

**Speech big data, Jacob Kohnstamm, Nationale Denktank  
Expertforum, 3 oktober 2014**

*Gesproken woord geldt*

Dames en heren,

Big data. Een van de meest besproken onderwerpen van dit moment als het gaat over privacy. Onnoemelijk veel conferenties, artikelen in zowel vak- als nieuwsmedia gaan over big data. Niet voor niets buigt de Nationale Denktank zich dit jaar over dit fenomeen. Overal worden big data enthousiast gepromoot en worden de kansen en de mogelijkheden in de schijnwerpers gezet.

En inderdaad, big data lijken prachtige kansen te bieden om uiteenlopende dreigingen en problemen het hoofd te bieden. Met behulp van big data is het waarschijnlijk mogelijk om epidemieën te voorspellen, verkeersstromen te regelen en onze energie-efficiency te verbeteren.

Maar toch vraag ik mijzelf af: zijn wij wel echt klaar voor dit nieuwe big data-tijdperk? En belangrijker, zijn wij voldoende toegerust om de problemen die met big data gepaard gaan, te adresseren?

Big data heten niet voor niets big: het gaat om het verzamelen en bewaren van zoveel mogelijk informatie uit almaar uitdijende databanken. Uit deze enorme hoeveelheid data worden vervolgens met behulp van onnavolgbare algoritmes allerlei nieuwe verbanden en onverwachte nieuwe informatie gedestilleerd. Dit kan inderdaad zeer waardevol zijn bij bijvoorbeeld het vroegtijdig ontdekken van georganiseerde misdaad of – zoals ik zojuist noemde – het voorspellen en vermijden van epidemieën.

Maar big data worden ook op individueel niveau gebruikt. En dat is waar

mijn zorgen zitten. Want daar waar het gaat om persoonsgegevens ontstaan de risico's voor de persoonlijke levenssfeer. Vraag is of de privacywaarborgen met de databanken meegroeien.

### **[Privacywetgeving]**

In onze privacywetgeving liggen principes verankerd die zeker ook in het licht van big data betekenisvol zijn. Dat is onlangs ook nog op internationaal niveau bevestigd in twee verschillende opinies, uitgebracht door privacytoezichthouders<sup>1</sup>, wetenschappers en internationale organisaties. De verwachting is dat de nationale privacytoezichthouders die binnenkort van over de hele wereld samenkomen voor de internationale privacyconferentie, hierover ook een resolutie zullen aannemen met als stelling: onze privacyprincipes hebben de tand des tijds goed doorstaan. En dan, gaat u rustig slapen.

Maar mijn angst is dat de principes in de privacywetgeving niet genoeg zullen zijn. Dat zij geen dijk zullen vormen die hoog genoeg is om ons te beschermen tegen de big data-tsunami.

De essentie van dataprotectie is, zoals een Amerikaanse deskundige die hier op bezoek was het eens uitdrukte, *surprise minimisation*. Zo min mogelijk verrassen door vast te houden aan de belangrijke privacyprincipes van doelbinding (gebruik data niet voor een ander doel dan waarvoor je deze hebt verzameld), dataminimalisatie (gebruik niet meer data dan noodzakelijk zijn voor het doel) en transparantie (wees transparant over de data die je verzamelt en gebruikt).

---

<sup>1</sup> In de recent aangenomen opinie van de zogeheten Berlijn Telecomgroep (de Internationale werkgroep voor gegevensbescherming en telecommunicatie) en de opinie van de Artikel 29 werkgroep waarin alle Europese privacytoezichthouders zijn verzameld.

De essentie van Big Data is juist *surprise maximalisation*: door massale hoeveelheden opgeslagen data te gebruiken en koppelen voor hele nieuwe doelen en zo nieuwe informatie en verbanden te ontdekken. Dit alles op basis van onnavolgbare wiskundige formules.

Om het maximale eruit te halen zullen zij die big data verzamelen en gebruiken worden verleid dit te doen door de zojuist genoemde basisprincipes van doelbinding, dataminimalisatie en transparantie te verwaarlozen en met voeten te treden.

### **[Profielen]**

Een ander beangstigend feit bij big data is dat de verzamelde statistische informatie, zelfs als deze volgens de regels van de kunst wordt geanonimiseerd, kan leiden tot zulke precieze/exacte resultaten dat het toch herleidbaar is tot een specifieke persoon. Dan gaat het in feite om re-identificatie. En dat is natuurlijk ook eigenlijk waar het de gebruikers van het fenomeen big data om te doen is.

Ik noem een voorbeeld. Stel, er is een licht kalende man, bijna 65 jaar oud, woonachtig in het centrum van Amsterdam. Hij reist elke dag met trein en metro op en neer naar Den Haag. Heeft geen auto, maar maakt soms gebruik van Green wheels. En pakt dan een auto die binnen een straal van een halve kilometer van een bepaalde locatie in het stadscentrum geparkeerd staat. Verder kijkt hij nauwelijks televisie. Maar checkt hij meerdere malen per dag een specifieke nieuwssite over internet en privacy.

Waarschijnlijk heeft ook u nu al een idee wie dat zou kunnen zijn?

Vorige week was in NRC te lezen welke digitale gegevens zoal over iemand digitaal circuleren. En wat voor beeld dat van die persoon kan opleveren. Bijna iedereen heeft tegenwoordig meerdere profielen. Op basis waarvan hij op een bepaalde manier benaderd en behandeld wordt. Anders dan anderen. Wat nu als het profiel niet blijkt te kloppen? De journalist van NRC voerde hiervan een treffend voorbeeld op.

Probleem is dat wij niet weten waar dat profiel op is gebaseerd, en hoe dat tot stand is gekomen, noch welke keuzes op basis daarvan worden gemaakt – bijvoorbeeld welke aanbiedingen wel en welke niet. Welke informatie ik wel, welke niet te zien krijg – alles op basis van mijn profiel.

Een van de belangrijkste kenmerken van onze democratische samenleving is dat een individu zich vrij moet kunnen ontwikkelen en ontplooiën. Dat mensen vrije keuze hebben om een eigen pad te bewandelen binnen de grenzen van onze rechtsorde.

Big data stelt precies dat in de waagschaal.

Volledige individuele ontplooiing en ontwikkeling is een illusie op het moment dat op basis van mijn profiel veel keuzes al vóór mij worden gemaakt in plaats van door mij. Zoals gezegd, zonder dat ik weet hoe het profiel eruit ziet, welke data hiervoor zijn gebruikt, laat staan dat ik overzie welke consequenties de profilering voor mij heeft.

En dit is mijn angstbeeld, deze vorm van wat ik 'digitale predestinatie' noem. Iets wat indruist tegen een van de meest fundamentele waarden in onze democratische samenleving, namelijk die van de vrije ontwikkeling en ontplooiing van het individu.

Het zal leiden tot een maatschappij waarin mensen worden geprofileerd en op een bepaalde manier behandeld op basis van hun profiel. Waar

publieke organen, overheden, je bijvoorbeeld met wantrouwen zullen benaderen omdat je op basis van je big dataprofiel valt in de categorie van mogelijke belastingfraudeurs. Waar je op basis van credit scoring veel duurder uit bent of bepaalde medische risico's niet meer kunt verzekeren en uiteindelijk tot ondermijning van solidariteit kan leiden.

### **[Voorbeelden risico's]**

Bekend is het voorbeeld van het Amerikaanse warenhuis Target dat op basis van bepaalde data - zoals de plotselinge aanschaf van ongeparfumeerde verzorgingsproducten - kortingsbonnen voor babyspullen aan een tienermeisje aanbood. Het bedrijf bleek eerder op de hoogte van de zwangerschap van het meisje dan haar over de kortingsbonnen klagende vader. En dit terwijl het meisje niet eens zwangerschap- of babyspullen had gekocht.

Maar van een andere orde is het voorbeeld van de familie Catalano die in eigen huis werd overvallen door een compleet SWAT-team en behandeld als terroristen. Naar later bleek omdat zij had gegoogeld op een snelkookpan voor het koken van linzen. Hij had zich die week ervoor op internet georiënteerd op een nieuwe rugzak. Die twee feiten bleken in de nadagen van de bomaanslagen tijdens de marathon in Boston genoeg voor een profiel van terrorist en een bijbehorende politie-inval.

Sommigen zullen misschien zeggen 'maar dat was een incident en zal niet zo snel aan de lopende band gebeuren'. Maar ik denk dat je dan onderschat wat de risico's en gevolgen van big data kunnen zijn.

### **[Risico's big data]**

Big data zal ons leiden naar een al voor ons geëffend – soms dus zelfs gevaarlijk - pad. Waarbij het moeilijk zal zijn om buiten de gebaande paden te treden. Big data zullen kunnen worden misbruikt door hen die op

zoek zijn naar macht ten koste van mensen die onder minder gelukkige omstandigheden leven.

### **[Toestemming]**

Als het gaat om big data wordt in het publieke debat vaak ten onrechte aangenomen dat individuen 'vrijwillig' hun persoonlijke gegevens afgeven. Dat mensen die data welbewust hebben gegeven en dus niet daarna moeten klagen en nadere informatie en uitleg moeten eisen over hoe en wanneer die data vervolgens worden gebruikt.

Als je je Facebook-pagina open stelt voor je 50.000 allerbeste vrienden is dat geen onzinnige gedachte.

Maar veel informatie wordt nu juist zonder actieve toestemming van het individu afgegeven. Zonder dat men eigenlijk een vrije keuze heeft. Het dataspoor dat wij achterlaten is in zekere zin vaak onvermijdelijk. De activiteiten die tot het dataspoor hebben geleid horen immers bij het normale leven. Telefoneren, reizen met het openbaar vervoer, geld pinnen of optimaal gebruik maken van de mogelijkheden die internet biedt. Mensen 'accepteren' het gebruik van hun persoonlijke data – vaak letterlijk door te klikken op de akkoordknop bij de algemene voorwaarden - omdat zij een bepaalde app nodig hebben of gebruik willen maken van een bepaalde dienst.

Zelfs al zou iedereen tijd en energie steken in het doorploegen van de vaak ellenlange algemene voorwaarden – sommige zelfs langer dan de toneelstukken van Shakespeare – dan nog zou het velen van ons waarschijnlijk niet lukken de juridische abracadabra te ontwarren. En werkelijk begrijpen wat er nu precies gebeurt met onze data.

Dit soort lange privacyvoorwaarden maken dat mensen door de bomen het bos niet meer zien.

Dames en heren, wij staan vandaag op een tweesprong: sluiten we onze ogen en laten we een deel van ons leven leiden op basis van de digitale predestinatie door big data? Of hijsen we de stormbal? Wat mij betreft is het tijd voor dat laatste.

### **[Vertrouwen]**

Wij weten allemaal dat vertrouwen, zowel in de private als in de publieke sector, een van de pilaren onder onze samenleving, economie en overheid, is. En ook weten we dat juist dat vertrouwen van burgers de laatste jaren forse klappen heeft gekregen. Ook in de sfeer van de bescherming van persoonsgegevens, onder meer door een gebrek aan transparantie. Zo bleken banken plannen te hebben om onze gegevens commercieel uit te buiten, bleek de grote sociale netwerksite Facebook data te verzamelen door ongevraagd psychologische experimenten met ons uit te voeren, was er het plan om onze pinggegevens door te verkopen en bleken de veiligheidsdiensten massale hoeveelheden gegevens van totaal onschuldige burgers te verzamelen. Deze acties hadden grote maatschappelijk ophef tot gevolg.

Transparantie van bedrijven en overheden is de eerste stap voor herstel van vertrouwen. Als mensen weten wie welke gegevens over hen verzamelt en met welk doel, kunnen zij daar zeggenschap over uitoefenen en zal hun vertrouwen groeien.

Maar zoals gezegd is er meer nodig dan het verhogen van transparantie en het hoog houden van de principes uit de privacywetgeving via effectief toezicht hierop.

Ik zou willen dat ik kon zeggen dat niemand zich zorgen hoeft te maken omdat de privacywetgeving met de daarin verankerde privacy-principes zoals doelbinding, dataminimalisatie en transparantie als dijken waken

over onze privacyrechten. Maar ik ben bang dat de privacytoezichthouders deze door mensenhanden veroorzaakte tsunami niet kunnen tegenhouden. En de datahonger niet binnen de sociaal geaccepteerde grenzen kunnen houden. Onze dijken zijn daarvoor simpelweg niet hoog en effectief genoeg.

Naast een vastberaden wetgever en een krachtige toezichthouder, hebben we een kritische massa nodig. Een die opstaat en zijn rechten die door big data worden bedreigd, opeist. Zodat mensen zich bewust zijn van het feit dat intrinsieke waarden in onze samenleving in het geding zijn door big data. Er moet een stevig maatschappelijk debat worden gevoerd over hoe wij de risico's en ongewenste gevolgen van big data effectief kunnen aanpakken.

Een debat waarin niet alleen de mooie toekomst van big data wordt geschetst, maar waarin ook de werkelijke risico's worden benoemd zoals het risico op discriminatie, ongelijke behandeling en het risico op machtsmisbruik. Een debat met als gedroomde uitkomst dat gezamenlijke eisen worden geformuleerd die zorgen voor een maatschappelijk verantwoorde toepassing van dit fenomeen.

Wij hebben daarvoor een vlugschrift 'Das Digital' nodig. Of een nieuwe film van Al Gore, 'an inconvenient truth 2.0'...

Want zonder dit, zonder een stevig maatschappelijk debat over big data, zullen wij, naar ik vrees, aan het kortste eind trekken.

Dank u.