

BIG DATA
EEN MODEKREET OF *THE NEXT BEST THING*?

*Chantal Bakermans**

Terwijl de Nederlandse overheid steeds meer data vrijgeeft in het kader van 'open data', met, onder meer, als doel om innovatie te stimuleren, waarschuwt onze privacy waakhond tegelijkertijd voor de nadelige gevolgen van het beschikbaar komen van steeds meer data. Hebben we hier te maken met de zoveelste privacy paradox?

Smart industry en big data

Tijdens de Hannover Messe, de grootste industriebeurs ter wereld, pleitte Minister Kamp van Economische Zaken er voor dat Europa meer werk moet maken van *smart industry*, oftewel de digitalisering van de industrie.[1] Dit om onze concurrentiepositie te versterken. Ook op Europees niveau wordt de (op)komst van smart technologieën en innovatieve big data toepassingen gestimuleerd. "*Privacy is essentieel, maar mag geen excuus zijn om niets met Big Data te doen*", aldus Neelie Kroes tijdens een ICT evenement in 2013.[2] Een jaar stelde de Europese Commissie 500 miljoen euro beschikbaar voor big data investeringen. Dit bovenop het bedrag van 2 miljard euro dat reeds door bedrijven was bijeengebracht.

Smart industry, *big data*, *Internet of things*, het zijn de nieuwste modekreten en ze hangen allemaal met elkaar samen. Het kenmerk van 'smart' initiatieven, zoals de smart phone, smart watch en smart meter, is dat ze niet enkel doen wat het tweede deel van het woord suggereert. Met de smart phone kan men naast bellen ook foto's en video's maken, Internetten, gamen en navigeren en de slimme meter verwerkt al lang niet meer enkel de meterstanden. Binnen een 'smart industry' draait het om slim en intensief gebruik maken van het Internet en ICT in combinatie met sensortechnologie. Het smart industry concept omarmt al deze en andere initiatieven op het gebied van innovatie, Internet of things, open- en big data.

Hoe zit het met de privacy?

Al die nieuwe technologieën en ontwikkelingen roepen op overheidsniveau tevens de vraag op of smart- en big data toepassingen geen negatieve invloed hebben op de persoonlijke levenssfeer van de burger. Hoewel Kroes een fervent voorstander is van innovatie, lijkt ook zij te worstelen met een innerlijke tegenstrijdigheid wat betreft (online) privacy. Zo sprak Kroes tijdens de Cyber security conferentie van 28 februari 2014: *"On the horizon stand new opportunities: big data, cloud computing, the Internet of things (…). People – including me – sometimes talk about our "digital rights". But I don't think that's quite right. These*

* **Chantal Bakermans**, Manager IP/IT bij Deloitte, Amsterdam.

1 http://www.emerce.nl/nieuws/kamp-europa-digitaliseer-industrie, 14 april 2015.

2 ICT 2013 Event – Session on Innovating by exploiting big and open data and digital content, 7 november 2013.

BIG DATA – EEN MODEKREET OF *THE NEXT BEST THING*?

are not digital rights, nor online rights: they are fundamental rights, and they apply just as much online as off. (…) New technology can enhance our humanity: it should not override our human rights."[3]

In een brief die Minister Kamp van Economische Zaken op 19 november 2014 aan het kabinet stuurde, werd een onderzoek aangekondigd waarbij wordt getoetst of de analyse en profilering van (klant)gegevens die steeds meer bedrijven inzetten, in strijd is met privacy wet- en regelgeving. Daarbij moet tevens worden gekeken naar de beveiligingsrisico's van de ontwikkelingen rondom big data. Een expertgroep met deelnemers uit de wetenschap, bij consumentenorganisaties, bedrijven en het maatschappelijk middenveld moet voor 2016 een visie geven die het kabinetsbeleid rondom big data gaat bepalen.

Het is duidelijk dat de overheid in een tweestrijd zit en niet goed weet voor welk anker zij nu moet gaan liggen. De vraag is echter of de genoemde onderzoeken daar enige verandering in kunnen en gaan brengen? Weten we immers niet allang wat het resultaat gaat zijn? Ons datagebruik en de ontwikkeling van nieuwe technologieën in dat verband gaan met sprongen vooruit. Het stopzetten van innovatie op grond van privacywetgeving, die toch nooit alomvattend zal zijn, zou wel eens kunnen betekenen dat we het kind met het badwater weggooien. Ik zou dan ook willen betogen dat de focus meer moet komen te liggen op de toepassing en randvoorwaarden van big data analyses.

Big data – wat is het?

"Big Data", om onbekende redenen veelal met hoofdletters geschreven. Dit duidt evenwel op een definitie en als er iets lastig is met big data, is het wel om het te definiëren. Althans om het eens te worden over een eenduidige definitie. In het vervolg van dit artikel zal ik dus gewoon "big data" gebruiken.

Wanneer wordt gesproken over big data gaat het veelal om datamaximalisatie. *Volume* (hoeveelheid), *velocity* (snelheid) en *variety* (verscheidenheid) maken data 'big'. Tegenwoordig wordt daaraan toegevoegd *veracity* (betrouwbaarheid). Volgens de jaarlijkse 'Digital Universe study' van IDC, zal de hoeveelheid data tot 2020 nog een tienvoud toenemen, tot ruim 40 zettabyte.[4] De meerderheid van die data zullen worden geproduceerd door machines die communiceren via data netwerken middels sensoren en smart apparaten en toepassingen. Big data kan van onschatbare waarde zijn voor onze economie. Door middel van big data toepassingen kunnen (bedrijfs) processen bijvoorbeeld efficiënter worden ingericht of kan een optimale beleidsstrategie op basis van analyses en voorspellingen al in vroegtijdig stadium worden bepaald. De Nationale Denktank 2014 schatte de potentiële waarde van big data op € 45 miljard euro.[5] Bedrijven in alle sectoren (zorg-, financiële- en mobiliteitssector) experimenteren er al vol op los en de vacatures voor (big) data scientists schieten bij bosjes uit de grond.

Zoals ook in de inleiding van dit artikel aan-

3 N. Kroes, "A secure online network for Europe", *Cyber security conferentie*, Brussel, 28 februari 2014, te vinden via: http://europa.eu/rapid/press-release_SPEECH-14-167_en.htm
4 http://www.computerweekly.com/news/2240217788/Data-set-to-grow-10-fold-by-2020-as-internet-of-things-takes-off
5 Eindrapport Nationale Denktank 2014.

gegeven, wordt innovatie gestimuleerd door de overheid. Het kabinet en een meerderheid in de Tweede Kamer zijn meer dan enthousiast over bedrijven als Uber en Airbnb en willen hier ook volop in investeren.[6] Onze informatiemaatschappij wordt steeds meer *data driven* en de wetten en regels zullen meer rekening moeten houden met de mogelijkheden die het Internet biedt. Volgens Minister Kamp kan de Nederlandse economie worden gestimuleerd door de wetgeving te vernieuwen. Daarbij zou het uitgangspunt moeten zijn: niet precies zeggen wat mag en niet mag, maar kaders schetsen en ondernemers binnen die kaders de ruimte geven. Een mooi streven, maar het blijft enigszins vaag. Ook op Europees niveau worden momenteel vooral de kaders geschetst en blijven de echte *do's* en *don'ts* achterwege. Kortom, tot die tijd, maar waarschijnlijk ook daarna, blijft het worstelen met de vraag hoe in het kader van big data moet worden omgegaan met privacygevoelige data.

Privacy risico's

De voordelen van big data op globaal, macro niveau zijn duidelijk: van het voorkomen van epidemieën, tot het optimaliseren van bedrijfsprocessen en zoveel meer. Veel big data analyses zijn mogelijk zonder daarbij persoonsgegevens te verwerken, maar het gaat natuurlijk om die toepassingen waarbij wél persoonsgegevens verwerkt worden. Het is evident dat daar – overigens net als bij andere verwerkingen van persoonsgegevens – privacy risico's aan kleven. Daar is geen nader onderzoek door de overheid voor nodig.

Jacob Kohnstamm, voorzitter van het College bescherming persoonsgegevens (Cbp), is vooral bang voor de gevolgen van big data op individueel niveau. Zo willen marketeers met big data juist inzicht in het (koop- en surf)gedrag van consumenten om gerichte (product)reclame aan te bieden. Een van de meest aangehaalde voorbeelden in dit verband is dat van de Amerikaanse supermarktketen Target. Op basis van data analyses constateerde Target dat zwangere vrouw ongeparfumeerde verzorgingsproducten (lotions en zeep) kochten. Vervolgens stuurde Target de vrouwen kortingsbonnen voor babyspullen. Op deze wijze ontdekte een vader dat zijn tienerdochter zwanger bleek te zijn. Het logo van Target is niet voor niets een *bulls-eye* die de consument duidelijk in het vizier houdt.

Toch klinkt het verhaal in eerste instantie ernstiger dan het is. Zo is dit voorbeeld lang niet meer zo beangstigend wanneer de tienerdochter niet enkel kortingsbonnen voor babyspullen had ontvangen, maar ook voor producten die zwangere vrouwen nooit zouden kopen. Bijvoorbeeld een barbecue naast luiers en wijnglazen naast een babyromper. Hoe dan ook, het Target voorbeeld omvat de voornaamste privacy risico's die in het kader van big data veelal worden opgeworpen:

- Een rechtmatige grondslag voor het verwerken van de persoonsgegevens ontbreekt;
- Er is sprake van datamaximalisatie in plaats van minimalisatie, waarbij het doelbindingsprincipe wordt losgelaten;
- Er worden (risico)profielen gemaakt, hetgeen kan leiden tot discriminatie.

6 Kamerstukken II 2014-2015, 32 761, nr. 78 Brief Minister Kamp d.d. 19-11-2014.

BIG DATA – EEN MODEKREET OF *THE NEXT BEST THING*?

Het bestaan van deze risico's kan moeilijk worden ontkend. Bij big data analyse geldt over het algemeen het principe van: hoe meer gegevens verzameld worden, des te beter de resultaten zijn. Dit betekent dat in eerste instantie data relatief onbeperkt verzameld worden en deze pas tijdens de analyse worden geselecteerd. Die onbegrensde verzameldrang brengt tevens met zich mee dat de doeleinden van big data van te voren niet altijd welbepaald en uitdrukkelijk omschreven zijn. Vanuit privacy oogpunt staat dit op gespannen voet met de door de Wbp voorgeschreven beginselen van doelbinding en dataminimalisatie.

Een ander punt van discussie is de rechtmatige grondslag voor de verwerking van persoonsgegevens bij big data toepassingen. Meestal wordt hiervoor verwezen naar de toestemming van de betrokkene, ofwel de gerechtvaardigde belangen van de verantwoordelijke. In de digitale wereld wordt het toestemmingsvereiste veelal ingevuld doordat voorafgaand aan het kunnen aanvragen of activeren van een bepaalde dienst, de algemene- en privacy voorwaarden moeten worden geaccepteerd. Dat hiermee de vereiste geïnformeerde, ondubbelzinnige toestemming is verkregen, lijkt evenwel een hypocriete gedachte. Het is immers geen geheim dat vrijwel niemand de privacy voorwaarden leest. De reden hiervoor is veelal dat ze te lang, te onduidelijk of te moeilijk zijn. Op de vraag: *Weet u wat er exact met uw gegevens (e-mail, adresboek, foto's) gebeurt als u deze met behulp van een online dienst (iCloud, Dropbox, SkyDrive) opslaat?*, antwoordde 86% van de ruim 1.000 ondervraagden in een onderzoek van Right Marktonderzoek in opdracht van Kaspersky Lab, "nee".[7]
Uit hetzelfde onderzoek volgt tevens dat de gemiddelde Nederlander best bereid is om een deel van zijn of haar privacy op te geven voor persoonlijk voordeel of gemak. Zo zou ruim 30% van de respondenten zijn of haar persoonsgegevens invullen voor het verkrijgen van extra korting bij een webshop en nog eens 36% staat hier neutraal tegenover.[8] Uit een onderzoek van PwC bleek zelfs dat de helft van de Nederlanders bereid is om persoonlijke gegevens beschikbaar te stellen voor een goedkopere polis en wereldwijd zou dit al ruim 65% zijn.[9]

Hieruit volgen in mijn optiek twee dingen. Ten eerste, toestemming vragen van de betrokkene is niet altijd toereikend om de privacy voldoende te waarborgen. Ten tweede, big data toepassingen, waarbij (ook) bepaalde profielen ontstaan, hebben niet per definitie een negatief effect op de persoonlijke levenssfeer. Dit is wellicht ook mede bepalend geweest bij de uiteindelijke vormgeving van de bepaling rondom profiling in de Europese privacy verordening. Uitgangspunt van die bepaling was in eerste instantie dat profiling niet is toegestaan, *tenzij*, bijvoorbeeld de toestemming van de betrokkene is verkregen

7 Dit onderzoek naar het bewustzijn onder de Nederlandse bevolking van het online beschikbaar stellen van privacygegevens, is uitgevoerd door Right Marktonderzoek, in opdracht van Kaspersky Lab. Het veldwerk vond plaats van 20 tot 27 januari 2015 en is gepubliceerd in februari 2015. Te raadplegen via: http://newsroom.kaspersky.eu/fileadmin/user_upload/nl/Downloads/PDFs/Kaspersky_Lab_Onderzoek_Online_privacy.pdf. ('Onderzoek Kaspersky Lab').

8 Onderzoek Kaspersky Lab.

9 Zie o.m.: http://www.dewaardevanadvies.nl/adviseur-van-de-toekomst-bewijst-wat-hij-waard-is/ en http://www.wegwijs.nl/artikel/2014/08/privacy-op-de-helling-bij-extra-korting-op-verzekering.

(opt-in). In de huidige tekst[10] van deze bepaling lijkt 'opt-out' evenwel meer het uitgangspunt en komt de opt-in pas om de hoek kijken wanneer profiling leidt tot: "*measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject.*" (artikel 20, lid 2). Dit is het kader, de invulling van de 'ernst van de inbreuk op de rechten en vrijheden van de betrokkene, wordt overgelaten aan de eigen beoordeling van het bedrijfsleven. Overweging 58 geeft ter illustratie enkel aan dat profiling op basis van pseudonieme data niet zal worden beschouwd als een inbreuk die dermate ernstig is dat de toestemming van de betrokkene benodigd is.

Voor de verdere invulling kan ook aansluiting worden gezocht bij de opinie omtrent doelbinding van de Artikel 29 Werkgroep ("Werkgroep") in het kader van de verdere verwerking van persoonsgegevens voor een ander doel dan het doel waarvoor de persoonsgegevens ooit zijn verkregen.[11] De Werkgroep bepaalde dat in voorkomend geval een beoordeling dient plaats te vinden van alle omstandigheden van het geval, in het bijzonder: i) de relatie tussen de doeleinden van de verwerkingen, ii) de context waarin de persoonsgegevens zijn verkregen en de verwachtingen die het individu heeft wat betreft verdere verwerkingen voor een ander doel, iii) het soort persoonsgegevens en de impact van de verdere verwerking op de persoonlijke levenssfeer van het individu, en iv) de beveiligingsmaatregelen die de verantwoordelijke heeft getroffen ter bescherming van de persoonsgegevens tegen verlies, ongeautoriseerde toegang of misbruik.

Kort en goed, de verwerking van persoonsgegevens in het kader van big data analyses kan risico's voor de privacy met zich meebrengen. Mede gelet op het handelen en de inzichten van de gemiddelde consument, vraag ik mijzelf echter af of het hardnekkig proberen te voldoen aan de privacy wet- en regelgeving ook daadwerkelijk het gewenste, acceptabele niveau van privacybescherming kan waarborgen. Ik meen dat de focus meer zou moeten liggen op het zoeken naar een werkbare oplossing voor alle partijen in plaats van vasthouden aan de stelling dat big data niet in overeenstemming is met onze wet- en regelgeving op het gebied van privacy.

Toekomst

De huidige nationale en Europese privacy regels, alsook de voorlopige tekst van de Europese privacy verordening, schetsen vooral de kaders die een verantwoordelijke in acht moet nemen bij het verwerken van persoonsgegevens. De invulling binnen die kaders is aan de verantwoordelijke zelf.

In de literatuur en de praktijk zijn de eerste (suggesties voor) oplossingen voor het big data versus privacy debat zichtbaar. Zo stelde hoogleraar Lokke Moerel in haar inauguratierede[12] bijvoorbeeld de '*harm*

10 European Parliament legislative resolution of 12 March 2014 on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) (COM(2012)0011 – C7-0025/2012 – 2012/0011(COD)).

11 Opinion 03/2013 (WP 203) on purpose limitation.

12 Oratie Prof. dr. mr. L. Moerel "Big Data Protection, How to Make the Draft EU Regulation on Data Protection Future Proof", 14 februari 2014, Tilburg University.

BIG DATA – EEN MODEKREET OF *THE NEXT BEST THING*?

Bron: http://actonline.org/projects/privacy-dashboard/

Bron: http://www.aiga.org/resources/app-design-mobile-privacy-laws/

based' aanpak voor. Het vertrekpunt is dat persoonsgegevens mogen worden verwerkt indien het recht op privacy van het individu niet zwaarder weegt dan de rechtmatige belangen van de verantwoordelijke. Die belangenafweging zou 'harm based' moeten zijn, waarbij kortgezegd pas toestemming wordt gevraagd indien een verwerking schadelijk is voor consumenten en de nadruk komt te liggen op de 'accountability' van degene die big data analyse toepast. Critici vragen zich evenwel af of het bedrijfsleven een sterke eigen verantwoordelijkheid wel aankan.[13]

De Europese toezichthouder voor gegevensbescherming (oftewel *European Data Protection Supervisor*) is ietwat gematigder en denkt dat onder meer het creëren van bewustwording bij alle betrokken partijen van belang is om te komen tot goede zogenaamde privacy-enhancing technologieën, alsook betere begeleiding bij de toepassing van privacy- en consumentenbeschermingsregels op online diensten, in het bijzonder wanneer het 'gratis' diensten betreffen.[14] In het buitenland ontwikkelde The App Association al het *privacy dashboard* voor smart phones en tablets, waarbij op eenvoudige wijze wordt weergegeven hoe een applicatie met je data omgaat. In de Privacy & Practice van 1 februari jl. werd hier ook al aan gerefereerd.[15]

Andere voorbeelden zijn het IRMA (I Reveal My Attributes)-pasje dat is ontwikkeld in samenwerking met de Radboud Universiteit Nijmegen en het online Qiy domein, waarbij de consument met een Qiy domein zelf bepaalt met wie de persoonsgegevens gedeeld mogen worden.

Al deze voorstellen en voorbeelden zijn een stap in de goede richting, maar dé oplossing lijkt er vooralsnog niet te zijn. Idealiter, moet worden gewerkt naar een situatie waarin het bedrijfsleven de druk van het 'accountability'-principe erkent en accepteert en het vertrouwen

13 O.a. Prof. Arno R. Lodder, Big data en privacy in: Financieel Dagblad, 21 januari 2015; C. J. Hoofnagle, The Potemkinism of Privacy Pragmatism, http://www.slate.com/articles/technology/future_tense/2014/09/data_use_regulation_the_libertarian_push_behind_a_new_take_on_privacy.html

14 Opinie van de EDPS, Privacy and competitiveness in the age of big data: the interplay between data protection, competition la wand consumer protection in the Digital Economy, 10 maart 2014.

15 J. Siljee, E. de Leede, "Privacy en Transparantie als Unique Selling Points".

BIG DATA – EEN MODEKREET OF *THE NEXT BEST THING*?

van de consument als onderscheidingsmiddel gaat zien. De consument is niet zonder meer tegen big data. Sterker nog, men is zelfs bereid om een deel van de privacy op te geven als daar een vorm van persoonlijk voordeel tegenover staat. Aan de andere kant zou diezelfde consument wel iets meer inzicht willen hebben in wat er wel en niet gebeurd met zijn of haar gegevens.

In de eerste plaats zou het uitgangspunt moeten zijn dat de gegevens in het kader van big data toepassingen zoveel mogelijk gepseudonimiseerd of geanonimiseerd worden (Privacy-Preserving Data Mining). Op die manier wordt het onmogelijk, althans moeilijker om gegevens terug te herleiden naar een individu. In de tweede plaats is een goede *data governance* van belang, waarbij kwaliteit van data, rollen en verantwoordelijkheden, consistente interne beleidsdocumenten en processen enzovoorts aan de orde komen. Tot slot dienen er heldere, overzichtelijke privacy verklaringen te komen. Transparantie is essentieel, maar dan wel in duidelijk leesbare vorm in plaats van juridische epistels van meer dan tien pagina's. Het gebruik van een gelaagde structuur en pictogrammen kan daarbij helpen. Daarbij moet 'zeggen wat je doet en doen wat je zegt' het uitgangspunt zijn. Handelt een bedrijf niet conform 'wat ie zegt', dan dienen er in de derde plaats, effectieve en controleerbare opt-out(s) te worden gehanteerd. De betrokkene moet kunnen aangeven dat hij of zij niet langer wil dat zijn of haar persoonsgegevens worden verwerkt. Om dit werkbaar te maken, dient de consument aldus ook kritisch te zijn richting bedrijven. Deze (ideale) eigen verantwoordelijkheid voor het bedrijfsleven, kan enkel bestaan wanneer de consument bereid is om bepaalde diensten niet meer af te nemen wanneer het bedrijf zich niet aan de eigen, mooie woorden houdt, of anderszins de privacybelangen van consumenten onevenredig schaadt. Pas dan is vertrouwen de sleutel en *enabler* voor effectieve en werkbare privacy waarborging. Tot die tijd, zou de onlangs uitgebreide boetebevoegdheid voor het Cbp die op korte termijn van kracht zal worden, net dat zetje in de rug kunnen zijn voor het bedrijfsleven om alvast wat aanpassingen en maatregelen te treffen en voor de consument om ook eens kritisch te zijn op bedrijven die worden beboet en zijn of haar eigen online gedrag daarop aan te passen.