



De Clercq Advocaten Notariaat

Oplossingen voor uw
vraagstukken van morgen

White Paper - Big Data

Januari 2017, Natascha van Duuren



Advocaten en notariaat
in Leiden en Den Haag

www.declercq.com

Inleiding

De meeste economen sluiten de rijen: data is het goud van de 21^{ste} eeuw.¹ Bedrijven kunnen op vele manieren geld verdienen aan het gebruik van grote hoeveelheden gegevens en de analyse daarvan. Een goed voorbeeld betreft *behavioural targeting*. Door klanten te benaderen met een product waarvan bekend is dat zij daar interesse in hebben, is de kans groter dat de klant daadwerkelijk het product koopt. Er zijn legio definities van het begrip big data. De klassieke omschrijving luidt dat sprake is van big data wanneer de dataset te groot is om met reguliere datamanagementsystemen te onderhouden en te gebruiken. Uit deze definitie volgt dat big data een relatief begrip is. Volgens de Wet van Moore verdubbelt het aantal transistors in computerprocessors iedere twee jaar. Deze voorspelling kwam uit van 1965 tot 2011. Omdat het aantal transistors van invloed is op de rekensnelheid van de processor, heeft deze ontwikkeling tevens invloed op het begrip big data. Ieder jaar moet een dataset groter zijn dan het vorige jaar, wil het (nog steeds) kwalificeren als big data.

Gebruik

Het meest bekende gebruik van big data is (vooralsnog) gericht op marketingdoeleinden. Winkelketen Target haalde al in 2012 het nieuws door het gebruik van big data. Een boze vader had Target erop aangesproken dat zijn dochter coupons ontving voor zwangerschap gerelateerde artikelen, terwijl zijn dochter helemaal niet zwanger was. Enkele dagen later bleek dat hij niet geheel op de hoogte was van hetgeen er in zijn huis afspeelde. Kortom: Target was eerder op de hoogte van een zwangerschap dan de aanstaande opa. Target was hiertoe in staat door de combinatie van grote hoeveelheden klanteninformatie. Zo kopen zwangere vrouwen in de eerste twintig weken van hun zwangerschap geurloze zalf en calcium-, magnesium- en zinksupplementen.

Het zou naïef zijn te denken dat big data slechts voor relatief onschuldige activiteiten als marketing wordt gebruikt. De op twee na hoogst gewaardeerde tech-startup ter wereld (na Uber en Airbnb) is Palantir. Over Palantir is verder weinig bekend, behalve dat het bedrijf zich specialiseert in data-analyse en onder andere de CIA, NSA en FBI tot haar klantenkring mag rekenen.

Risico's

Wanneer naar de risico's van het gebruik van big data wordt gekeken blijkt al snel dat de vergelijking met goud scheef is. De vergelijking met olie is meer op zijn plaats.² Het grootschalig gebruik van big data heeft als potentieel risico dat de privacy voorgoed verloren gaat. Veel mensen realiseren zich op het moment dat zij data delen niet welk (toekomstig) gebruik hiervan zal worden gemaakt. Een goed voorbeeld hiervan is het gebruik van gratis diensten, waarvoor veelal geldt: "If you're not paying for it, you're the product". Het verdienmodel van veel gratis diensten zoals Gmail bestaat uit het vergaren van data, waar vervolgens aan wordt verdiend. Mocht een gebruiker later spijt krijgen van het verstrekken van de gegevens, dan is het schier onmogelijk om deze data uit alle databases verwijderd te krijgen.

¹ Zie bijvoorbeeld <http://zembra.vara.nl/dossier/uitzending/data-het-nieuwe-goud>.

² Zie bijvoorbeeld <http://fd.nl/morgen/1153166/big-data-is-geen-goud-maar-olie>.

Voor big data-bedrijven geldt, net als voor oliemaatschappijen, dat de maatschappelijke schade die zij aanrichten (in dit geval in de vorm van het verlies van privacy) niet door henzelf gedragen wordt. Het beschermen van de privacy is daarom niet alleen een individueel belang, maar vooral ook een maatschappelijk belang. Voor het functioneren van een democratie is privacy immers essentieel. Mensen moeten zelf een mening kunnen vormen, ook meningen die tegen de stroom ingaan. Grootschalig gebruik van big data-analyses kan ervoor zorgen dat het bestaan van privacy wordt bedreigd.

Privacy

Veel gebruikers van big data menen dat de privacywetgeving niet op hen van toepassing is. Zij menen dat hun big dataset geen persoonsgegevens bevat, omdat de data geaggregeerd is en daardoor niet langer herleidbaar tot een individueel persoon. In sommige gevallen is dit juist, maar het risico van identificatie ligt bij een grote dataset altijd op de loer. Veel big datasets worden gecreëerd door het opkopen van andere datasets, wat het risico creëert dat geanonimiseerde gegevens plots weer identificeerbaar zijn (de zogenaamde re-identificatie).

Hiervan kan bijvoorbeeld sprake zijn indien data wordt toegevoegd die door middel van een cookie-id kan worden gekoppeld aan bestaande gegevens, en daardoor identificeerbaar is. In zijn algemeenheid geldt dat, hoe meer data wordt verzameld, hoe meer gegevens kwalificeren als persoonsgegevens. Met meer gegevens is het immers gemakkelijker een individu te identificeren. De toepasselijkheid van de privacywetgeving ligt dan ook voortdurend op de loer.

Door de toepassing van geavanceerde anonimiserings- of pseudonimiseringstechnieken kunnen de persoonsgegevens uit de big data-collecties worden verwijderd of vervangen door een niet-identificerend gegeven. Echter, de waarde creatie door middel van big data is veelal gebaseerd op de mogelijkheid om een voorspelling te kunnen doen over het gedrag van of de risico's voor een specifieke persoon. Bij een geheel geanonimiseerd of gepseudonimiseerd proces, kan deze waardecreatie niet plaatsvinden, en verliest de big data een groot gedeelte van zijn waarde. Kortom, het gebruik van big data zonder gevolgen voor de privacy is geen reële mogelijkheid, omdat de waardecreatie in dat geval niet of nauwelijks kan plaatsvinden.

Strengere regulering

De Europese beleidsbepalers hebben het risico van innovaties zoals big data ingezien en meegenomen bij de modernisering van het gegevensbeschermingsrecht in de EU. In de Algemene Verordening Gegevensbescherming ('AVG') is het recht opgenomen om niet aan geautomatiseerde individuele besluitvorming, waaronder profilering, te worden onderworpen. Daarnaast is in de AVG het 'recht om vergeten te worden' opgenomen. Met een beroep hierop zou het mogelijk zijn gegevens die op deze persoon slaan te laten verwijderen uit de big data collectie.

Zowel het recht om niet aan profilering te worden onderworpen als het recht om vergeten te worden hebben één manco, wat ervoor zorgt dat de gevolgen van deze 'nieuwe' rechten voor gebruikers van big data waarschijnlijk beperkt zullen blijven: de toepassing is afhankelijk van de wil van de 'betrokkene' en de actie die de betrokkene hierop neemt.³ Het recht om niet aan profilering te onderworpen geldt immers niet voor profileringen waarvoor de betrokkene zijn toestemming heeft gegeven. Het recht om vergeten te worden vindt alleen toepassing indien de betrokkene er een beroep op doet. Kortom, de strengere regulering van big data toepassingen is afhankelijk van de waarde die de betrokkene aan zijn privacy hecht. Wanneer de vergelijking wordt getrokken met de gebruiksvoorwaarden van apps en software, waarbij de gebruikers (vrijwel) altijd toestemming gegeven voor de vermindering van hun privacy zonder de voorwaarden te lezen, dan ligt het niet voor de hand dat het gebruik van big data toepassingen op korte termijn hinder zal ondervinden van betrokkenen die gebruik maken van hun recht op privacy.

Conclusie

Big data biedt voor zowel bedrijven als overheden prachtige kansen om toegevoegde waarde te creëren door het gebruik van beschikbare data. Echter, voor de privacygevolgen van het gebruik van big data toepassing is tot op heden nog weinig aandacht.

Jacob Kohnstamm, toenmalig voorzitter van het CBP (tegenwoordig de Autoriteit Persoonsgegevens) verwoordde dit mooi in 2014: *"Big data lijken prachtige kansen te bieden om uiteenlopende dreigingen en problemen het hoofd te bieden. [...] 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."*⁴

Kohnstamm gaf daarnaast aan dat zijn 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."*⁵ Met de Algemene Verordening Gegevensbescherming is een privacykader gecreëerd voor (ten minste) het komende decennium. Hieruit blijkt dat Kohnstamm op dit punt gelijk lijkt te krijgen. De mogelijkheden voor de bescherming tegen de 'big data-tsunami' zijn in het leven geroepen, maar zijn allen afhankelijk van (pro-actief) handelen door de betrokkene. De praktijk heeft echter geleerd dat de gemiddelde internetgebruiker nonchalant omgaat met het achterlaten van persoonsgegevens en het accepteren van privacyvoorwaarden. Het is dus maar zeer de vraag of gebruikers van big data toepassingen op korte termijn in hun mogelijkheden zullen worden beperkt.

³ De betrokkene is de persoon op wie de persoonsgegevens betrekking hebben.

⁴ Speech big data, Jacob Kohnstamm, Nationale Denktank Expertforum, 3 oktober 2014, <
https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/speech_big_data_nationale_denktank_versie_3_okt_2014_website.pdf>

⁵ Idem.

Overwegingen

- De verwachting is dat big data de komende jaren een enorme vlucht zal nemen. De voordelen voor bedrijven zijn evident: zij kunnen hun doelgroep beter in kaart brengen en gericht benaderen. Ook individuen kunnen profiteren van big data, met name door het gebruik van big data in het onderwijs en de zorg.
- De keerzijde is dat big data de privacy in ernstige mate kan schaden. Veel mensen beseffen zich niet dat hun gegevens worden doorverkocht aan andere bedrijven. Dit met alle gevolgen van dien. Het is dan ook zaak dat iedereen zich ervan bewust is dat je op Internet digitale sporen achterlaat en dat er vanuit de overheid meer voorlichting komt over de mogelijkheden om digitale sporen te beheren.
- In de Algemene Verordening Gegevensbescherming zijn rechten opgenomen waarmee betrokkenen hun privacy kunnen waarborgen en kunnen voorkomen dat zij aan big data toepassingen worden onderworpen. Deze rechten van de betrokkene zijn echter afhankelijk van pro-actief handelen van de betrokkene. Zolang betrokkenen zelf geen actie ondernemen hebben gebruikers van big data toepassingen vrij spel.



Advocaten en notariaat
in Leiden en Den Haag

Hoge Rijndijk 306, 2314 AM Leiden
WTC The Hague, Prinses Margrietplantsoen 33, 2595 AM Den Haag
www.declercq.com, info@declercq.com, 071-5815300