijke technologie²⁸ en in methoden om het gedrag van individuen en groepen nog sterker en onmerkbaar te beïnvloeden²⁹. Op veel terreinen maken digitale platforms in steeds grotere mate de dienst uit³⁰. Het lijkt er steeds meer op dat de digitalisering ook de economische spelregels verandert: ‘the winner takes all’³¹: “As a result, the internet has become increasingly concentrated, less open and growingly hostile to innovation and entrepreneurship”³².



3. Toenemende druk op (digitale) autonomie, een ‘tipping point’

De voordelen én de risico’s van de digitalisering nemen toe door de ontwikkeling van AI³³. De combinatie van AI, big data en quantum computing³⁴ levert in enkele jaren heel veel extra mogelijkheden op voor zowel positief gebruik als misbruik. Terwijl er nu al zoveel kan...

In het project ‘Database verkeersgegevens’ wordt bijgehouden wat de intensiteit is van het wegverkeer op verschillende soorten wegen (rijks- en provinciale). Voorheen gebeurde dat altijd met telcamera’s, maar nu steeds meer geautomatiseerd doordat auto’s permanent verbonden zijn met het internet en zelf direct data doorgeven. Dat is goedkoper en levert meer bruikbare gegevens op die gebruikt kunnen worden om het vervoer slimmer te organiseren. Deze data worden bovendien in principe allemaal geaggregeerd als open data beschikbaar gesteld.

Dilemma: onlangs bleek tijdens een hackathon dat een paar studenten door het slim combineren van verschillende open datasets in vier uur precies konden achterhalen wie wanneer in welke auto had gezeten. Als grapje hebben ze dat uitgezocht voor de hoogste baas. Maar als zij het kunnen, kan iedereen het.

Uit: Engers-Lobrecht, R. van; Hooghiemstra, T.; Ligthart, J.; Haans, E.; Wassing, A.; (2018: 26)

Anno 2019 is er meer AI dan menigeen denkt. De zoekresultaten in Google, de aanbevelingen in de webwinkels Amazon en Bol.com en het scannen van kentekens bij parkeerhandhaving van de gemeente Amsterdam zijn bijvoorbeeld allemaal op AI gebaseerd. De toepassing van AI zal verder toenemen. Bijvoorbeeld bij de Johan Cruijff Arena in Amsterdam. Daar wordt AI verder ontwikkeld om de klantreis (*customer journey*) van bezoekers te optimaliseren. De gemeente Rotterdam test een andere toepassing van AI: een object herkende camera om de stad schoner te maken en om overlast van groepen op locaties tegen te gaan. De Belastingdienst zet AI in om het risico op fraude te analyseren en voor automatische verwerkingen.



In China gebruikt men AI om een aantal sociale media te scannen op berichten die kunnen wijzen op een voornemen tot zelfdoding. Op basis daarvan worden vrijwilligers ingeschakeld om contact op te nemen met de mensen die die berichten gestuurd hebben³⁵.

Anno 2019 lijkt er nog geen Artificial Intelligence te zijn die de *driver* van ons gedrag en emoties per individu kan beïnvloeden. Ons individuele gedrag lijkt nog de kern van het maatschappelijk denken en werken. Er zijn echter al *bots* gesignaleerd die ons gedrag kunnen analyseren om het meest geschikte moment van beïnvloeding te bepalen³⁶. Niemand weet hoe snel de ontwikkeling van AI zal gaan. Onder deskundigen is er discussie over hoe snel en hoe vergaand de ontwikkeling zal zijn³⁷. Dat AI steeds meer en meer ingrijpende toepassingen zal krijgen is echter vrij zeker. Verwacht mag worden dat AI steeds meer zelfstandig zal kunnen opereren en zich steeds meer zelfstandig zal kunnen door-ontwikkelen. Naarmate AI meer volwassen wordt, en meer functionaliteiten krijgt voor *deep learning*³⁸, zal het ook steeds moeilijker worden om de kwaliteit van de gebruikte databronnen en algoritmen te kunnen monitoren. Die veranderen dan immers zelfstandig. Het wordt in dat geval moeilijk om de kwaliteit van de analyse en de bijhorende oordeelsvorming door de AI te beoordelen.

Dat AI steeds meer en meer ingrijpende toepassingen zal krijgen is echter vrij zeker

‘We’ zitten op een *tipping point*. De komende jaren kunnen er nog effectief maatregelen genomen worden om te voorkomen dat grote organisaties en technologie stap voor stap in toenemende mate ons leven gaan sturen. Bianca Wylie zegt hierover: “*Dystopia doesn’t happen with a big bang. It’s little stuff happening over and over again, until we find ourselves in the positions thinking: Is this really happening? Are we now depending on a corporation to deliver our basic needs like housing, education and safety?*”³⁹. Zuboff (2019) bevestigt deze zienswijze. Stap voor stap leren data scientists hoe emoties en gedrag steeds beter en onmerkbaar beïnvloed kunnen worden. Voordat de technologie hiervoor is uitontwikkeld en in gebruik is genomen, moet er houvast zijn in het fundament van het maatschappelijk bestel én in instrumenten die het individu in staat stellen om haar (digitale) autonomie zelf in te vullen en te borgen. De tijd dringt.

4. (Inter)nationale initiatieven voor houvast (digitale) autonomie

Er zijn verschillende initiatieven die houvast bieden aan de (digitale) autonomie van het individu. Diverse (inter)nationale partijen werken bijvoorbeeld aan de ontwikkeling van een persoonlijke datakluis, zoals in het Europese project DECODE⁴⁰. Ook zijn er veel initiatieven om (gedrags)regels te stellen aan de digitale omgeving. Enkele voorbeelden hiervan zijn:

Internationaal

- Cities4digital rights, een initiatief van Barcelona, New York en Amsterdam, nov 2018⁴¹
- 'The Institute for ethical AI and machine learning' (UK), de acht principes voor AI⁴²
- Ethische principes van AI van de High-Level Expert Group on AI⁴³
- Het Rathenau Instituut, dat in opdracht van PACE (Parlementaire Assemblée van de Raad van Europa) een aantal mensenrechten formuleerde, passend bij de gedigitaliseerde wereld⁴⁴

Rijksoverheid

- BZK, de code goed digitaal bestuur, 2019⁴⁵
- Kamerbrief visie Regie op Gegevens, 2019, waarin o.a. wordt aangegeven dat een burger niet alleen recht heeft te weten waarom welke data bij de overheid over haar zijn vastgelegd, maar ook het recht heeft om die data te gebruiken door die te delen met anderen, op een vertrouwde manier zodat de burger de regie houdt⁴⁶
- Het initiatief om een Agenda Digitale Overheid te ontwikkelen (2018)⁴⁷

Gemeenten

- VNG Realisatie, principes voor de digitale samenleving, 2019⁴⁸
- Amsterdam en Eindhoven: principes digitale stad, jan 2017
- Utrecht met het instrument DEDA⁴⁹
- 'Economische Board (regio) Amsterdam' met TADA, nov 2018⁵⁰

Diverse initiatieven van anderen

- <https://www.teachthought.com/the-future-of-learning/examples-of-good-digital-citizenship>
- 'Taskforce Behoorlijk Datagebruik' van de Kamer van Koophandel⁵¹
- 'A vision for a shared digitale Europe' uit een samenwerking van Kennisland, Centrum Cyfrowe en Commons Network⁵²
- 'Fact, Fair, Robust, Shared principes'⁵³ van digital society (VSNU)
- 'Gedragscode ICT. Voor de digitale economie' & de 'Ethische code Artificial Intelligence (AI)' van Nederland ICT (2019)⁵⁴
- Manifest (2018) van 'Public Spaces', samenwerkende organisaties voor een veilige digitale omgeving⁵⁵

Men vindt: er ontbreken (rechts)normen in de digitale samenleving.

Enkele citaten:

- *In de huidige digitale samenleving is de balans zoek tussen het individuele en het collectieve belang, de private en de publieke sfeer. Er is sprake van twee dimensies digitale vrijheid versus digitale veiligheid, die op elkaar inwerken en waarvan de rechtsnormen ontbreken.*
- *Nieuwe (sociale) protocollen zijn nodig en misschien wel nieuwe instituties om toezicht te houden op het opereren in een digitale wereld. In die zin was de invoering van de GDPR in Europa een goede eerste interventie vanuit het perspectief van de individuele burger en een voorbeeld voor de rest van de wereld. Maar er is nog meer nodig en daar ligt nog wel een opgave.*
- *Er moet een nieuw denkkader komen waarin we meer hybride geïntegreerd kijken: analoog en digitaal.*

Uit: dialoogtafel Amsterdam; 20 mei 2019

De Nederlandse overheid heeft maatregelen genomen om de risico's van de digitale wereld het hoofd te bieden, onder andere met een digitaliseringsstrategie⁵⁶. Deze leidde inmiddels tot bijvoorbeeld: de Nederlandse Cyber Security Agenda (NCSA), het Actieplan Digitale Connectiviteit, Smart City NL Strategie, NL DIGIbeter: Agenda Digitale Overheid, NL DIGITAAL: Data Agenda Overheid, het Actieplan mkb en het Actieplan Digitale Inclusie⁵⁷.

In de EU worden digitale gevaren actief bestreden. In 2018 werd bijvoorbeeld met enkele grote digitale platforms een gedragscode getekend over online desinformatie⁵⁸. De EU is eveneens scherp op de (data)monopolieposities van grote digitale platforms, net als Nederland⁵⁹. Er zijn diverse hoge boetes opgelegd. Ook wordt er gewerkt aan diverse standaarden en voorschriften voor de digitale wereld. Denk aan de regel dat u moet kunnen kiezen welke data er over u verzameld mogen worden op websites, via cookies, en aan de regels over privacy (GDPR/AVG). Eerder is geïllustreerd dat deze wetgeving nog niet sluitend is.

5. Wat nodig is voor versterking (digitale) autonomie

Het is belangrijk dat de komende jaren een effectieve basis wordt gelegd voor behoud van de (digitale) autonomie. Dat wil zeggen: de vrijheid van individuen en groepen om activiteiten, meningen en voorkeuren van welke aard dan ook op eigen kracht vorm te geven en de vrijheid van individuen om ‘geheimen’ te kunnen hebben. Naast dat mensen zelf andere keuzes moeten maken, vraagt dat om Europese democratisch gelegitimeerde maatregelen om de vrijheid te beschermen. Dat vraagt om een breed maatschappelijk gedragen verandering in ‘denken en werken’ (Rathenau, 2018). Dat neemt veel tijd in beslag. Op korte termijn moet gewerkt worden aan de realisatie van enkele belangrijke voorwaarden, namelijk: versterking van de digitale weerbaarheid van burgers; concretisering van de rol van de overheid in de digitale wereld en een digitale omgeving waarin het individu de regie kan voeren over haar data.

5.1 Werken aan het centraal stellen van mensen en waarden

Volgens het Rathenau Instituut⁶⁰ moeten de overheid, het bedrijfsleven, de kennisinstellingen en maatschappelijke organisaties samenwerken om de digitale samenleving doelgericht in te richten vanuit mensen en waarden. Dat traject vergt een lange adem. Het vraagt van iedereen een omslag van bijvoorbeeld een overwegende oriëntatie op individueel, economisch nut naar een overwegende oriëntatie op internationaal, maatschappelijk nut; van meer ‘ik’ naar meer ‘we’, van meer welvaart naar meer welzijn en van beslissingen over business cases met een accent op financiële argumenten naar beslissingen met een accent op

maatschappelijk relevante argumenten. Een digitaliseringsakkoord als voorgesteld door het Rathenau Instituut (2017) kan de beweging naar het meer handelen vanuit publieke waarden stimuleren bij de digitalisering. Dat handelen vraagt van ondernemers onder andere het gebruik van een ander, financieel vaak minder gunstig, op mensen georiënteerd en meer duurzaam businessmodel. Van de overheid vraagt het dat zij het belang van het individu en de groep meer centraal stelt in de democratische context. Het vraagt ook de ruimte om te experimenteren op basis van een idee, met het risico dat dit idee toch niet helemaal oplevert wat er van verwacht werd; daar van te leren en het zo nodig nog een keer te proberen. Van het individu vraagt het een actieve participatie. Gedrag waarvan de historie laat zien dat het maar een beperkte groep mensen gegeven is. Actieve participatie wordt bovendien steeds moeilijker door de toenemende maatschappelijke complexiteit, verwachtingen en de grote druk om te presteren en ‘iemand’ te zijn.

Een digitaliseringsakkoord als voorgesteld door het Rathenau Instituut (2017) kan de beweging naar het meer handelen vanuit publieke waarden stimuleren bij de digitalisering

Bij dit alles gaat het om grote veranderingen in gedrag en in *mindset*. Hoewel er hoopgevende signalen zijn, lijkt de algemene trend anders⁶¹. Volgens het Sociaal en Cultureel Planbureau (2019) denken de Nederlandse burgers op hoofdlijnen min of meer hetzelfde over wat de Nederlandse identiteit omvat, verschillen zijn er wel degelijk. Deze worden versterkt door de sociale media. Deze stimuleren het ontstaan van een gepolariseerd land⁶². Het toenemende individualisme en argwaan tegen democratie en politiek⁶³ versterken samen de ontwikkeling van ‘*cocooning*’⁶⁴. Mensen lijken steeds meer geïnteresseerd te zijn in een kleine kring om hen heen van min of meer gelijk denkenden. Men lijkt steeds minder begaan met het welzijn van anderen, steeds minder geïnteresseerd in andere meningen en argumenten. Het is een vruchtbare bodem voor tunnelvisie en voor polarisatie en radicalisatie.

Leven in een informatietunnel wordt versterkt door onzekerheid en de (algoritmen van) digitale platforms. ROB (2019) nuanceert dit fenomeen; zij meent dat de publieke omroepen bijdragen aan een gemeenschappelijk beeld van de werkelijkheid. Dat de anonimiteit en het grote bereik (de zichtbaarheid) van de digitale wereld echter polarisatie kunnen stimuleren, daar zijn inmiddels veel, ook recente, voorbeelden van, o.a. in Amerika, Egypte, Nieuw Zeeland en Nederland.

De aanbeveling van het Rathenau Instituut stelt daardoor veel meer eisen dan waar op korte termijn aan kan worden voldaan. Vaak vraagt het veranderen van gedrag en mindset immers veel tijd⁶⁵. Hoe groter en diverser de groep, hoe moeilijker het is. Laat staan als het gaat om maatschappelijke veranderingen. Het duurde bijvoorbeeld decennia voordat niet-roken min of meer de norm was en klimaatneutraal denken een breder geaccepteerd idee. Ook het maken, vaststellen en aanpassen van wet- en regelgeving vergt veel tijd. Niet alleen door het maatschappelijke draagvlak en de democratische legitimatie die ervoor nodig zijn, ook vanwege de noodzakelijke zorgvuldigheid. Dat laat onverlet dat er aan wet- en regelgeving gewerkt moet worden.

5.2 Voorwaarden: digitale weerbaarheid, concretisering rol overheid, regie over data

Met de ROB zijn wij van mening dat de digitale weerbaarheid van burgers vergroot moet worden. Dat begint bij het bewustzijn wat het betekent - en in de toekomst kan betekenen - wanneer je niet voldoende kritisch bent in het gebruik van digitale diensten of wanneer je niet voldoende kritisch kijkt naar de realiteitswaarde van digitale berichten. En, als je dat niet doet, het bewustzijn over hoezeer je beïnvloed kunt worden en waarschijnlijk al wordt. De kans bestaat dat er groepen mensen zijn die niet erg geïnteresseerd zijn in deze risico's: *“The masses, in general, are addicted to marketing gibberish! Why? Well, it is much easier than figuring things out for yourself”*⁶⁶. Toch moet er wat aan gedaan worden. Dat past bij de rol van de overheid. Een rol die geconcretiseerd moet worden, evenals haar instrumentarium en de ruimte voor de nodige dataverzameling en het datagebruik door de overheid en instituties met een publieke functie.

Het is verder nodig dat individuen in staat zijn goede data te verzamelen en te gebruiken. Toegang tot betrouwbare data - de basis voor informatie, meningen en kennis - is een basisrecht, net als vrijheid van meningsuiting.

Op zijn minst moet men de kwaliteit van data en informatie kunnen duiden. Informatietunnels moeten voorkomen of goed herkend kunnen worden en worden doorbroken. Die combinatie geeft een goede informatiepositie. Het stimuleert persoonlijke ontwikkeling, toegang tot geld en kwaliteit van leven. Een gelijke, goede informatiebasis voor iedereen was ooit de belofte van het internet (Berners-Lee; 2018). Het waarmaken van die belofte wordt echter in toenemende mate bedreigd door digitale en maatschappelijke ontwikkelingen.

Tenslotte is het nodig dat elk individu beschikkingsrecht heeft over haar persoonsgebonden data, zodat iedereen zelf uit kan maken wie welke data hoe lang en waarvoor mag gebruiken. Ook moet er borging zijn van de identiteit van de afzender, van de verzamelaar van data, en van de identiteit van de objecten waar individuen het over hebben of waar men toegang toe wil verlenen of krijgen. Daarnaast moet geborgd worden dat data ingezet worden conform de afspraken en dat men de data kan corrigeren als men dat nodig acht. Verder moet elk individu inzicht kunnen hebben welke data over haar verzameld worden, door wie en waarvoor. Tenslotte moet geborgd worden dat digitale platforms en slimme apparaten niet meer persoonsgebonden data opnemen dan toegestaan in de wet- en regelgeving, dan toegestaan door het individu en/of dan nodig is voor de veiligheid van maatschappij en stad.

De bescherming van persoonsgebonden data is een grondrecht in Nederland (grondwet, artikel 10, lid 1) en in de EU^{viii}, terwijl 'eigen kracht' al langere tijd een politiek-maatschappelijk issue is. Het is dus zaak om zo snel mogelijk werk te maken van een instrumentarium dat het individu achter het stuur zet bij het verzamelen en het gebruiken van de persoonsgebonden data van zichzelf en haar naasten.



6. Conclusie

De voorgaande schets van de gedigitaliseerde wereld anno 2019 leidt tot de conclusie dat het nu nodig is om enkele basisvoorwaarden te realiseren waarmee het groeiende data-monopolie van digitale platforms wordt tegengegaan en waarmee de verwante risico's voor de (digitale) autonomie van het individu verminderen. Gegeven de voorgaande beschrijving denken wij aan basisvoorwaarden waarmee:

1. elk individu de regie heeft over de persoonsgebonden data, wat er wordt verzameld, wanneer en door wie en wie het gebruikt, waarvoor en hoe lang. Met waarborgen voor de identiteit van personen en objecten en voor de kwaliteit van de data die worden uitgewisseld. Hiervoor is een nationale, onafhankelijke, open vertrouwensinfrastructuur nodig.
2. elk individu kan kiezen welke data over haar verzameld mogen worden in de openbare ruimte, wellicht met een 'datakill-switch'.
3. wordt gestimuleerd dat slimme apparaten werken en gebruikt worden conform de regelgeving. Bijvoorbeeld met een 'datasafety-keurmerk' voor slimme apparaten, toegekend door een democratisch geborgd onafhankelijk instituut met een wettelijke grondslag voor het actief voorkomen en zo nodig aanpakken van datadiefstal en -misbruik, in aanvulling op de GDPR (AVG).

Een vertrouwensinfrastructuur “biedt voorzieningen waarmee data veilig kunnen worden uitgewisseld tussen burgers en aanbieders van diensten waarbij toestemming en gebruik van de eigen data onder controle blijven van de burger. Een dergelijke infrastructuur biedt ook kerndiensten als identificatie, authenticatie, toestemmingen en beveiliging. Publieke en private dienstverleners kunnen daar gebruik van maken voor hun aanbod van diensten waarbij persoonlijke data nodig zijn”⁶⁷. Een vertrouwensinfrastructuur werpt een drempel op voor de macht en invloed van digitale platforms. Het faciliteert dat organisaties en individuen regie hebben over hun persoonsgebonden data. Het borgt de rechtmatigheid, identiteit en authenticiteit van deze data. Onderdelen van zo'n vertrouwensinfrastructuur zijn al beschikbaar. Bestaande initiatieven om data te delen in een vertrouwensinfrastructuur zijn bijvoorbeeld MedMij (voor de zorgsector), Joindata (agrarische sector) en iShare (logistieke sector).



Voorbeelden van faciliteiten voor een vertrouwensinfrastructuur zijn het Tippiq-platform (open source, gefinancierd door Alliander), DigiD en Irma⁶⁸. Daarmee is de vertrouwensinfrastructuur nog niet af. Deze moet worden uitgewerkt en ingevuld.

Overheid, bedrijfsleven en burgers installeren steeds meer camera's en andere sensoren⁶⁹. Veel van die slimme apparaten nemen gegevens op uit de openbare ruimte en over de daar aanwezige personen, direct (bijvoorbeeld met camera-beelden) of indirect (bijvoorbeeld bij het tellen van vervoersmiddelen). De groei van het aantal slimme apparaten in de openbare ruimte zet door. De gemeente kan het gebruik van slimme apparaten beperken via bestaande instrumenten, zoals een vergunning. Daarop is het echter moeilijk handhaven. Het wordt bovendien moeilijker omdat slimme apparaten steeds kleiner worden. Meer praktisch is het om het verzamelen van data met een persoonsgebonden karakter preventief te regelen. Wij denken bijvoorbeeld aan een technologische voorziening en/of een bestuurlijke voorziening, respectievelijk een datakill-switch en/of een datasafety-keurmerk.

Datakill-switch

Bij een datakill-switch wordt een afspraak gemaakt met producenten van slimme apparaten dat zij in hun apparaat een 'hardware embedded' voorziening opnemen waarmee het verzamelen van persoonsgebonden data door het slimme apparaat kan worden aangezet - standaard staat hij op 'uit' - via een app waar het individu beschikking over heeft. Zo kan elk individu die dat wil de app downloaden op de smartphone en aangeven of zij het oké vinden dat er data over hen wordt verzameld. Zo ja, dan kan het slimme apparaat alles opnemen. Zo niet, dan kan het slimme apparaat alleen de geanonimiseerde data opnemen, waarvan democratisch is vastgesteld dat die nodig zijn om de openbare orde in en de veiligheid en het beheer van de stad te borgen.

Datasafety-keurmerk

Met een datasafety-keurmerk wordt aangegeven dat een product voldoet aan specifieke betrouwbaarheids- en veiligheidseisen. Producten mogen niet zonder dit keurmerk de markt op. Uiteraard hoort hier een democratisch gestuurd en onafhankelijk instituut bij met een wettelijke grondslag dat zorgt voor toekenning, monitoring en handhaving. De democratische sturing door burgers^{ix} van zo een instituut is essentieel⁷⁰.

Een wettelijk verbod op datadiefstal of -misbruik, in aanvulling op de AVG, geeft dit instituut een nog betere basis om te handhaven en tot vervolging over te gaan wanneer pogingen tot diefstal en/of misbruik van data bekend worden. Om die reden stellen we voor het keurmerk, het instituut en de bijhorende wet tegelijk te ontwikkelen. Met als voorwaarde dat de mogelijk langere doorlooptijd voor de ontwikkeling van een wettelijke grondslag geen belemmering mag vormen voor de ontwikkeling en realisatie van het keurmerk en het onafhankelijk instituut. Verder bevelen wij aan primair te kiezen voor een nationaal keurmerk en instituut en gelijktijdig in Europees verband het initiatief te ondernemen voor een Europees datasafety-keurmerk.

Zowel de datakill-switch als het datasafety-kenmerk moeten onderzocht worden op haalbaarheid, uitvoerbaarheid en effect.

Dankwoord

Dit essay is mede tot stand gekomen met de waardevolle opmerkingen en suggesties van enkele mensen, waaronder: Jim Janssen (Maastricht); Frans Jorna (Almere); Jos Maessen (Rotterdam); Hans Nouwens (ext. adv.); Patrick Reijnen (Amsterdam); Wim Schimmel (CIO a.i.); Robert de Snoo (Taskforce Behoorlijk Datagebruik, Kamer van Koophandel); Frank Vieveen (Rotterdam); Aantink Yeh (VNG Realisatie) en collega's van het CTO Innovatieteam Amsterdam (o.a. Burcu, Nathalie, Tamas). Manon den Dunnen (politie) gaf inbreng vanuit haar activiteiten rond de vertrouwensinfrastructuur (aanbeveling 1). De whitepaper kan een kapstok zijn bij een verdere ontwikkeling en realisatie er van. Marleen Stikker (Waag) en Mark van den Broek (Eindhoven) inspireerden o.a. tot aanbeveling 3.

Over de auteurs



Theo Veltman werkt vanaf september 1979 in de wereld van organisatie, mens en technologie. Nu bij de gemeente Amsterdam, voorheen was hij o.a. CTO Amstelveen, CIO ministerie en directeur bij en van enkele agencies. Hij had diverse andere management- en adviesfuncties in het onderwijs, de sociale zekerheid, de industrie en de overheid, landelijk en lokaal, meestal a.i.



Rob van Kranenburg schreef in 2007 een artikel over RFID en Internet of Things, uitgegeven door het Instituut voor Netwerkcultuur. Hij werkte bij de UvA, Balie, Doos of Perceptiën (Flow 2003) en Waag Society. In 2009 richtte hij het expertnetwerk #IoT Council (theinternetofthings.eu) op en een jaar later iotday.org om een brede maatschappelijke discussie aan te zwengelen. Hij werkt in de Coordinated Support Action NGI.eu FORWARD, de strategiegroep van het Next Generation Internet Programma.



Referenties

- 1 https://www.youtube.com/watch?v=PIAXG_QcQNU
- 2 Soules, M. (2015). *Media, persuasion and propaganda*. Edinburgh: Edinburgh University. In: Veltman, T. (2017). Waarheidsvinding met Artificial Intelligence en Big Data? Het ontmaskeren van 'framend shamen'. <https://platformoverheid.nl>. Zie ook: Thaler, R. & Sunstein, C. (2008). *Nudge: Improving decisions about Health, Wealth and Happiness*. New York: Penguin Group.
- 3 EU. (2018). Bestrijding van desinformatie, een Europese benadering [Mededeling]. Brussel: EUCommissie 2018/236.
- 4 Casper, T. (2019). Surveillancekapitalisme: Een parasitaire economie. De mens als grondstof. *De Groene Amsterdammer*.
- 5 Garvey, J. (2016) *The Persuaders: The Hidden Industry That Wants To Change Your Mind*. London: Icon Books. En: Veltman, T. (2017). De vrijheid van (dis)connectiviteit, meningsvorming en bewegen: Digitale mensenrechten als waarborg. Geraadpleegd van <https://platformoverheid.nl/> En: Kool, L., J. Timmer, Royakkers, L. & Est van, R. (2017). *Opwaarderen - Borgen van publieke waarden in de digitale samenleving*. Den Haag, Rathenau Instituut. En: Heezen, M., Louwerse, D. & Riedstra, E. (2018). *Smart City? Graag. Maar dan wel met bewuste burgers!* Amsterdam: Platform31.
- 6 Berlinquette, P. (2018). Deel 1: How Google Tracks Your Personal Information. Geraadpleegd in maart 2019, van <https://medium.com/s/story> En: Berlinquette, P. (2018). Deel 2: How Google Tracks marketeers exploit your discomfort. Geraadpleegd in maart 2019, van <https://medium.com/s/story>
- 7 <https://www.pri.org/stories/2019-06-13/internet-deepfakes-threaten-truth-and-reality> Geraadpleegd in juni 2019. Ter illustratie.
- 8 Kouwenhoven, A., Poel van der, R. (2019). Hoezo ben ik niet echt? Ik chat toch met je? NRC.
- 9 Calabresi, M. (2017). Inside Russia's Social Media War on America Geraadpleegd van <https://time.com/> En: Abrams, A. (2019). This is what we know so far about Russia's 2016 meddler. Geraadpleegd in met 2019 van <https://time.com/> En: Check Point Research. (2019). *Security Report 2019*. Tel Aviv: Check Point Research. En: Karim, A. & Noujaim, N. (2019). *The Great Hack* [Netflix]. Documentaire over de use case van het datalek bij Facebook en het gebruik ervan door Cambridge Analytica in 2015-2016. Gegevens van tenminste 87 miljoen Facebook-klanten werden zonder hun toestemming gebruikt om verkiezingen in de VS te beïnvloeden, waaronder de presidentsverkiezingen in 2016.
- 10 Hsu, T. (2019). These Influencers Aren't Flesh and Blood, Yet Millions Follow Them. *New York Times*.
- 11 Check Point Research. (2019). *Security Report 2019*. Tel Aviv: Check Point Research: "In late January 2018, the 'IoTroop' botnet, discovered by Check Point researchers in October 2017, launched its first attack against the financial sector. IoTroop is a powerful internet of things (IoT) botnet comprised primarily of compromised home routers, TVs, DVRs, and IP cameras. The first attack used 13,000 IoT devices across 139 countries to target a financial organization with a DDoS attack, followed by two more attacks against similar targets within 48 hours." Ter illustratie.
- 12 Munnichs, G., Kouw, M. & Kool, L. (2017). Een nooit gelopen race - Over cyberdreigingen en versterking van weerbaarheid. Den Haag: Rathenau Instituut. Ter illustratie: 'de live cyber attack threat map', te raadplegen via <https://threatmap.checkpoint.com/ThreatPortal/livemap.html>
- 13 Sanger, D. E. & Perloth, N. (2019). U.S. Escalates Online Attacks on Russia's Power Grid. *New York Times*. En: Satariano, A. (2019). Russia Sought to Use Social Media to Influence E.U. Vote, Report Finds. *New York Times*. En: Plicht van der, J. (2019). Interview met landmachtcommandant: 'Feitelijk is oorlog al aan de gang'; nu.nl, gezien 29 augustus 2019: "Als de kans op oorlog nul was, dan hadden we geen leger nodig. De kans is nooit nul. Ik kan alleen niet inschatten hoe groot de kans is. Maar we moeten er wel op voorbereid zijn dat er oorlog uitbreekt. Feitelijk is die ook al aan de gang, met spionage en cyberaanvallen."
- 14 Raad voor het Openbaar Bestuur. (2019). *Zoeken naar waarheid*. Den Haag: ROB.
- 15 Kahneman, D. (2016). *Ons feilbare denken. Thinking, fast and slow*. Antwerpen: Business Contact. En: Senge, P. M. (1990). *The fifth discipline: The Art and Practice of the Learning Organization*. NYC: Doubleday/Currency.
- 16 Lange de, P. (2019). DUO overtreedt privacyregels met volgsoftware in e-mails aan studenten. *De Volkskrant*. En: Outeren van, E. (2018). Overheid stelt met big data fraudebescherming boven privacy. NRC. En: Martijn, D., Tokmetzis, D. (2016). Deze wetenschappers luiden de noodklok: Big data heeft een tegenmacht nodig. *De Correspondent*. En: Hartholt, S. (2016). Trackers Gelderse provinciewebsites in kaart gebracht. *Binnenlands Bestuur*. Ter illustratie.
- 17 AVG. (2016, 27 april). VERORDENING (EU) 2016/679 VAN HET EUROPEES PARLEMENT EN DE RAAD betreffende de bescherming van natuurlijke personen in verband met de verwerking van persoonsgegevens en betreffende het vrije verkeer van die gegevens en tot intrekking van Richtlijn 95/46/EG. Brussel: EU parlement en Raad. Geraadpleegd in augustus 2019 van <https://www.privacy-regulation.eu/nl/index.htm> Dit is de 'Algemene Verordening Gegevensbescherming' (AVG) of 'General Data Protection Regulation' (GDPR) met onder andere artikel 5: de 'beginselen inzake verwerking van persoonsgegevens' en artikel 6: gronden voor 'rechtmatigheid van de verwerking', namelijk: toestemming van de persoon, het algemeen belang, vitale belangen van personen, behartiging van gerechtvaardigd belangen (bv een kind), een wettelijke verplichting en een overeenkomst.
- 18 Croon, T., Boogaard van den, L. & Parie, J. (2019). Geef bankzaken niet te makkelijk prijs. NRC.
- 19 Raad van State. (2018). *Ongevraagd advies over de effecten van de digitalisering voor de rechtsstatelijke verhoudingen*. Den Haag: Raad van State. En: Franzke, A. & Kloe de, C. (2018). Dringende data verhalen. Bewustwording van knelpunten in dataprojecten. Utrecht: Gemeente Utrecht en Universiteit Utrecht. En: Leeuw, A. (2018). Drukte binnenstad bestrijden met nieuwe technologieën: Binnenlands Bestuur. En: Heezen; M., Louwerse, D. & Riedstra, E. (2018). *Smart City? Graag. Maar dan wel met bewuste burgers!* Amsterdam: Platform31. En: Studiegroep Informatiesamenleving en overheid. (2017). *Maak waar!* Den Haag: ministerie van BZK. En: Kool, L., Timmer, J.; Royakkers, L. & Est van, R. (2017). *Opwaarderen - Borgen van publieke waarden in de digitale samenleving*; Den Haag, Rathenau Instituut. En: Shaw, J., & Graham, M. (redactie) (2017). *Our digital rights to the city*. Meatspace Press.
- 20 Martijn, M. (2019). De datahonger van staten en bedrijven zet veel meer op het spel dan uw privacy alleen. *De Correspondent*. Met een samenvatting van de Godwin-lezing van Maxim Februari uit 2017.
- 21 Engers-Lobrecht, R. van, Hooghiemstra, T., Ligthart, J., Haans, E. & Wassing, A. (2018). *Inventarisatie dilemma's bij behoorlijk datagebruik in de openbare ruimte in Amsterdam*; Den Haag: PBLQ. En: Edwards, A.R., Kool de, D. (2015). *Kansen en dilemma's van digitale democratie - Wat kan digitale burgerbetrokkenheid betekenen voor het Nederlandse parlement?* Den Haag: Rathenau Instituut.
- 22 Weinberg, G. (2019). What are the biggest tracker networks and what can I do about them? Geraadpleegd in juli 2019 van www.quora.com En: Tokmetzis, D., & Martijn, M.; (2016). Dit zijn de stalkers gliuipers en snelle jongens die je de hele dag achtervolgen. Geraadpleegd in januari 2017 van www.decorrespondent.nl
- 23 Mettau, P. (2017). Nederland en de Amerikaanse Platform Bubbels. Geraadpleegd in juli 2019 van platformoverheid.nl. "Platformen met een technologische, economische en sociaal-culturele infrastructuur voor het faciliteren en organiseren van online sociaal en economisch verkeer tussen gebruikers en aanbieders, met (gebruikers) data als brandstof. Het zijn infrastructuren die, als digitale marktplaatsen, gebruikers, makers, aanbieders en afnemers met elkaar verbinden. Waarbij hun grootste kapitaal informatie en interacties zijn in plaats van gebouwen en grondstoffen."

- 24 Bullguard. (2019). New Survey Results Show 76% of Americans Blindly Access Public Wi-Fi. <https://www.bullguard.com>, aangehaald door: Warzel, C. (2019). The Privacy Project. *New York Times*.
- 25 <https://duckduckgo.com>: "Ons privacybeleid is eenvoudig: we verzamelen of delen geen enkele persoonlijke gegevens". Ook geeft men aan dat de zoekgeschiedenis niet wordt opgeslagen en dat de activiteiten van het individu niet gevolgd worden. En: www.signal.org: "Signal messages and calls are always end-to-end encrypted <...> We can't read your messages or see your calls and no one else can either."
- 26 <https://mobileecosystemforum.com/programmes/personal-data/global-consumer-trust-report-2018/>. Geraadpleegd in januari 2019.
- 27 Jameson, S., Richter, C. & Taylor, L. (2018). People's strategies for perceived surveillance in Amsterdam Smart City. *Urban Geography*. DOI: 10.1080/02723638.2019.1614369.
- 28 Slijper, F., Beck, A., Kayser, D. & Beenes, M. (2019). Don't be evil. A survey of the tech sector's stance on lethal autonomous weapons. Utrecht: PAX. Aangehaald in: Berkhout, K., Dupuy, L., (2019). Die 'gamebril' maakt militairen dodelijker. NRC. En: Metz C. (2019). With \$1 Billion From Microsoft, an A.I. Lab Wants to Mimic the Brain. *New York Times*.
- 29 Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books Ltd. En: Zuboff in Naughton, J., (2019): 'The goal is to automate us: welcome to the age of surveillance capitalism'; London: The Guardian, interview: "It is no longer enough to automate information flows about us; the goal now is to automate us. These processes are meticulously designed to produce ignorance by circumventing individual awareness and thus eliminate any possibility of self-determination". En: Popcorn, F. (januari 2018): "Popcorn sees Emo-Surveillance rising, as our moods are mined by AI. Facial recognition will track our emotions, and be used to customize both messaging and products and services in real time to our moods. This constant scanning will also lead to private, AI-free get-togethers, where people can emote freely, without Big Brother watching". Geraadpleegd in juli 2019: <https://www.faithpopcorn.com/press/the-shocking-future-faith-popcorn-predicts-the-year-ahead.html>
- 30 Berners-Lee, T. (2018). The web is under threat. Join us and fight for it. Geraadpleegd in augustus 2019 van <https://webfoundation.org/2018/03/web-birthday-29/>
- 31 Gils van, S., Bos, J. & Betlem, R. (2019). Takeaway grijpt met meesterzet de macht in Europa. *Financieel Dagblad*. Ter illustratie.
- 32 Tijdens het schrijven van dit essay kondigde de VS een 'antitrust review' aan van de grote digitale platforms <> Wakabayashi, D., Benner, K., & Lohr, S. (2019). Justice Department Opens Antitrust Review of Big Tech Companies. *New York Times*.
- 33 Ratcliff, C. & Bianchini, R. (2019). *Artificial Intelligence: Challenges for EU Citizens and Consumers*. Brussel: EU Parlement, briefing. En: Raad van State. (2018). *Ongevraagd advies over de effecten van de digitalisering voor de rechtsstatelijke verhoudingen*. Den Haag: Raad van State.
- 34 https://en.wikipedia.org/wiki/Quantum_computing
- 35 Vervaeke, L. (2019). Levens redden met een algoritme. *De Volkskrant*. Ter illustratie.
- 36 Gesprek met Chris Kubecka, CEO HypaSec, cybersecurity onderzoeker; gesproken in juli 2019.
- 37 Meulder, M. (2019). Machines die slimmer worden dan wij? Kunstmatige intelligentie is een grote belofte, maar overdrijven we niet een beetje? *Financieel Dagblad*.
- 38 Metz, C. (2019). With \$1 Billion From Microsoft, an A.I. Lab Wants to Mimic the Brain.; *New York Times*. Geraadpleegd in juli 2019 van <https://www.nytimes.com/2019/07/22/technology/open-ai-microsoft.html> Ter illustratie dat het denkbaar is dat het er komt. Er wordt flink in geïnvesteerd.
- 39 Bianca Wylie (digital civil rights activist bij Digital Rights Now); uitgesproken tijdens de conferentie d.d. 19 juni 2019; <https://wemakethe.city/production/cities-for-digital-rights?ref=17238&lang=en&q=digital> %3B; gezien op Waag (2019): <https://waag.org/en/article/thats-not-business-model-thats-government-function>
- 40 <https://decodeproject.eu/> Geraadpleegd in 2018.
- 41 <https://citiesfordigitalrights.org/>
- 42 <https://ethical.institute/index.html> Geraadpleegd in juli 2019.
- 43 High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. Brussel: European Commission.
- 44 Est van, R., & Gerritsen, J.B.A. met assistentie van Kool, L. (2017). *Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality – Expert report written for the Committee on Culture, Science, Education and Media of the Parliamentary Assembly of the Council of Europe (PACE)*. Den Haag: Rathenau Instituut. Zie ook het overzicht van voorgestelde ethiek rond AI op <https://www.rathenau.nl/digitale-samenleving/overzicht-van-ethische-codes-en-principes-voor-ai>
- 45 <https://www.digitaleoverheid.nl/actielijn/code-goed-digitaal-bestuur/>
- 46 <https://www.rijksoverheid.nl/documenten/kamerstukken/2019/07/11/kamerbrief-visie-regie-op-gegevens> Geraadpleegd in 2019.
- 47 Overheidsbrede Beleidsoverleg Digitale Overheid. (2018). *NL DIGibeter. Een digitale overheid*. Geraadpleegd in 2018 van <https://www.digitaleoverheid.nl/nl-digibeter2019/>
- 48 <https://www.vngrealisatie.nl/nieuws/uitvraag-principes-voor-de-digitale-samenleving>
- 49 <https://dataschool.nl/deda/> Geraadpleegd in 2018.
- 50 <https://tada.city/en/home-en/>
- 51 <https://www.kvk.nl/over-kvk/media-en-pers/nieuws-en-persberichten/eerste-resultaten-taskforce-maatschappelijk-verantwoord-datagebruik/> Geraadpleegd in juli 2019.
- 52 Bloemen, S., Tarkowski, A., & Keller, P. (2019). A vision for a shared digital Europe. Geraadpleegd in augustus 2019 van <https://shared-digital.eu/vision/>
- 53 <https://www.thedigitalsociety.info/nl/over/dataprincipes> Geraadpleegd in mei 2019.
- 54 <https://www.nederlandict.nl/> Geraadpleegd in juli 2019.
- 55 <https://publicspaces.net/manifesto/>
- 56 EZK. (2018). *Resultaten en opbrengsten van de conferentie Nederland digitaal 2019*. Den Haag: EZK. En: EZK. (2018). *Nederlandse digitaliseringsstrategie 2018-2021*. Nederland digitaal. Hier kan het. Hier gebeurt het. Den Haag: EZK: Geraadpleegd van <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/publicaties/2019>
- 57 EZK. (2019). *Nederlandse Digitaliseringsstrategie 2.0*. Den Haag: EZK.
- 58 EU. (2018). *Bestrijding van desinformatie, een Europese benadering*. Brussel: EUCommissie 2018/236, mededeling. En: EU. (2018). https://europa.eu/rapid/press-release_STATEMENT-18-5914_en.htm Geraadpleegd in juli 2019.
- 59 Broekhuizen, K.; (2019); 'Monica Keijzer wil dominantie van digitale platforms indammen'; Amsterdam: Financieel Dagblad.
- 60 Kool, L. E. Duijso en Est van, R. (2018). *Doelgericht digitaliseren – Hoe Nederland werkt aan een digitale transitie waarin mensen en waarden centraal staan*. Den Haag: Rathenau Instituut. En: Kool, L., Timmer, J., Royakkers, L. & Est van, R. (2017). *Opwaarderen - Borgen van publieke waarden in de digitale samenleving*. Den Haag: Rathenau Instituut.
- 61 Eisinga, R., Scheepers, P. & Bles, P. (2012). *Waardenverandering in Nederland: resultaten van de SOCON-enquête: 1980-2011*. Den Haag: CBS. Met BZK. (2019). *Trends en ontwikkelingen*. Den Haag. Geraadpleegd in augustus 2019 van <https://kennisopenbaarbestuur.nl/thema> En: SCPB. (2019). *Denkend aan Nederland. Sociaal en cultureel rapport 2019*. Den Haag. In 30 jaar van een oriëntatie op overwegend familialisme naar

- een meer hedonistische oriëntatie ('het dikke ik'; Kunneman) en naar meer maatschappijkritisch, ook als men familialistisch is (mondige burgers); van een redelijk homogene, geordende en stabiele samenleving naar één anno 2018/2019 met een grote variëteit en diversiteit en een toenemende polarisatie en populisme.
- 62 SCPB, Beugelsdijk, S. (RUG) (Red.); Hart de, J., Houwelings van, P. & Versantvoort, M. (2019). *Denkend aan Nederland. Sociaal en cultureel rapport 2019*. Den Haag: SCPB.
- 63 Lange de, R. (2019). *De onvervulde belofte. sociale, economische en politieke analyse van de tijd waarin wij leven*. Den Bosch: Gompel&Svacina Bvba. En: kranten als het NRC, Financieel dagblad, The Guardian, The New York Times.
- 64 De term is in 1981 gemunt door Faith Popcorn.
- 65 Senge, P. M. (1990). *The fifth discipline: The Art and Practice of the Learning Organization*. NYC: Doubleday/Currency.
- 66 Sabirović, A. (2018, Juni). Artificial Intelligence, AI will turn into aiaiai. Geraadpleegd in december 2018 van [linkedin.com](https://www.linkedin.com).
- 67 Dunnen den, M., Kronenburg, B. & Zalm van der, E. (2019). *Een publieke vertrouwensinfrastructuur voor persoonlijke data*. Amsterdam. Dit is een whitepaper over dit onderwerp waar een aantal publieke instellingen aan meewerkt.
- 68 Irma (<https://privacybydesign.foundation/irma/>)
- 69 McCabe, B. (2019). What is a Digital Twin and How is it Changing AI and IoT? Geraadpleegd van www.linkedin.com. Gartner haalt aan dat hij in 2018 verwachtte dat er in 2020 wereldwijd rond de 21 biljoen sensoren zullen zijn in de openbare ruimte. En: Naafs, S. (2017). De muren hebben sensoren. *De Groene Amsterdammer*.
- 70 Mounk, Y. (2018). The undemocratic dilemma. Geraadpleegd in juli 2019 van <https://www.journalofdemocracy.org> En: speech burgemeester Amsterdam; NYC: Columbia University; 9 april 2019. En: publiekdenken.nl (2019, juli). Nederlanders willen begripvolle overheid. Geraadpleegd in juli 2019. Interview met Reinier Van Zutphen, Nationale Ombudsman, n.a.v. het onderzoek naar de relatie burger/overheid, waarin staat: "Burgers maken zich bijvoorbeeld zorgen over een overheid die zich steeds verder terug trekt. Ze moeten het zelf oplossen of kunnen terecht bij particuliere organisaties die publieke taken uitvoeren."

- I (Digitale) autonomie: de vrijheid van individu en groepen om activiteiten, meningsvorming en voorkeuren van welke aard dan ook op eigen kracht vorm te geven en de vrijheid van individuen om die activiteiten bewegingen, meningen, voorkeuren e.d. voor zichzelf of in kleine kring te houden.
- II 'De Algemene verordening gegevensbescherming (AVG) geeft aan dat een persoonsgegeven <<ook: persoonsgebonden data>> alle informatie is over een geïdentificeerde of identificeerbare natuurlijke persoon. Dit betekent dat informatie ofwel direct over iemand gaat, ofwel naar deze persoon te herleiden is. Gegevens van overleden personen of van organisaties zijn geen persoonsgegevens volgens de AVG'; autoriteitpersoonsgegevens.nl. Bron: artikel 4, AVG; <http://www.privacy-regulation.eu/nl/4.htm>
- III (digitale) autonomie: de vrijheid van individu en groepen om activiteiten, meningsvorming en voorkeuren van welke aard dan ook op eigen kracht vorm te geven en de vrijheid van individuen om die activiteiten bewegingen, meningen, voorkeuren e.d. voor zichzelf of in kleine kring te houden.
- IV Met wij worden de auteurs van dit essay bedoeld met 'we', het collectief van individuen.
- V 'De Algemene verordening gegevensbescherming (AVG) geeft aan dat een persoonsgegeven <<ook: persoonsgebonden data>> alle informatie is over een geïdentificeerde of identificeerbare natuurlijke persoon. Dit betekent dat informatie ofwel direct over iemand gaat, ofwel naar deze persoon te herleiden is. Gegevens van overleden personen of van organisaties zijn geen persoonsgegevens volgens de AVG'; autoriteitpersoonsgegevens.nl. Bron: artikel 4, AVG; <http://www.privacy-regulation.eu/nl/4.htm>
- VI Informatietunnel = echokamer en informatiefilter; ROB, 2019
- VII Bij Google komt ca 85% van de omzet uit advertenties; bij Facebook ca 98%. Uit: Hueck, H.; Financieel Dagblad, 25 juni 2019
- VIII "Krachtens artikel 8, lid 1, van het Handvest van de grondrechten van de Europese Unie (het „Handvest”) en artikel 16, lid 1, van het Verdrag betreffende de werking van de Europese Unie (VWEU) heeft eenieder recht op bescherming van zijn persoonsgegevens"; zie: <https://www.europarl.europa.eu/factsheets/nl/sheet/157/bescherming-van-persoonsgegevens>
- IX Bij de democratische sturing bedoelen wij niet sturing via het model van de vertegenwoordigende democratie, maar een directe vorm van democratische beïnvloeding, niet zijnde referenda. Zie ook o.a. Mounk, Yascha; ([journalofdemocracy.org](https://www.journalofdemocracy.org); 2018); het interview met Van Zutphen, Reinier; (publiekdenken.nl; 2019)

Overzicht

Beleidsmaatregelen uit essays

Hergebruik van gegevens in smart cities

Juridische en ethische kaders voor big data in de openbare ruimte

Prof. dr. mr. ir. Bart Custers

1. Gemeentelijke overheden moeten bij smart cities inzetten op behoorlijk hergebruik van gegevens via publiek-private samenwerkingen.
2. Gemeentelijke overheden moeten bij smart cities burgers actief betrekken bij afwegingen.
3. Gemeentelijke overheden moeten bij smart cities zorgen voor transparantie met betrekking tot gegevens en afwegingen.
4. Gemeentelijke overheden moeten via pilots smart cities verder ontwikkelen.

Digitalisering, gedragsbeïnvloeding en de overheid

Het gladde ijs tussen nudging en überveillance

Prof. dr. Wijnand IJsselsteijn

1. Wanneer de overheid gebruik wil maken van technologie ten behoeve van gedragsverandering in het publieke domein, dan dient dit transparant te gebeuren, op basis van democratische besluitvorming, ethische toetsing en wetenschappelijke evidentie, en met expliciete betrokkenheid en instemming van partijen die deze technologie direct of indirect zal gaan raken. De overheid dient hierbij te allen tijde proactief, open, begrijpelijk, en eerlijk te communiceren over digitalisering in het publiek domein, of plannen daartoe, op alle niveaus – van landelijk en provinciaal tot gemeentelijk en wijkniveau.
2. Betrek vroegtijdig en continu alle partijen die, direct of indirect, in aanraking komen met digitale gedragsbeïnvloeding in de openbare ruimte, door middel van een proces van participatoir ontwerp.

Neem extra moeite om ook de meest kwetsbare groepen uit de maatschappij een actieve stem te geven in dit proces. Neem ook ethische waardes mee in het ontwerp. De overheid heeft de primaire verantwoordelijkheid om dit proces te organiseren, samen met haar maatschappelijke partners.

3. De overheid moet haar eigen digitale competenties vergroten, en dient zich hierbij gedegen op de hoogte te stellen van de mogelijkheden en beperkingen van de meest recente vormen van digitale technologie. Een technologisch naïeve overheid is een makkelijke prooi voor kwaadwillende partijen, een slechte (onderhandelings)partner voor de industrie, en een tekortschietende hoeder van het publiek belang.
4. Data-eigendom en privacy van burgers in de publieke ruimte dienen te worden gerespecteerd en beschermd. De overheid dient geen gebruik te maken van onbewezen, onveilige of te weinig robuuste technologie. Alternatief, onvoorzien, of illegaal gebruik van de technologie en de data moet in kaart worden gebracht en gereguleerd.
5. De overheid moet een constructief-kritisch maatschappelijk (tegen)geluid organiseren. Hierbij moet fors geïnvesteerd worden in de digitale geletterdheid van burgers, met daarbij meer aandacht voor de nieuwste technologische ontwikkelingen en hoe deze de fundamentele ethische waardens binnen onze digitale maatschappij raken. De versterking van technologisch burgerschap kan vele vormen aannemen, bijvoorbeeld debatten, cursussen, living lab-demonstraties, kritische culturele projecten, of design fictions.
6. De overheid moet voldoende en onafhankelijke kritische controle en toezicht organiseren en verankeren in de wet. Hierbij moeten we denken aan continu aanzwengelen en organiseren van het waarden-debat, bespreking van relevante ethische en juridische casuïstiek, en toetsing van governance, juridische en ethische kaders via volksvertegenwoordiging op alle niveaus (gemeentelijke en provinciale raden, Eerste en Tweede kamer). Daarnaast is het van belang het maatschappelijk middenveld en relevante adviesorganen continu te betrekken, en de rol en positie van toezichthouders te versterken. Hierbij moet ook onderlinge afstemming over handavingsdomeinen heen worden georganiseerd.

De bestuurlijke regie over experimentele data en algoritmen

Aanbevelingen voor een nationaal algoritmeregister en rapportagekader

Prof. dr. Gerd Kortuem

1. Richt een Algoritme Rapportage Initiatief op, dat tot doel heeft de maatschappelijke waarde en risico's van algoritmeprojecten in Nederland te documenteren en te volgen.
2. Ontwikkel een Nationaal Algoritme Register dat een effectieve vergelijking en beoordeling van initiatieven op het gebied van data en algoritmen in Nederland mogelijk moet maken.
3. Ontwikkel een Publiek Algoritme Forum om publieke partijen in staat te stellen de werking, het gebruik en de uitkomsten van initiatieven op het gebied van data en algoritmen te bespreken, te bekritisieren en te betwisten.
4. Richt een Data en Algoritme Instituut op ter bevordering van de ontwikkeling van kennis, benaderingen, tools, infrastructuren en normen voor ethisch en verantwoord gebruik van algoritmen.
5. Ontwikkel een Nationale Skills Agenda voor Data en Algoritmen om ervoor te zorgen dat alle belanghebbenden (van burgers tot organisaties) over de juiste vaardigheden beschikken om deel te nemen aan de nieuwe data- en algoritme-economie.

De noodzaak van een digitale omgevingsvisie

Van principes naar praktijk

Prof. dr. Liesbet van Zoonen en dr. Jiska Engelbert

1. Laat gemeenten een digitale omgevingsvisie ontwikkelen in samenspraak met burgers, bestuurders, bedrijven en andere direct betrokkenen, naar directe analogie met de fysieke omgevingsvisie, vóórdat ze aan daadwerkelijke digitale of dataprojecten beginnen.

AI in de digitale samenleving

Kwaliteit van algoritmen en besluitvorming

Prof. dr. Eric Postma

1. School politici en beleidsmakers regelmatig bij op het gebied van datagebruik en de laatste mogelijkheden van AI in relatie tot sociale en juridische aspecten.

2. De digitale samenleving vereist professionals die thuis zijn in zowel de technische als de menselijke en sociale kant van AI. Investeer in opleidingen die dergelijke professionals afleveren, om zodoende het tekort aan professionals aan te zuiveren.
3. Probeer de brain drain van onderzoekers te stoppen door de omstandigheden voor AI-onderzoekers te verbeteren.
4. Investeer in onderzoek om de transitie naar een goede digitale samenleving te faciliteren.
5. Een onafhankelijk CBS-achtig orgaan dat AI-technologie screent op genoemde statistische principes voorkomt mistanden. Gezien de voortdurende upgrades van machinelearning-algoritmen, is een periodieke toetsing noodzakelijk. De overheid zou toetsing verplicht kunnen stellen om de kans op misstanden te minimaliseren.

Data makes the world go round

Voorstel voor onderzoek naar drie instrumenten ter versterking van de (digitale) autonomie

Theo Veltman en Rob van Kranenburg

1. Realiseer een nationale, onafhankelijke en open digitale vertrouwensinfrastructuur voor identificatie, authenticatie en autorisatie van persoonlijke data inclusief de bijbehorende governance op basis van de bestaande bouwstenen zoals Tippiq en Irma. Gebruik de whitepaper (Dunnen c.s.; 2019) als vertrekpunt.
2. Onderzoek het nut van en de mogelijkheden voor de realisatie van een hardware embedded voorziening, een 'datakill-switch' in slimme apparaten zoals sensoren, waarmee het verzamelen van persoonsgebonden data standaard op 'uit' staat. Standaard kan dan alleen data verzameld worden, waarvan democratisch besloten is dat deze anoniem zijn en nodig voor de borging van de openbare orde en de veiligheid en voor het beheer van de openbare ruimte van de stad.
3. Onderzoek het nut van en de mogelijkheden voor de realisatie van een landelijk 'datasafety-keurmerk' door een democratisch geborgd onafhankelijk instituut, op basis van een wettelijke grondslag voor het voorkomen en zo nodig aanpakken van datadiefstal en -misbruik, in aanvulling op de privacywetgeving AVG (Algemene Verordening Gegevensbescherming).



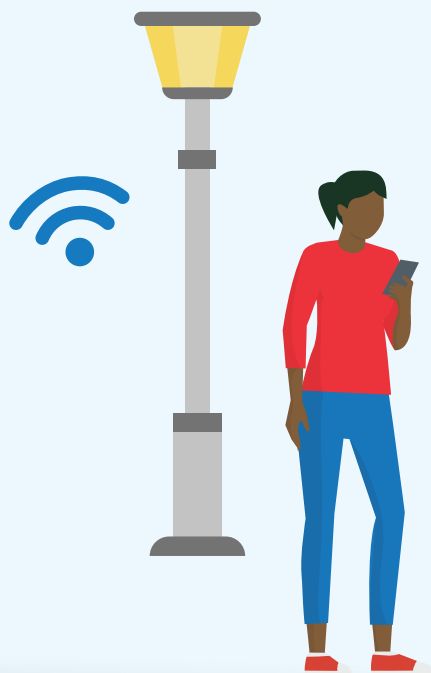
Epiloog: Beleidsuitdagingen bij datagebruik

*Dr. Sarah Giest
Universitair docent*

*Prof. dr. ing. Bram Klievink
Hoogleraar bestuurskunde*

*Prof. dr. Bernard Steunenberg
Hoogleraar bestuurskunde*

*Instituut Bestuurskunde
Universiteit Leiden*



Samenvatting

Beleid maken voor een dynamisch en sterk ontwikkeld terrein als datagebruik door de overheid is uitdagend maar lastig. In deze epiloog kijken wij terug op de verschillende aanbevelingen die gedaan worden in de essays en bespreken wij de mogelijkheden om nieuw beleid te formuleren. Daarbij geven wij aan dat soms beleidsaanbevelingen bredere consequenties hebben dan wordt verwacht en dat het terrein 'datagebruik' voor het formuleren van beleid verder ingekaderd moet gaan worden. In het licht van die onduidelijkheden en onzekerheden bepleiten wij adaptief beleid dat is gebaseerd op een polycentrische co-reguleringsaanpak. Deze aanpak, die in feite een combinatie van verschillende analytische kaders is, wordt in deze epiloog verder uitgewerkt. Dat levert een voorstel op voor een procedureel beleidskader dat de mogelijkheid biedt om in de komende jaren een inhoudelijk beleidskader te gaan ontwikkelen. De belangrijkste stakeholders krijgen een rol in het omzetten van ervaringen met datagebruik in beleid.

1. Inleiding

De verschillende essays geven mooie voorzetten voor het verder ontwikkelen van beleid. Ze gaan over een actueel onderwerp met vele facetten, wat te zien is aan de diversiteit van perspectieven, analyses en voorgestelde maatregelen. In deze epilooog willen wij een verdere handreiking doen ten behoeve van beleidsvorming op het terrein van datagebruik door de overheid. Dan doen wij op de volgende wijze.

In de eerste plaats bespreken we het kader om de verschillende suggesties te structureren, waarmee het duidelijk wordt of de suggesties het beleidsveld in voldoende mate dekken. Terwijl suggesties vaak naar gesignaleerde problemen worden geordend, is het ook belangrijk te bezien hoe beleid ingrijpt in de *'world of action'* (O'Toole 2000: 273). Dat biedt alvast een doorkijkje naar de mogelijke beleidsuitvoering die na de beleidsformulering belangrijk wordt.

In de tweede plaats gaan we in op het zich ontwikkelende beleidsterrein van datagebruik bij de overheid. Vanwege de digitalisering van onze samenleving en de overheid, en de snelle ontwikkeling van de bijbehorende technologie, is het lastig te bepalen wat het probleemveld of domein van dit beleid zou moeten zijn. Datagebruik door de overheid vloeit naadloos over naar de samenleving en het bedrijfsleven, omdat data vooral betrekking hebben op onze samenleving en ook door private partijen worden verzameld.

In de derde plaats gaan wij in op de vraag hoe wij beleid kunnen ontwikkelen in een dynamisch veld dat nog sterk in ontwikkeling is en waarbij het onduidelijk en onzeker is hoe datagebruik (in de brede zin van het woord) zich in de komende jaren verder ontwikkelt. Beleid zal, onder deze omstandigheden, ook nadat het is vastgesteld, open moeten staan voor verdere aanpassingen op basis van leren. In deze bijdrage geven wij aan hoe dat kan, mede gebaseerd op de inzichten vanuit verschillende analytische kaders die zijn ontwikkeld.

Op basis van de inzichten uit de literatuur geven wij een voorbeeld van een meer open beleidsontwerp met het doel daarmee de beleidsmakers te inspireren. We besluiten onze bijdrage met enkele suggesties voor beleidsontwikkeling in een onzekere context, zoals dat het geval is voor datagebruik bij de overheid.

2. Het kader: het structureren van de verschillende suggesties

De essays leveren een veelheid aan boeiende suggesties op voor het ontwikkelen van beleid ten aanzien van het datagebruik door de overheid. Ter voorbereiding op het uitzetten van de verschillende essays heeft het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) met stakeholders knelpunten geïnterpreteerd over het huidige datagebruik door overheden. Dat heeft uitgemond in een knelpuntennotitie waarin problemen zijn benoemd waarmee overheden op dit moment worstelen en waarbij het nog te ontwikkelen beleid een rol zou moeten spelen. Deze knelpunten zijn (BZK 2019):

1. De spanning tussen privacy en techniek;
2. De spanning tussen bedrijven en burgers en tussen het individuele en publieke belang;
3. De spanning met transparantie en uitlegbaarheid van datagebruik in het publieke domein;
4. De spanning tussen doelbinding (naar de voorwaarde die de AVG aanreikt) versus hergebruik van data;
5. De spanning in de samenwerking tussen publieke en private partijen; en
6. De vraag wat de rol van de overheid zou moeten zijn binnen de *smart society*.

Dit zijn punten waarop het beleid zich idealiter zou moeten richten.

De verschillende suggesties in de essays zouden dus een oplossing moeten bieden voor deze problemen. De algemene vraag is of dat ook het geval is, waarbij wij vooral kijken naar het algemene kader en niet zozeer naar de concrete acties die worden voorgesteld.

Om overzicht te krijgen is een eerste stap om de suggesties naar problemen te ordenen. Dat geeft antwoord op de vraag of ieder probleem met een aanbeveling kan worden verbonden. Dat is belangrijk, maar maakt nog niet duidelijk of we ook een redelijk overzicht hebben van alle mogelijke manieren om deze problemen aan te gaan pakken. Daarvoor zijn ook acties van belang die wellicht nog niet zijn opgenomen in de essays. Een structurering naar knelpunten is dus niet zonder meer gelijk aan een structurering naar maatregelen of acties die in het kader van beleid van belang kunnen zijn.

Daarbij zijn andere indelingen van belang, die bijvoorbeeld uitgaan van typologieën van mogelijke maatregelen of manieren waarop beleid invloed kan hebben op de samenleving. Een indeling richt zich op de verschillende niveaus van handelen waarvan beleid uitgaat: een wat aangepaste indeling gebaseerd op de indeling van Kiser en Ostrom (1982) naar ‘niveaus van actie’.¹ Deze niveaus zijn onderling verschillend, maar spelen allemaal een rol wanneer het gaat om uitvoering. De indeling maakt duidelijk dat acties die in het kader van beleid worden voorgesteld op verschillende niveaus kunnen plaatsvinden, maar benadrukt ook dat de gevolgen van die acties en beleid het product zijn van keuzes die op elk van die niveaus worden gemaakt.

De verschillende niveaus van acties zijn:

1. **Het constitutionele niveau:** het kiezen van basisregels die een grondslag moeten zijn voor een maatschappelijke orde (in het geval van datagebruik zou dat basisregels of grondrechten rondom datagebruik kunnen betreffen). Sommigen hebben een voorkeur om hierbij nog een meta-niveau te onderscheiden, wat enigszins overeenkomt met het onderscheid dat Custers maakt tussen juridische en ethische overwegingen bij datagebruik.
2. **Het politieke niveau:** regels over het maken van keuzes binnen de politiek. Hoe nemen we nu beslissingen in een politiek systeem? Met andere woorden, hoe gaan we procedureel met bepaalde vraagstukken om, bijvoorbeeld wie erover gaat en hoe voorstellen worden voorbereid?
3. **Het organisatieniveau:** hoe er door (overheids)organisaties wordt gehandeld en welke regels op dit niveau gehanteerd worden, inclusief de samenwerking met anderen. Op dit vlak is ook de bestuurlijke verhouding relevant, aangezien voor het reguleren van datagebruik de afstemming tussen gemeenten, provincies en het Rijk belangrijk is.
4. **Het operationele niveau:** hoe in de praktijk door verschillende functionarissen en andere stakeholders daadwerkelijk met de materie wordt omgegaan. We zijn daarmee terechtgekomen op het terrein van het feitelijke gedrag. Hier zit veel complexiteit ten aanzien van onder andere de data zelf, de techniek en algoritmen, de verantwoording over inzet daarvan en de waarde-afwegingen die in de uitvoering worden gemaakt.

Wanneer we de verschillende suggesties uit de essays naast deze indeling leggen, valt een aantal zaken op. In de eerste plaats komen *alle* niveaus van actie in de beleidsaanbevelingen naar voren. De relatieve accenten variëren per essay, maar over het geheel bestaat een zekere balans. Wat verder opvalt is dat de aanbevelingen variëren naar hun mate van abstractie: terwijl soms heel concrete aanbevelingen worden gedaan, zijn andere aanbevelingen veel minder concreet. Dat heeft waarschijnlijk te maken met het feit dat het beleidsterrein van datagebruik bij de overheid veel dynamiek kent, terwijl er betrekkelijk weinig ervaring is die als ankerpunt voor beleidsontwikkeling kan worden gebruikt.

Vooraf op het *constitutionele niveau* komt die spanning in verschillende essays naar voren. Om het datagebruik door de overheid in goede banen te leiden en te verbinden met publieke waarden, wordt gezocht naar principes voor publiek datagebruik. Daarbij wordt soms teruggevallen op algemene principes die ook voor andere vormen van overheidshandelen zijn ontwikkeld (zie bijvoorbeeld het essay van Custers), of worden verschillende en soms vergelijkbare waarden genoemd die voor overheden belangrijk zouden moeten zijn. Dat is een boeiende constatering die duidelijk maakt dat wij dit terrein nog verder moeten ontwikkelen. Voor de vormgeving van de verdere zoektocht, waarbij experimenteren een belangrijk onderdeel is (zie het essay van Kortuem), spelen ook waarden een rol. Engelbert en Van Zoonen beargumenteren dit in hun essay en benadrukken in dat verband het belang van de SHARED-waarden.

Om het datagebruik door de overheid in goede banen te leiden en te verbinden met publieke waarden, wordt gezocht naar principes voor publiek datagebruik

Op het *politieke niveau* wordt in de aanbevelingen relatief vaak de burger benoemd, die meer betrokken zou moeten worden bij de besluitvorming over datagebruik (zie bijvoorbeeld de essays van IJsselstein en Custers).

Daarmee hoopt men niet alleen op een vermindering van hun informatieachterstand, maar ook op de vorming van draagvlak bij de burger en wellicht zelfs het ontstaan van een meer proactieve houding. Verdere betrokkenheid van burgers is belangrijk. De vraag is wel of alle gehoopte effecten optreden. Bij veel eerdere experimenten met burgerparticipatie bleek dat niet alle burgers in gelijke mate meedoen, wat tot ongelijkheid leidt. Verder hebben burgers soms zeer verschillende opvattingen, wat niet bijdraagt aan duidelijkheid over de te kiezen handelingsrichting.

De essays bevatten ook beleidsaanbevelingen op *organisatieniveau*, wanneer het gaat om de interactie tussen overheid en private partijen. Omdat data door zowel private als publieke partijen worden verzameld en geanalyseerd, is die samenwerking onvermijdelijk. Ook worden aanbevelingen gedaan over het instellen van nieuwe organisaties met een taak in dit veld. Dat leidt tot een interessante discussie, waarbij ook ingegaan moet worden op de vraag hoe die nieuwe taken en bevoegdheden passen bij die van andere organisaties. Wat verder opvalt is dat samenwerking tussen overheden en de gevolgen van meer datagebruik binnen de overheid in de essays minder aandacht krijgen, terwijl die juist zeer van belang zijn.

In aanbevelingen op het *operationele niveau* wordt aandacht besteed aan onderwijs en het kennis laten maken van overheidsfunctionarissen met nieuwe technologieën, bijvoorbeeld in de vorm van pilots die nu al door decentrale overheden worden uitgevoerd.

Naast het feit dat alle niveaus van actie in de beleidsaanbevelingen naar voren komen, is een tweede constatering dat in de analyses nog beperkt aandacht is voor de *wisselwerking* tussen de vier niveaus. Ook die moet een rol spelen bij beleidsontwikkeling, omdat maatregelen op niveau 1 impact hebben op maatregelen op niveau 2, en die weer impact hebben op maatregelen op niveau 3, enzovoort. Het zou goed zijn om de aanbevelingen meer vanuit dit samenhangende perspectief te benaderen. Met andere woorden: het kiezen van algemene uitgangspunten op het constitutionele niveau heeft gevolgen voor aanbevelingen op de onderliggende niveaus, waaronder het organisatie- en het operationele niveau.

De samenhang en wisselwerking tussen verschillende niveaus van actie moet natuurlijk een rol spelen in het ontwikkelen van beleid, waarbij de verschillende niveaus met elkaar moeten worden verbonden. Bij de bespreking van de aanbevelingen in de essays valt op dat de wisselwerking tussen niveaus niet altijd expliciet wordt, terwijl een voorstel voor een bepaalde beleidsmaatregel op één niveau vaak ook keuzes op een ander niveau impliceert. Weliswaar appelleren de essays aan waarden, maar vaak gaat het om een afweging tussen verschillende waarden en de vraag wie daarover zou moeten beslissen. De voorgestelde maatregelen herbergen – soms impliciet – een keuze ten aanzien van deze afweging tussen waarden. Het ontrafelen welke, vaak concurrerende, waarden van toepassing zijn, kan helpen om inzichtelijk te maken dat bijvoorbeeld maatregelen op operationeel niveau een relatie hebben met (nu nog impliciete) keuzes op constitutioneel of politiek niveau.

Een voorbeeld is het idee van *'data-kill-switch'* uit het essay van Veltman en Kranenburg. Terwijl een *'kill-switch'* gezien zou kunnen worden als een vrij operationele maatregel, bevinden de achterliggende waarden zich op constitutioneel niveau: eigenaarschap en regie over eigen data als een recht, in lijn met het recht om vergeten te worden zoals in de AVG is vastgelegd. Een mogelijk concurrerende waarde is bijvoorbeeld de effectiviteit van dienstverlening of risico-inschattingen, die negatief beïnvloed kunnen worden wanneer iedereen een *'data-opt-out'* heeft. Als het idee van een kill-switch breed toegepast zou worden, maakt dat sommige toepassingen van datagebruik vrijwel onmogelijk.



Wanneer bijvoorbeeld iedereen ervoor kan kiezen data niet beschikbaar te stellen voor data-ondersteunde fraude-detectie, dan heeft fraude-detectie geen zin meer. Het interessante is dat fraude-detectie ook een appel doet op constitutionele en politieke waarden: fraude laten bestaan is onrechtvaardig en in gevallen zelfs gevaarlijk. Soms hebben burgers wellicht zelf voor een data-ondersteunende fraude-detectie gepleit. Zo bezien gaat het bij de keuze van een concrete maatregel opeens om een fundamentele spanning tussen waarden als rechtvaardigheid en veiligheid versus zelfbeschikking over eigen data. Daarnaast is het denkbaar dat burgers, via hun input in het politieke proces, beide maatregelen omarmen, wat vervolgens een duidelijke en zorgvuldige afweging vraagt bij het nemen van een politieke (bindende) beslissing. Er valt weliswaar veel af te dingen op dit voorbeeld, maar het illustreert dat het zinnig kan zijn om na te denken over de wisselwerking tussen de niveaus van actie. Bij de discussie over datagebruik is het belangrijk ons te blijven realiseren dat achter iedere concrete beleidsmaatregel een vaak impliciete waardenafweging schuilgaat. Dit betekent ook dat we de vraag moeten blijven stellen hoe concrete maatregelen invloed hebben op andere niveaus, die op eerste gezicht niet aan de orde lijken te zijn.

Een laatste opmerking is dat de indeling van Kiser en Ostrom (1982), die we nu hebben gehanteerd, geen aandacht schenkt aan een belangrijk aandachtspunt bij beleidsontwikkeling op een nieuw en dynamisch terrein, namelijk hoe wij beleid kunnen ontwikkelen onder onzekerheid waarbij rekening moet worden gehouden met beleidsleren.

3. Data als beleidsdomein

Kijken we naar datagebruik als beleidsvraagstuk, dan moeten we in ieder geval constateren dat het niet zonder meer evident is waar we het dan over hebben. De essays geven gewaar van een erg breed concept, met vele facetten, waardoor het beleid hierop welhaast moet differentiëren. In het verlengde van die essays, onderscheiden wij vier brede onderwerpen waarop beleid zich kan richten.

In de eerste plaats zijn dat *algemene kaders voor gebruik van data*. Hoe passen bestaande wetten en regels in het datatijdperk? Hoe zorg je ervoor dat deze adaptief zijn voor toekomstige innovaties? Van wie zijn data eigenlijk? Hoe zijn

rechten en beginselen als non-discriminatie, transparantie en privacy te borgen? Voor de beantwoording van al deze vragen kan de overheid kaders uitzetten. Het onderwerp data is echter zo algemeen, dat een overkoepelende vraag is, waar de overheid precies voor aan de lat staat en waarop realistisch te sturen valt.

In de tweede plaats kan beleid zich richten op *het gebruik van data in/door de samenleving*. Dit valt uiteen in diverse aspecten. Data bieden kansen voor innovatie, efficiëntie en nieuwe economische activiteiten. Het creëren van die kansen kan een beleidsdoel zijn. Data herbergen ook uitdagingen en bedreigingen voor democratische waarden, toegankelijkheid, inclusie en vele andere publieke waarden. Beleid kan helpen maar is niet noodzakelijk overal effectief. Complicerende factoren hier zijn dat tech-bedrijven veelal internationaal opereren, dat de verdeling van rollen en verantwoordelijkheden soms ambigu is en dat jurisdictie niet volledig past bij hoe en waar data worden opgeslagen en gebruikt.

Beleid kan zich ook richten op *het gebruik van data in/door overheden*, een centraal element in de knelpunteninventarisatie waarvan het ministerie uitgaat (BZK 2019). De overheid produceert en gebruikt data. De complexiteit hiervan is erg hoog in termen van bronnen, (mogelijke) toepassingen, belangen, gebruik en soorten overheidsorganisaties en afdelingen. Er is sprake van fragmentatie van datahuishouding en een functionele organisatie in silo's die lastig op elkaar aan te sluiten zijn. Dit vraagt om *data governance* met aandacht voor het delen, beheren en gebruiken van data. Hierbij gaat het om toezicht, om het goed gebruik en hergebruik van data als kapitaalgoed, transparantie, verantwoording en democratische controle. Ook gaat het om goed opdrachtgeverschap en controle op derden die een rol spelen in verzameling, verwerking of gebruik van data. Soms gaat het om een organisatievraagstuk, in andere gevallen om een beleidsvraagstuk. Daarbij vraagt de diversiteit in het gebruik om beleid dat ruimte biedt zonder nietszeggend te worden.

In de vierde plaats kan beleid zich richten op de bouwstenen om bovenstaande te faciliteren. Gelet op het belang van data voor de samenleving en de overheid, zijn er ook infrastructurele aspecten te onderkennen. Data kan als publiek goed worden gezien, waarbij het beschikbaar stellen en faciliteren van het gebruik van data vraagt om (meta)standaarden, technische infrastructuur, beheer en ontsluiting.

In het Verenigd Koninkrijk heeft de *National Infrastructure Commission* een ontwikkelpad vastgesteld naar het creëren van een 'digital twin' (een digitale replica) voor de gehele nationale infrastructuur. Zeker wanneer derden hiermee aan de slag kunnen, is regie op tal van voorzieningen nodig, zoals op het gebied van autorisatie, authenticatie en beveiligde uitwisseling.

4. Adaptieve beleidsontwikkeling en leren: hoe om te gaan met onzekerheid?

Het ontwikkelen van beleid voor nieuwe en dynamische verschijnselen, vraagt om een andere opzet dan meer 'statische' problemen (de 'klassieke' methode). In essentie komt de aanpak neer op de methodes die al door Lindblom (1959) worden beschreven, waarbij de hoeveelheid informatie over het beleidsprobleem de mogelijkheden van beleidvorming bepaalt. Lindblom onderscheidt voor beleidsvorming een *alomvattende methode* en een *incrementele methode*. Terwijl de alomvattende methode een beleidsinterventie voorstelt op basis van een uitgebreide en veelal uitputtende analyse van het beleidsprobleem, gaat de incrementele methode uit van kleinere stapjes; het uitproberen van een aanpak, het analyseren van de gevolgen daarvan in termen van het reduceren van het probleem en het bijstellen van de aanpak in het licht van deze ervaring. De incrementele methode levert een 'minder ambitieuze' blauwdruk van het beleid op, maar bij deze methode staat het leren over het beleidsprobleem en hoe dat kan worden aangepakt centraal.

In situaties van onduidelijkheid en onzekerheid over het beleidsprobleem (wanneer niet alles bekend is in termen van mogelijke interventies en hun effecten, terwijl het gissen is naar de waarschijnlijkheid van het optreden van de bekende effecten) heeft de incrementele methode voordelen. In sommige essays over het datagebruik door de overheid wordt dit ook benadrukt, door te wijzen op het experimentele karakter van datagebruik. Doordat technologie steeds in ontwikkeling blijft, is het niet mogelijk te overzien wat 'morgen' mogelijk is en welke effecten dat zal hebben. Tegelijkertijd is het belangrijk dat de overheid proactief beleid ontwikkelt om datagebruik nu al te stroomlijnen en om te voorkomen dat burgers te maken krijgen met ongewenste effecten.

Nu is er vooral sprake van wildgroei aan initiatieven en toepassingen, die bovendien verspreid zijn over verschillende overheidslagen en tussen verschillende organisaties. Pogingen om tot beleid te komen zijn wellicht veeleer als reactief te bestempelen omdat vooral wordt gereageerd op ontwikkelingen die tot problemen leiden of als ongewenst worden gezien. Die problemen, zoals de essays dat ook laten zien, liggen vooral op het vlak van onder meer privacy, ongelijkheid, discriminatie en toenemende dominantie van private partijen bij de toegang tot en het gebruik van voor de overheid belangrijke data.

Voor een meer dynamische aanpak van beleid zijn er in de loop van de tijd verschillende opties ontwikkeld en uitgetoet. Een aantal belangrijke voorbeelden van modellen die relevant kunnen zijn voor beleid ten aanzien van datagebruik, zullen we bespreken. Daarbij wordt vooral gelet op de mate waarin van eerdere ervaringen met de uitvoering van beleid kan worden geleerd.

De eerste groep modellen betreft de meer statische beleidsmodellen die rekenschap geven aan het feit dat beleid slechts een beperkte houdbaarheid heeft. Hierbij weet men dat er sprake is van een veranderlijke omgeving, maar is men niet in staat of bereid om de gevolgen daarvan volledig in het beleidsontwerp mee te nemen.



Een voorbeeld is de ‘*sunset-wetgeving*’ waarmee de looptijd van beleid is begrensd zodat daarna een nieuw ontwerp kan worden gekozen. Een tweede voorbeeld is het werken met verplichte evaluatiebepalingen na een periode van bijvoorbeeld vijf jaar. Voor veel Europese wetgeving wordt met dit model gewerkt. Beleid is in dit geval nog steeds statisch qua opzet, waarbij de hoop is dat na een redelijke periode verbeteringen kunnen worden voorgesteld. Voor beleid over datagebruik zijn deze modellen minder aantrekkelijk, omdat ze het leren van de gevolgen van technologische verandering nauwelijks stimuleren.

De tweede groep modellen voor beleidsvorming zijn modellen waarin nadruk wordt gelegd op onzekerheid en in het proces van beleidsvorming al rekening wordt gehouden met verschillende opties die men vooraf in kaart heeft gebracht. Een voorbeeld hiervan is *adaptieve beleidsplanning* waarmee een aantal belangrijke elementen aan het beleid worden toegevoegd (zie Van der Plas e.a. 2013; Haasnoot e.a. 2013). Allereerst moeten in dit model de voornaamste onzekerheden van de voorgestelde beleidsmaatregelen in kaart worden gebracht, waarna de beleidsvoerders vervolgens nadenken over mogelijke correcties in de vorm van flankerend beleid. Tot slot wordt de uitvoering voortdurend gemonitord, zodat, indien nodig, aanvullende maatregelen tijdig kunnen worden ingezet. In vergelijking met statische modellen zijn adaptieve modellen een belangrijke stap voorwaarts omdat over onverwachte en onbedoelde effecten wordt nagedacht. Tegelijkertijd bieden deze modellen weinig houvast op het moment dat die effecten nog onvoldoende duidelijk zijn, laat staan dat ze duidelijk maken wat een aanpak kan zijn om die effecten te verminderen. Voor het vormen van beleid over datagebruik, waarbij het belangrijk is om te leren van de mogelijkheden die zich in de komende jaren zullen voordoen, zijn deze modellen daarom minder geschikt.

Een derde groep modellen benadrukt leren door middel van de uitwisseling van ervaringen met het doel *best practices* te identificeren. De ervaringen kunnen aanleiding zijn om beleid verder te specificeren. Een bekend voorbeeld is de Open Coördinatiemethode (OCM) die binnen de EU wordt gebruikt. Daarbij wordt op basis van vrijwilligheid getracht te leren door de ervaringen met de uitvoering van beleidsambities te delen en uit te wisselen. Elementen van deze methode zijn te vinden in het door BZK gestarte proces om ervaringen en knelpunten bij datagebruik te delen. Bij toepassing van dit model wordt duidelijk dat het delen van ervaringen belangrijk is, maar, zoals de uitwisseling over datagebruik in de

publieke ruimte laat zien, nog niet hetzelfde is als het formuleren van nieuw of ander beleid. Op dat punt is de OCM nog te ongestructureerd.

Een vierde groep modellen voor beleidsvorming gaat uit van verdere beleidsontwikkeling binnen een vooraf vastgesteld kader. Een bekend voorbeeld is het Lamfalussy-procedure zoals dat voor technisch ingewikkelde financiële vraagstukken binnen de EU is ontwikkeld. Deze methode kent een aantal lagen waarin, stap-voor-stap, het beleid wordt uitgewerkt. Terwijl de wetgever het algemene beleidskader vaststelt, inclusief de voornaamste uitgangspunten bij de uitvoering, wordt het kader door belanghebbenden verder uitgewerkt. Hun voorstellen moeten door een commissie van nationale vertegenwoordigers worden geaccordeerd (een vorm van gedelegeerde wetgeving binnen de EU-context). Het interessante van dit model is het betrekken van deskundigen in een gelaagd model van beleidsontwikkeling. Daarmee kunnen nieuwe inzichten in het beleid worden verwerkt, wat met name voor meer complexe problemen belangrijk is. De keerzijde is dat de politieke (de ‘democratische’) grip op de lagere niveaus in dit proces minder is.

Het gaat om een **continu proces**, waarbij het beleid voortdurend kan worden bijgesteld aan de ervaringen van nu

De genoemde modellen benadrukken verschillende aspecten van het proces dat gebruikt wordt om tot een succesvol beleid te komen. In alle gevallen is er sprake van een dynamische component in de zin dat men zich realiseert dat beleid met nieuwe ervaringen moet mee bewegen. Dat meebewegen staat centraal in het concept van adaptief beleid (Steunenberg 2018), waarbij de nadruk ligt op beleidsleren. Dit leren kan op verschillende manieren gebeuren, variërend van leren achteraf (wanneer er gewerkt wordt met evaluatiebepalingen) tot leren tijdens het proces (bijvoorbeeld bij de OCM). Belangrijk voor adaptief beleid is dat het om een continu proces gaat, waarbij het beleid in de uitvoering voortdurend kan worden bijgesteld aan de ervaringen van nu. Voor het ontwerpen van beleid over datagebruik door de overheid, waarbij de wens is om nu een beleid te formuleren, is het de uitdaging een kader voor te stellen dat het mogelijk maakt dit beleid te blijven aanpassen aan inzichten die we nu nog niet hebben.



5. Een voorstel voor adaptief beleid

Vanuit analytisch oogpunt zal bij het maken van een beleidsontwerp dat adaptief is (door zich aan te kunnen passen aan een sterk ontwikkelende techniek en gebruik), ingezet moeten worden op het leren en de ervaringen die bij de uitvoering worden opgedaan. Daarin moet een balans worden gezocht tussen de *centraal* gekozen uitgangspunten of principes en *decentraal* beschikbare ervaringen en kennis. Een balans tussen *top-down* geformuleerde uitgangspunten en *bottom-up* beschikbare kennis over knelpunten, mogelijke beleidsinterventies en hun effecten.

Bij het vinden van die balans in het geval van datagebruik zijn decentrale overheden van groot belang vanwege de beleidsexperimenten die zij uitvoeren. Op het niveau van uitvoering moeten immers tal van afwegingen worden gemaakt tussen concurrerende waarden, divergerende perspectieven en diverse belangen van betrokken actoren. Technische expertise moet aansluiting vinden bij domeinexpertise en beleidsexpertise. Daar spelen ook private partijen een rol in. Voor productief beleidsleren zijn de ervaringen van decentrale overheden en private partijen dan ook essentieel bij het verder vormgeven van het beleid. Adaptief beleid in deze context vraagt om herhaalde dialoog tussen stakeholders en experts, vanuit het adagio *'making progress rather than solutions'* (Head 2017), waarbij het beleid publieke waarden moet borgen.

Zoals gezegd, moet bij een beleid over datagebruik voldoende ruimte zijn om te experimenteren en te leren. De kennis die dat oplevert moet input zijn voor verdere aanpassing en uitwerking van het beleid. We stellen daarom voor een *polycentrische co-reguleringsaanpak* te hanteren. Deze term, zoals wij die gebruiken, vormt een combinatie van drie analytische raamwerken. Ten eerste moet **polycentrische besluitvorming** de 'informatieasymmetrie' reduceren door kennis samen te brengen die verspreid over de samenleving aanwezig is (Dorsch en Flachsland 2017). Ten tweede gaat het bij **co-regulering** over een bredere groep van stakeholders en besluitvormers die worden betrokken in het opstellen, uitvoeren en verder ontwikkelen van beleid. Dat kan betrekking hebben op zowel de regelgeving als het ontwikkelen van normen die in regulering of het toezicht een rol moeten gaan spelen. Ten derde is sprake van een **cyclische benadering**, waarbij wordt geleerd door de voortdurende confrontatie van eerdere keuzes met nieuwe inzichten.

Het **polycentrische perspectief** past bij beleidsdomeinen waarin sprake is van zeer complexe, of zelfs 'wicked' problemen, waarin beleidsmakers te maken hebben met een grote diversiteit aan doelen, belangen, sectoren en niveaus die relevant zijn voor het beleid en de uitvoering. Dit perspectief biedt ruimte voor context-specificiteit en voor experimenten en leren op decentraal niveau. Het stelt een overkoepelende regulering voor, in combinatie met zelforganisatie in specifieke context.

Het polycentrische perspectief biedt ruimte voor context-specificiteit en voor experimenten en leren op decentraal niveau

Co-regulering gaat in op het probleem dat geen enkele individuele partij alle kennis heeft die nodig is om complexe, veelzijdige en dynamische problemen op te lossen, laat staan dat er overzicht is en grip op de instrumenten die nodig zijn om tot effectief beleid te komen (Finck, 2017). Gegeven het internationale karakter van de IT-sector lijkt zelf- of co-regulering hier goed te passen. De afspraken tussen de gemeente Amsterdam en Airbnb ten aanzien van toeristenbelasting en een 60-dagen-termijn zijn hier een voorbeeld van. Dit model wordt *'regulated self-regulation'* genoemd. Tegelijkertijd laat het Airbnb-voorbeeld ook zien dat private partijen slechts beperkte prikkels krijgen om voor een dergelijke aanpak te kiezen en dat toezicht uitdagend is.

Regulering en beleid ten aanzien van de private sector is dan toch van belang, waarbij het *'really responsive regulation'*-concept (Baldwin en Black, 2008; Black en Baldwin, 2010) relevant lijkt. Dat concept impliceert een strategie waarin een grote diversiteit aan instrumenten (waaronder wet- en regelgeving) op een flexibele wijze wordt gebruikt, zodat ruimte wordt geboden aan diverse factoren, waaronder gedrag, houding en cultuur van het bedrijf. Ook de institutionele omgeving, de interactie tussen instrumenten, veranderende prioriteiten en doelen en de prestaties van middelen kunnen worden meegewogen (Baldwin en Black, 2008).

Wat betreft de aanpak die binnen een polycentrisch perspectief wordt gekozen, blijft er een inherente spanning bestaan tussen enerzijds het bieden van 'lokale' flexibiliteit of decentrale verantwoordelijkheid en anderzijds het doel van meer coherente regulering en standaardisatie van technologie in het kader van het voorgestane beleid. Deze spanning komt voort uit het willen ontwikkelen van regulering die aangepast kan worden aan de context en verschillende lagen van de overheid, terwijl die regulering tegelijkertijd ook als brede basis voor besluitvormingsprocessen dient. Complex en vrijblijvend beleid valt er soms moeilijk door te krijgen, terwijl te nauw en inflexibel beleid het risico loopt niet alle nuances aan te kunnen die inherent zijn aan datagebruik in de publieke sector. Een focus op communicatie, afstemming en onderhandeling tussen publieke, private en maatschappelijke stakeholders is nodig om tussen de uitersten een weg te vinden en te komen tot regulering die aansluit op de werkelijkheid waarin die regulering moet werken.

In de wetenschappelijke literatuur wordt dit omschreven als een twee-richtingen-stroom van opgaande en neergerende communicatie. Die laatste omvat het communiceren van een implementatiestrategie en implementatiedoelen, tezamen met de procedures en middelen die bij die implementatie horen. Opgaande communicatie omvat het betrekken van belanghebbenden bij afwegingen, bezwaren en beoogd nut van nieuwe of aangepaste regulering, en bij het inventariseren van behoeften ten aanzien van de implementatie daarvan. Aandachtspunt hierbij is het risico op een gebrek aan aansluiting bij bestaande procedures en wetgeving. Vooral bij opgaande communicatie zijn er voorbeelden te vinden waaruit blijkt dat het nodig is meerdere communicatiekanalen eropna te houden, zonder dat de feedback die via die kanalen terugkomt versplinterd mag raken. Stakeholders hebben verschillende wijzen van en behoeften bij communicatie, die ook nog eens wijzigen als de omstandigheden veranderen. Zo verloopt (opgaande) communicatie soms via de feedback vanuit koepels (zoals de VNG), terwijl andere belangrijke input kan komen uit bijeenkomsten. Ook kunnen teams van betrokkenen die in een specifieke context werken, vanuit hun lokale rol aanbevelingen doen en uitdagingen identificeren.

Dit polycentrisch perspectief is een stap verwijderd van een reactieve reguleringsaanpak, waarbij de overheid reageert op de opvolging van de regulering of de mate waarin die plaatsvindt. In plaats daarvan stelt het een flexibele aanpak

voor waarbij er continu communicatie en feedback wordt georganiseerd waarmee aanpassingen aan specifieke context en plaatsen mogelijk wordt.

De **cyclische benadering** grijpt terug naar vormen van beleidsleren, dat al in meer klassieke studies binnen de bestuurskunde naar voren komt. Het incrementalisme dat Lindblom propageerde, ging al uit van combinaties van doen en reflecteren, zodat kennis over hoe beleid op dat moment uitwerkt bij vervolgstappen kon worden gebruikt. Dit idee van korte leercycli komt in verschillende aanpakken en concepten terug, waaronder in de bekende 'plan-do-check-act'-cyclus. Belangrijk is wel dat de 'check' op een verantwoorde wijze wordt aangepakt om er zeker van te zijn wat de beleidseffecten zijn en in hoeverre die effecten het resultaat zijn van het beleid. Daarbij kan een aanpak die gebaseerd is op wetenschappelijke inzichten ('*evidence based*') een belangrijke rol spelen

De combinatie van polycentrische besluitvorming, co-regulering en een cyclische benadering maakt dat meer belanghebbenden een plek aan tafel kunnen krijgen, terwijl het tegelijkertijd flexibiliteit biedt om technische, context- en locatie-specifieke oplossingen te ontwikkelen binnen een bredere set van regels.

6. Hoe kunnen we adaptief beleid toepassen?

Om onze polycentrische co-reguleringsaanpak voor adaptief beleid te illustreren, hebben wij een voorbeeld uitgewerkt hoe beleid ten aanzien van het datagebruik door de overheid eruit zou kunnen zien. In lijn met wat benadrukt wordt in de aanpak, gaat het in dit voorbeeld om een combinatie van beleid en proces, omdat er sprake moet zijn van de mogelijkheid om beleid op basis van eerdere ervaringen verder te ontwikkelen en aan te passen.

We geven eerst ons voorbeeld, dat we daarna zullen toelichten:

1. De wetgever stelt, indien mogelijk of gewenst, een algemeen *inhoudelijk beleidskader* voor met verschillende principes die centraal moeten staan. Dit kunnen principes zijn die in verschillende essays worden benoemd, zoals privacy, gelijkheid, non-discriminatie en evenredigheid.

2. De wetgever geeft ook een *procedureel beleidskader* voor hoe het beleid in de tijd verder wordt ontwikkeld.
Daarbij wordt uitgegaan van *werkgroepen* waarin bepaalde onderwerpen of thema's verder worden uitgewerkt. De verantwoordelijkheid voor die uitwerking wordt bijvoorbeeld aan de verantwoordelijke minister gedelegeerd. (Er zijn ook andere modellen denkbaar, maar in dit voorbeeld wordt uitgegaan van een beperkte mate van het 'op afstand' plaatsen van de verantwoordelijkheid voor dit terrein).
3. De werkgroepen bestaan uit *stakeholders* betrokken bij een thema, waarbij het belangrijk is dat ook de private sector wordt betrokken. Zoals uit de essays blijkt, is medewerking van partijen in de markt essentieel voor het welslagen van het datagebruik door de overheid. De verschillende partijen uit de overheidssector (inclusief decentrale overheden) en de markt brengen relevante kennis op het thema van de werkgroep in.
4. De werkgroepen *formuleren de regels* die ten aanzien van een thema of probleem worden voorgesteld. Zo'n voorstel wordt door de verantwoordelijke minister vastgesteld.
5. De werkgroepen *monitoren* de uitvoering van de vastgestelde regels in het licht van de actuele ontwikkelingen (inclusief *compliance* en handhaving) en nieuwe kennis (*evidence based*).
6. Nieuwe voorstellen worden vooral beoordeeld naar beschikbare '*evidence*' over werking en effecten.
7. De monitoring vormt de basis voor een *korte (technische) cyclus* om beleid aan te passen, bijvoorbeeld eens per een of twee jaar. De minister stelt de nieuwe voorstellen voor aanpassing vast.
Dat levert flexibiliteit op, wat op dit terrein van groot belang is. (Natuurlijk moet men zich realiseren dat daarmee de rechtszekerheid in de tijd minder goed is geborgd, maar daarbij is sprake van een afruil met flexibiliteit).
8. Over een *langere (politiek) cyclus*, bijvoorbeeld eens per vijf jaar, wordt het inhoudelijke beleidskader, dat onder punt 1 is vastgesteld, met de Kamer besproken. Dat kan leiden tot bijstellingen.

Graag plaatsen we verschillende kanttekeningen bij deze proeve.

In de eerste plaats gaat het voorbeeld uit van leren door te werken met cycli van ontdekken, uitproberen, ervaren en aanpassen; deze cycli zijn essentieel zijn voor

beleidsleren. Dat geldt voor de korte cycli op het niveau van de werkgroep met stakeholders, die moeten komen tot meer concrete uitwerkingen. Het geldt ook voor de cyclus van aanpassing van het regulerende kader waarbij de minister betrokken is. Tot slot geldt dit ook voor de politieke cyclus waarin het beleid met de Kamer wordt besproken en, in algemene zin, op punten kan worden bijgesteld.

Het werken met deze leercycli op verschillende niveaus biedt mogelijkheden om ervaringen in het beleid door te laten klinken, naast flexibiliteit om, op termijn, met onverwachte effecten of problemen om te gaan die wij nu nog onvoldoende kennen.

In de tweede plaats verbindt het voorbeeld de verschillende actieniveaus uit de vorige paragraaf met het beleidskader. Binnen het kader van algemene normen, die overigens in het beleidsproces ter discussie kunnen worden gesteld, wordt gewerkt aan een verdere uitwerking naar inrichting en operationalisering. Daarmee is in het ontwerp sprake van een verbinding tussen de genoemde niveaus van actie.

In de derde plaats gaat het voorbeeld uit van een organisatie van beleid waarbij politieke verantwoordelijkheid bij het Rijk en in het bijzonder bij de minister is belegd. Wat betreft de rol van het Rijk, is het belangrijk te constateren dat het streven naar een gemeenschappelijk kader belangrijke voordelen heeft, zoals schaalvoordelen en de borging van rechtsgelijkheid en rechtszekerheid.



Tegelijkertijd moet de mogelijkheid van experimenteren en daarmee leren blijven bestaan bij de verschillende stakeholders, waaronder decentrale overheden. Daarom zijn verschillende overheden, naast de private sector, belangrijk in de werkgroepen.

7. Conclusie

Data is een complex en niet zelden een rommelig veld om beleid in en voor te maken. Onvermijdelijk is er sprake van informatie-asymmetrie met betrekking tot verschillende vormen van het verzamelen, gebruiken en analyseren van data. Als gevolg hiervan zijn er twee zaken nodig. Ten eerste is er een bepaalde mate van standaardisatie en coherentie nodig van beleidsbegrippen, terwijl tegelijkertijd aansluiting op specifieke context moet worden georganiseerd. Ten tweede moet een brede groep aan stakeholders worden betrokken, waarbij zij continu input en feedback kunnen geven middels daarvoor geschikte communicatiekanalen. Dit is een cyclisch proces, waarbij ruimte is voor leren. Dit alles vereist een brede blik op het proces, zonder de specifieke context van implementatie uit het oog te verliezen.

Meer algemeen is het van belang om regelmatig stil te staan bij de vraag wat beleid op dit domein precies vermag; waarvoor de overheid en beleid precies aan de lat staan. Voor het gebruik van data in, voor en door de overheid zelf is dat duidelijker dan wanneer het gaat om toepassingen waarbij (open) data uit publieke bronnen gecombineerd worden met data uit andere bronnen en private toepassingen. En nog weer anders is het voor volledig private toepassingen die de publieke ruimte raken. In hoeverre valt data-beleid te maken dat de complexiteit aankan die voortkomt uit het alomtegenwoordige en diverse karakter van data die we genereren bij vrijwel alles wat we doen? Die algemeenheid en alomtegenwoordigheid zouden allerm minst een reden moeten zijn om niets of weinig te doen, maar wel om voor iedereen helder te krijgen waar beleid wel en waar het niet over kan gaan.

Gerelateerd daaraan is de constatering dat, door de vele facetten, ambiguïteit haast onvermijdelijk is. Dat heeft ook betrekking op de in deze epiloog onderscheiden niveaus.

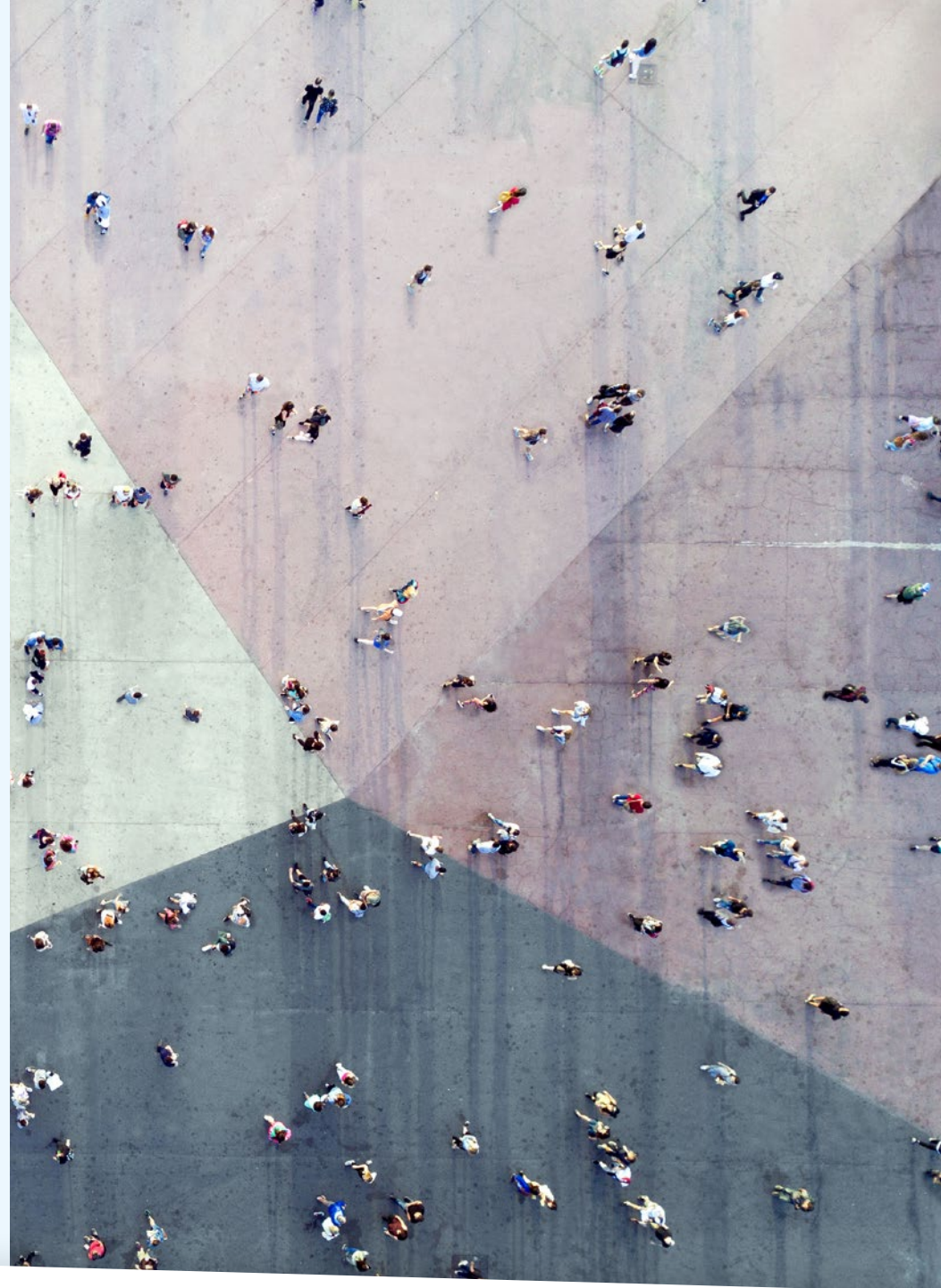
Intuïtief hebben we snel een beeld bij het niveau waarop bepaalde maatregelen spelen. Het is van belang dat soort aannames steeds ter discussie te stellen. Een conflict tussen operationele instrumenten en democratische waarden lijkt een ongelijke strijd. Echter, achter veel van ons denken en praten over data, gaan al tal van waarden-afwegingen schuil. Het onderkennen en expliciteren daarvan kan helpen om voorbij al te algemene (in)richtingsprincipes te komen.

Om een begin te maken met adaptief beleid op basis van onze polycentrische co-reguleringsaanpak valt te overwegen te starten met het vormen van werkgroepen (punten 3 en 4 in ons voorbeeld) en het formuleren van en discussiëren over mogelijke regels. Dat is sterk vergelijkbaar met de decentrale aanpak die bij het opstellen van de knelpunteninventarisatie door het ministerie van BZK is gekozen. Deze aanpak vraagt wel om een duidelijk procedureel beleidskader waarin de taken en verantwoordelijkheden zijn verdeeld (punt 2). Vervolgens levert dat input op voor het inhoudelijke beleidskader (punt 1), voor zover dat nu nog moeilijk is vast te stellen. Dit kader kan vervolgens, via de andere elementen in onze aanpak, gaan 'groeien' tot een meer afgewogen pakket dat is gebaseerd op de problemen maar ook de mogelijke oplossingen die we in de komende tijd zullen ervaren.

Concluderend schetsen we hier de uitdaging om beleid voor te stellen dat meer op leren gericht is en waarin de inzichten die we nu nog niet hebben, straks wel aan het kader kunnen worden toegevoegd. Dat vraagt om een beleid dat 'evolutionair stabiel' is, dat wil zeggen, in staat is om met de dynamiek van de techniek om te gaan. Dat vraagt ook beleid dat prikkelt tot beleidsleren en dus beleids-experimenten aanmoedigt, waarbij de resultaten daarvan weer in het beleidskader worden opgenomen. Lastig, maar wel een mooie uitdaging!

Referenties

- 1 Zie ook Ostrom (2004: 58) waarin zij deze niveaus en verwante indelingen bespreekt. Elk van deze indelingen kent een bepaalde focus die past bij de vragen die men met die classificatie zou willen beantwoorden. In ons geval voegen we het organisatieniveau aan de indeling toe, omdat die binnen bestuurlijke stelsels belangrijk is.
- Baldwin, R. en J. Black (2008) Really responsive regulation, *Modern Law Review* 71: 59.
 - Black, J. en R. Baldwin (2010) Really responsive risk based regulation, *Law & Policy* 32: 181–213. doi/10.1111/j.1467-9930.2010.00318.x/full
 - Dorsch, M.J. en C. Flachsland (2017) A Polycentric Approach to Global Climate Governance, *Global Environmental Politics* 17: 45–64. doi/10.1162/GLEP_a_00400.
 - Finck, M. (2017) Digital Regulation: Designing a Supranational Legal Framework for the Platform Economy, London: LSE Law, Society and Economy Working Papers 15/2017 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2990043)
 - Haasnoot, M, J.H. Kwakker, W.E. Walker en J. Ter Maat (2013) Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world, *Global Environmental Change* 23: 485–98.
 - Head, B. (2017) New agendas for the study of wicked problems, 3rd International Conference on Public Policy (ICPP3)
 - Kiser, L.L. en E. Ostrom (1982) The three worlds of action: A metatheoretical synthesis of institutional approaches, in: E. Ostrom (ed.) *Strategies of political inquiry*, Beverly Hills: Sage. 179–222.
 - Lindblom, C.E. (1959) The Science of 'Muddling Through', *Public Administration Review* 19: 79.
 - BZK (2019) Nota Proeftuin Behoorlijk datagebruik in de openbare ruimte: knelpunteninventarisatie, Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Directoraat-Generaal Overheidsorganisatie.
 - Ostrom, E. (2005) *Understanding institutional diversity*, Princeton, New Jersey: Princeton University Press.
 - O'Toole, L.J. (2000) Research on policy implementation: Assessment and prospects, *Journal of Public Administration Research and Theory* 10: 263–88.
 - Steunenberg, B (2018) Adaptieve beleidsontwikkeling: zoeken naar nieuwe vormen van beleidsanalyse voor de digitale overheid, *Beleidsonderzoek Online*, februari 2018. doi: 10.5553/BO/221335502018000001001.
 - Van der Pas, J.W., B. van Wee en J. Van der Waard (2013) Mobiliteitstrends en implicaties voor beleid: Dynamisch adaptief beleid, Delft: TUD (Programma Duurzame Bereikbaarheid Randstad: Notities voor wetenschap en praktijk nr. 8; zie: <http://dbr.verdus.nl/mobiliteitstrends-en-implicaties-voor-beleid>).



Voorstel voor beleidsmaatregelen en acties

De beleidsaanbevelingen uit de essays hebben als basis gediend voor onderstaande beleidsmaatregelen en acties. Hierbij is gebruik gemaakt van het wetenschappelijk kader van het Instituut Bestuurskunde van Universiteit Leiden. Op het moment van schrijven zijn de beleidsmaatregelen en acties op een longlist geplaatst ter voorbereiding voor de Data Agenda Overheid 2020. De beleidsaanbevelingen en acties zullen ter beoordeling worden voorgelegd aan de klankbord- en stuurgroep Data Agenda Overheid.

Beleidsvoorstellen

1. *Onderzoek of het Integraal Afwegingskader voor beleid en regelgeving (IAK) actueel is op het punt van digitalisering in het publiek domein.*

Actuele afwegingskaders zijn randvoorwaardelijk voor het actief betrekken van stakeholders (waaronder burgers) bij afwegingen en het communiceren erover. Hierbij kan een link worden gelegd met de Code Goed Digitaal Bestuur, de toolbox Verantwoord Ondernemen of het programma Maatschappelijke Dialoog. Het veld van stakeholders in het publieke domein is veranderd met digitalisering, met name als het gaat om ethische aspecten. Waar voorheen tech-bedrijven digitaal waren, begeven ze zich nu in het publieke domein. De ethische dilemma's lijken hierbij de AVG of andere wetten te overstijgen. Herijking van het IAK kan concrete handvatten bieden, rekening houdend met voldoende speelruimte om op lokaal niveau de diversiteit van dilemma's te borgen¹. In het verlengde daarvan is er op uitvoerend (gemeentelijk) niveau een gemis aan afwegingskaders door digitalisering in (gemeente)werken zoals bruggen en tunnels. Ook op dit punt zou het IAK getoetst moeten worden op actualiteit.

2. *Organiseer een maatschappelijke dialoog over AI waarbij een samenwerkingsverband wordt ingezet van onderzoekers, ontwikkelaars, beleidsmakers, burgers, etc.*

Onderwerpen die ter discussie gesteld zouden kunnen worden, zijn:

- Een Algorithm Reporting Framework ontwikkelen (alsmede een hotline)
- Een Nationaal Algoritmeregister neerzetten, gekoppeld aan een Algoritme Forum
- Een agenda AI-vaardigheden opstellen voor opleiders, politici en beleidsmakers³
- Verkennen van de impact van digitalisering in de publieke ruimte op de grondrechten van burgers
- Verkennen van de mogelijkheden voor certificering van AI-algoritmen (datasafety-keurmerk)

3. *A. Onderzoek hoe het debat met 'de burger' over datagebruik in de openbare ruimte in het proces van innovatie (in beleidsontwikkeling en/of in uitvoering van publieke taken) aan de voorkant van het traject geïncorporeerd kan worden.*

Denk hierbij aan versterking van het lokaal bestuur en het gebruik van design principles (waaronder de principes van user-centred én value-sensitive design) aan de voorkant van het traject. Periodieke toetsing met burgers is hierbij evident om mee te nemen in het proces van innovatie.

Uitvoering van dit onderzoek zou een samenspel kunnen zijn van het Rijk, lokale overheden, private partijen en de wetenschap. Belangrijk is om de praktijk aan de theorie te koppelen en dit in living labs (met experimenteeruimte) te bewerkstelligen. Het is de moeite waard om te onderzoeken of hierbij kan worden aangesloten bij bestaande initiatieven of nog te starten initiatieven. Living labs, met een casusgerichte setting en de betrokkenheid van stakeholders met verschillende belangen, lenen zich goed hiervoor.



3. B. Organiseer een platform om pilots in totaliteit te bezien.

De organisatie van een platform sluit aan op maatregel 3A. De functies van het platform zouden vierledig zijn:

- Een onderzoekende functie, onder meer op de ethische kant (zoals gedragscodes die toezien op het verantwoord verwerken van persoonsgegevens in de openbare ruimte) en bevoegdheden van (lokale) overheden voor (her)gebruik van (gecombineerde) data.
- Een uitvoerende functie, onder meer via het toepassen van de toolbox Verantwoord Innoveren, het toepassen van ontwikkelde gedragscodes en de uitwerking van het voorstel om te komen tot digitale en data-omgevingsvisies (in analogie met de op handen zijnde omgevingsvisies voor de lokale fysieke ruimtes).
- Een faciliterende functie wat betreft kennisdeling (d.m.v. bijvoorbeeld draaiboeken voor gezamenlijke pilots).
- Een adviserende functie (die zich richt op waarden in de digitale samenleving).

Acties

1. *Onderzoek doen naar technische dan wel organisatorische oplossingen waarmee anonimiteit van gebruikers alsmede de veiligheid in de openbare ruimte wordt geborgd.*

Onderzoek zou kunnen plaatsvinden in samenwerking met wetenschap, bedrijfsleven, maar zeker ook organisaties uit het maatschappelijk middenveld (Stichting Bits of Freedom, Waag Society e.d.)

2. *Onderzoek doen naar een generieke vertrouwensinfrastructuur in de publieke ruimte.*

Naast inventariseren en evalueren van de digitale basisinfrastructuur in de publieke ruimte, het onderzoeken van mogelijkheden voor realisatie van een nationale, onafhankelijke en open digitale vertrouwensinfrastructuur voor identificatie, authenticatie en autorisatie van persoonlijke data, inclusief de bijbehorende governance.

3. *Haalbaarheidsonderzoek doen naar registratie van sensoren. We zien op dit moment verschillende initiatieven; is het mogelijk om deze op te schalen?*

Op lokaal niveau bestaan er verschillende initiatieven voor registratie van sensoren. Een haalbaarheidsonderzoek naar het mogelijk samenvoegen dan wel (federatief) centraal onderbrengen bij een partij (publiek, privaat of publiek-privaat) of juist decentraal houden, zou de mogelijkheden in kaart brengen. De focus zal dan liggen op de mogelijkheid van een landelijk wetgevend kader voor aanmelding van sensoren (zoals de Wet informatie-uitwisseling bovengrondse en ondergrondse netten en netwerken (WIBON) hierin voorziet voor kabels en leidingen).

De organisatie hiervan zou kunnen plaatsvinden in samenwerking met het ministerie van BZK, (lokale) overheden en private partijen.

4. *Zodra traject modelovereenkomst Slimme Toepassingen (VNG/BZK) opgeleverd is, bezien of verbreding/opschaling mogelijk is.*

Bij verbreding kan gedacht worden aan het gebruik van de modelovereenkomst door alle 355 gemeenten. Bij opschaling kan gedacht worden aan het opschalen van de modelovereenkomst naar gebruik door andere overheden of gebruik door het Rijk voor overeenkomsten met commerciële partijen. Naast modelovereenkomsten zou een dergelijk traject ook een bredere scope kunnen hebben met het oog op standaardisatie, certificering of validaties en bijbehorende normen (creëren van ‘algemeen verbindende voorschriften’).



Referenties

- 1 Mogelijk kan bij het IAK een paragraaf opgenomen worden over verantwoorde experimenteerruimte (legal sandbox) ten behoeve van living labs.
- 2 Hiervoor zijn individuele startende initiatieven, zowel op rijksniveau (bij BZK RADIO) als op gemeentelijk niveau, die met elkaar verbonden zouden kunnen worden.



Colofon

Deze essaybundel is mede tot stand gekomen dankzij de onvoorwaardelijke inzet van de begeleidingscommissie 'Behoorlijk datagebruik in de openbare ruimte', bestaande uit de volgende personen:

Prof. dr. Frans Feldberg (Vrije Universiteit Amsterdam)

Ran Haase (Gemeente Eindhoven)

Emeritus hoogleraar bestuurskunde Peter van Hoesel (Erasmus Universiteit Rotterdam) (Voorzitter)

Marcel Hopman (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties)

Prof. dr. Maurits Kaptein (JADS Instituut)

Klaas Kloosterman (Gemeente Groningen)

Prof. mr. dr. Sofia Ranchordas (Rijksuniversiteit Groningen)

Prof. dr. ir. Arjan van Timmeren (AMS Instituut)

Theo Veltman (Gemeente Amsterdam)

Jeanine Vosselman (Gemeente Groningen)

Aantink Yeh (VNG Realisatie)

www.digitaleoverheid.nl/behoorlijk-datagebruik

www.rijksoverheid.nl

Oktober 2019