Big Data –Fundamentele rechten in een fundamenteel veranderende wereld

mr. A. de Jong[1]

1.1 Inleiding

De ontwikkeling van onze informatiemaatschappij gaat in een hoog tempo en lijkt te versnellen. Dit is duidelijk te zien aan snel opeenvolgende innovaties op het vlak van ICT, maar ook bij andere steeds meer met ICT convergerende disciplines, zoals bio-, nano- en cognitieve technologie.[2] Alhoewel de aanduiding van deze ontwikkelingen, zoals *Cloud computing* en het *Internet of Things*, uit commercieel oogpunt door het bedrijfsleven met *buzz words* en hypes worden omgeven, zijn het ontwikkelingen die wezenlijke gevolgen voor onze maatschappij met zich meebrengen. Een laatste ontwikkeling die op steeds meer maatschappelijke belangstelling heeft mogen rekenen en dat in de bredere trend van ontwikkelingen past is Big Data.[3]

Big Data is, in het kort, een verzamelnaam voor het met hoge snelheid verzamelen en het analyseren van grote hoeveelheden diverse gegevens.[4] Op de resultaten hiervan kan door mensen of geautomatiseerd actie ondernomen worden. Big Data wordt mogelijk gemaakt door het feit dat steeds meer gegevens verzameld en gedeeld (kunnen) worden, de kosten van opslag en het verwerken van deze gegevens drastisch daalt en de methoden van analyse verbeteren. Denk hierbij niet alleen aan het verzamelen van persoonsgegevens op het *world wide web,* bijvoorbeeld met behulp van cookies, maar ook aan het verzamelen van persoonsgegevens in de fysieke wereld, zoals het bijhouden van hartslag en bloeddruk door nieuwe "slimme" fitnessbandjes. Het gaat echter niet alleen om persoonsgegevens maar ook over bijvoorbeeld meteorologische informatie, informatie over de staat van rivier- of zeedijken of andere infrastructuur en bedrijfsprocessen die gemonitord worden met behulp van sensoren. De hoeveelheid digitale informatie die elk jaar wordt gemaakt neemt enorm toe. In 2013 werd 90% van alle informatie ter wereld in de laatste twee jaar gecreëerd en zeker is dat deze groei zich doorzet.[5]

---

[1] mr. Arjan de Jong is werkzaam bij de directie burgerschap en informatiebeleid van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Dit preadvies is op persoonlijke titel geschreven.

[2] Rathenau Instituut, *Intieme technologie: de slag om ons lichaam en gedrag,* Den Haag: Rathenau Instituut 2014.

[3] Volgens de 2014 Gartner Hype Cycle Special Report is Big Data in 2014 op zijn top qua hype, waarna ontnuchtering plaatsvindt. B. Burton, D.A. Willis, *Gartner's Hype Cycle Special Report for 2014*, www.gartner.com.

[4] Een nadere definiëring wordt in hoofdstuk twee gegeven.

[5] Åse Dragland, 'Big Data – for better or worse', http://www.sintef.no/home/Press-Room/Research-News/Big-Data--for-better-or-worse/ (geraadpleegd op 3 september 2014).

Economische machtsblokken als de Verenigde Staten[6] en de Europese Unie[7] zien grote maatschappelijke en economische kansen voor de inzet van Big Data. Tegelijkertijd brengt het, zoals elke technologische ontwikkeling, ook mogelijke risico's met zich mee. Deze spelen in het bijzonder ten aanzien van de bescherming van fundamentele rechten.[8] Dit roept fundamentele vragen op ten aanzien van de adequaatheid van huidige regelgeving en de invulling daarvan. De herziening van privacyrichtlijn 95/46/EG in de vorm van de voorgestelde Algemene Verordening Gegevensbescherming moet dan ook mede in dat licht bekeken worden.[9]

1.2 Onderzoek

In dit preadvies wordt onderzocht wat de gevolgen van het inzetten van Big Data zijn op fundamentele rechten, in het bijzonder op de bescherming van de persoonlijke levenssfeer[10], de bescherming van persoonsgegevens[11], het non-discriminatiebeginsel[12], de vrijheid van gedachte[13] en de vrijheid van meningsuiting.[14] Centraal in dit preadvies staat hierbij dat de technologische ontwikkelingen niet uitsluitend in de sleutel van privacyvraagstukken geformuleerd dienen te worden, maar juist in de sleutel van een samenspel van fundamentele rechten en de onderliggende ethische overwegingen. Voor de bescherming van deze fundamentele rechten wordt het gegevensbeschermingsrecht als belangrijk instrument gezien.

Om een antwoord op deze te vraag te kunnen formuleren zal in hoofdstuk twee het begrip Big Data nader onderzocht en gedefinieerd worden. Aansluitend wordt Big Data in relatie gebracht tot enkele andere technologische ontwikkelingen die van belang zijn om de gevolgen van Big Data voor fundamentele rechten te duiden. Vervolgens zal in hoofdstuk drie een overzicht gegeven worden van de (internationaal) juridische grondslag voor de genoemde

---

[6] Executive Committee of the White House, *Big Data: Seizing Opportunities, preserving values*, 1 mei 2014.
[7] Mededeling van de Europese Commissie, *Naar een bloeiende data-economie*, 2 juli 2014.
[8] Big Data kan echter ook gevolgen hebben op het vlak van mededinging en consumentenbescherming. Zie o.a. European Data Protection Supervisor, *Preliminary opinion on privacy and competitiveness in the age of big data:*
The interplay between data protection, competition law and consumer protection in the Digital Economy, mei 2014.
[9] Voorstel voor een verordening betreffende de bescherming van natuurlijke personen in verband met de verwerking van persoonsgegevens en betreffende het vrije verkeer van die gegevens (Algemene Verordening Gegevensbescherming), COM(2012)/11 final, 25 januari 2012.
[10] Artikel 8 EVRM, artikel 7 Handvest van de Grondrechten van de EU, artikel 16 Verdrag van de Werking van de Europese Unie.
[11] Artikel 8 Handvest van de Grondrechten van de EU.
[12] Artikel 14 EVRM, artikel 1 GW, Algemene Wet Gelijke behandeling, artikel 21 Handvest van de Grondrechten van de EU.
[13] Artikel 9 EVRM, artikel 10 Handvesten van de grondrechten van de EU, artikel 18 Universele verklaring van de rechten van de mens.
[14] Artikel 10 EVRM en artikel 11 Handvest van de Grondrechten van de EU.

fundamentele rechten worden in relatie tot technologische ontwikkelingen in de maatschappij. In de vier opvolgende hoofdstukken zullen vervolgens een viertal thema's behandeld worden om de gevolgen van Big Data voor de persoonlijke levenssfeer, de bescherming van persoonsgegevens, het non-discriminatie beginsel, de vrijheid van gedachte en de vrijheid van meningsuiting te duiden. Allereerst zal worden geanalyseerd wat Big Data betekent voor de positie van het individu te midden van Big Data aan de hand van het in het gegevensbeschermingsrecht centrale begrip persoonsgegeven. In hoofdstuk vijf wordt ingegaan op transparantie van en voor het individu. In hoofdstuk zes worden de gevolgen van Big Data in de context van geautomatiseerde besluitvorming onderzocht. Ten slotte zal het gebruik van Big Data door de overheid besproken worden en wordt afgesloten met een conclusie.

2.1 Big Data

In de inleiding is Big Data kort beschreven. Voor een goed begrip hiervan is echter een nadere analyse vereist. Informatie wordt inmiddels namelijk al decennia met behulp van ICT verwerkt, hierbij gebruik makend van statistiek, in bijvoorbeeld de financiële en retail sector. Verwante begrippen zijn datamining en business intelligence.

Wanneer naar een definitie van Big Data wordt gevraagd lopen de antwoorden uiteen, vaak afhankelijk van het perspectief van de geïnterviewde.[15] Big Data wordt door de Europese Commissie gedefinieerd als "[…] grote hoeveelheden data van diverse aard die met hoge snelheid uit een groot aantal bronnen van diverse aard worden gehaald. Voor het verwerken van de zeer uiteenlopende realtime datasets die momenteel ter beschikking staan, zijn nieuwe instrumenten en methoden nodig, zoals sterke processors, software en algoritmen."[16] Gartner omschrijft het als "*Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making*." De centrale eigenschappen zijn derhalve snelheid, volume en verscheidenheid van de data en de verwerking hiervan en het kosteneffectief verkrijgen van nieuwe inzichten en het maken van keuzes op basis van deze data.[17] Wat onder "groot" of "big" moet worden verstaan is gezien de snelle ontwikkelingen op het vlak van verwerkings- en opslagcapaciteit relatief. Dit zal veranderen met de tijd en is sectorafhankelijk.[18] Om de gevolgen van Big Data te kunnen duiden valt binnen

---

[15] Een onderzoek van de University Berkely leverde 42 definities op. http://datascience.berkeley.edu/what-is-big-data/ (geraadpleegd op 10 september 2014).

[16] Mededeling van de Europese Commissie, *Naar een bloeiende data-economie*, 2 juli 2014.

[17] In het Engels wordt verwezen naar de "3 V's" van velocity, volume en variety". Zie Executive Office of the President, *Big data:*
Seizing opportunities, preserving values, mei 2014, p. 4.

[18] Zie de definitie van McKinsey "[…] This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data […]Also note that the definition can vary by sector […]". McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and productivity,*

dit preadvies het gehele proces van verzamelen, opslaan, verwerken/analyseren van en het maken van keuzes op basis van de data binnen de definitie van Big Data. Met name de laatste onderscheidde stap is van belang, omdat op deze wijze met Big Data direct invloed op individuen en onze maatschappij kan worden uitgeoefend.

De enorme groei aan hoeveelheid data wordt voornamelijk door twee ontwikkelingen mogelijk gemaakt. Allereerst dalen de kosten van het opslaan van gegevens die (steeds meer) standaard door organisaties verzameld en verwerkt worden. Zo kunnen deze tegen lagere kosten worden opgeslagen.[19] Ten tweede kan de toename worden toegeschreven aan de grote stappen die gemaakt zijn op het gebied van sensortechnologie. Denk aan een smartphone die tegenwoordig standaard met zo'n dertien sensoren en vele verbindingsmogelijkheden uitgerust is.[20] De productie- en inzetkosten van sensoren lopen snel terug en het formaat wordt steeds kleiner, waardoor gebruik van sensoren binnenkort niet meer is voorbehouden aan high-tech industrie.[21] Ze worden al steeds breder ingezet bij domotica (het automatiseren van woningen).[22] De inzet van sensoren zal de komende jaren doordringen in alle sectoren van de economie, zowel voor consumenten als bedrijven en overheden.

De data kan continu (in *real-time*) en met grote snelheid worden verzameld door de toegenomen mogelijkheden om (draadloos) netwerk- en internetverbindingen te gebruiken. Dit wordt ook wel omschreven als het *internet of things* dat uitgebreide machine naar machine communicatie mogelijk maakt. Ook hierdoor kunnen sensoren tegen lage kosten breed ingezet worden.[23]

Bij Big Data verschilt de aard van de gegevens in grote mate, het gaat om heterogene data. Het kunnen foto's, audiobestanden, video, tekst, binaire data, geolocatiegegevens of simpelweg andersoortige gegevens uit sensoren zijn en dat alles gestructureerd of ongestructureerd, persoonsgegevens en niet-persoonsgegevens.

Nieuwe analysetechnieken maken het mogelijk om deze grote hoeveelheden data die snel verzameld worden en die sterk gevarieerd zijn te analyseren en hier informatie en nieuwe kennis uit te halen. Waar echter in het verleden gewerkt werd met een hypothese en (representatieve) steekproef is bij Big Data het

---

juni 2011, p.1.

[19] http://www.zdnet.com/storage-in-2014-an-overview-7000024712/ (geraadpleegd op 15 september 2014).

[20] Denk aan microfoon, camera, GPS, nabijheidssensor, lichtmeter, gyroscoop, magnetometer (kompas), versnellingsmeter, thermometer, touchscreen, NFC, WiFi, Bluetooth, vingerafdrukscanner.

[21] http://www.cio.com/article/2606003/consumer-technology/stanfords-ant-sized-radios-could-connect-the-world.html (geraadpleegd op 13 september 2014)

[22] Verwacht wordt dat in 2022 elke woning gemiddeld 500 "slimme" apparaten bevat. http://www.gartner.com/newsroom/id/2839717 (geraadpleegd op 13 september).

[23] http://www.cio.com/article/2604262/firechats-developer-wants-to-give-iot-a-meshnetwork-boost.html (geraadpleegd op 13 september 2014).

"gebruik alles en zoek de verbanden"-principe leidend. Hierbij gaat het derhalve om het zoeken van verbanden in enorme hoeveelheden data die op zichzelf weinig informatie prijsgeven, maar met behulp van geavanceerde algoritmen nieuwe inzichten bieden.[24] Hierbij staan correlaties centraal, niet het vaststellen van causaliteit.[25] Op basis van deze correlaties kunnen modellen worden gemaakt waarmee toekomstige ontwikkelingen voorspeld kunnen worden.

Data spreekt niet voor zichzelf. Data moet verzameld worden, verwerkt worden met algoritmen en vervolgens dient de uitkomst van de analyse geïnterpreteerd te worden. In alle drie stadia kunnen mogelijk vertekeningen optreden.[26] Zo zal zelfs als er sprake is van een grote hoeveelheid verzamelde data slechts een gedeelte van de werkelijkheid verzameld worden. Ten tweede zal van deze gegevens een representatie van deze werkelijkheid worden gemaakt. Er wordt, in andere woorden, een model of abstractie van de werkelijkheid gemaakt. Vervolgens komen met behulp van algoritmen tijdens de analysefase verschillende (soorten) verbanden boven drijven. Ten slotte dienen de uitkomsten hiervan weer geïnterpreteerd te worden door mens of machine. Het gebruik van Big Data als instrument voor het genereren van nieuwe inzichten is derhalve geen sinecure en vereist een grote hoeveelheid kennis van statistiek en kennisrepresentatie.

2.2 Big Data als onderdeel van grotere technologische verandering

In paragraaf 2.1 is beschreven wat Big Data omhelst en welke technologische ontwikkelingen Big Data mogelijk maken. Andersom is Big Data ook een *enabler* die andere technologische ontwikkeling mogelijk maakt en voortstuwt. Om de ontwikkeling van Big Data en de gevolgen ervan voor fundamentele rechten te doorgronden zal Big Data dan ook in context van deze toekomstige ontwikkelingen geplaatst moeten worden.

In het rapport "Intieme Technologie" van het Rathenau Instituut wordt deze ontwikkeling als volgt verwoord: "We beleven thans het historische omslagpunt waarop de afstand tussen technologie en onszelf in hoog tempo kleiner wordt."[27] Dit wordt veroorzaakt door een convergentie van nano-, biomedische-,

---

[24] Een algoritme is een verzameling van stappen die gevolgd worden om een probleem op te lossen. Zie http://www.merriam-webster.com/dictionary/algorithm.

[25] Correlatie geeft aan dat er een statistisch significant verband is tussen X en Y. Als bijvoorbeeld X zich voordoet, doet Y zich ook voor. Correlatie is probabilistisch van aard. Causaliteit betekent een duidelijk oorzaak en gevolg. X veroorzaakt Y. Deze natuurwetenschappelijke causaliteit is niet altijd hetzelfde als juridische causaliteit. De criteria voor de vaststelling van de laatste verschilt per rechtsgebied, zoals strafrecht, civielrecht, bestuursrecht. Zo is in het civielrecht toerekening een belangrijk element. Zie ook A.J. Akkermans, 'Causaliteit bij letselschade en medische expertise', *Tijdschrift voor vergoeding personenschade* 2003-4, p. 93-104.

[26] http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/ (geraadpleegd 1 november 2014).

[27] Rathenau Instituut, *Intieme technologie: de slag om ons lichaam en gedrag,* Den Haag: Rathenau Instituut 2014, p. 8.

informatie- en cognitieve technologie (NBIC). Hiermee vervagen steeds meer de grenzen tussen de digitale (ICT) en de biologische en fysiologische wereld. Er treedt een versmelting tussen NBIC-technologie en de mens op, steeds meer technologie komt in ons, tussen ons, over ons en wordt als ons. De eerste ontwikkelingen hiervan zijn al zichtbaar bij de snelle ontwikkeling van *wearables*, slimme apparatuur die ons niet alleen informatie over de staat van de wereld geven, maar vooral over onze interne staat, zoals hartslag, glucoseniveaus, stressniveau en zelfs hersenactiviteit.[28] Er wordt wel gesproken over de *quantified self*.[29] Daarnaast maakt deze NBIC convergentie mogelijk dat onze leefwereld steeds adaptiever gemaakt kan worden, steeds meer aanpasbaar, al dan niet aangestuurd door kunstmatige intelligentie. Dit betreft niet alleen bekende voorbeelden als thermostaten[30] die zich aanpassen aan het leefpatroon van bewoners van een huis, maar eveneens zichzelf organiserende of zelfs bouwende (mini-)robots.[31] Ten slotte wordt een steeds nauwere verbondenheid van ICT en levende wezens, inclusief mensen, mogelijk. Een van de verstrekkendste ontwikkelingen hiervan is machine-hersencommunicatie , ook wel *mind-machine communication* genoemd, die mogelijk gemaakt wordt door cognitieve technologie en waar nu al praktische toepassingen mogelijk zijn.

Big Data is zowel een gevolg als een voorwaarde voor deze vergaande ontwikkelingen. De versmelting van mens en machine en versmelting van de omgeving met technologie brengt toenemende mogelijkheden voor het verzamelen van informatie met zich mee. Ten tweede kan met deze informatie steeds verdergaande analyses uitgevoerd kunnen worden. Ten slotte kan er door de versmelting van technologie in onszelf en onze omgeving ook steeds meer controle over onze digitale en fysieke wereld en onszelf uitgeoefend worden. Deze onderdelen staan centraal in de hierop volgende hoofdstukken.

3. Samenspel van fundamentele rechten in een wereld met Big Data

Evenals dat Big Data niet alleen op zichzelf bekeken dient te worden, geldt dit ook voor de fundamentele rechten die door deze ontwikkelingen worden geraakt. Zoals in de introductie aangegeven gaat het hierbij in het bijzonder om de nauw met elkaar verbonden bescherming van de persoonlijke levenssfeer, de bescherming van persoonsgegevens, non-discriminatie en de vrijheid van gedachte en de vrijheid van meningsuiting. Deze fundamentele rechten zullen in samenhang besproken worden, omdat ze elkaar ondersteunen en versterken. De

---

[28] Waar EEG-scanners voorheen voorbehouden waren aan ziekenhuizen zijn ze tegenwoordig ook binnen bereik van consumenten. Zie http://interaxon.ca. Ethici vrezen dat hiermee de laatste privacybarriere op termijn geslecht gaat worden, namelijk de beslotenheid van onze gedachten. Zie D.J. Church, 'Neuroscience in the Courtroom: An International Concern', *William & Mary Law Review* 53-5, p. 1826-1853.

[29] http://ecp.nl/item/4158 (geraadpleegd op 21 oktober 2014).

[30] Denk aan het bedrijf NEST dat slimme, zichzelf automatisch configurerende thermostaten en rookalarmen produceert.

[31] http://wyss.harvard.edu/viewpressrelease/162 en http://www.seas.harvard.edu/news/2014/08/self-organizing-thousand-robot-swarm (geraadpleegd op 21 oktober 2014).

fundamentele rechten hebben, in andere woorden, naast een eigenstandig ook zeker een complementair en zelfs instrumenteel karakter ten opzichte van elkaar. Zo is de bescherming van persoonsgegevens instrumenteel bij het beschermen van de persoonlijke levenssfeer, non-discriminatie en de vrijheid van gedachte.[32] Tegelijkertijd kan er ook spanning ontstaan, zoals de bescherming van persoonsgegevens en de persoonlijke levenssfeer versus vrijheid van meningsuiting.[33] Uit de grote hoeveelheid van jurisprudentie die zich rond deze fundamentele rechten heeft ontwikkeld kunnen de bescherming van de persoonlijke autonomie[34], keuzevrijheid[35], fysieke en psychische integriteit van het individu[36] en een vrije stroom van denkbeelden als basis voor zelfontwikkeling[37] als kernpunten worden gezien. Naast het belang van de ontwikkeling van het individu en het waarborgen van zijn vrijheden worden deze rechten ook als waarborgen voor het goed functioneren van een pluriforme democratische samenleving gezien.

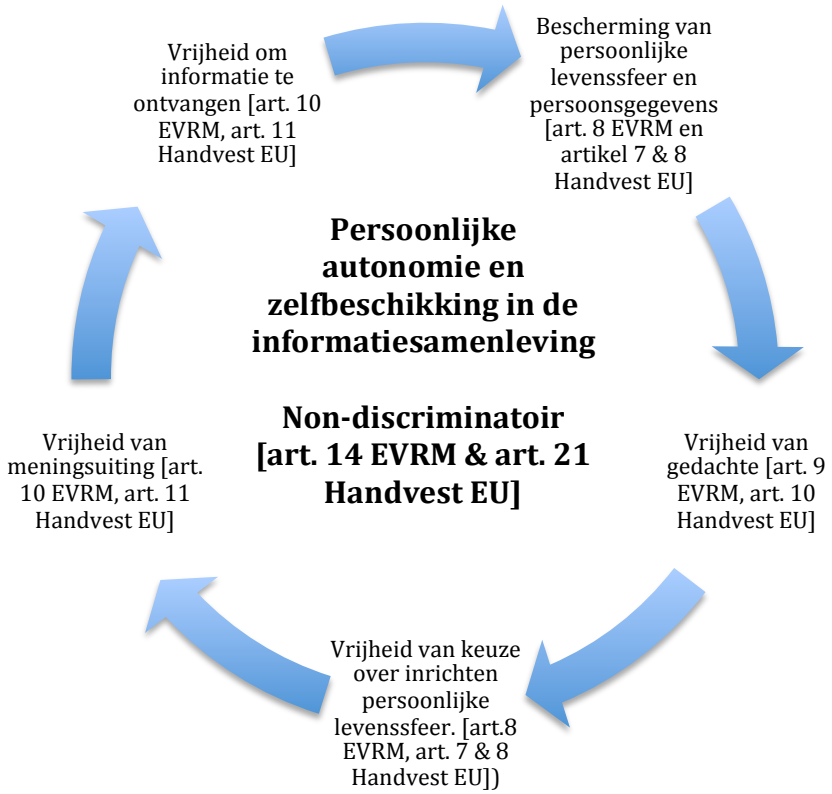
De hiervoor genoemde fundamentele rechten zijn niet van absolute aard. In casu dient gekeken te worden naar de functie ervan in de samenleving en dient er een afweging gemaakt te worden tussen (individuele en publieke) conflicterende rechten en belangen.[38]

Ondanks dat fundamentele rechten voortvloeiend uit het EVRM in principe tussen de staat en burger gelden, wordt in de Straatsburgse jurisprudentie eveneens in gevallen een positieve verplichting aan de verdragsstaten opgelegd om de in het verdrag neergelegde rechten ook in horizontale verhoudingen, tussen burgers onderling en daarmee ook burgers en niet-statelijke actoren, te beschermen.[39] De mate van bescherming die geboden moet worden tegen bijvoorbeeld inbreuken op de persoonlijke levenssfeer kan met de tijd veranderen en is onder andere afhankelijk van toekomstige geavanceerde technologieën die inbreuken mogelijk maken.[40]

---

[32] Dit gegeven komt expliciet naar voren in de considerans van richtlijn 95/46/EG. Het Duitse Bundesvervassungsgericht maakte deze verbinding al in 1983. Zie BVerfG 15 december 1983, 65.
[33] EHRM 24 juni 2004, 59320/00 (Von Hannover).
[34] EHRM 12 juni 2014, 56030/07 (Fernández Martínez v. Spain); EHRM 29 april 2002, 2346/02 (Pretty v. the United Kingdom), EHRM 15 januari 2009, 1234/05 (Reklos and Davouris v. Greece), EVRM 10 april 2007, 6339/05 (Evans v. the United Kingdom). Zie ook N.R. Koffeman, '(The right to) personal autonomy in the case law of the European Court of Human Rights', http://bit.ly/1xevIe6.
[35] EVRM 18 mei 1976, 6825/74 (X. v. Iceland). Zie T. Gomez-Arostequi, 'Defining private life under the European Convention on Human Rights by referring to reasonable expectations', *California Western International Law Journal,* Vol. 35-2 mei 2005, p. 161.
[36] EHRM 4 december 2008, *NJ* 2009, 410 (S. and Marper v. United Kingdom), r.o. 66.
[37] EVRM 8 juli 1986, 9815/82 (Lingens v. Austria), r.o. 41 en EVRM 18 juli 2000, 26680/25 (Şener v. Turkey), r.o. 39.
[38] HvJ 9 november 2010, zaken C-92/09 en C93/09 (Volker and Markus Schecke v Land Hessen), r.o. 48.
[39] EVRM 8 juli 2003, 36022/97 (Hatton et al. v. United Kingdom) en EVRM 24 juni 2004, 59320/00 (Von Hanover), r.o. 57.
[40] EHRM 4 december 2008, *NJ* 2009, 410 (S. en Marper v. Verenigd Koninkrijk), r.o. 71.

De verhouding tussen de fundamentele rechten kan op velerlei wijze gekenschetst worden. Wanneer de nadruk echter op de drie hiervoor genoemde kernpunten van persoonlijke autonomie, keuzevrijheid, integriteit en vrije stroom van denkbeelden gelegd wordt kan het volgende proces onderscheiden worden. Hierbij gaat het in het bijzonder om de onderlinge verhoudingen en niet of ze in een specifieke context succesvol ingeroepen zouden kunnen worden jegens de staat of een derde.



Een burger heeft in beginsel de vrijheid om een mening te koesteren en inlichtingen en denkbeelden te ontvangen of te uiten, tenzij de beperking in wet voorzien is en noodzakelijk is in een democratische samenleving en in het belang is van de nauw omschreven gronden van het tweede lid van artikel 10 EVRM. De vrijheid om informatie te ontvangen is een belangrijk aspect van de zelfontwikkeling van het individu en een pluriforme democratische samenleving. Zonder vrije nieuwsgaring wordt de stroom van ideeën binnen de maatschappij en daarmee naar individuen bemoeilijkt.[41]

In de eigen persoonlijke levenssfeer kan het individu zonder interferentie van buitenaf zijn eigen persoonlijkheid, identiteit en opvattingen, mede gebaseerd op

[41] De vrije stroom van ideeën kan ook in economisch opzicht van groot belang worden geacht. Recente sociaalpsychologische studies, uitgevoerd met moderne technologie en Big Data analyse waarmee de interactie tussen personen en groepen kan worden geanalyseerd, tonen een sterke relatie tussen een vrije stroom van ideeën en innovatie en hiermee economisch succes aan. Zie A. Pentland, *Social Physics: How Good Ideas Spread - The Lessons from a New Science,* New York: Penguin Press 2014.

denkbeelden ontvangen van andere partijen, vormgeven.[42] Dit aspect kan gekwalificeerd worden als negatieve vrijheid, vrij zijn van inmenging. Deze persoonlijke levenssfeer wordt beschermd door artikel 8 EVRM dat als gekwalificeerd recht wederom alleen mag worden ingeperkt onder de condities van het tweede lid. In het verlengde van de bescherming van de persoonlijke levenssfeer ligt de bescherming van persoonsgegevens, om zo (volledige) transparantie van de persoonlijke levenssfeer te voorkomen.[43] De verwerking van persoonsgegevens is nodig voor het goed functioneren van de maatschappij, maar de verwerking ervan is gereguleerd. Een nadere invulling van het recht op bescherming van persoonsgegevens wordt momenteel gegeven door Conventie 108[44], richtlijn 95/46/EG en de hiervan afgeleide Wet bescherming persoonsgegevens (Wbp). Zowel de richtlijn als de conventie ondergaan momenteel een vernieuwingsslag. Ter vervanging van de richtlijn is de Algemene Verordening Gegevensbescherming in januari 2012 door de Europese Commissie gepresenteerd. Het Europese Parlement heeft in eerste lezing haar positie bepaald, de onderhandelingen in de Raad duren nog voort. Over modernisering van Conventie 108 wordt momenteel onderhandeld in Straatsburg.

In het verlengde van vrije vorming van persoonlijkheid, identiteit en opvattingen ligt de vrijheid van gedachte die zijn grondslag vindt in artikel 9 EVRM waarin in het eerste lid wordt gesteld: “een ieder heeft recht op vrijheid van gedachte, geweten en godsdienst […]” . De meeste jurisprudentie omtrent artikel 9 EVRM ziet op de vrijheid van godsdienst, maar ook het hebben van principes of overtuigingen zoals veganisme[45], pacifisme[46] en atheïsme[47] vallen onder dit recht. Het recht kan in twee delen worden onderscheiden, namelijk ten aanzien van het *forum internum* en het *forum externum*.[48] Het *forum internum* betreft de eigenlijke gedachte of overtuiging van een persoon, zonder dat deze gedachte zich door een handeling in de buitenwereld gemanifesteerd heeft.[49] Zo kan iemand een voorkeur voor een politicus hebben die pas geopenbaard wordt en

---

[42] In de literatuur wordt hierbij een onderscheid gemaakt tussen de *ipse* identiteit, het reflectieve zelfbewustzijn een uniek individu te zijn en de *idem* identiteit, waarmee we ons positioneren in de wereld, bijvoorbeeld sociaal, cultureel, economisch en juridisch. Zie o.a. P.J.A. de Hert, *A right to identity in the face of the internet of things,* Straatsburg: UNESCO 2008, p.1.

[43] In feite reguleert dit de vrijgave en het verwerken van attributen van de *idem* identiteit aan de buitenwereld, de persoonsgegevens.

[44] Conventie 108 voor de bescherming van individuen in relatie tot de geautomatiseerde verwerking van persoonsgegevens, ook wel verdrag van Straatsburg van 1981 genoemd, valt binnen het juridisch kader van de Raad van Europa.
http://conventions.coe.int/Treaty/en/Treaties/html/108.htm

[45] EHRM 10 februari 1993, 18187/91 (W. v. the United Kingdom).

[46] EHRM 12 oktober 1978, 7050/75 (Arrowsmith v. the United Kingdom).

[47] EHRM 3 december 1986, 10491/83 (Angeleni v. Sweden).

[48] J. Murdoch, *Protecting the right to freedom of thought, conscience and religion under the European Convention on Human Rights,* Straatsburg: 2012, p. 18.

[49] EHRM 22 februari 1995, 22838/93 (Van den Dungen v. the Netherlands.)

bewezen kan worden door het uitbrengen van een stem.[50] De bescherming van het *forum internum* is hoofdzakelijk een negatieve vrijheid. Het *forum externum* betreft het handelen in overeenstemming met de overtuiging., gedachte of het geloof, een vorm van positieve vrijheid. Dit aspect heeft met name betrekking op het tweede deel van het eerste lid van het artikel. "[…] dit recht omvat tevens de vrijheid om van godsdienst of overtuiging te veranderen, alsmede de vrijheid hetzij alleen, hetzij met anderen, zowel in het openbaar als privé zijn godsdienst te belijden of overtuiging tot uitdrukking te brengen in erediensten, in onderricht, in praktische toepassing ervan en in het onderhouden van geboden en voorschriften." Alleen het belijden van een godsdienst of het uiten van een overtuiging kan aan beperkingen worden onderworpen, het *forum internum* niet.[51] In jurisprudentie die zich tot op heden heeft ontwikkeld omtrent artikel 9 EVRM waren de eisen van de eiser met name gericht op het kunnen uitoefenen van activiteiten op het *forum externum* en niet inmenging in het *forum internum.* De jurisprudentie van het EHRM heeft zich dan ook gericht op rechtsvragen met betrekking tot het *forum externum,* wat de vraag oproept hoe het Europese Hof voor de Rechten van de Mens een beroep op uitsluitend het *forum internum* zal behandelen, in het bijzonder in de context van horizontale verhoudingen.[52] De mogelijkheden om direct en indirect invloed uit te oefenen op cognitieve processen van personen nemen namelijk sterk toe, wat een bedreiging kan vormen voor de vrijheid van gedacht in het *forum internum*. Zo kan op basaal niveau al met fMRI techniek gezichten die mensen zien of zich voorstellen uitgelezen worden uit de hersenen.[53] Ook werden de hersenen van twee ratten via het internet met elkaar verbonden, waardoor ze over grote afstand konden samenwerken om een actie uit te voeren en zo een beloning te krijgen.[54]

De vrijheid van het maken van keuzes en uitingen gebaseerd op de gedachten en overtuigingen van het *forum internum* wordt, naast artikel 9, eerste lid, EVRM, met name gebaseerd op de notie van persoonlijke autonomie beschermd onder artikel 8 EVRM.[55] Deze keuzevrijheid kan alleen worden ingeperkt op basis en onder de voorwaarden van artikel 8, tweede lid, EVRM.

Het uiten van gedachten en meningen wordt zoals eerder aangegeven beschermd door artikel 10 EVRM, de vrijheid van meningsuiting. Deze gedachten en meningen kunnen vervolgens weer door anderen binnen de maatschappij ontvangen worden. Hiermee wordt de cirkel rond. Alhoewel niet uitputtend, zorgen de genoemde fundamentele rechten voor het waarborgen van de

---

[50] EHRM 8 juli 2008, 9103/04 (Georgian Labour Party v. Georgia), als genoemd in J. Murdoch, *Protecting the right to freedom of thought, conscience and religion under the European Convention on Human Rights,* Straatsburg: 2012, p. 19.
[51] Artikel 9, tweede lid, EVRM.
[52] Deze vraag geldt uiteraard ook voor de uitleg van artikel 10 Handvest van de grondrechten van de Europese Unie door het Hof van Justitie.
[53] A. S. Cowen, 'Neural portraits of perception: Reconstructing face images from evoked brain activity', *NeuroImage* 2014-94, p. 12-22.
[54] http://www.nature.com/news/intercontinental-mind-meld-unites-two-rats-1.12522 (Geraadpleegd 20 november 2014).
[55] EHRM 29 april 2002, 2346/02 (Pretty v. The United Kingdom), r.o. 61 en 82.

persoonlijke autonomie, in de zin van vorming van persoonlijkheid, identiteit en opvattingen.

Het recht op non-discriminatie ex artikel 14 EVRM zorgt er hierbij voor dat rechten krachtens het EVRM voor een ieder gelden en bijvoorbeeld niet op basis van geslacht, etniciteit of ras kunnen worden onthouden.[56] Artikel 21 Handvest EU gaat hierbij nog verder met een algeheel verbod op discriminatie, met name op grond van geslacht, ras, kleur, etnische of sociale afkomst, genetische kenmerken, taal, godsdienst of overtuigingen, politieke of andere denkbeelden, het behoren tot een nationale minderheid, vermogen, geboorte, een handicap, leeftijd of seksuele geaardheid.

4. Het individu te midden van Big Data

In de enorme hoeveelheid digitale data die vandaag de dag gegenereerd wordt zijn veel gegevens te vinden die betrekking hebben op natuurlijke personen. Deze gegevens worden verzameld voor velerlei administratieve processen, marketing en het verlenen van diensten aan individuen. Daarnaast worden ze ook gebruikt om beter inzicht te krijgen in het gedrag van individuen of groepen van mensen. Dit roept de vraag op welke positie het individu te midden van deze Big Data inneemt en hoe deze gekenschetst kan worden.

Het verwerken van gegevens van individuen ligt gevoelig, omdat het kan leiden tot een inbreuk op de informationele privacy van het individu.[57] Dit omvat zeggenschap over welke informatie wanneer, hoe en in hoeverre over hen gedeeld worden. Daarnaast kan het ook om relationele privacy gaan, omdat uit gegevens de sociale context van een persoon te herleiden valt, namelijk met wie hij omgaat. Met de aan Big Data gelieerde ontwikkelingen zoals geschetst in hoofdstuk twee kan het ook de ruimtelijke privacy beïnvloeden, aangezien individuen in bijvoorbeeld hun eigen huis steeds meer geobserveerd kunnen worden. Ten slotte kan met NBIC ook de lichamelijke privacy steeds meer geraakt worden, bijvoorbeeld door *wearables* of zelfs technologie in het lichaam.

4.1 Identiteit

Het gegevensbeschermingsrecht, uitgewerkt in onder andere de Wbp, tracht het verwerken van gegevens die individuen raken te reguleren, de zogenaamde persoonsgegevens. Artikel 1 sub a Wbp definieert een persoonsgegeven als “elk gegeven betreffende een geïdentificeerde of identificeerbare natuurlijke persoon.” De definitie in artikel 2 sub a richtlijn 95/46/EG voegt daarbij aan toe “als identificeerbaar wordt beschouwd een persoon die direct of indirect kan worden geïdentificeerd, met name aan de hand van een identificatienummer of

[56] De reikwijdte van het non-discriminatiebeginsel binnen het EVRM is met de het twaalfde protocol van het EVRM uitgebreid. Waar artikel 14 beperkt was tot rechten krachtens het EVRM is onder protocol 12 het gelijkheidsbeginsel van toepassing voor alle rechten. Nederland heeft dit protocol in 2004 geratificeerd, waarna het in 2005 in werking is getreden.

[57] A. Westin, *Privacy and Freedom*, New York: 1967.

van een of meer specifieke elementen die kenmerkend zijn voor zijn of haar fysieke, fysiologische, psychische, economische, culturele of sociale identiteit." Gegevens over personen zullen op velerlei wijze in de Big Data terecht komen. Wanneer er geen sprake is van persoonsgegevens is het gegevensbeschermingsrecht niet van toepassing. Gezien de grote belangen die gemoeid zijn bij het verwerken van gegevens over individuen is het niet verrassend dat er veelvuldig gediscussieerd wordt over de reikwijdte van dit begrip.

Het dient derhalve te gaan om een geïdentificeerd of identificeerbaar natuurlijk persoon. Wanneer een verzameling gegevens de naam, woonplaats, geboortedatum, nationaliteit en taal omvat is er meestal geen twijfel over mogelijk dat het gegevens betreffende een identificeerbaar natuurlijk persoon betreft.[58] Dit ligt anders wanneer er sprake is van andersoortige gegevens die in het normale maatschappelijk verkeer niet direct een associatie met een aanwijsbaar individu opleveren, zoals een (cookie-)code, waarmee bijvoorbeeld (koop-)gedrag van een persoon wordt bijgehouden en advertenties worden getoond. Betreft dit wel een identificeerbaar natuurlijk persoon?

De computerwetenschap maakt niet alleen de verwerking van deze gegevens mogelijk, maar biedt ook een analytisch instrumentarium om deze vraag uit te werken.[59] De natuurlijke persoon kan omschreven worden als "entiteit", de mens van vlees en bloed. Deze entiteit kan vervolgens aangeduid worden met verschillende identiteiten, die bestaan uit een aantal attributen die gezamenlijk deze identiteit vorm geven. In het normale taalgebruik gaan aanduiding van de entiteit en de burgerlijke identiteit vaak hand in hand. Identiteit heeft in deze zin een sociale functie, namelijk een permanente verwijzing naar een persoon.[60] Als we spreken over Willem-Alexander, wonende te Wassenaar weet iedereen wie we bedoelen. Dezelfde persoon zouden we echter ook kunnen aanduiden als "vader van drie kinderen, wonende te Wassenaar en liefhebber van voetbal". Deze aanduiding zal in andere situaties toepasselijk kunnen zijn. Dit voorbeeld toont aan dat een identiteit contextueel is.[61] Beide identiteiten refereren naar dezelfde persoon (entiteit) en zijn hier een representatie van. De identiteit bestaat uit een aantal attributen of kenmerken. Of deze identiteit (verzameling attributen of één uniek attribuut zoals BSN of een apparaatnummer) uniek is hangt af van de vraag of er binnen de groep die de verantwoordelijke[62] kan overzien niet meer dan één persoon (entiteit) aan deze set van attributen

---

[58] HvJ 17 juli 2014, C-141/12 en C-372/12, r.o 38

[59] J. Talburt, 'Entity and identity resolution', http://mitiq.mit.edu/IQIS/2010/Addenda/T2A%20-%20JohnTalburt.pdf (Geraadpleegd 31 oktober 2014).

[60] STI Workingpaper 2007/7, At a Crossroads: "personhood" and Digital Identity in the Information Society (OECD), p. 27.

[61] STI Workingpaper 2007/7, At a Crossroads: "personhood" and Digital Identity in the Information Society (OECD), p. 26.

[62] De verantwoordelijke wordt in Artikel 1 sub d Wbp gedefinieerd als "de natuurlijke persoon, rechtspersoon of ieder ander die of het bestuursorgaan dat, alleen of te zamen met anderen, het doel van en de middelen voor de verwerking van persoonsgegevens vaststelt."

(identiteit) voldoet.[63] De samenstelling van deze groep kan wijzigen. Identiteit is in die zin onzeker of kansgebaseerd en contextgebonden.

4.2 Identiteit in Big Data

De identiteit van een persoon kan geautomatiseerd, bijvoorbeeld met behulp van Big Data analyse, tot stand worden gebracht. Deze praktijk wordt ook wel als *profiling* omschreven. Theoretisch maar ook praktisch gezien zou er een profiel van een persoon kunnen worden opgesteld, bestaande uit kenmerken die gezamenlijk uniek zijn voor die persoon. In gevallen kan met behulp van dit profiel een persoon fysiek terug gevonden worden, in andere gevallen niet. In andere woorden, een identiteit kan niet altijd terug herleid worden naar de achterliggende entiteit in de fysieke wereld.

Dit leidt tot een onderscheid in wat Leenes noemt, opzoek-identiteiten en herken-identiteiten.[64] Bij de eerste variant is iemand terug te herleiden naar zijn burgerlijke identiteit, dan wel direct door NAW gegevens, dan wel indirect via bijvoorbeeld zijn telefoonnummer of kentekennummer waar een derde partij NAW gegevens aan kan koppelen. Bij de tweede variant is een persoon onderscheidbaar van anderen, maar is er geen link te leggen naar de burgerlijke identiteit van de persoon. Naast deze twee vormen onderscheidt Leenes eveneens een categorische-identiteit die aansluit bij de Big Data praktijk, waarbij personen op basis van attributen die bij de verantwoordelijke bekend zijn in bepaalde klassen (collectieve identiteiten of groepsprofielen) worden ingedeeld. Het behoren tot een bepaalde klasse impliceert meestal dat de betreffende persoon ook over andere kenmerken beschikt, zelfs indien die niet door de verantwoordelijke zelf zijn vastgesteld. Het hebben van een bepaalde identiteit kan dan ook, soms onverwachte, gevolgen met zich meebrengen.[65]

Vaak zal binnen Big Data analyse sprake zijn van een combinatie van een herken-identiteit gecombineerd met een categorische identiteit. Dat betekent dat de persoon binnen de dataset uniek te herkennen is en dat hem op basis van de klasseindeling extra kernmerken worden toegedicht, zonder dat zijn burgerlijke identiteit te achterhalen is. Een identiteit is in die zin samengesteld. Enerzijds vormt de betrokkene[66] zelf deel van zijn identiteit, anderzijds wordt deze door anderen vormgegeven.[67]

Bij de behandeling van de Algemene Verordening Gegevensbescherming is door het Europees parlement het onderscheid in verschillende vormen van

---

[63] Dit wordt ook wel k-Anonimity genoemd, waarbij k in dit geval 1 moet zijn. http://www.springerreference.com/docs/html/chapterdbid/64340.html

[64] R. Leenes, 'Do They Know Me? Deconstructing Identifiability', *University of Ottawa Law & Technology Journal* 2007-4, p. 146.

[65] STI Workingpaper 2007/7, At a Crossroads: "personhood" and Digital Identity in the Information Society (OECD), p. 26.

[66] Artikel 1 sub f Wbp. De betrokkene is degene op wie een persoonsgegeven betrekking heeft.

[67] STI Workingpaper 2007/7, At a Crossroads: "personhood" and Digital Identity in the Information Society (OECD), p. 26.

identiteiten in zekere mate erkend. In artikel 4 sub 2a wordt het begrip "pseudonieme gegevens" gedefinieerd. Dit zijn "persoonsgegevens die niet kunnen worden geattribueerd aan een specifieke betrokkene zonder het gebruik van aanvullende informatie, zo lang deze informatie gescheiden wordt bewaard en onderworpen is aan technische en organisatorische maatregelen om non-attributie te verzekeren." In feite wordt er een informationele machtenscheiding gemaakt.[68] De opzoek-identiteit wordt gescheiden van de herken-identiteit. Als het dit type gegeven betreft, wordt de verantwoordelijke in gevallen meer ruimte gegeven om gegevens te verwerken. De Artikel 29 Werkgroep ziet dit niet als een anonimiseringstechniek, maar als een beveiligingsmaatregel.[69] De Artikel 29 Werkgroep ziet effectieve anonimisering als een situatie waarin het voor alle partijen onmogelijk is om een specifiek persoon uit de dataset te lichten, gegevens binnen een set of meerdere sets van gegevens aan een persoon te linken of extra informatie over een persoon uit de dataset af te leiden.[70]

Het maken van onderscheid tussen de beschreven vormen van identiteiten is van belang in het debat over Big Data, omdat enerzijds langs deze lijnen het begrip persoonsgegeven in de toekomst verder zal moeten worden ingevuld. Anderzijds dwingt het tot een precieze definiëring van begrippen als pseudonimisering en ook anonimisering. Als een persoon herkend kan worden aan een van zijn identiteiten en basis daarvan anders beoordeeld of behandeld wordt[71], is het moeilijk vol te houden dat het om anonieme gegevens, niet-zijnde persoonsgegevens gaat.[72]

Het gebruik van verschillende soorten identiteiten kan gebruikt worden om mede invulling te gegeven aan de beginselen van proportionaliteit en subsidiariteit waaraan moet worden voldaan bij het verwerken van persoonsgegevens. In veel gevallen is het namelijk niet nodig om de burgerlijke identiteit van iemand te gebruiken, maar kan van een andere vorm van identiteit, bijvoorbeeld herken-identiteit, gebruik gemaakt worden die voor de betrokkene minder risico's met zich meebrengt. Deze risicovermindering ziet echter zoals gezegd voornamelijk op het waarborgen van de beveiliging of vertrouwelijkheid van de gegevens en niet op de gevolgen die de gegevens met zich meebrengen, bijvoorbeeld in het geval van geautomatiseerde besluitvorming. De link tussen individu en gegeven kan immers nog steeds gelegd worden.

---

[68] G. Hornung en C. Schnabel, 'Data protection in Germany I: The population census decision and the right to informational self-determination', *Computer Law & Security Report* 25-1, 2009, p. 85.
[69] Artikel 29 Werkgroep 10 april 2014, Opinion 05/2014 on Anonymisation Techniques, p. 3. De Artikel 29 Werkgroep is het samenwerkingsverband van onafhankelijke toezichthouders dat conform artikel 29 van privacyrichtlijn 95/46/EG ingesteld is.
[70] Artikel 29 Werkgroep 10 april 2014, Opinion 05/2014 on Anonymisation Techniques, p. 9.
[71] *Kamerstukken II* 1997/1998, 25 892, nr. 3. p. 46-47.
[72] Zie in lijn hiermee Artikel 29 Werkgroep 10 april 2014, Opinion 05/2014 on Anonymisation Techniques, p. 9 en 11.

5. Transparantie over of aan het individu

In het vorige hoofdstuk is beschreven op welke wijze een individu aangeduid kan worden te midden van Big Data. Doordat gegevens op steeds grotere schaal verwerkt kunnen worden en steeds meer gegevens over individuele personen verzameld kunnen worden, lijkt het individu steeds transparanter te worden. De vraag is de betrokkene ook voldoende terug kan kijken, of dat hij geconfronteerd wordt met een situatie van een eenrichtingsspiegel waarbij transparantie maar eenzijdig is.

Transparantie en informatieplichten richting de betrokkene, zoals beschreven in artikel 33 t/m 35 Wbp, worden gezien als belangrijke instrumenten om diens positie en autonomie te versterken en een eerlijke verwerking van persoonsgegevens mogelijk te maken. Het wordt ook wel als basisbeginsel gezien, want een betrokkene kan immers pas zijn rechten laten gelden als hij überhaupt weet dat er gegevens over hem verwerkt worden. Veronderstelt wordt dat dit weten de autonomie van het individu versterkt en hem en betere controlepositie brengt. In deze tijd van Big Data en hieraan gelieerde technologische ontwikkelingen lijkt het instrument transparantie echter soms aan kracht in te boeten.

5.1 Taxonomie van gegevens

Binnen Big Data zijn gegevens afkomstig uit een veelvoud aan verschillende bronnen. Een taxonomie van deze bronnen van gegevens is een goed instrument om de verhouding tussen transparantie en gegevensverzameling scherp te krijgen.[73] Gegevens kunnen allereerst verstrekt worden door de persoon zelf waarop ze betrekking hebben. Deze informatie kan expliciet verstrekt worden, bijvoorbeeld NAW gegevens, of impliciet door bijvoorbeeld een telefoonnummer te bellen. Ten tweede kunnen gegevens geobserveerd worden, bijvoorbeeld door het inzetten van sensoren of het krijgen van gegevens uit een derde bron. De betrokkene zal in veel gevallen niet op de hoogte zijn van het feit dat deze gegevens verzameld worden. Ten derde kunnen gededuceerde gegevens onderscheiden worden, die tot stand komen door gegevens te combineren en daar een logische gevolgtrekking uit te maken. Een voorbeeld hiervan is het uitzoeken uit welke leeftijdscategorie de meeste kopers van een product komen. Ten slotte is er nog de categorie van geïnduceerde gegevens die een probabilistisch karakter hebben. Een inmiddels bekend voorbeeld hiervan is de Amerikaanse supermarktketen Target die op basis van de kooppatronen van een tiener afleidde dat zij waarschijnlijk zwanger was en haar op basis daarvan gerichte aanbiedingen deed, zelfs voordat haar eigen vader van het feit op de hoogte was.[74] Een ander voorbeeld is het afleiden van iemands seksuele

---

[73] Deze taxonomie is ontleent aan OECD, Summary of the OECD Expert Roundtable Discussion'Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking', DSTI/ICCP/REG(2014)4, p. 3

[74] C. Duhigg, 'How companies learn your secrets', *The New York Times* 16 februari 2012, http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&_r=0

geaardheid op basis van zijn of haar Facebook activiteiten.[75] Dit zijn beide voorbeelden van indirecte bijzondere persoonsgegevens.[76]

Binnen de bovenstaande taxonomie neemt de betrokkenheid van de persoon bij en de voorzienbaarheid van het verwerken van deze gegevens steeds verder af. Juist in de laatste twee categorieën zit echter het zwaartepunt van Big Data. Met het toenemend aantal sensoren neemt de hoeveelheid geobserveerde data toe en vervolgens kan daar met Big Data technieken met behulp van deductie, inductie en andere statistiek nieuwe gegevens worden verkregen en zo nieuwe identiteiten of profielen van personen worden gecreëerd. Dit zorgt in potentie voor een sterke informatie-asymmetrie tussen de verantwoordelijke en betrokkene. Bovendien heeft de betrokkene mogelijk weinig invloed op de identiteit of het profiel dat zo van hem gecreëerd wordt.

Bij het verkrijgen van de gegevens bij de betrokkene zelf is de verantwoordelijke ex artikel 33 Wbp verplicht om wanneer gegevens van de betrokkene worden verkregen hem voor het moment van verkrijging van deze informatie te informeren over zijn identiteit, de doeleinden van de verwerking en nadere informatie die een zorgvuldige verwerking waarborgen, tenzij de betrokkene hiervan al op de hoogte is. Dit is volgens de memorie van toelichting bij de Wbp alleen van toepassing als de betrokkene zijn informatie zelf actief verstrekt.[77]

Bij de andere vier onderscheiden bronnen, observeren, deductie en inductie, dient terug gevallen te worden op artikel 34 Wbp. De verantwoordelijke dient dan de hierboven eerder genoemde informatie te verstrekken op het moment van vastleggen van de gegevens of, indien van toepassing, op het moment van eerste verstrekking aan een derde. Verstrekking van deze informatie is uit hoofde van artikel 34 lid 4 Wbp niet verplicht indien dit onevenredig inspanning kost of onmogelijk blijkt te zijn. In dat geval hoeft alleen de herkomst van de gegevens vastgelegd te worden. Artikel 14 van de voorgestelde Algemene Verordening Gegevensbescherming dat dezelfde materie regelt brengt geen wezenlijke verandering mee op dit vlak.

5.2 Transparantie bij Big Data

Bij Big Data waarbij sprake is van veelvoudig gebruik van data voor verschillende doeleinden,[78] met een veelvoud aan spelers, complexe waardenketens, grote snelheid en omvang van gegevensverwerking[79] en diversiteit van soorten gegevens, lijkt het bijzonder lastig een goede invulling te

[75] Y. Bachrach, e.a., 'Personality and Patterns of Facebook Usage', https://research.microsoft.com/pubs/163535/FacebookPersonality_michal_29_04_12.pdf
[76] Artikel 16 Wbp.
[77] *Kamerstukken II* 1997/1998, 25 892, nr. 3, p. 155-156.
[78] Dit levert spanning op met artikel 9 Wbp waarin gesteld wordt dat persoonsgegevens niet verwerkt mogen worden voor een doel dat onverenigbaar is met waarvoor de gegevens zijn verkregen.
[79] Dit levert spanning op met artikel 11 Wbp waarin het beginsel van gegevensminimalisatie is beschreven.

geven aan informatieverplichtingen uit de Wbp. De gegevens zijn immers vaak geobserveerd, gededuceerd of geïnduceerd en verbonden aan een herken-identiteit van de betrokkene.[80] Bij *Internet of Things*-toepassingen mist bovendien vaak de mogelijkheid voor uitgebreide communicatie naar de betrokkene. Veel van de apparatuur beschikt niet over een scherm om informatie over te dragen en verzamelt bovendien informatie over personen die de apparatuur niet hebben aangeschaft. In veel gevallen zal dit in de praktijk tot de conclusie kunnen leiden dat communicatie naar de betrokkene onevenredige inspanning vergt. Een actieve verstrekking van informatie conform artikel 34 Wbp bij het verzamelen van gegevens is daarom problematisch.

Een inzageverzoek in zijn persoonsgegevens door de betrokkene op basis van artikel 35 Wbp lijkt meer mogelijkheden te bieden. Op basis hiervan kan de betrokkene zich met redelijke tussenpozen richten tot de verantwoordelijke met de vraag of persoonsgegevens over hem worden verwerkt. Indien dit het geval is dient ingevolge het tweede lid van het artikel een overzicht daarvan in begrijpelijke vorm te worden gegeven. Tevens doet de verantwoordelijke op verzoek mededeling over de logica die ten grondslag ligt aan de geautomatiseerde verwerking van de gegevens.[81]

Zoals in omschreven in hoofdstuk vier zal bij Big Data veel gebruik kunnen worden gemaakt van herken-identiteiten. De burgerlijke identiteit met NAW gegevens van een persoon zijn niet bekend, maar hij heeft wel een unieke identiteit binnen de gegevensverzameling en kan hier aan herkend worden. Indien de gegevens die betrekking hebben op de herkenidentiteit als persoonsgegevens worden gezien zou de betrokkene op verzoek hier inzage in moeten krijgen. Hiervoor is het wel noodzakelijk dat de verantwoordelijke in staat is om de betrokkene voldoende betrouwbaar te herkennen aan zijn herken-identiteit. Identificatie zal dan waarschijnlijk ook moeten plaatsvinden binnen de context waarbinnen de identiteit is gevormd, zodat deze daar kan worden herkend. Een eenvoudig voorbeeld is het vormen van een herken-identiteit op basis van het surfgedrag van een gebruiker op internet. Zodra betrouwbaar genoeg is vastgesteld dat het om wederom om dezelfde unieke internetsurfer gaat zou op basis hiervan aan hem inzage in de over hem verzamelde gegevens geboden kunnen worden.

In artikel 10 van de voorgestelde Algemene verordening gegevensbescherming wordt gesteld dat wanneer de gegevens die door de verantwoordelijke verwerkt worden de verantwoordelijke niet in staat stellen om de natuurlijke persoon te identificeren, de verantwoordelijke niet verplicht is om extra informatie in te winnen om de betrokkene te identificeren om te voldoen aan regels uit de verordening. In feite wordt hiermee gesteld dat een verantwoordelijke van niet-persoonsgegevens geen persoonsgegevens hoeft te maken.

Dit is echter tweeërlei bijzonder. Allereerst impliceert het hebben van de mogelijkheid om gegevens te herleiden naar een persoon dat het al om

[80] Zie voor een beschrijving van de herken-identiteit hoofdstuk vier van dit preadvies.
[81] Artikel 35, vierde lid, Wbp.

persoonsgegevens gaat. Ten tweede is het bijzonder dat dit zo absoluut gesteld is, omdat het in gevallen de betrokkene de mogelijkheid ontneemt om via het gegevensbeschermingsrecht invloed op zijn gegevens uit te oefenen. In principe kan een herken-identiteit immers voldoende zijn om de betrokkene, in ieder geval op verzoek, van informatie te kunnen voorzien.

Bij het bieden van een betekenisvolle manier aan de betrokkene om kennis te nemen van persoonsgegevens en zo inzage te krijgen in de digitale identiteiten die rondom hem zijn gevormd ligt derhalve een grote uitdaging. Deze uitdaging moet echter wel worden aangegaan en met een goede invulling hiervan zou Big Data juist voordelen voor het individu met zich mee kunnen brengen. Door inzage te geven in hoe een individu wordt gezien door de verantwoordelijke, inclusief betekenisvolle inzage in hoe dit tot stand kwam, kan de informatie-asymmetrie opgeheven worden en de positie van de betrokkene juist worden versterkt. Als de betrokkene immers weet hoe naar hem wordt gekeken kan hij hier op reageren en eventueel ageren. Dit is bovendien van bijzonder belang voor het kunnen ontdekken van ongeoorloofde discriminatie. In de complexe ketens van Big Data vraagt dit echter wel veel van de infrastructuur van de verantwoordelijke. Het gaat immers om grote hoeveelheden heterogene data die op zichzelf niet veelzeggend hoeven te zijn, maar juist door analyse belangrijke inzichten opleveren. Werkelijk betekenisvol inzicht geven in de gevormde identiteiten is daarmee een uitdaging. Privacy by design, het al in het ontwerpstadium van een systeem implementeren van de normen van het gegevensbeschermingsrecht, wordt vaak in de sleutel van doelbinding en gegevensminimalisatie gezet, maar is zo gezien ook een basisvoorwaarde om in de context van Big Data transparantie aan de betrokkene te bieden.[82]

6. Het autonome individu

Het controleren van het individu en het daarmee beperken van zijn autonomie is een terugkerend thema rondom Big Data. In verhalen en fictie gaat hierbij vaak de dreiging en kwade opzet uit van één partij. Vaak is de staat die partij.[83] Door de discussie hiertoe te beperken wordt tekort gedaan aan de complexe sociale, economische en technische systemen die onze samenleving en daarmee ook onszelf als individu beheersen.

6.1 Identiteitsvorming

Zoals omschreven in hoofdstukken vier en vijf maken Big Data en gerelateerde ontwikkelingen het mogelijk om een veelvoud aan gegevens over een individu vast te leggen en uit te breiden met gededuceerde of geïnduceerde gegevens. Deze gegevens vormen een (categorische-)identiteit of profiel die aan een individu wordt verbonden.

Deze identiteit wordt derhalve door de verantwoordelijke (deels) samengesteld. De eerste vraag die hierbij speelt is of deze aangemeten identiteit wel klopt. De

[82] Zie artikel 23 van de voorgestelde Algemene Verordening Gegevensbescherming.
[83] Denk aan George Orwells 1984.

verantwoordelijke heeft de plicht om ervoor te zorgen dat de persoonsgegevens juist en nauwkeurig zijn, gelet op het doel van de verwerking.[84] Problematisch hierbij is dat een identiteit in wezen vaak een subjectief gegeven is.[85] Dit geldt evenzeer voor identiteiten die binnen Big Data bestaan. De verantwoordelijke kan namelijk besluiten dat persoon X tot categorie xyz behoord en daarom anders behandeld wordt. Deze categorie xyz kan echter volgens lijnen vastgesteld zijn die in het normale maatschappelijk verkeer geen duidelijke betekenis of classificatie zou hebben. Het recht van de betrokkene om ex artikel 36 Wbp persoonsgegevens over hem "te verbeteren, aan te vullen, te verwijderen, of af te schermen indien deze feitelijk onjuist zijn, voor het doel of de doeleinden van de verwerking onvolledig of niet ter zake dienend zijn" is daarmee voor dergelijke categorisatie maar van beperkte waarde. De controle van het individu over zijn identiteit(-en) is derhalve maar gering.

De logica op basis waarvan nieuwe informatie over een individu en daarmee diens identiteit gecreëerd wordt, vormt een belangrijk startpunt om toch meer controle over de vorming van identiteiten of profielen te krijgen. Op basis van artikel 35, vierde lid, Wbp kan de betrokkene op verzoek inzage krijgen in de onderliggende logica van de geautomatiseerde verwerking. Uit de memorie van toelichting volgt dat dit artikel van toepassing is "[…]in geval bijzondere computerprogrammatuur een wijze van verwerking mogelijk maakt die de betrokkene niet reeds duidelijk is uit de mededeling ingevolge het tweede lid" en dat deze mededeling in algemene bewoording kan worden gedaan.[86] Deze inzage mag geen afbreuk doen aan zakengeheim of intellectueel eigendomsrecht, maar hierdoor mag ook niet totale inzage geweigerd worden. Deze bepaling, die zijn oorsprong vindt in artikel 12a van richtlijn 95/46/EG, is niet als zodanig opgenomen in het commissievoorstel voor de AVG. Het Europees Parlement heeft echter voorgesteld deze wederom op te nemen.[87] Inzage in de logica, als onderliggend mechanisme voor identiteitscreatie, is een belangrijke voorwaarde om betekenisvol andere rechten, zoals het correctierecht, uit te kunnen oefenen. Anders kan het tot een situatie leiden waarbij de betrokkene maar zeer weinig controle over de (geautomatiseerd) aan hem toegemeten identiteiten heeft, wat al *prima facie* als beperking van zijn autonomie kan worden gezien.

Deze identiteiten en de logica op basis waarvan de gegevens verwerkt worden, kunnen echter buiten de context waarin ze gebruikt worden weinig betekenisvol zijn. Dit wordt veroorzaakt doordat ook veel van deze logica wordt doorontwikkeld door *machine learning* dat er voor zorgt dat vaak ook de verantwoordelijke geen volledig zicht heeft op de reden waarom iemand een bepaalde identiteit toegewezen heeft gekregen. Het is daarom van belang dat de verantwoordelijke altijd aanspreekbaar blijft voor de identiteit die wordt

---

[84] Artikel 11, tweede lid, Wbp.
[85] STI Workingpaper 2007/7, At a Crossroads: "personhood" and Digital Identity in the Information Society (OECD), p. 26.
[86] *Kamerstukken II* 1997/1998, 25 892, nr. 3.
[87] http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2F%2FEP%2F%2FTEXT%2BREPORT%2BA7-2013-0402%2B0%2BDOC%2BXML%2BV0%2F%2FEN&language=EN

gecreëerd. Bij onvoldoende onderbouwing van de juistheid en relevantie hiervan zou de betrokkene mogelijk verwijdering van het profiel kunnen eisen.[88]

6.2 Persoonlijke autonomie en sturing

Op basis van de identiteiten van een betrokkene kunnen besluiten ten aanzien van hem worden genomen. Belangrijk hierbij is dat dit zowel ten aanzien van een opzoek-identiteit (burgerlijke identiteit) als een herken-identiteit kan plaatsvinden. Iemand kan dus nog steeds geraakt worden door een besluit, zelfs wanneer de verantwoordelijke niet de naam, adres, geboorteplaats of andere meer gangbare kenmerken de van de betrokkene weet.

Geautomatiseerde besluitvorming over het individu op basis van kenmerken van hem wordt ook wel aangeduid als *profiling*. Dit wordt door het Europees Parlement gedefinieerd als "elke vorm van automatische verwerking van persoonsgegevens met als doel het evalueren van bepaalde persoonlijke aspecten van een natuurlijk person of het analyseren of voorspellen van deze natuurlijke persoon zijn prestatie op het werk, zijn economische situatie, locatie, gezondheid, persoonlijke voorkeuren, betrouwbaarheid of gedrag."[89] Dit is een zeer breed begrip en ziet in feite op het automatisch beslissen over personen op basis van door technologie geobserveerde, geïnduceerde of gededuceerde kenmerken op basis waarvan identiteiten (profielen) van mensen worden gecreëerd. Artikel 42 Wbp normeert deze situatie, waarbij aangetekend moet worden dat dit artikel alleen van toepassing is als er rechtgevolgen aan het besluit zijn verbonden of wanneer deze de betrokkene in aanmerkelijke mate treft. In de Algemene verordening gegevensbescherming wordt in artikel 20 een soortgelijke regeling voorgesteld.

Automatische besluitvorming op basis van inzichten verkregen uit Big Data-analyse kan gevolgen meebrengen voor de autonomie van het individu en diens fundamentele rechten. Deze gevolgen kunnen het resultaat zijn van veranderingen in informatievoorziening en goederen en diensten. Deze zullen namelijk tot in het extreme gepersonaliseerd kunnen worden op basis van data-analyse. Dit kan zorgen voor padafhankelijkheid of een zelfversterkend effect dat veroorzaakt wordt doordat op basis van het profiel een persoon anders behandeld wordt, op basis waarvan hij zich anders zal gedragen en waar vervolgens het profiel weer nader wordt verfijnt.

Een bekend voorbeeld hiervan is de door Eli Pariser gepopulariseerde *filter bubble*.[90] Zoekmachines passen hun zoekresultaten steeds meer aan op de

---

[88] Artikel 36 Wbp.

[89] Artikel 4 sub 3a AVG (Europees Parlement, eerste lezing). http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2F%2FEP%2F%2FTEXT%2BREPORT%2BA7-2013-0402%2B0%2BDOC%2BXML%2BV0%2F%2FEN&language=EN

[90] E. Pariser, *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*, Pengiun Books: 2012. De filter bubble wordt ook wel als *echo chamber* aangeduid. Zie International Working Group

zoekende persoon, gebaseerd op informatie over hem die binnen of buiten de zoekmachine is verzameld. Eenzelfde constructie wordt gebruikt door Facebook voor het bepalen welke items in het nieuwsoverzicht moeten worden getoond op basis van het EdgeRank algoritme.[91] Aangezien het wereldbeeld van mensen in sterke mate bepaald wordt door wat zij in hun omgeving waarnemen kan aanpassing van de informatievoorziening aan mensen grote gevolgen hebben. Zo experimenteerde Facebook eerder dit jaar door een groep van ongeveer zevenhonderdduizend gebruikers uitsluitend positieve berichten van vrienden en familie te tonen.[92] Dit zorgde ervoor dat deze groep ook steeds positievere informatie ging plaatsen. Dit toont duidelijk de invloed van deze mechanismen aan. De wijze van selectieve informatievoorziening kan natuurlijk via een veelvoud aan scheidslijnen, waardoor steeds verdergaande personalisering, of negatief gesteld, manipulatie, mogelijk wordt.

Sturing kan zich echter ook via andere wegen dan informatievoorziening voordoen. Lessig onderscheidt in zijn theorie van regulering een viertal modaliteiten die invloed op de autonomie van het individu uitoefenen.[93] Dit betreft het recht, (sociale) normen, de markt en ten slotte architectuur. Het recht en rechtshandhaving beperken de autonomie van het individu of beschermen juist deze door de autonomie van derden te beperken. Sociale normen beperken eveneens de bewegingsvrijheid van het individu, evenals de markt waar door vraag en aanbod en prijzen gestuurd wordt. Ten slotte is er architectuur. Een goed voorbeeld in de fysieke wereld is de planologie van een stad die, naast verkeersregels (recht), in sterke mate het verkeer reguleert. Veel hiervan is statisch, zoals de ligging van de weg, maar ook dynamische elementen zoals verkeerslichten zijn mogelijk. Architectuur ligt echter ook verborgen in de technologie waarmee onze wereld steeds meer doordrongen raakt. Ook hierdoor kan het individu gestuurd worden.

Wanneer dagelijkse objecten en structuren mensen kunnen herkennen en hierop kunnen reageren wordt de vergaande personalisering in het fysieke domein werkelijkheid. Het functioneren van fysieke *smart objects* zal voor een ieder met hetzelfde profiel gelijk zijn, maar aangezien het profiel per persoon nooit helemaal gelijk zal zijn, zal de uitkomst tussen personen altijd verschillen. De omgeving en objecten zullen zich aanpassen aan de persoon waarmee geïnteracteerd wordt.

Dit kan ook op een wijze dat een persoon wordt aangezet om bepaald gedrag te vertonen of bepaalde acties te ondernemen, ook wel *pursuasion profiling*

---

on Data Protection in Telecommunications 5-6 mei 2014, Working Paper on Big Data and Privacy Privacy principles under pressure in the age of Big Data analytics, p. 11.

[91] https://econsultancy.com/blog/7885-the-ultimate-guide-to-the-facebook-edgerank-algorithm en http://techcrunch.com/2010/04/22/facebook-edgerank/. (Geraadpleegd 2 november 2014).

[92] http://www.nytimes.com/2014/06/30/technology/facebook-tinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html?_r=0 (Geraadpleegd 2 november 2014).

[93] L. Lessig, *Code version 2.0,* New York: Basic Books 2006, p. 123.

genoemd.[94] Deze *persuasion profiles* geven informatie over welke (soorten) informatie of interactie overtuigend is voor het individu dat door het profiel wordt beschreven. Uit onderzoek blijkt dat deze eigenschappen, bijvoorbeeld gevoeligheid voor autoriteit, schaarste (koopjes) of een streven naar consensus zeer contextonafhankelijk zijn.[95] Met andere woorden, personen zijn in verschillende situaties gevoelig voor dezelfde soort overtuigingen.[96] Het is niet verwonderlijk dat marketeers op zoek zijn naar het gebruik van psychologische *triggers* om personen aan te zetten tot het aanschaffen van diensten en producten.[97] In de praktijk zijn dergelijke profielen echter ook al ingezet bij het overtuigen van kiezers, zoals tijdens de Amerikaanse presidentiele verkiezingen in 2008.[98]

Het gevaar bestaat dat het individu weinig tot geen controle heeft over de vorming van zijn identiteiten en de reactie daarop door zijn steeds slimmere omgeving en uiteindelijk het vormen van zijn eigen keuzes. Ook kan het leiden tot een *chilling effect* op de vrijheid van meningsuiting, omdat niet goed voorzienbaar is welke gevolgen de verzameling en verwerking van gegevens in de toekomst met zich meebrengt.[99] Het analyseren van alle activiteiten en gedrag maken een persoon inzichtelijk en voorspelbaar. Daarnaast zou er ongeoorloofd onderscheid tussen personen kunnen worden gemaakt. Op de 36e internationale conferentie van toezichthouders gegevensbescherming zijn deze gevaren benoemd in de context van Big Data.[100] Om negatieve gevolgen zoveel mogelijk te beperken wordt onder andere opgeroepen doelbinding en gegevensminimalisatie te respecteren en waar passend toestemming te vragen. Daarnaast wordt nadruk gelegd op transparantie over gegevensverwerking, besluiten die genomen worden en profielen die samengesteld worden. In het bijzonder wordt opgeroepen om Big Data eerlijk, transparant en verantwoordelijk in te zetten en profielen en beslisregels continu te onderhouden, zodat de profielen eerlijk, ethisch en proportioneel zijn en er geen onterechte beslissingen worden genomen. Als er significante effecten voor het individu zijn, zou menselijke tussenkomst altijd mogelijk moeten zijn.[101]

---

[94] M.C. Kaptein, *Personalized Persuasion in Ambient Intelligence (diss. TU Eindhoven),* Eindhoven: Technische Universiteit Eindhoven 2012, p. 101.
[95] M. C. Kaptein en D. Eckles, 'Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling', http://www.persuasion-profiling.com/wp-content/uploads/2010/04/EffectiveMeans.pdf, p. 2.
[96] Dit brengt wederom een spanning met het beginsel van doelbinding van artikel 9 Wbp.
[97] http://www.persuasionapi.com
[98] http://www.technologyreview.com/featuredstory/508836/how-obama-used-big-data-to-rally-voters-part-1/ (Geraadpleegd 2 november 2014).
[99] P. de Hert, e.a., 'Legal safeguards for privacy and data protection in ambient intelligence', *Personal and Ubiquitous Computing* 2009-13, p. 436.
[100] http://www.cbpweb.nl/downloads_med/Resolution%20Big%20Data.pdf (Geraadpleegd 2 november 2014).
[101] Zie huidige artikel 42, tweede lid, Wbp.

6.3 Kansen ten aanzien van individuele autonomie

In de discussie is het duidelijk dat er steeds meer nadruk wordt gelegd op open begrippen als 'eerlijk', ' transparant', 'verantwoordelijk' en zelfs ethisch. Gezien de snelle technologische ontwikkelingen wordt ook in de voorgestelde Algemene verordening Gegevensbescherming zoveel mogelijk met techniek- en implementatieneutrale open normen gewerkt. Dit lijkt mede ingegeven door onzekerheid over de effectiviteit van instrumenten als toestemming en transparantie voor het beschermen van het individu. In die zin zou juist binnen de innovatie op het terrein van Big Data onderzoek gedaan moeten worden naar welke instrumenten in welke context effectief zijn in het beschermen van het individu en diens autonomie. Big Data kan namelijk ook juist de autonomie van het individu versterken. Zo kan het allereerst verborgen vormen van discriminatie die bewust of onbewust binnen de maatschappij plaatsvinden inzichtelijk maken.[102] Daarnaast kunnen profielen, wanneer vrijgegeven aan de betrokkene, hem juist ook beter inzicht geven in zijn eigen besluitvormingsproces en hem zo weerbaarder maken voor subtiele overtuigingen of manipulatie.[103] Inzet in de praktijk hiervan zal echter afhankelijk zijn van de belangen van de verantwoordelijke, die er soms (commercieel) belang bij kan hebben juist niet het zelfinzicht van de betrokkene te versterken.

7. Big data en de overheid

Naast bedrijven zien ook overheden de potentiele voordelen en kansen van Big Data. Een groot deel van deze voordelen hebben betrekking op het in kaart brengen van zaken als milieuvervuiling en gezondheidsrisico's in steden, bijvoorbeeld door het op grote schaal meten van fijnstofconcentraties. Tevens kunnen gegevens over de hoeveelheid wegverkeer op het wegennet worden gebruikt om de inrichting van de infrastructuur te verbeteren.

Tegelijkertijd zijn er uiteraard ook Big Data gerelateerde toepassingen die voor controverse zorgen. Denk aan het gebruik van diverse gegevensbronnen voor het opsporen van fraudegevallen, bijvoorbeeld op basis van de wijziging op 1 januari 2014 van de Wet structuur uitvoeringsorganisatie werk en inkomen en enige andere wetten in verband met fraudeaanpak door gegevensuitwisselingen en het effectief gebruik van binnen de overheid bekende zijnde gegevens. De discussie over het gebruik van dergelijke technieken worden meestal langs de lijn van bescherming van de persoonlijke levenssfeer tegen inmenging van de overheid gevoerd. Alhoewel dit een wezenlijke en fundamentele discussie betreft zal in dit preadvies de discussie langs een andere lijn gevoerd worden, namelijk onvoorziene gevolgen die het inzetten van deze technologie voor uitvoering en beleidsvorming met zich mee kan brengen.

---

[102] http://thehill.com/blogs/pundits-blog/technology/221583-big-data-is-a-powerful-weapon-in-the-fight-for-equality (Geraadpleegd op 2 november 2014).
[103] M.C. Kaptein, *Personalized Persuasion in Ambient Intelligence (diss. TU Eindhoven)*, Eindhoven: Technische Universiteit Eindhoven 2012, p. 179.

7.1 Aandachtspunten in de context van de overheid

Big Data brengt op diverse wijzen kansen en risico's met zich mee, doordat het als instrument voor het maken van onderscheid kan worden gebruikt, op basis waarvan beleid vorm kan worden gegeven. In een paper over datamining analyseert Barocas risico's die onder andere relevant zijn in een overheidscontext.[104] Allereerst kan beleidsvorming op basis van Big Data tot ongewenste gevolgen leiden indien er sprake is van een statistische tekortkoming. Als voorbeeld wordt het inzetten van een app op moderne smartphones voor het opsporen van gaten in het wegdek benoemd.[105] Indien onderhoudswerkzaamheden vervolgens voornamelijk op basis van deze gegevens worden ingepland bestaat er een risico dat bijvoorbeeld arme wijken worden achtergesteld, omdat daar bijvoorbeeld minder bewoners gebruik maken van een moderne smartphone met de app. Indien beleidsmakers deze verborgen afwijkingen onvoldoende onderkennen kan dit onbedoeld voor groeiende ongelijkheid zorgen.

Tevens kunnen beslissingen genomen worden over individuen of groepen van individuen op basis van onvolledige gegevens. De door Big Data afgeleide gegevens zijn immers maar een beperkte representatie van de werkelijkheid. Daar tegenover bestaat het gevaar dat juist met Big Data heel precies individuen of groepen van individuen onderscheiden kunnen worden en voorspellingen over de toekomst kunnen worden gedaan. Doordat steeds preciezer inzichtelijk is hoeveel personen en groepen aan de maatschappij in sociale of economische zin bijdragen of kosten kan dit de sociale cohesie, of *social fabric,* binnen de maatschappij verminderen en de druk op specifieke groepen verhogen.[106] Zo kunnen mensen bijvoorbeeld onder druk worden gezet om gezonder te leven. Als zij hier geen gevolg aan geven kunnen zij gedwongen worden om bijvoorbeeld meer belasting of premie te betalen. Controle op hoe veilig en zuinig iemand rijdt is een ander voorbeeld. Ten slotte kan een toegenomen controle op bepaalde groepen in de maatschappij *an* sich al met zich meebrengen dat vanzelf meer gevallen van normafwijkend gedrag worden geconstateerd, simpelweg, omdat deze groepen vaker wordt gecontroleerd.

De gemene deler van de bovenstaande noties is dat er de nodige mogelijke valkuilen zijn bij het inzetten van Big Data voor beleidsvorming. Het is daarom belangrijk om altijd de beperkingen van de technieken in ogenschouw te houden. Resultaten uit Big Data analyse kunnen immers gepercipieerd worden als volstrekt objectief, terwijl ook daar op velerlei wijze vertekende beelden kunnen optreden.

---

[104] S. Barocas, 'Data Mining and the Discourse on Discrimination', *Proceedings of Data Ethics Workshop*, 24 augustus 2014, p. 1-4.
[105] Zie http://www.streetbump.org (Geraadpleegd 2 november 2014)
[106] S. Barocas, 'Data Mining and the Discourse on Discrimination', *Proceedings of Data Ethics Workshop*, 24 augustus 2014, p. 3.

7.2 Nieuwe vormen van regulering

Het gebruik van grote databronnen en data-analyse voor het reguleren van de maatschappij wordt ook wel aangeduid als *algorithmic regulation*.[107] Het idee is dat technieken die bijvoorbeeld ook door grote IT-bedrijven gebruikt worden om spam te filteren of zoekresultaten te filteren ook gebruikt kunnen worden voor regulering in meer traditionele zin. Een bedrijf als Google kan bijvoorbeeld niet handmatig continu nieuwe regels opstellen om spammails te detecteren.[108] Door continu de situatie te meten, automatisch regels (algoritmes) toe te passen en deze regels continu op basis van feedback van gebruikers en het systeem aan te passen aan de gewenste uitkomsten kan snel ingespeeld worden op nieuwe situaties en kan stabiliteit van de gewenste uitkomsten verzekerd worden. In casu het juist aanduiden of verwijderen van spammails. Regulering wordt in die zin adaptief, dat de regels continu worden aangepast om tot de gewenste uitkomst te komen.

Een eerste voorbeeld in de context van de fysieke wereld en de overheid is het dynamisch aanpassen van de maximale snelheid op snelwegen (dynamax), gebaseerd op de gemeten drukte op de weg en de hoeveelheid luchtvervuiling.[109] In 2009 zijn hier al uitgebreide proeven mee ondernomen, wat in 2011 tot een eindrapport met positieve resultaten heeft geleid.[110] Met behulp van een veelvoud aan metingen wordt duidelijk welke snelheid voorgeschreven dient te worden om de gewenste optimale reistijd, verkeersdrukte en hoeveelheid luchtvervuiling te krijgen.[111] Dynamische snelheden zijn een fenomeen waar we al reeds aan gewend zijn, maar met de snelle ontwikkeling van het *Internet of Things* waarmee onze volledige omgeving uitgerust kan worden met slimme technologie nemen de mogelijkheden enorm toe om ook andere aspecten te monitoren en zelfs automatisch aan te passen. De eerste tekenen zijn al te zien, bijvoorbeeld bij het Living Lab in uitgaansgebied Stratumseind in Eindhoven, waar gebruik gemaakt wordt van een grote hoeveelheid sensoren, zoals voor geluid, temperatuur, bezoekersaantallen etc.[112] Op basis hiervan kunnen (automatisch) acties ondernomen worden om voor een goede sociale omgeving te komen. Uiteindelijk kan dit leiden tot steeds verdergaande *real-time* adaptieve (invulling van) regelgeving, uitvoering en handhaving. In feite maakt dit meer maatwerk mogelijk, om zo het gewenste einddoel te behalen. Hier is een parallel te vinden met de inzet van open normen in het gegevensbeschermingsrecht. In plaats van materiële normen voor te schrijven hangt de invulling van deze open

---

[107] http://beyondtransparency.org/chapters/part-5/open-data-and-algorithmic-regulation/ (Geraadpleegd 2 november 2014).
[108] http://www.theguardian.com/technology/2014/jul/20/rise-of-data-death-of-politics-evgeny-morozov-algorithmic-regulation (Geraadpleegd 3 november 2014).
[109] http://www.wegenwiki.nl/Dynamische_maximumsnelheid (Geraadpleegd 3 november 2014).
[110] http://www.rijksoverheid.nl/ministeries/ienm/documenten-en-publicaties/rapporten/2011/02/11/bijlage-2---rapport-dynamische-maximumsnelheden-evaluatie-praktijkproeven.html (Geraadpleegd 3 november 2014).
[111] Op basis van artikel 186 Wegenverkeerswet 1994 kunnen experimenten uitgevoerd worden. Zie ook aanwijzing 10b van de Aanwijzingen voor regelgeving.
[112] http://ditss.nl/nieuws/ditss-workshops-op-industria-congres-big-data-big-business-or-big-brother/ (Geraadpleegd 3 november 2014).

normen af van de omstandigheden van het geval en de ontwikkelingen in de jurisprudentie. Het verschil met *algoritmic regulation* is echter dat hierbij de omstandigheden van het geval geautomatiseerd worden gewogen en tot een vastgestelde uitkomst leiden.

Algorithmic regulation kan gevolgen met zich meebrengen voor de (rechts-)positie van de individuele burgers, met name wanneer de uitkomsten van deze *algorithmic regulation* voor iedere burger anders kunnen zijn. Bij het invoeren van een dynamische maximum snelheid is hier maar beperkt sprake van, omdat deze voor alle personen die op een bepaald moment op een specifieke weg rijden hetzelfde is. Echter, wanneer er maar rekening gehouden wordt met genoeg omstandigheden en de toepassing van de norm met zich meebrengt dat niet één burger gelijk is of zich op het hetzelfde moment in gelijke omstandigheden bevindt, zal dit voor iedere burger feitelijk tot een andere uitkomst leiden. Dit gegeven zal een dan ook een belangrijk aandachtspunt zijn in debatten over gelijke behandeling.

De voorgaand geschetste ontwikkelingen zullen mogelijk tot efficiëntieverbeteringen kunnen leiden, maar deze zullen wel binnen de grenzen van een goede vervulling van de algemene beginselen van behoorlijk bestuur moeten blijven, zoals het al eerder genoemde gelijkheidsbeginsel. Daarnaast kan het rechtszekerheidsbeginsel onder druk kunnen komen te staan. De kern van rechtszekerheid is voorzienbaarheid en kenbaarheid.[113] Dit kan onder druk komen te staan, omdat de complexiteit van de regels vergroot wordt en de variabelen die gebruikt worden voor toepassing van de regel voor de burger weinig inzichtelijk kunnen zijn. Bij *algorithmic regulation* zal de uitkomst van de regels bovendien steeds meer afhankelijk kunnen zijn van factoren die buiten de burger liggen, maar die wel zijn positie beïnvloeden. In het geval van besluiten zal er in het bijzonder aandacht moeten zijn voor het motiveringsbeginsel. Uitkomsten van beleid en wetgeving kunnen immers uiteraard niet gebaseerd worden op de uitkomst van een *black box*. Voor het gebruik, de inhoud en de uitkomst van algoritmes dient in beginsel ook rekenschap te worden afgelegd, zogenaamde *algorithmic accountability*.[114] Dit kan gaan om informatie over op welke wijze gegevens gebruikt worden als invoer, hoe deze worden gestructureerd en geclassificeerd, welke afwijkingen er zijn en hoe doorlopend de juistheid van het algoritme wordt verzekerd, om zo een eerlijke en inhoudelijke juiste toepassing te verzekeren.[115]

---

[113] R. Koning, 'De crisis- en herstelwet: nood, spoed of experiment', in A. G. Bregman, H.E. Bröring en K.J. de Graaf (red.), *Onbegrensde rechtsbeoefening* (Lubach bundel), Den Haag: IBR 2014, p. 42.

[114] https://freedom-to-tinker.com/blog/felten/accountable-algorithms/ (Geraadpleegd op 3 november 2014).

[115] N. Diakopolous, 'Algorithmic accountability reporting: on the investigation of black boxes', http://towcenter.org/wp-content/uploads/2014/02/78524_Tow-Center-Report-WEB-1.pdf , p. 3-10.

8. Conclusie

De snelle ontwikkelingen op het gebied van (informatie-)technologie stuwen elkaar vooruit. Afnemende kosten en toenemende mogelijkheden van sensoren, gegevensopslag, gegevensverwerking en het integreren van elektronica in alle dagelijkse objecten en convergentie tussen de neuro-, bio-, informatie- en cognitieve technologie maken dit mogelijk. Big Data, de verwerking van grote hoeveelheden informatie, met hoge snelheid van verschillende aard ondersteunt deze ontwikkeling.

Met deze ontwikkelingen kunnen fundamentele rechten onder druk komen te staan, maar in gevallen ook versterkt worden. Hierbij dient niet alleen gekeken te worden naar de rechten op bescherming van de persoonlijke levenssfeer en persoonsgegevens, maar ook de vrijheid van gedachte, de vrijheid van meningsuiting en het recht op non-discriminatie. In de context van Big Data en aanverwante technologische ontwikkelingen liggen de kansen en risico's in samenhang ten aanzien van deze rechten met name op het vlak van persoonlijke autonomie, keuzevrijheid, integriteit en een vrije stroom van denkbeelden. De bescherming van de persoonlijke levenssfeer en persoonsgegevens maken dat het individu zichzelf kan vormen, onder meer op basis van denkbeelden die hij ontvangen heeft van derden. In de vrijheid van zijn eigen gedachte kan hij keuzes maken over het inrichten van zijn persoonlijke levenssfeer en daar vervolgens zijn gedachten over uiten.

Met de toegenomen mogelijkheden van data-analyse zal anders naar de positie van het individu te midden van Big Data moeten worden gekeken. De burgerlijke identiteit zoals naam, adres en woonplaats zal steeds minder een belangrijke rol vervullen en tegelijkertijd zal de rol van herken-identiteiten waarmee een persoon onderscheiden kan worden van anderen toenemen. Bij de interpretatie van het gegevensbeschermingsrecht zal hier rekening mee moeten worden gehouden, om zo de beschermingsomvang van dit recht niet te ver in te perken.

De vorming van de identiteit van het individu zal steeds meer geautomatiseerd en buiten het zicht van hem plaatsvinden. Waar in het verleden door de betrokkene zelf gegeven informatie en door de verantwoordelijke geobserveerde informatie het zwaartepunt van gegevensverwerking was, wordt nu het deduceren, induceren en andersoortig afleiden van informatie van het grootste belang. Aangezien de betrokkene niet direct betrokken is bij deze verwerkingen zal het belang van het verstrekken van informatie over de verwerking op verzoek in belang toenemen, om zo een te grote informatie-asymmetrie tussen betrokkene en verantwoordelijke tegen te gaan.

Met de steeds verdere inbedding van informatietechnologie in onze alledaagse omgeving wordt de mate van mogelijke sturing via architectuur steeds groter. Zaken zijn immers steeds meer tot in het extreme aanpasbaar en personaliseerbaar. In feite gaat het om geautomatiseerde besluitvorming op basis van geobserveerde, gededuceerde of geïnduceerde gegevens. Vergaande personalisering op basis van profielen brengt het gevaar van inperking van de autonomie van individuen met zich mee. Tegelijkertijd kan echter betekenisvol inzicht in deze profielen juist een versterking van de autonomie van het individu

met zich meebrengen, omdat hij op deze wijze inzage krijgt in zijn eigen oordeelsvorming of besluitvormingsproces.

Inzet van Big Data in de context van de overheid moet breder beschouwd worden dan alleen in het kader van privacybescherming. Beleidsvorming op basis van meer (empirische) gegevens kan veel voordelen met zich meebrengen op het vlak van efficiëntie en effectiviteit, maar de grenzen en beperkingen van deze gegevens moeten ook duidelijk in ogenschouw genomen worden. De gegevens zijn immers een model of representatie van de werkelijkheid, waarbij er zich veel vertekeningen kunnen voordoen die, wanneer er blind op geacteerd wordt, ongewenste gevolgen met zich mee kunnen brengen.

Big Data technieken kunnen, behalve als instrument voor beleidsvorming, ook dienen als instrument voor nieuwe vormen van regulering, zoals *algorithmic regulation*. Hierbij staat niet de materiële rechtsregel centraal, maar het bereiken van een bepaald einddoel. De gewenste uitkomst staat zo centraal en de invulling van de normen om tot die uitkomst te komen kan variëren.

Aangezien Big Data doorwerkt in alle aspecten van onze maatschappij, zelfs in de wijze waarop deze gereguleerd kan worden, zullen discussies hierover aan de hand van fundamentele rechten moeten worden gevoerd. Zet fundamentele rechten en beginselen daarom centraal, niet de technologische ontwikkeling zelf. De implementatie van techniek is uiteindelijk niet waardenvrij, zelfs als de ontwikkeling niet door een specifieke agenda wordt bepaald.

9. Aanbevelingen / stellingen:

- Big Data als techniek brengt zowel kansen als risico's ten aanzien van de bescherming van fundamentele rechten met zich mee.
- Big Data en aanverwante technologische ontwikkelingen werken door in alle aspecten van de maatschappij. Stel daarom de fundamentele rechten centraal en niet de technologische ontwikkeling.
- De bescherming van fundamentele rechten in een wereld met Big Data kan niet alleen door het recht en juristen opgelost worden. In het bijzonder ethici en statistici zullen een belangrijke rol moeten vervullen, evenals bestuurders bij overheid en bedrijfsleven en politici.