PREADVIES VERENIGING GEZONDHEIDSRECHT 2017

Privacyrisico's en -waarborgen bij het gebruik van big data tegen zorgfraude: een verkenning

Gerrit-Jan Zwenne & Wilfred Steenbruggen[1]

1. INLEIDING

Er is veel aandacht voor zorgfraude. In de landelijke media, zowel online als offline, zien we vrijwel dagelijks soms wat tendentieus of schreeuwerig getoonzette berichten, zoals dat 'artsen voor miljard euro frauderen',[2] 'miljoenenfraude in de thuiszorg',[3] 'zorgfraude ziekenhuizen voor tientallen miljoenen'[4], '2,5 miljoen boete voor St. Antonius Ziekenhuis wegens foute declaraties'[5], of meer anekdotisch 'tandartsrekening voor niet bestaande cliënt[6], 'rekening voor afgezegd consult',[7] 'Nijverdaller Wesseling krijgt dubbele rekening ziekenhuis'[8], 'frauderen blijkt een fluitje van een cent',[9] 'VVD wil frauderende pgb-bemiddelaars op zwarte lijst'[10], 'Opnieuw geknoei met declaraties in de ggz'[11], 'Zorgverzekeraars hebben geen idee welke psychische zorg zij vergoeden'[12] en 'Miljoenenverlies zorgverzekeraars door EuroPsyche'[13]

Het is een willekeurige greep uit de vele nieuwsberichten die ons in de laatste jaren hebben bereikt. We moeten daarbij natuurlijk onderkennen dat de keuze voor het taalgebruik in deze berichten, en de redactie van de koppen erboven, ook verband houdt met de commerciële wens om meer lezers te bereiken en meer clicks te genereren. En ook dat het helemaal niet is gezegd dat in al deze gevallen ook daadwerkelijk sprake is van opzettelijk frauduleus handelen. Toch maakt het wel duidelijk dat er het nodige mis gaat bij het declareren door zorgaanbieders en dat daarmee grote bedragen zijn gemoeid. Hoe groot precies is niet bekend. Schattingen lopen uiteen van tientallen miljoenen euro's tot enkele miljarden euro's per

---

[1] Prof mr. Gerrit-Jan Zwenne is hoogleraar Recht en de informatiemaatschappij in Leiden en advocaat in Amsterdam, mr. dr. drs. Wilfred Steenbruggen is advocaat in Rotterdam.

[2] Nu.nl 4 november 2011 <www.nu.nl/economie/2659067/artsen-frauderen-miljard-euro.html>

[3] Een vandaag 12 juni 2009 <www.eenvandaag.nl/binnenland/34816/miljoenenfraude_in_de_thuiszorg>

[4] Nu.nl 16 mei 2013 <www.nu.nl/binnenland/3476441/zorgfraude-ziekenhuizen-tientallen-miljoenen.html>

[5] Nu.nl 11 februari 2014 <www.nu.nl/economie/3698935/25-miljoen-boete-st-antonius-ziekenhuis.html>

[6] NRC 4 september 2013.

[7] Tubantia 12 augustus 2014.

[8] Tubantia 1 augustus 2014.

[9] AD 14 februari 2013.

[10] Nu.nl juli 2015 <www.nu.nl/politiek/4079090/vvd-wil-frauderende-pgb-bemiddelaars-zwarte-lijst.html>

[11] Nos.nl 27 december 2015 <http://nos.nl/artikel/2077357-opnieuw-geknoei-met-declaraties-in-de-ggz.html>

[12] Volkskrant 4 mei 2012.

[13] Volkskrant 14 februari 2013.

jaar.[14]

Dit soort bedragen zet de betaalbaarheid van ons zorgstelsel onder druk en leidt tot erosie van solidariteit, zeker in tijden waarin de zorgkosten toch al sterk oplopen door de vergrijzing. Het verbaast dan ook niet dat wordt aangedrongen op een stevige aanpak van zorgfraude. Vanaf 2011 wordt er hard geroepen om handhaving en *naming-and-shaming* van 'sjoemelende zorginstellingen, frauderende ziekenhuizen en malafide pgb-bemiddelingsbureaus'. Daarbij wordt veel verwacht van de inzet van big data. In de zorgsector worden enorme hoeveelheden gegevens verzameld en bij elkaar gebracht, bij zorgverzekeraars en bij toezichthouders. Deze gegevens zouden met behulp van slimme data-analyses (*big data predictive analytics*) mogelijk onregelmatigheden en fraude sneller en beter kunnen herkennen, zodat deze kan worden voorkomen of in elk geval dat daartegen tijdig kan worden opgetreden

Voor de meeste juristen en vele anderen is het thema big data onlosmakelijk verbonden met vragen over privacy- en gegevensbescherming. Onbegrijpelijk is dat niet. Als het gaat om big data zijn er wel meer belangrijke rechtsvragen, bijvoorbeeld op het gebied van intellectuele eigendom,[15] mededingingsrecht[16] of consumentenbescherming,[17] maar de op dit moment meest urgente vragen, althans de vragen waarover de grootste zorgen zijn, hebben vooral betrekking op de risico's ervan in termen van bedreigingen van de persoonlijke levenssfeer en de bescherming van persoonsgegevens. Deze vragen zijn ook pregnant aanwezig bij de inzet van big data in de context van de bestrijding van zorgfraude, zeker indien gezondheidsgegevens bij de analyse worden betrokken.

In deze bijdrage hebben wij er daarom voor gekozen om het gebruik van big data in het kader van het tegengaan van zorgfraude vooral vanuit het perspectief van privacy- en gegevensbescherming te bezien. Eerst beschrijven we op welke wijze nu al gebruik wordt gemaakt van data-analyses om zorgfraude tegen te gaan (§2). Vervolgens gaan we in op de risico's die we daarbij zien (§3), waarna we bij wijze van vingeroefening bespreken enkele waarborgen bespreken waarin wet- en regelgeving voorziet om de verschillende rechtssubjecten (zoals verzekeringnemers en verzekerden, patiënten, zorgaanbieders etc.) te beschermen (§4). We sluiten af met een korte beschouwing over de vraag of dat genoeg is (§5).

---

[14] PWC (2013), Naar een fraudebeeld Nederland.

[15] Bijv. Lundqvist, B., (2016). Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet of Things World, Faculty of Law, University of Stockholm Research Paper No. 1. <https://ssrn.com/abstract=2891484>

[16] Zie bijv. Kerber, W. (2016). Digital Markets, Data, and Privacy: Competition Law, Consumer Law, and Data Protection, MAGKS, Joint Discussion Paper Series in Economics, No. 14-2016. <https://ssrn.com/abstract=2770479>

[17] Bijv. Rhoen, M., (2015). Big Data and Consumer Participation in Privacy Contracts: Deciding who Decides on Privacy. Utrecht Journal of International and European Law 31(80), pp.51–71. <DOI: http://doi.org/10.5334/ujiel.cu>

2. BESTRIJDING VAN ZORGFRAUDE DOOR MIDDEL VAN (BIG) DATA-ANALYSES

2.1 Aanloop

Fraude is van alle tijden, ook in de zorg.[18] De invoering van het nieuwe zorgstelsel in 2006 leidde evenwel tot een hernieuwde aandacht voor het onderwerp. In de aanloop daarheen werd duidelijk dat er niet alleen frauderisico's zijn aan de zijde van de zorgvrager, maar ook aan de zijde van de zorgaanbieders.[19] En daarbij speelt vooral de aard van het bekostigingssysteem een belangrijke rol. Hoe complexer namelijk de tarieven zijn, hoe groter de risico's op fouten en fraude aan de kant van de zorgaanbieder.[20] En dat het systeem van prestatiebekostiging op basis van Diagnose Behandel Combinaties (DBC's) complex is, behoeft weinig betoog. Weliswaar is het DBC systeem in 2012 aanzienlijk vereenvoudigd met de invoering van de DOT-methodiek (DBC's op weg naar transparantie) die het aantal zorgproducten en tarieven terugbracht van 30.000 naar zo'n 4.400, maar het is nog steeds erg ingewikkeld en dus foutgevoelig en daarmee fraudegevoelig.

Onder de Wet Tarieven Gezondheidszorg (WTG)[21] werden er daarom al maatregelen genomen om fouten en fraude door zorgaanbieders te voorkomen, met name door de transparantie rondom prestatie en tarief te vergroten en het bestuursrechtelijk toezicht te verstevigen. Het wetsvoorstel WTG ExPres[22] voorzag bijvoorbeeld in algemene administratievoorschriften voor zorgaanbieders en verzekeraars, zodat deze gehouden zouden zijn om een administratie te voeren waarin niet alleen de geleverde prestaties zichtbaar zijn, maar ook de daarvoor in rekening gebrachte tarieven en de in verband daarmee ontvangen of verrichtte betalingen en vergoedingen aan derden. Daarbij kwamen nieuwe verplichtingen ten aanzien van tijdige patiënteninformatie en de wijze waarop wordt gedeclareerd. Daarnaast werd er een centraal Meldpunt Onregelmatige Declaraties ingesteld bij het CTG. Ook kreeg de toezichthouder, indertijd het College tarieven gezondheidszorg (CTG), meer mogelijkheden om informatie te vorderen en uit te wisselen met andere toezichthouders en kon hij voortaan een last onder bestuursdwang of een last onder dwangsom bij overtreding van deze administratie-, declaratie- en informatieverplichtingen.

In de Wet marktordening gezondheidszorg (Wmg[23]) is ingezet op een ver-

[18] Zie uitgebreid T.A.M. van der Ende, 'Zorgfraude: van handhaving en hoe systemen hun eigen ongelukken creëren', in: J.G. Sijmons e.a., Op weg naar 10 jaar nieuw zorgstelsel: terug- en vooruitblik, Den Haag: Sdu Uitgevers 2015, p. 174 e.v.

[19] Kamerstukken II 2002–2003, 28 828, nr. 1 en bijlagen.

[20] Vgl. ook Van der Ende 2015, p. 188.

[21] Wet van 20 november 1980, houdende regelen met betrekking tot de tarieven van organen voor gezondheidszorg, Stb. 1980, 646 (ingetrokken Stb. 2006, 415)

[22] Kamerstukken II 2003/2004, 29379, nrs. 1-2.

[23] Wet van 7 juli 2006, houdende regels inzake marktordening, doelmatigheid en beheerste kostenontwikkeling op het gebied van de gezondheidszorg (Wet marktordening gezondheidszorg), Stb. 2006, 415

dere versteviging van het bestuursrechtelijk toezicht op zorgverzekeraars en zorgaanbieders. Onder de WTG kon het CTG niet punitief optreden tegen een overtreding van artikel 2 WTG op grond waarvan zorgaanbieders geen tarieven voor prestaties in rekening mogen brengen die afwijken van of hoger zijn dan het gereguleerde tarief. Opzettelijke overtredingen waren echter wel via de band van de Wet economische delicten (Wed) strafrechtelijke gesanctioneerd. De rechtsopvolger van het CTG, de Nederlandse Zorgautoriteit (NZa), kan echter wel (forse) bestuurlijke boetes op te leggen, tot maximaal EUR 500.000 of, als dat meer is, tot maximaal 10% van de omzet van de onderneming in Nederland, ook bij overtreding van artikel 35 WMG dat – net als voorheen artikel 2 WTG – zorgaanbieders verbiedt om tarieven in rekening te brengen die in strijd zijn met de tarief- of prestatieregulering, en zorgverzekeraars om die tarieven te betalen. Voor zware gevallen van fraude is strafrechtelijke sanctionering mogelijk gebleven.[24]

Desondanks kwam de fraudebestrijding in de zorg in de eerste jaren na invoering van het nieuwe zorgstelsel nog niet echt van de grond. Wij kunnen slechts speculeren over de oorzaken, maar het lijkt erop dat de transitie naar het nieuwe en complexe zorgstelsel gewoon veel tijd vergde -- tijd die de stakeholders (politiek, toezichthouders en veldpartijen) nodig hadden om uit te vinden wat precies de juiste invulling van hun rol moet zijn binnen het nieuwe stelsel, ook ten opzichte van andere partijen. Misschien koos de NZa daarom aanvankelijk ook wel voor een 'zachtere' toezichtsstijl waarbij de nadruk vooral op informeel en preventief toezicht lag (door middel van informatieverzoeken en zogenaamde 'wenkbrauwgesprekken').

Van het gericht opsporen van fraude en repressief optreden door de NZa was in die eerste jaren in ieder geval geen sprake. Bij de eerste evaluatie van de Wmg in 2009 werd zelfs de vraag opgeworpen of de NZa die omslag naar meer repressief toezicht wel zou kunnen maken. [25] Inmiddels is duidelijk geworden dat de NZa wel degelijk repressief *kan* optreden: in de afgelopen jaren heeft de NZa een aantal malen forse boetes opgelegd aan zorgaanbieders in verband met onjuist declareren.[26] Dit laat onverlet dat de NZa nog steeds terughoudend lijkt bij de inzet van het boete-instrument. Het aantal boetes sinds de invoering van het nieuwe zorgstelsel is nog steeds op twee handen te tellen: sinds 2011 heeft de NZa nog maar 8 boetes opgelegd. Dat is erg weinig als wordt gekeken naar vele onregelmatigheden en daarmee gemoeide miljardenbedragen die de afgelopen jaren zijn langsgekomen onder de noemer fraude in de zorg.

[24] Kamerstukken II 2004/05, 30186, nr. 3, p. 25.

[25] Zie R.D. Friele e.a., Evaluatie Wet Marktordening Gezondheidszorg, Den Haag: ZonMw 2009, p. 147 e.v.

[26] Zo kreeg de Ommelander Ziekenhuis Groep wegens niet correct declareren in 2011 een boete van EUR 500.000, het St. Antonius Ziekenhuis in 2014 een boete van EUR 2.500.000 en Stichting Altrecht in 2015 een boete van EUR 700.000.

Mede daardoor staat, met name sinds 2013, de bestrijding van zorgfraude hoog op de politieke agenda. Zorgaanbieders, zorgverzekeraars en handhavingspartners hebben, op aandringen van het parlement, allerlei initiatieven genomen en samenwerkingsverbanden opgezet om de kwaliteit van declaraties te vergroten en fouten en fraude te voorkomen. Ook wordt ingezet op een meer integrale handhaving en nauwe samenwerking van handhavingspartners bij de opsporing en vervolging van onrechtmatigheden. Deze (publiek-private) samenwerking vormt de opmaat naar verder gezamenlijk gebruik van databestanden voor fraudebestrijding. Zowel ingeval van verdenkingen als voor het analyseren van risico's vanuit grote hoeveelheden data.

2.2 Big data en fraudebestrijding

De term big data wordt in veel verschillende betekenissen gebruikt.[27] De gemeenschappelijke deler is dat het gaat om het verzamelen en gebruiken van grote volumes data afkomstig uit diverse bronnen die niet altijd eenvoudig doorzoekbaar of koppelbaar zijn. De desbetreffende gegevensverzamelingen laten zich hierdoor lastig gericht, op basis van vooraf opgestelde hypotheses en vraagstellingen, doorzoeken. Daarom wordt met complexe zoekalgoritmes gezocht naar mogelijk interessante verbanden en patronen in de data. Dit kan leiden tot nieuwe en soms volstrekte onverwachte inzichten die kunnen worden gebruikt voor het opstellen van risico- en andersoortige profielen.

In de private sector kennen we verschillende voorbeelden van organisaties die op basis van aankoopgegevens of gegevens over wat wij doen op internet (surfgedrag) klant- of gebruikersprofielen opstellen aan de hand waarvan ze vervolgens gerichte aanbiedingen kunnen doen ('behavorial advertizing' of 'behavorial targeting').[28] Al wat langer kennen we de risicoprofielen met betrekking tot kredietwaardigheid en betalingsgedrag ('credit score') op basis waarvan wordt bepaald of klanten al dan niet in aanmerking komen voor een krediet of voor een bepaalde betaalwijze (denk aan: kopen op afbetaling). Van sommige vliegmaatschappijen wordt wel gezegd dat op basis van dergelijke profielen de door een klant te betalen ticketprijs wordt gepersonaliseerd, wat naar verluidt met zich kan brengen dat de gebruikers van dure computers (zeg: een Apple Macbook) een hogere prijs betalen dan degenen die gebruik maken van een goedkopere Windows PC.[29]

---

[27] Zie voor een overzicht WRR, Big data in een vrije en veilige samenleving, Amsterdam: Amsterdam University Press 2016, p. 33 e.v.

[28] Zo kon de Amerikaanse supermarkt Target aan de hand van een dergelijke analyse voorspellen of iemand zwanger was en daarop inspelen door aanbiedingen. Zie C. Duhigg, How Companies Learn Your Secrets, New York Times 16 februari 2012. Zie ook uitgebreid over behavioral targeting: F. Zuiderveen Borgesius, Improving Privacy Protection in the Area of Behavioural Targeting, Deventer: Kluwer Law International 2015.

[29] J. Angwin & D. Mattioli, 'Coming Soon: Toilet Paper Priced Like Airline Tickets', Wall Street Journal 2 september 2012. Over personalised pricing zie ook OFT, Personalised Pricing, May 2013, OFT 1489.

Ook in de publieke sector zijn tal van voorbeelden te vinden van het gebruik van big data. Vaak wordt daarbij direct gedacht aan de onthullingen van Snowden over de Amerikaanse veiligheidsdienst NSA die op ongekend grote schaal data verzamelt over het wereldwijde telecommunicatie- en internetgebruik, maar ook in Nederland wordt veelvuldig gebruik gemaakt van data-analyses en profilering. Zo maakt de Belastingdienst bijvoorbeeld al veelvuldig gebruik van big data-analyses om belastingfraude te kunnen opsporen,[30] de politie gebruikt data-analyses om criminele hotspots en veelplegers te vinden om preventief misdaad te bestrijden,[31] gemeenten, UWV en andere instanties maken gebruik van het Systeem Risico Identificatie (SyRI) om risicoanalyses uit te voeren met het oog op de opsporing van sociale zekerheidsfraude,[32] de Marechaussee kijkt naar opvallende reisbewegingen van voertuigen of personen en financiële toezichthouders kijken naar afwijkende transacties, boven een bepaald bedrag of naar een bepaald land, om te bepalen hoe en waar zij hun toezichtsbevoegdheden inzetten.[33]

Als het gaat om fraudebestrijding, betekent het gebruik van big data vooral het kunnen voorspellen welke patronen (zoals combinaties van eigenschappen en kenmerken van personen, gedragingen geldstromen etc.) naar verwachting aanduiden waar de grootste statistische kans is dat er sprake is van fraude of onregelmatigheden. Dit kan samengaan met profilering van situaties en personen. Het is de fraudebestrijders echter niet zozeer te doen om het vinden van patronen – dat is een tussenstap – maar om het zo snel mogelijk in beeld brengen van fraudeurs op naam. Het gaat hen om het in beeld brengen van concrete frauderisico's, het vergroten van pakkansen van fraudeurs, en uiteindelijk het identificeren van de fraudeurs, zodat die daadwerkelijk 'gepakt' kunnen worden. Voor het oplossen van concrete zaken wordt dan ook niet alleen gebruik gemaakt van grote databestanden (big data). Veel vaker is sprake van analyse van een relatief beperkte set data over een persoon of zaak (little data), veelal al met een min of meer concrete verdenking.

2.3 Voorbeelden van (big) data-analyse bij de aanpak van zorgfraude

Fraudebestrijding door middel van data-analyse is in de zorg nog een relatief nieuw fenomeen. De meeste gevallen van fraude komen nog steeds aan het licht door middel van tips, bijvoorbeeld van patiënten die een factuur zien en vaststellen dat de gefactureerde behandeling niet overeenkomt met de ondergane behandeling en dit melden bij hun zorgverzekeraar of de

[30] P. Olsthoorn, Big Data voor Fraudebestrijding, WRR: Den Haag 2016.

[31] Zie D. Willems & R. Doelemans, 'Predictive Policing – wens of werkelijkheid?', Tijdschrift voor de Politie, 2014-4/5, p. 39 e.v.

[32] Olsthoorn 2016, p. 99 e.v.; zie over SyRI ook G-J. Zwenne & A.H.J. Schmidt, Wordt de homo digitalis bestuursrechtelijk beschermd? in: Homo Digitalis, NJV-preadvies 2016; Deventer 2016, p. 310 en 339-341.

[33] B.H. Custers, 'Risicogericht toezicht, profiling en Big Data', Tijdschrift voor Toezicht 2014/5, p. 9-16.

NZa. Maar dat laat onverlet dat er al wel degelijk gebruik wordt gemaakt van data-analyse en patroonherkenning bij de bestrijding van zorgfraude. Dit zal naar verwachting in de toekomst alleen maar verder toenemen.

We geven een aantal voorbeelden gegeven van data-analyse bij de bestrijding van zorgfraude. De voorbeelden maken vooral duidelijk dat, zeker in samenwerkingsverbanden, de succesvolle inzet van big data bij de bestrijding van zorgfraude in de praktijk nog niet zo eenvoudig is. En dat is eigenlijk ook wel het meer algemene beeld bij gegevensuitwisseling en data-analyse in de toezichts- en handhavingspraktijk. Hiervan wordt, vaak niet ten onrechte, veel verwacht, maar in de praktijk blijkt deze gegevensuitwisseling en data-analyse dusdanig omgeven met lastige technische, organisatorische en juridische vraagstukken dat deze in het geheel niet van de grond komt of, als dat wel het geval is, niet voldoet aan de verwachtingen.

Dat blijkt bijvoorbeeld ook uit de gang van zaken rondom het Expertisecentrum Zorgfraudebestrijding (EZB). Dit EZB is in 2013 opgericht door de Taskforce Integriteit Zorgsector (TIZ) ter uitvoering van het convenant ter verbetering van de bestrijding van zorgfraude tussen de NZa, IGZ, Zorgverzekeraars Nederland, Centrum Indicatiestelling Zorg (CIZ), Inspectie SZW, FIOD, Belastingdienst en OM. Het convenant is tot stand gekomen onder regie van VWS en heeft tot doel de samenwerking in de keten te verbeteren bij de aanpak van zorgfraude, met name door meer informatie-uitwisseling (waaronder persoonsgegevens) over ernstige zorgfraudes en gezamenlijk gebruik van databestanden voor zover mogelijk binnen de wettelijke kaders. Daarnaast voorzag het convenant in de oprichting van projectorganisatie en werkgroepen voor 'specifieke risico-analyse'.[34] Hiervoor werd in 2013 het EZB opgericht, in eerste instantie bij wijze van pilot.

Het EZB had zowel een strategische als operationele functie. De operationele functie omvatte de centrale aanpak van meervoudige zorgfraudezaken op terreinen waar de grootste risico's liggen. Bij de strategische functie ging het om het maken van analyses met het doel om meer zicht te krijgen op waar zorgfraude vooral speelt en hoe dat te herkennen. Intelligence speelde daarbij een grote rol. Het was expliciet de bedoeling om gegevens uit bestaande grote databases, zoals de Vektis database en het DBC-informatiesysteem (DIS), te koppelen en door middel van bestandsanalyse patronen te zoeken die een aanwijzing zouden opleveren voor frauduleus handelen. Vektis beheert als uitvoeringsorganisatie van zorgverzekeraars een omvangrijke database met de declaratiegegevens van nagenoeg alle zorgverzekeraars. En het DIS dat inmiddels is ondergebracht bij de NZa, bevat alle gegevens van zorgaanbieders over afgesloten DBC-trajecten in de ziekenhuiszorg, GGZ en forensische zorg en over trajecten in de basis-GGZ. Ook zou het EZB gevoed worden met fraudesignalen vanuit het Verzamelpunt Zorgfraude dat ook op grond van het al genoemde convenant

[34] Kamerstukken II 2012/2013, 28828, nr. 50.

was ingesteld.

Dit klinkt allemaal dus al wel echt als big data. Er werd dan ook veel van het EZB verwacht. In de praktijk bleek het echter geen succes. Volgens Olsthoorn die daarbij een woordvoerder van de NZa aanhaalt, werkte het in de vorm van het experiment niet. 'Vooral de algemene analyses waren niet goed genoeg om te worden toegepast.[35] In 2015 is het EZB dan ook een stille dood gestorven.

Ook zorgverzekeraars zijn bezig met data-analyses in de strijd tegen zorgfraude. Zij zijn wettelijk verplicht na te gaan of ingediende declaraties voor vergoeding in aanmerking komen en moeten in dat verband controleren of (1) patiënten verzekerd zijn voor geleverde zorg en of gehanteerde tarieven kloppen (formele controle)[36] en (2) of door de zorgaanbieder in rekening gebrachte zorg ook daadwerkelijk is geleverd en het meest was aangewezen in het licht van de gezondheidstoestand van de patiënt (materiële controle).[37] Om dit effectief te kunnen doen, hebben zorgverzekeraars in de afgelopen jaren elk apart veel geïnvesteerd in de verbetering van hun processen en systemen voor controles en fraudebeheersing. Data-analyse door patroonherkenning speelt daarin een steeds belangrijkere rol om (risico's op) onrechtmatigheden te detecteren.[38]

Deze investeringen hebben ook wat opgeleverd. Uit cijfers van het Kenniscentrum Fraudebeheersing in de Zorg (onderdeel van Zorgverzekeraars Nederland) blijkt bijvoorbeeld dat zorgverzekeraars in 2014 2 miljard aan declaraties hebben afgewezen bij controle voor betaling (formele controle voor betaling). EUR 449 miljoen euro is teruggevorderd door betaalde declaraties achteraf te controleren (formele controles na betaling en materiele controles). Er is voor EUR 18,7 miljoen aan fraude geconstateerd.[39] En in 2015 is bij controle voor betaling voor een bedrag van EUR 2,4 miljard aan declaraties door zorgverzekeraars afgewezen. Via controles na betaling is er EUR 485 miljoen teruggevorderd. In 2015 is er voor EUR 11,1 miljoen aan fraude vastgesteld.[40]

De cijfers maken echter ook duidelijk dat er een groot verschil in omvang bestaat tussen fouten en fraude. Het bedrag aan geconstateerde fraude bedraagt minder dan 1% van het totale bedrag aan onjuist geachte declaraties. Dat wil niet per se zeggen dat het dus met de fraude in de zorg wel meevalt. Fraude is een heftige beschuldiging en moet dus zorgvuldig wor-

[35] Olsthoorn 2016, p. 37.

[36] Art. 1, aanhef, en onder t, Rzv.

[37] Art. 1, aanhef, en onder u, Rzv.

[38] Zo beschikt bijvoorbeeld Achmea al sinds 2011 over een eigen Kenniscentrum voor onderzoek en analyse van de data in haar eigen Achmea Health database. Dit Kenniscentrum doet niet alleen onderzoek met het oog op fraudebestrijding, maar ook met het oog op de verbetering van zorginkoop of –verlening. Zie Olsthoorn 2016 p. 46 e.v.

[39] Zorgverzekeraars Nederland, Rapportage controle en fraude 2014 <www.zn.nl>.

[40] Zorgverzekeraars Nederland, Rapportage controle en fraude 2015 <www.zn.nl>.

den onderbouwd. Juist omdat nu eenmaal veel fouten worden gemaakt en vaak lastig te achterhalen is of de boel bewust wordt geflest, zijn zorgverzekeraars terughoudend met de kwalificatie fraude en kiezen zij er veelal voor om hun geld terug te vorderen, desnoods via de rechter[41], en eventueel om het contract met een bepaalde zorgaanbieder te beëindigen.[42] Het lijkt dan ook aannemelijk dat veel gevallen waarin sprake is van fraude, niet worden opgepikt en (tucht-, bestuurs- of strafrechtelijk) vervolgd, omdat niet bewezen kan worden dat sprake is van fraude of het daarmee gemoeide belang eenvoudig te gering is en de kosten dus niet tegen de baten opwegen.

Het is goed mogelijk dat meer fraudegevallen zouden kunnen worden aangetoond, indien zorgverzekeraars nauwer zouden samenwerken en gebruik zouden maken van elkaars data en analyses. De NZa en de politiek willen graag dat zorgverzekeraars meer samen optrekken bij de bestrijding van fraude. Van gezamenlijk fraudeonderzoek door zorgverzekeraars is tot nog toe echter geen sprake. Het al genoemde Kenniscentrum Fraudebeheersing in de Zorg doet niet aan data-analyse, maar rapporteert enkel over de fraudemeldingen van de afzonderlijke verzekeraars. Wel hebben de zorgverzekeraars in 2015 aangegeven dat zij de mogelijkheden zullen verkennen om gezamenlijk fraudeonderzoek te doen door middel van een onafhankelijke, gezamenlijke onderzoeksunit.[43] Volgens Olsthoorn heeft gezamenlijk fraudeonderzoek voor een aantal grote zorgverzekeraars echter geen prioriteit, omdat zij de kostenbesparingen als gevolg van fraudebestrijding als concurrentievoordeel zien.[44]

Ook de NZa houdt zich bezig met big data, samen met handhavingspartners en veldpartijen (vgl. het reeds genoemde EZB), maar ook op eigen initiatief. Zo heeft de NZa in 2013 en 2014 uitgebreid onderzoek verricht en laten verrichten om zicht te krijgen op de aard en omvang van onregelmatigheden in een aantal verschillende zorgsegmenten (huisartsen, mondzorg, farmacie, GGZ, paramedische zorg en de medisch specialistische zorg).

In de eerste plaats heeft het Fraude Detectie Expertise Centrum (FDEC) in opdracht van de NZa met behulp van data-mining een enorme hoeveelheid declaratiebestanden van zorgverzekeraars, zorgkantoren, CAK en CIZ (verzameld en aangeleverd door het reeds eerder genoemde Vektis) doorgelicht op zoek naar onregelmatigheden in declaraties. Voor deze analyse zijn 881 miljoen records gebruikt van declaraties van huisartsen, 190 miljoen records met betrekking tot mondzorg, 619 miljoen records met betrekking tot farmaceutische zorg, 4 miljoen records voor GGZ-zorg, 166 miljoen voor paramedische zorg en 49 miljoen records voor medisch specialisti-

[41] Zie bijv. Rb. Rotterdam Midden-Nederland 1 mei 2013, ECLI:NL:RBMNE:2013:BZ9057; Rb. Rotterdam 17 juli 2013, ECLI:NL:RBROT:2013:5587 en Rb. Midden-Nederland 9 juli 2014, ECLI:NL:RBMNE:2014:2742.

[42] Olsthoorn 2016, p. 35.

[43] Zorgverzekeraars Nederland, Plan van aanpak fraudebeheersing 2015-2017

[44] Olsthoorn 2016, p. 38..

sche zorg. In totaal zijn maar liefst 1,9 miljard records die betrekking hebben op de jaren 2010-2012 door het FDEC geanalyseerd. Voor de GGZ waren geen gegevens beschikbaar over 2012.

Het FDEC heeft op deze data een uitgebreide data-analyse gedaan om onregelmatigheden op te sporen in declaraties. Daarbij is in de eerste plaats gekeken naar harde overtredingen van declaratieregels en daarnaast naar opvallende afwijkingen ten opzichte van het gemiddelde declaratiegedrag (door het FDEC aangeduid als anomalieën).[45] Bij geconstateerde overtredingen is het daadwerkelijk vergoede bedrag vergeleken met het bedrag dat het volgens de declaratieregels zou moeten zijn. Bij anomalieën is het bedrag vergeleken met het bedrag dat hoort bij het gehanteerde referentiepunt.

Uit het onderzoek blijkt dat in de onderzochte zorgsegmenten de declaratieregels op grote schaal worden overtreden en/of sprake is van anomalieen. Die onregelmatigheden kunnen allerlei vormen aannemen, zoals ongeldige combinaties van declaraties, dubbele declaraties, upcoding, declaraties nadat de patiënt is overleden enz. enz. Het FDEC doet geen uitspraken over of in deze gevallen sprake is van frauduleus handelen, maar adviseert in deze gevallen nader onderzoek te doen om na te gaan of sprake is van oprechte vergissingen, opzet of dat de regels wellicht niet helder (genoeg) zijn. Het FDEC merkt daarnaast op dat ook de kwaliteit van de brongegevens in voorkomende gevallen een oorzaak kan voor het signaleren van merkwaardigheden.

Op basis van de gevonden overtredingen en anomalieën concludeert het FDEC dat in de onderzochte segmenten in 2010 minimaal een bedrag van EUR 113 miljoen is gemoeid met onregelmatigheden en in 2011 minimaal EUR 117 miljoen. Voor 2012 kon door het FDEC geen totaalcijfer worden gegeven, omdat over dat jaar GGZ-gegevens ontbraken en bovendien ook de gegevens met betrekking tot de medisch specialistische zorg niet compleet waren.

De NZa heeft vervolgens ook nog zelf aanvullend onderzoek op basis van het DBC-informatiesysteem (DIS) gedaan voor de GGZ.[46] Dit onderzoek is uitgevoerd als aanvulling op de analyse van de declaratiegegevens voor de GGZ, omdat de declaratiegegevens niet alle informatie bevatten die nodig is voor een volledige analyse. In het bijzonder is de daadwerkelijk bestede tijd (lees: geschreven tijd) niet bekend is in de Vektis bestanden. In dit onderzoek heeft de NZa in het bijzonder gekeken naar anomalieën op basis van de bestede tijd ten opzichte van de tijd die een zorgaanbieder gemiddeld besteedt aan een patiënt. Op basis van deze analyse heeft de NZa vastgesteld dat voor 2011 EUR 54,6 miljoen en in 2012 voor EUR 50,8 miljoen

[45] Zie FDEC, Onregelmatigheden in declaratiebestanden bij huisartsen, mondzorg, farmacie, GGZ, paramedische zorg en medisch specialistische zorg, 28 augustus 2014. Beschikbaar via <www.nza.nl>.

[46] NZa, Tijdschrijven, verblijfsdagen en diagnoses in de GGZ. Een beeld van onregelmatigheden in de DIS data, november 2014. Beschikbaar via <www.nza.nl>.

teveel is uitbetaald. Ook heeft de NZa gekeken naar overlap in verblijfdagen. Op basis van deze analyse komt de NZa nog eens tot een te veel vergoed bedrag van EUR 155 miljoen euro in 2012.

In haar eindrapport[47] is de NZa erg voorzichtig met het trekken van harde conclusies over zorgfraude op basis van deze onderzoeken, omdat de onderzochte bestanden niet voor het opsporen van fraude bedoeld zijn. De gegevens zijn niet altijd volledig en deels nog niet gecontroleerd door zorgverzekeraars. Wanneer ze wel zijn gecontroleerd en er correctie door de verzekeraar heeft plaatsgevonden, dan is de correctie niet altijd vastgelegd in de bestanden van Vektis. Ook voor de geconstateerde anomalieën geldt dat ze niet met zekerheid aan te merken zijn als een fout of fraude. Ze wijken weliswaar opvallend veel af van gemiddelden, maar daar kan in nader onderzoek een verklaring voor blijken te zijn. Andersom zijn overtredingen die niet afwijken, niet zichtbaar via deze methode. Spookzorg, het in rekening brengen van zorg die niet is geleverd, blijft bijvoorbeeld geheel buiten beeld, indien de declaraties geen fouten bevatten. De NZa geeft dan ook aan dat zij, de beperkingen van het onderzoek wegende, niet op een verantwoorde en betrouwbare wijze een totaalcijfer voor zorgfraude kan bepalen. Maar de analyse biedt naar het oordeel van de NZa wel inzicht in gevallen waar het mis kan zijn en geeft daarmee belangrijke input voor toezicht, nader onderzoek en mogelijkheden voor zorgverzekeraars om hun controleprocessen te verbeteren. Op grond hiervan beveelt de NZa dan ook onder meer aan dat zorgverzekeraars meer met data-analyse gaan doen om fraude te herkennen en aan te pakken.[48]

2.4 Wat brengt de toekomst?

Het doen van voorspellingen is lastig. Zeker als het om de toekomst gaat.[49] Maar, afgaand op ontwikkelingen in andere domeinen, is het meer dan aannemelijk dat er ook in het zorgdomein veel gebruik gaat worden gemaakt van data-analyse en big data. In de fiscale vakliteratuur worden bijvoorbeeld al methodes besproken gericht op het 'construeren van patronen op basis van (potentiële) toekomstige gedragingen' en het 'conceptualiseren van toekomstige non-compliante gedragspatronen'.[50] Ook in de toelichting bij het welbekende (beruchte) SyRI-besluit[51] wordt voorgesorteerd op scenario's die veel weg hebben van wat er in de scifi-speelfilm Mi-

[47] NZa, Onderzoek naar kwetsbaarheden en financiële onregelmatigheden in de zorg, november 2014. Beschikbaar via <www.nza.nl>.

[48] Zie NZa, Rapport Onderzoek zorgfraude: Tussenrapport (update), februari 2014, p. 13-17, 30-31, 38. Zie ook NZa, Onderzoek naar kwetsbaarheden en financiële onregelmatigheden in de zorg, november 2014. Beschikbaar via <www.nza.nl>.

[49] De uitspraak 'prediction is very difficult, especially about the future' wordt wel toegeschreven aan Niels Bohr.

[50] T. van Berkhout & T. van Engers 'Onderzoeksmethodologie voor informatiegestuurd sociaal toezicht', WFR 2012/824.

[51] Besluit van 1 september 2014 tot wijziging van het Besluit SUWI in verband met regels voor fraudeaanpak door gegevensuitwisselingen en het effectief gebruik van binnen de overheid bekend zijnde gegevens met inzet van SyRI, Stb. 2014, 320.

nority Report wordt aangekondigd.[52]

Ook het buitenland biedt inzichten in wat er wellicht gaat komen. In de Verenigde Staten (waar anders) zien we dat Centers for Medicare and Medicaid Services (CMS), een federale verzekeringsorganisatie, data-analyses gebruikt om zorgfraude te voorkomen. Voorheen werden zorgdeclaraties digitaal en handmatig vergeleken met aangeleverde documentatie van de zorgprofessionals. Inmiddels vergelijkt het 'Fraude Prevention System' declaratiepatronen met profielen van fraudeurs. Zo worden de zorgaanbieders die meer uren declareren dan dat er in een dag passen eruit gepikt. En er wordt er met behulp van gegevens uit sociale netwerken rekening gehouden met samenwerkingspartners van eerdere fraudeurs. Op deze manier worden naar verluidt honderden miljoenen dollars per jaar aan fraude opgespoord, significant meer dan de jaren voordat dit nieuwe systeem in werking werd gesteld.[53]

Duidelijk is in ieder geval dat, als het aan VWS en de NZa ligt, in de komende jaren in de strijd tegen zorgfraude steeds meer gebruik zal worden gemaakt van data-analyse en big data.[54] Door zorgverzekeraars in het kader van hun wettelijke controletaak, maar ook door de NZa en haar handhavingspartners in het kader van gericht toezicht en handhaving.

3. BEPERKINGEN VAN, EN RISICO'S BIJ, GEBRUIK VAN BIG DATA

Het voorgaande maakt duidelijk dat in de strijd tegen zorgfraude veel van big data wordt verwacht en dat we daar dus in de toekomst steeds meer mee geconfronteerd zullen worden.

Zorgverzekeraars en de NZa kunnen al beschikken over uitgebreide databases met heel veel, gedetailleerde gegevens en hebben daarnaast uitgebreide wettelijke bevoegdheden om informatie op te vragen bij zorgaanbieders. Als deze gegevens naast elkaar gelegd worden en/of met andere bestanden (bijvoorbeeld die van de Belastingdienst of de Kamer van Koophandel) gekoppeld worden en door middel van geavanceerde technieken 'slim' geanalyseerd, kan dat inzichten opleveren op grond waarvan fouten en fraude veel gerichter en effectiever kunnen worden voorkomen, aangetoond en aangepakt. Daarmee kan in potentie heel veel geld worden bespaard, en dat is relevant in deze tijden waarin de zorgkosten als gevolg van de vergrijzing een steeds groter beslag leggen op de publieke middelen.

---

[52] Zwenne & Schmidt 2016, p. 339; Custers 2014.

[53] Zie http://www.modernhealthcare.com/article/20160524/NEWS/160529960/commentary-medicares-big-data-tools-to-fight-and-prevent-fraud-yield; https://www.stopmedicarefraud.gov/fraud-rtc12142012.pdf

[54] Zie meer uitvoerig het Programmaplan Rechtmatige zorg en het Plan van aanpak Fouten en Fraude 2015-2018 (Kamerstukken II 2014/2015, 28828, nr. 29) dat in april 2015 door de Minister van VWS aan de Kamer is aangeboden en de daaropvolgende voortgangsrapportages.

Toch is bij de inzet van big data ook terughoudendheid op zijn plaats. Ook big data kent namelijk zo zijn beperkingen en risico's.

3.1 Beperkingen van big data

Dat big data zo zijn beperkingen kent, blijkt bijvoorbeeld uit Google Flu Trends, een applicatie waarvan werd geclaimd dat die op basis van 45 zoektermen de opkomst van griepepidemieën realtime te kunnen volgen en in kaart te kunnen brengen. Ruim twee weken sneller dan de officiële overheidskanalen. Toch bleek Google Flu Trends de griep structureel te overschatten en miste de applicatie de buiten het winterseizoen optredende A-H1-N1 pandemie in 2009 volledig.[55]

Dit voorbeeld illustreert dat ook bij de analyse van grote hoeveelheden data de resultaten niet automatisch correct zijn. Maar er zijn nog andere beperkingen verbonden aan big data-analyses, zoals statistische tekortkomingen en beperkingen in het toepassingsbereik. Wij noemen de belangrijkste aandachtspunten bij het gebruik van big data.

Bias

In de eerste plaats is er het bias-probleem. Gegevens worden altijd in een specifieke context verzameld. Daardoor zit in bijna iedere dataset een specifieke bias. Als die niet wordt gecorrigeerd, kan dat makkelijk leiden tot onjuiste of anderszins problematische uitkomsten. Wanneer de politie bijvoorbeeld enkel in wijken met veel allochtonen ('mensen met een migrantenachtergrond') surveilleert, zullen politiedatabanken vooral gevuld zijn met juist dergelijke personen. En dat zal doorwerken in de gemodelleerde risicoprofielen die dan onevenredig veel van deze allochtonen bevatten. Vooral bij het combineren en hergebruiken van databronnen kan het lastig zijn om te achterhalen hoe de dataset tot stand zijn gekomen en wat dus de precieze bias is die in data zit.

Correlaties

Het is in de tweede plaats van belang te beseffen dat de verbanden die met behulp van big data analyses worden gelegd, niet noodzakelijk causaal van aard zijn. Het gaat om correlaties, dat wil zeggen om statistische verbanden. En dat maakt dat het enkele feit dat iemand past binnen een profiel nog niet automatisch betekent dat hij dus ook het aan dat profiel gekoppelde gedrag vertoont. Het verband kan ook indirect zijn of zelfs berusten op louter toeval. Zonder een betrouwbare theoretische basis over het hoe en waarom van de gevonden correlaties, en zonder goede aanwijzingen dat de verbanden causaal van aard zijn, kunnen daarop gebaseerde maatregelen

[55] D. Lazer, R. Kennedy, G. King, en A. Vespignani. 2014. 'The Parable of Google Flu: Traps in Big Data Analysis,' Science 343 (6176): 1203–1205, beschikbaar op: https://dash.harvard.edu/bitstream/handle/1/12016836/The%20Parable%20of%20Google%20Flu%20(WP-Final).pdf?sequence=1 .

de plank volledig misslaan. Het maakt daarbij nogal uit of de informatie enkel aanleiding is voor verder onderzoek of dat er hardere consequenties aan verbonden worden.

Foutmarges

In de derde plaats zijn profielen, en dat geldt ook voor profielen die zijn opgesteld op basis van *big data predictive analytics*, sowieso altijd een benadering van de werkelijkheid en bevatten zij dus foutmarges. Het gevaar bestaat daarmee dat op basis van profielen de verkeerde conclusies worden getrokken, bijvoorbeeld omdat ten onrechte wordt aangenomen dat iemand binnen het profiel valt (false positive) valt of dat ten onrechte wordt aangenomen dat iemand er buiten valt (false negative). Dat sprake is van een foutmarge is voor het aanraden van een boek op Amazon of een film op Netflix niet erg. Maar het wordt wel problematisch als iemand op basis van analyses bijvoorbeeld ten onrechte op een no fly-list wordt gezet zoals de Amerikaanse senator Ted Kennedy[56] of Europees Parlementariër Sophie In 't Veld overkwam[57] of, in het kader van dit preadvies, relevant een zorgaanbieder ineens op een zwarte lijst van zorgverzekeraars komt en niet langer in aanmerking komt voor een bepaalde aanvullende verzekering.

Kwaliteit data

In de vierde plaats kan ook de data zelf onvolledig of onjuist zijn, met als gevolg dat ook de op basis daarvan bepaalde risicoprofielen onjuist kunnen zijn. FDEC en de NZa wijzen er in de door hen uitgevoerde analyses op de databestanden van Vektis niet voor niks op dat de onderzochte bestanden niet voor het opsporen van fraude bedoeld zijn en de gegevens niet altijd volledig zijn en deels nog niet gecontroleerd door zorgverzekeraars. Er is dan alle aanleiding voor voorzichtigheid met betrekking tot het trekken van harde conclusies. Het werken met risicoprofielen vereist bovendien ook dat profielen voortdurend worden geactualiseerd. Profielen raken namelijk na verloop van tijd 'uitgewerkt', bijvoorbeeld omdat zorgaanbieders die aan het profiel voldoen, al zijn aangepakt of omdat de frauderende aanbieders hun gedrag aanpassen en daardoor buiten het profiel vallen.[58]

Incidenten

In de vijfde plaats is van belang te beseffen dat big data-analyses niet altijd de meest geschikte oplossing zijn voor een vraagstuk. Big data-analyses ontlenen hun waarde aan patroonherkenning in grote hoeveelheden data. Big data-analyses zijn dus vooral nuttig als er sprake is van een zekere regelmaat of in ieder geval bepaalde terugkerende kenmerken. Als iemand

[56] V. Mayer-Schonberger & K. Cukier, Big data, Boston: Mariner Books 2014, p. 166-167.

[57] E. Nakashima, 'European Lawmaker To Sue U.S. Over Data', Washington Post, 1 juli 2008.

[58] Custers 2014, p. 12.

eenmalig of incidenteel een spooknota verstuurt, zal dat waarschijnlijk niet uit de big data-analyse blijken. Uiteraard is het wel mogelijk dat deze persoon dan op een andere manier tegen de lamp loopt, bijvoorbeeld omdat de betreffende 'patient' meldt dat hij geen zorg heeft genoten.

Het voorgaande maakt duidelijk dat big data de nodige beperkingen kent die geadresseerd moeten worden indien wordt overwogen gebruik te maken van big data, zeker als daarop beslissingen worden gebaseerd die een serieus te nemen impact op personen hebben.

Een zorgvuldig gebruik van big data veronderstelt wat ons betreft dat zorgverzekeraars en toezichthouders die daarvan gebruik willen maken ervoor zorgen dat zij beschikken over de voor hun doeleinden benodigde gegevens en dat deze van voldoende kwaliteit is. Uiteraard zullen zij ook moeten beschikken over de nodige expertise om goede analyses uit te voeren en de uitkomsten op een passende en controleerbare wijze te duiden. Een ander belangrijk aandachtspunt in dit verband wordt daarnaast gevormd door de beveiliging van de data. Met het verzamelen, opslaan, uitwisselen en koppelen van grote hoeveelheden data wordt de beveiliging daarvan steeds belangrijker. Ook goed beveiligde systemen kunnen worden aangevallen, door codes te kraken te kraken, hardware te manipuleren of anderszins de beveiliging te omzeilen. Dat het niet gaat om een theoretisch risico, blijkt wel uit de vele berichten over beveiligingsincidenten die bijna dagelijks in de media langskomen. Naarmate de databestanden groter zijn en er meer partijen betrokken zijn bij de data-analyse, neemt het belang van goede beveiligingsmaatregelen verder toe.

3.2 Risico's bij big data analyses

Er is bovendien ook op meer principiële gronden aanleiding voor een terughoudend en zorgvuldig gebruik van big data, in zijn algemeenheid, maar ook zeker in de strijd tegen zorgfraude. In de literatuur wordt in dit verband een aantal risico's geïdentificeerd die geadresseerd moeten worden, wil het gebruik van big data maatschappelijk en juridisch aanvaardbaar zijn.[59]

Sommige van deze risico's vloeien voort uit het niet adequaat omgaan met een aantal van de bovengenoemde beperkingen. Andere risico's hebben te maken met de verzameling en analyse van grote hoeveelheden data die voor meerdere doeleinden worden gebruikt. Wij vatten de belangrijkste risico's als volgt samen:

Stigmatisering en discriminatie

Data-gedreven oplossingen werken toe naar wat wel wordt aangeduid als '*social sorting'*, dat wil zeggen: het indelen van mensen in specifieke groepen of categorieën, waaraan in allerlei contexten een betekenis kan wor-

[59] Zie uitvoerig WRR 2016, p. 88 e.v.

den toegekend die niet altijd positief is. Bij big data gebeurt dat op basis van correlaties die in de data te vinden zijn. De afbakening van groepen op basis van big data kan evenwel problematisch zijn als de *bias* die in elke dataset zit, niet goed wordt geadresseerd. De conclusies kunnen dan niet perfect passen op de groep en leiden tot ongerechtvaardigde stigmatisering en discriminatie, met grote negatieve gevolgen voor individuen. Een voorbeeld is te vinden in een studie uit de VS waarin werd vastgesteld dat zoekacties met zogeheten 'black-identifying names' (bijv. "Jermaine") aanleiding gaf tot webadvertenties gerelateerd aan "arrest" terwijl zoekacties op 'white-identifying names' (zoals "Geoffrey") heel andere, veel minder diskwalificerende resultaten opleverden. Waarom precies dit gebeurde was onduidelijk. Wel was duidelijk dat er sprake was van onwenselijke 'racially biased results'.[60]

Blackbox en dehumanization

Het gaat hier om het ontbreken of wegvallen van de mogelijkheden (voor zowel degenen die beslissingen nemen als degenen die daardoor worden geraakt) om te begrijpen waarom beslissingen worden genomen. We denken dan vooral aan de situaties die worden beschreven in het werk van Kafka. Of, wat luchtiger, de sketch 'Computer Says No' uit Britse comedy Little Britain, waarin iemand zich in een ziekenhuis moet verweren tegen een administratiesysteem dat een heel andere behandeling suggereert dan waarvoor de patiënt is gekomen.[61]

Onschuldpresumptie en fair trial

Big data processen kunnen daarnaast de onschuldpresumptie onder druk zetten en raken daarmee aan het recht op een eerlijk proces dat onder meer is gewaarborgd in artikel 6 EVRM.[62] Een aspect van de onschuldpresumptie is dat opsporingsbevoegdheden niet zomaar jegens een ieder mogen worden toegepast, maar dat er een redelijk vermoeden van het plegen van een strafbaar feit moet zijn of tenminste aanwijzingen voor de betrokkenheid bij strafbare feiten. Aan de verdachte komt bovendien een beschermde status toe om een eerlijk proces te waarborgen (denk aan het zwijgrecht en andere rechten die voortvloeien uit het nemo teneturbeginsel, de inzage in processtukken en equality of arms). Degene die op basis van een data-analyse tot een risicogroep behoort die bijzondere aandacht krijgt, heeft die speciale status vaak (nog) niet en kan zich dus niet hiertegen verweren, als hij daarvan al op de hoogte is. Daarbij komt dat diegene ook vaak helemaal niet weet hoe de algoritmes, computersystemen en profielen werken, zodat een effectieve verdediging sowieso lastig

[60] White House, Big Data: Seizing Opportunities, Preserving Values, May 2014, p. 7

[61] Een eenvoudige internetzoekactie op zoekwoorden <computer says no> en <little brittain> levert de vindplaats op van deze hilarische episode.

[62] M. Hildebrandt, 'Data-gestuurde intelligentie in het strafrecht', in: Homo Digitalis, NJV-preadvies 2016; Deventer 2016, p. 188 e.v.

is, ook nadat hij in een later stadium alsnog tot verdachte is 'gepromoveerd. Dit maakt het eens te meer van belang dat zorgverzekeraars en toezichthouders zorgvuldig omgaan met big data en niet zomaar op basis daarvan overgaan tot het opleggen van sancties of andere ingrijpende maatregelen.

Individuele en collectieve privacy

De combinatie van volume en variëteit maakt dat big data processen drijven op een overvloed aan informatie die op verschillende manieren aan personen gekoppeld is of kan worden. Spanningen met privacy en het gegevensbeschermingsrecht zijn dan nooit ver. Het ongelimiteerd verzamelen, bijeenbrengen en hergebruiken van persoons- en andere gegevens, omdat die ooit wellicht van pas komen, staat op gespannen voet met kernprincipes van het gegevensbeschermingsrecht zoals gegevensminimalisatie en doelbinding, en proportionaliteit. Ook roept de mogelijkheid van heridentificatie van geanonimiseerde gegevens, bijvoorbeeld op basis van een koppeling van eigen data aan data uit heel andere bronnen, ook weer allerlei privacyvraagstukken op. Big data zet daarmee met name de collectieve privacy onder druk, dat wil zeggen: de waarde van privacy als collectief of maatschappelijk goed.

De 'function creep' of 'mission creep'

de term ziet op het sluipenderwijs (op een zgn. glijdende schaal) uitbreiden of verbreden van de doeleinden waarvoor gegevens worden gebruikt. Soms is daarbij sprake van (al dan niet) voorziene of beoogde neveneffecten. Zo is er sprake van function creep als videocameratoezicht eerst alleen werd ingezet voor het vergroten van de veiligheid in de openbare ruimte en vervolgens ook bruikbaar blijken als opsporingsmiddel.[63] Een ander bekend voorbeeld betreft het geval waarin uit een data-analyse bleek dat gegevens betreffende onderwijsprestaties van kinderen (zgn. CITO-toets) een indicatie kunnen geven van problemen bij gezinnen, die in aanmerking komen voor bemoeizorg.[64] Dat function creep geen theorie is, blijkt reeds uit het gegeven dat de NZa voor haar onderzoek naar onregelmatigheden bij het declareren gebruik heeft gemaakt van Vektis databestanden die daarvoor niet zijn bedoeld. Dat laat onverlet dat een big data-analyse op zo'n bestand nuttige inzichten kan opleveren, maar daarbij moet men zich wel realiseren dat het bestand voor een ander doel is aangelegd en er dus een bepaalde bias in kan zitten of relevante gegevens kunnen ontbreken.

3.3 Dus...

Het is niet gezegd dat al deze beperkingen en risico's altijd relevant zijn bij

[63] J.E.J. Prins, 'Function creep: over het wegen van risico's en kansen', in Justitiële verkenningen: Function creep en privacy, Boom: Den Haag 2011/8, p. 8 e.v.

[64] Zie F. Kleeman, B. Gritter & J. Bouma, Hoe preventiever, hoe liever!: Predictie en preventie van probleemgedrag bij 6/7-jarige kinderen in Coevorden, RUG/Wetenschapswinkel April 2005, p. 18.

de inzet van big data in het kader van de strijd tegen zorgfraude, maar het maakt wel duidelijk dat er terughoudend en zorgvuldig te werk moet worden gegaan bij het gebruik van big data, zeker als daarop ingrijpende beslissingen worden gebaseerd. Er zijn dus waarborgen nodig en daarop gaan we in de volgende paragraaf in.

4. WELKE WAARBORGEN BIEDT PRIVACY- EN GEGEVENSBESCHERMINGSWETGEVING?

4.1 Aanloop

Het gebruik van big data in de strijd tegen zorgfraude kent dus de nodige haken en ogen. Dat maakt het van belang dat er een effectieve rechtsbescherming is tegen besluitvorming op basis van big data en risicoprofielen.

Afhankelijk van de specifieke situatie zijn er allerlei regels en normen waaraan rechtssubjecten rechtsbescherming kunnen ontlenen tegen de geschetste risico's. Als zorgtoezichthouders of gemeenten gezamenlijk hun gegevens gaan analyseren denken we dan aan de algemene zorgvuldigheidsnormen uit de Algemene wet bestuursrecht (Awb). Als het gaat om analyses door zorgaanbieders of –verzekeraars denken we aan contractuele en zorgvuldigheids- en betamelijkheidsnormen. Eerder nog denken we natuurlijk nog aan de fundamentele rechten en vrijheden, zoals vastgelegd in de Grondwet (GW), het Handvest voor de grondrechten van de EU (HvEU) en de mensenrechtenverdragen (EVRM en IVBPR).[65] En bij het gebruik van big data gaat het dan vooral om het recht op bescherming van de persoonlijke levenssfeer (art. 10, eerste lid, GW, art. 7 HvEU, art. 8 EVRM, art. 17 IVBPR) en het daaraan verwante recht op bescherming van persoonsgegevens (art. 10, tweede lid GW, art. 8 HvEU). Als het gaat om dataanalyses en big data komt evenwel ook (of: juist) veel betekenis toe aan de onschuldpresumptie (art. 48 HvEU, art. 6, tweede lid, EVRM, art. 14 IVBPR) en natuurlijk ook het discriminatieverbod (art. 1 GW, art 21 HvEU, art. 14 EVRM, art. 2 IVBPR).

Verder denken we vooral aan de wetgeving waaraan nadere regels wordt gegeven ter bescherming van deze rechten en vrijheden, zoals dat voor het discriminatieverbod wordt gedaan in de Algemene wet gelijke behandeling en op termijn voor de onschuldpresumptie, zij het in de context van opsporing en vervolging, mogelijk in de wetgeving ter implementatie van de Onschuldpresumptierichtlijn.[66]

Verreweg het meest relevant voor big data, in elk geval het meest uitvoerig en gedetailleerd, zijn op dit moment de nadere regels die worden gesteld

[65] G.J. Zwenne & WA.M. Steenbruggen, 'Privacyvoorwaarden voor de iOverheid. Vuistregels voor wet- en regelgevers met betrekking tot overheidsinformatiesystemen', Regelmaat 2015/1, p. 19-36.

[66] Zie ook: Proposal for a Directive of the European Parliament and of the Council on the strengthening of certain aspects of the presumption of innocence and of the right to be present at trial in criminal proceedings, 27.11.2013, COM (2013) 821 final.

ter bescherming van het recht op bescherming van persoonsgegevens. Deze regels vinden we op dit moment de Wet bescherming persoonsgegevens (Wbp) en straks, vanaf 25 mei 2018, in de Algemene Verordening Gegevensbescherming (AVG) in combinatie met de Uitvoeringswet AVG (Uw AVG).[67]

Voor de specifieke verhoudingen van zorgaanbieder en -afnemers zijn er natuurlijk ook wetten als de Wet op de geneeskundige behandelingsovereenkomst (art. 7:456 t/m 7:468 BW) en de Wet op de beroepen in de individuele gezondheidszorg (Wet BIG) van belang. Echter, omdat we verwachten dat in deze verhoudingen (nog?) niet direct sprake zal zijn van data-analyses in het kader van de aanpak van zorgfraude, gaan we daarop in deze bijdrage niet verder in.[68]

Ook kiezen we ervoor om voor dit moment nog even voorbij te gaan aan wetswijzigingen en wetsvoorstellen waarmee wordt beoogd de mogelijkheden om te komen tot een effectieve aanpak te verruimen. Dit omdat het, gelet op de heftigheid van de discussies daarover, het nog te speculatief lijkt om op basis daarvan harde conclusies te trekken.

Wel is er een duidelijke trend waarin de aanpak van zorgfraude door middel van (gezamenlijke) data-analyse meer aandacht krijgt en daarin zien wij een bevestiging dat big data veel meer betekenis zal krijgen, en ook dat de beperkingen op het gebruik ervan niet zozeer in de zorgspecifieke wetgeving moeten worden gezocht, maar in de privacy- en gegevensbeschermingswetgeving. En daaruit dan vooral (1) de regels voor de verwerking van gezondheidsgegevens en (2) de doelbindingsvereisten, (3) de regels voor profilering en ten slotte (4) de transparantieverplichtingen, dat wil zeggen de verplichtingen om degenen over wie gegevens worden verwerkt (de zgn. betrokkenen of datasubjecten) daarover op begrijpelijke wijze te informeren. Voordat we komen tot enkele afsluitende woorden gaan we, bij wijze van vingeroefening, in op de voor deze vier thema's gestelde regels en de daaruit voortvloeiende waarborgen.

4.2 Gezondheidsgegevens

De wetgever gaat uit van een ruim begrip van 'gegevens betreffende iemands gezondheid' of 'gezondheidsgegevens'. Eronder vallen niet alleen gegevens over ziektes, aandoeningen en stoornissen en de behandeling daarvan, maar ook het gegeven dat iemand ziek is zonder dat bekend is wat eraan mankeert.[69] Dit betekent dat een enkele ziekmelding of een on-

[67] Voor de aanpak van zorgfraude zal deze uitvoeringswet van belang zijn omdat daarin de regels voor het gebruik van gezondheidsgegevens te vinden zullen zijn.

[68] Daarmee willen we niet zeggen dat de in deze verhoudingen verwerkte gezondheidsgegevens geen betekenis hebben in dergelijke analyses, maar waarschijnlijk eerder in relatie tot de in de privacy- en gegevensbeschermingswetgeving gestelde regels (zoals het doelbindingsvereiste).

[69] Kamerstukken II 1997/98, 25 892, nr. 3, p. 109; Autoriteit persoonsgegevens, De zieke werknemer: beleidsregels voor de verwerking van persoonsgegevens over de gezondheid van zieke werknemers, z2015-00774, 23 februari 2016, p. 10; zie ook overw. 35 van de AVG.

gespecificeerde tandartsrekening al onder het begrip kan vallen. En, hoewel vergezocht, is ook voorstelbaar dat onschuldige gegevens over geboortejaar of leeftijd, of woonplaats, kunnen worden opgevat als gezondheidsgegevens, aangezien ook dergelijke gegevens inzicht kunnen geven in de gezondheidstoestand van de betrokkene.[70]

Voor gezondheidsgegevens geldt een verwerkingsverbod. Op grond van artikel 16 Wbp respectievelijk art. 9, eerste lid, AVG mogen dergelijke gegevens niet worden verwerkt (verzameld, vastgelegd, gebruikt etc.), tenzij sprake is van een *lex specialis* die dit wel uitdrukkelijk toestaat of als er gebruik kan worden gemaakt van een in de verordening of wet genoemde uitzondering. Deze uitzonderingen vinden we in artikel 21 Wbp en het daarmee overeenkomende artikel 23 UwAVG (ontwerp), en in artikel 23 Wbp respectievelijk art. 9, tweede lid, onder a t/m j AVG. Van deze uitzonderingen biedt er naar onze indruk maar één duidelijk de ruimte om gezondheidsgegevens te verwerken in verband met een big data analyse ten behoeve van de aanpak van zorgfraude.

En dat is de uitzondering van artikel 23, eerste lid, onder g, Wbp, die ziet op de gevallen waarin de gegevensverwerking nodig is met het oog op een zwaarwegend algemeen belang, waarbij er passende waarborgen worden geboden ter bescherming van de persoonlijke levenssfeer en dit bij wet wordt bepaald dan wel door Ap ontheffing is verleend. Een voorwaarde die deze toezichthouder tegenwoordig[71] wel stelt aan het verlenen van zo een ontheffing is echter dat er reeds een wetsvoorstel aanhangig is waarin wordt voorzien in een wettelijke regeling die de verwerking mogelijk maakt.[72] Voor de verwerking van gezondheidsgegevens in het kader van de aanpak van zorgfraude lijkt en wettelijke regeling de enige echt goed begaanbare weg te zijn.

4.3 Doelbinding

Een belangrijk uitgangspunt uit de privacy- en gegevensbeschermingswetgeving is het doelbindings- of doelverenigbaarheidsvereiste. Dit verlangt enerzijds dat persoonsgegevens alleen mogen worden verzameld voor

[70] Een hoge of lage leeftijd zegt immers iets over de kans om last te krijgen van bepaalde kinder- of ouderdomsziektes. En wie woont in Amsterdam of Utrecht heeft veel meer dan een inwoner van Zeeland of Drenthe, een risico om klachten te krijgen over fijnstof.

[71] Het beleid van de Ap daarover is nog betrekkelijk onduidelijk. In de zaak die uiteindelijk leidde tot ABRvS 3 september 2008, ECLI:NL:RVS:2008:BE9698, AB 2008, 335, werd door de toezichthouder zelf betoogd dat een ontheffing ook kan worden verleend om een pilot mogelijk te maken, zodat aan de hand van de uitkomsten daarvan kan worden besloten om al dan niet zo een wetsvoorstel voor te bereiden. In latere besluiten lijkt de toezichthouder daarvan afstand te hebben nemen, maar zonder uit te leggen wat de overwegingen daarvoor zijn.

[72] In de context van een publiekrechtelijk samenwerkingsverband kunnen gezondheidsgegevens op grond van art. 22, vijfde en zesde lid, Wbp 'meeliften' met de verwerking van strafrechtelijke gegevens die nodig is om de publiekrechtelijke taken van de deelnemers aan dat samenwerkingsverband uit te voeren. Daarvan kan dan wellicht gebruik worden gemaakt door toezichthouders, zoal NZa en IGZ, en organen als de FIOD of gemeenten, maar niet door zorgverzekeraars en zorgaanbieders.

welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde verzameldoelen (art. 7 Wbp en art. 5, eerste lid, onder, b, AVG) en anderzijds dat deze gegevens, nadat ze zijn verzameld, alleen verder mogen worden verwerkt voor doeleinden die niet onverenigbaar zijn met de doeleinden waarvoor deze zijn verzameld (art. 9, eerste lid, Wbp en art. 5, eerste lid, onder, b, AVG). Of er sprake is van 'niet onverenigbaarheid' wordt bepaald aan de hand van de verwantschap tussen de onderscheiden doeleinden, de gevoeligheid van de gegevens, de gevolgen voor de betrokkenen en of de gegevens al dan niet van de betrokkenen zelf zijn verkregen, en de overige waarborgen die zijn getroffen ter bescherming van de betrokkenen. Daarbij komt veel betekenis toe aan de redelijke verwachtingen die de betrokkenen hebben met betrekking tot de verdere verwerking van hun gegevens (art. 9, tweede lid, Wbp, art. 6, vierde lid, AVG).[73]

Deze doelverenigbaarheidsvereisten komen in verschillende andere gedaanten terug in de wetgeving. We zien het onder andere in het zgn. opslagbeperkingsvereiste dat verlangt dat gegevens niet langer worden bewaard dan nodig voor het doel waarvoor deze zijn verzameld (art. 10 Wbp, art. 5, eerste lid, onder e, AVG) en in het gegevensminimalisatiebeginsel, op grond waarvan niet meer gegevens mogen worden verzameld en verwerkt dan nodig voor het doel waarvoor deze zijn verzameld (art. 11, eerste lid, Wbp, art. 5, eerste lid, onder a, AVG).

Voor het doelbindingsvereiste in het algemeen, en voor het opslagbeperkings- en gegevensminimalisatievereiste in het bijzonder, is wel duidelijk dat er op zijn minst sprake is van een gespannen verhouding met de toepassing van big data analyses. Vanuit het perspectief van big data ligt voor de hand om zoveel mogelijk gegevens zolang mogelijk te bewaren -- wie weet levert het nog bruikbare correlaties, patronen en trends op -- terwijl deze gegevensbeschermingsvereisten juist erop zijn gericht om over zo min mogelijk gegevens te beschikken. Als zodanig lijken deze vereisten dus vergaande beperkingen op te leggen. Echter, zoals dat gaat, juist dan voorziet de wet ook in uitzonderingsmogelijkheden. Zo kan voorbij worden gegaan aan het doelbindingsvereiste als dat nodig is voor, onder andere, de voorkoming, opsporing en vervolging van strafbare feiten, of gewichtige economische en financiële belangen van de staat en andere openbare lichamen, of het toezicht op de naleving van wettelijke voorschriften die zijn gesteld ten behoeve van deze gewichtige economische en financiële belangen (art. 43, onder b t/m d, Wbp, art. 23, eerste lid, onder d, e, g en h, AVG). Ook de bescherming van de belangen van de betrokkenen of anderen kan een reden zijn om voorbij te gaan aan het doelbindingsvereiste (art. 43, onder e, Wbp, art. 23, eerste lid, onder i, AVG).

Voor het optuigen van arrangementen gericht op de aanpak van zorgfraude zal het probleem, of de uitdaging, zijn dat er verzameldoelen worden vast-

[73] Zie overw. 50 AVG.

gesteld die voldoende welbepaald zijn, en tegelijkertijd ook de mogelijkheid van data-analyse niet uitsluiten. De verzameldoelen moeten een kader kan bieden waaraan getoetst kan worden of de gegevens nodig zijn voor die doeleinden of niet. En als we moeten vaststellen dat de verzameldoelen wat dat betreft geen onderscheidend vermogen hebben en eigenlijk alles toelaten, betekent dat die onvoldoende welbepaald zijn.[74]

Om deze reden volstaat het waarschijnlijk niet dat zorgaanbieders of -verzekeraars in hun privacyverklaringen opnemen dat de gegevens van hun cliënten kunnen worden gebruikt voor 'analyse-doeleinden' of 'de aanpak van fraude'. In plaats daarvan zullen de doeleinden nader moeten worden gespecificeerd, bijvoorbeeld door uiteen te zetten wat voor soort fraude het betreft en welke organisaties daarbij nog meer over de gegevens kunnen beschikken, enz. Waar het gaat om gegevens die in een heel andere context worden verzameld, zoals de in de vorige paragraaf genoemde onderwijsresultaten, zal het gebruik ervan niet snel te verenigen zijn met het doelbindingsvereiste. Er zal dan dus gebruik moeten worden gemaakt van een van de uitzonderingen waarin de wet voorziet. Onze indruk is dat die, waar het gaat om de aanpak van zorgfraude, daarvoor wel de nodige ruimte bieden.

4.4 Profilering

Voor big data analyses zijn verder ook de regels voor profilering van belang. We vinden deze regels met zoveel woorden vooral in de AVG, die wat dit betreft evenwel voortbouwt op regels die in de Wbp werden gesteld voor zogeheten geautomatiseerde besluitvorming (art. 42 Wbp). Het begrip 'profilering' is ruim en omvat elke vorm van geautomatiseerde gegevensverwerking waarbij aan de hand van persoonsgegevens bepaalde persoonlijke aspecten van een natuurlijke persoon worden geëvalueerd. Het gaat daarbij met name om het te analyseren of te voorspellen van de beroepsprestaties van de betrokkene, diens economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie of verplaatsingen (art. 4, onder 4, AVG).

Een betrokkene heeft het recht niet te worden onderworpen aan een uitsluitend op geautomatiseerde verwerking, waaronder profilering, gebaseerd besluit waaraan voor hem rechtsgevolgen zijn verbonden of dat hem of haar anderszins in aanmerkelijke mate treft (art. 42, eerste lid, Wbp, art. 22, eerste lid AVG). In drie gevallen zijn er uitzonderingen mogelijk.

- Een betrokkene heeft dit recht niet (meer) als de profilering nodig is voor de uitvoering van een overeenkomst met de betrokkene: we kunnen daarbij denken aan de situatie waarin een aanbieder van streaming video-on-demand-diensten op basis van profilering, en met gebruik making van big data analyses, aan haar abonnees bepaalde program-

[74] Kamerstukken II 1997/98, 25982, nr. 3, p. 79.

ma's of series aanbeveelt (art. 22, tweede lid, onder a, AVG).

- En ook niet als er een wettelijke regeling is die de profilering mogelijk maakt en die voorziet in passende maatregelen ter bescherming van de rechten en vrijheden en gerechtvaardigde belangen van de betrokkene (art. 22, tweede lid, onder b, AVG). Een voorbeeld zou hier het SyRI-besluit[75] kunnen zijn dat in het voorgaande al werd genoemd en dat data-analyse ten behoeve van de aanpak van fraude in het sociale domein mogelijk maakt.

- En ten slotte, als restcategorie, ook niet als de gegevensverwerking voor profileringsdoeleinden gebeurt op basis van de uitdrukkelijke toestemming van de betrokkene (art. 22, tweede lid, onder c, AVG).

Voor de aanpak van zorgfraude is verder nog relevant dat in de wetgeving is bepaald dat er bij de gegevensverwerking voor profileringsdoeleinden alleen gebruik mag worden gemaakt van bijzondere gegevens, zoals gezondheidsgegevens, als de betrokkene daarvoor uitdrukkelijke toestemming heeft gegeven of als er passende maatregelen ter bescherming van de gerechtvaardigde belangen van de betrokkene zijn getroffen (art. 22, vierde lid, jo. art. 9 tweede lid, onder a of g, AVG).

Deze regels zijn streng en, zoals we ook al zagen bij de bespreking van de doelbindingsvereisten, voorziet de wetgever dan al snel in uitzonderingen. Interessant is dat dit alleen in de AVG is gebeurd. In dezelfde bepaling als die voorziet in uitzonderingen op het doelbindingsvereiste is ook voorzien in de mogelijkheid om voorbij te gaan aan de beperkingen die de wet stelt met betrekking tot profilering (art. 23, eerste lid, AVG).[76]

Wat betekent dit voor het gebruik van big data analyses ten behoeve van de aanpak van zorgfraude? Wij denken wel wat, maar uiteindelijk misschien niet heel veel. Immers, er zal zonder twijfel sprake zijn van profilering in de zin van de wet als er gegevens worden verwerkt voor het opstellen van risicoprofielen, bedoeld om bijvoorbeeld de controle van zorgdeclaraties, te intensiveren waar dat het meest oplevert. En ook al leidt die profilering niet meteen tot besluiten met rechtsgevolgen voor de betrokkenen, dan nog zal er al snel sprake zijn van dat hem of haar anderszins in aanmerkelijke mate treft.

---

[75] Besluit van 1 september 2014 tot wijziging van het Besluit SUWI in verband met regels voor fraudeaanpak door gegevensuitwisselingen en het effectief gebruik van binnen de overheid bekend zijnde gegevens met inzet van SyRI, Stb. 2014, 320; zie daarover Zwenne & Schmidt 2016, p. 310 en 339-341.

[76] In het ontwerp wetsvoorstel voor de UwAVG, dat eind 2016 via www.internetconsultatie.nl beschikbaar werd gesteld, zien we dat de wetgever voornemens is van deze uitzonderingsmogelijkheid gebruik te maken, maar alleen voor de gevallen waarin er sprake is van geautomatiseerde individuele besluitvorming, anders dan op basis van profilering, die noodzakelijk is om te voldoen aan een wettelijke verplichting die op de verwerkingsverantwoordelijke rust of ter uitvoering van een taak van algemeen belang. Zie art. 30, eerste lid, UwAVG (ontwerp) en MvT (ontwerp), p. 53.

Om een dergelijke vorm van profilering dan mogelijk te maken, zal dan gebruik moeten worden gemaakt van een wettelijke regeling die de profilering mogelijk maakt en die voorziet in passende maatregelen ter bescherming van de rechten en vrijheden en gerechtvaardigde belangen van de betrokkene. En daarin moet dan ook uitdrukkelijk worden voorzien in mogelijkheden om gezondheidsgegevens, waarvan zoals gezegd al snel sprake zal zijn, te betrekken in de data-analyse. Echter, doordat de AVG ook voorziet in uitzonderingen, kan het zomaar zijn dat uiteindelijk de wet niet heel veel beperkingen oplegt met betrekking tot de profilering door middel van big data analyses ten behoeve van de aanpak van zorgfraude.

4.5 Transparantie

De privacy- en gegevensbeschermingswetgeving verlangt verder dat de betrokkenen worden geïnformeerd over de verwerking van de hen betreffende persoonsgegevens. Onder Wbp moeten betrokkenen worden geïnformeerd over de identiteit van degene die verantwoordelijk is voor de gegevensverwerking (zgn. 'verantwoordelijke') en de doeleinden waarvoor de gegevens worden verwerkt, alsmede 'nadere informatie voor zover dat gelet op de aard van de gegevens, de omstandigheden waaronder zij worden verkregen of het gebruik dat ervan wordt gemaakt, nodig is om tegenover de betrokkene een behoorlijke en zorgvuldige verwerking te waarborgen (art. 33, eerste t/m derde lid, art 34, eerste t/m derde lid, Wbp). In het geval de gegevens verkregen bij iemand anders dan de betrokkenen zelf worden verkregen, behoeft er evenwel niet te worden geïnformeerd, als dat onmogelijk blijkt of een onevenredige inspanning kost (art. 34, vierde lid, Wbp). Evenmin behoeft er te worden geïnformeerd als de vastlegging of verstrekking van de gegevens bij of krachtens wet is voorgeschreven. In dat geval moeten betrokkenen desgevraagd wel worden geïnformeerd over dat wettelijk voorschrift (art. 34, vijfde lid, Wbp).

Onder de AVG is de regeling niet heel anders, zij het wel dat betrokkenen wel veel uitgebreider en gedetailleerder moeten worden geïnformeerd. Zo moet ook mededeling worden gedaan van de contactgegevens van degene die verantwoordelijk is voor de verwerking, en over klachtmogelijkheden, ontvangers van de gegevens, eventuele doorgiftes van gegevens naar landen buiten de EU, etc. (art. 13, tweede lid, en 14, tweede lid, AVG).

Voor big data analyses is daarbij in het bijzonder van belang dat zowel de Wbp als straks de AVG ook verlangen dat, in het geval er sprake is van profilering, de betrokkenen ook mededeling wordt gedaan van 'de logica mee die ten grondslag ligt aan de geautomatiseerde verwerking van de hen betreffende gegevens' (art. 42, vierde lid, Wbp, art. 13, tweede lid, onder f, AVG, art. 14, tweede lid, onder g, AVG). Wat daaronder precies moet worden verstaan is nog niet heel duidelijk. Uit de parlementaire geschiedenis van de Wbp kan het voorbeeld worden ontleend van het geval aan een naam bepaalde gegevens worden toegevoegd dat is bepaald op basis van 'postcodesegmentatie of van statistisch onderzoek' (zeg 'verhoogd incasso-

risico'). In zo een geval kan de betrokkene erop aanspraak maken dat hem wordt medegedeeld welke methoden zijn gebruikt om daartoe te komen. Als het gaat om algemeen toegankelijke kennis geldt dit echter niet. Zo is algemeen bekend dat een levensverzekeraar het overlijdensrisico op kortere termijn van een tachtigjarige hoger inschat dan dat van een veertigjarige. De verzekeraar is niet gehouden daarover nadere uitleg te geven. Dat is anders als het gaat het om minder voor de hand liggende verbanden bijvoorbeeld op grond van eigen ongepubliceerde statistische recherches.[77]

Met een en ander wordt beoogd de betrokkene, over wie in het kader van de aanpak van zorgfraude met big data analyses gegevens worden verwerkt, in staat te stellen ten minste kennis te hebben van wat er over hem bekend is en wat daarvan de gevolgen kunnen zijn. Echter, evenals bij het doelbindingsvereiste (zie §3.2) voorziet de wet ook hier in betrekkelijk algemeen geformuleerde de uitzonderingen (art. 43, onder b t/m e, Wbp, art. 23, eerste lid, onder d, e, g t/m i, AVG). Het is onze indruk dat daarvan al snel gebruik kan worden gemaakt als het gaat om big data analyses en de aanpak van zorgfraude. Als het al niet gaat om de voorkoming, opsporing en vervolging van strafbare feiten dan kan er zomaar sprake zijn van gewichtige economische en financiële belangen van de staat en andere openbare lichamen of het toezicht op de naleving van wettelijke voorschriften die zijn gesteld ten behoeve van deze belangen. Ook zal er al snel sprake kunnen zijn van de bescherming van belangen of rechten en vrijheden van anderen.

5. AFSLUITING: IS HET GENOEG ALLEMAAL?

De vraag is of de privacy- en gegevensbeschermingswetgeving in voldoende mate in staat is de waarborgen te bieden die nodig zijn om de rechten van betrokkenen goed te beschermen. De beantwoording van deze vraag veronderstelt een raamwerk waarbinnen kan worden bepaald wat genoeg is, en wat niet. De wetgeving biedt enig houvast bij de constructie daarvan. Van de in de vorige paragraaf verkende regels zien we dat die voor het gezondheidsgegevens en profileren (resp. par. 4.1 en 4.3) erop neer komen, of in elk geval moeilijk te vermijden maken, dat er voor big data analyses ten behoeve van de aanpak van zorgfraude wordt voorzien in een wettelijke regeling. In die regeling moet dan ook zijn voorzien in passende waarborgen voor de bescherming van deze rechten. En passend betekent in dat verband dat in die regeling de wezenlijke inhoud van de grondrechten en fundamentele vrijheden onverlet wordt gelaten en in een democratische samenleving een noodzakelijke en evenredige maatregel (vgl. art 23, eerste lid, AVG). We komen dan vanzelf uit bij het EVRM, IVBPR en HvEU. Veel houvast kan dan worden gevonden in de rechtspraak van het Europees Hof voor de Rechten van de Mens en het Hof van Justitie van de Europese

[77] Kamerstukken I 1999/2000, 25892, nr. 92c, p. 13.

Unie.[78]

[78] Zwenne & Steenbruggen 2015.

LITERATUUR

Van Berkhout & Van Engers 2012

T van Berkhout & T van Engers 'Onderzoeksmethodologie voor informatiegestuurd sociaal toezicht', *WFR* 2012/824

Custers 2014

B.H. Custers, 'Risicogericht toezicht, profiling en Big Data', *Tijdschrift voor Toezicht* 2014/5

Van der Ende 2015

T.A.M. van der Ende, 'Zorgfraude: van handhaving en hoe systemen hun eigen ongelukken creëren', in: J.G. Sijmons e.a., *Op weg naar 10 jaar nieuw zorgstelsel: terug- en vooruitblik*, Preadvies VGR 2015, Den Haag: Sdu Uitgevers 2015, p. 174 e.v.

Hildebrandt (2012)

M. Hildebrandt 'The Dawn of a Critical Transparency Right for the Profiling Era' in J. Bus et al. (Eds.) *Digital Enlightenment Yearbook* 2012

Hildebrandt (2016)

M. Hildebrandt, Data-gestuurde intelligentie in het strafrecht, in: *Homo Digitalis*, NJV-preadvies 2016: Deventer 2016

Olsthoorn 2016

P. Olsthoorn, *Big Data voor Fraudebestrijding*, WRR: Den Haag 2016,

WRR 2016

WRR, *Big data in een vrije en veilige samenleving*, Amsterdam: Amsterdam University Press 2016

Zorgverzekeraars Nederland 2014

Zorgverzekeraars Nederland, Rapportage controle en fraude 2014

Zorgverzekeraars Nederland 2015

Zorgverzekeraars Nederland, Rapportage controle en fraude 2015

Zwenne en Steenbruggen 2016

G.J. Zwenne & W. Steenbruggen, 'Privacyvoorwaarden voor de iOverheid. Vuistregels voor wet- en regelgevers met betrekking tot overheidsinformatiesystemen', *Regelmaat* 2015/1, p. 19-36

Zwenne & Schmidt 2016

G-J. Zwenne & A.H.J. Schmidt, Wordt de homo digitalis bestuursrechtelijk beschermd? in: *Homo Digitalis*, NJV-preadvies 2016; Deventer 2016