Iedereen wil uiteindelijk die voorspellende glazen bol

Laat Big Data de maatschappelijke veiligheid vergroten

November 2015

kpmg.nl

Big Data is als tandpasta: als het eenmaal uit de tube is, kun je het nooit meer terugduwen. De publieke sector reageert met iets meer voorzichtigheid dan het bedrijfsleven en dat is begrijpelijk vanuit zijn maatschappelijke verantwoordelijkheid. Niettemin is het zaak dat ook overheidsorganisaties het potentieel van Big Data niet onbenut laten. Dit is een verkennende analyse van zowel het potentieel als de stand van zaken in het veiligheidsdomein.

People are more predictable than particles

Impossible is nothing. Althans, daar lijkt het op zodra we praten over de toepassingsmogelijkheden van Big Data. Er zijn scenario's denkbaar waarin we met data-analyse misdaden voorspellen (predictive policing). Dat roept natuurlijk direct de vraag op of dat maatschappelijk gewenst is. Feit is dat diverse Amerikaanse steden al enige jaren succes boeken met 'voorspellend surveilleren'. Op basis van analyse van grote hoeveelheden data uit diverse bronnen bepalen algoritmes welke straten extra aandacht verdienen van de agenten. Het effect: meer arrestaties, maar vooral ook lagere misdaadcijfers. Er is dan ook geen twijfel over mogelijk: de voorspellende mogelijkheden van data-analyse zijn groot. Uit diverse (praktijk)onderzoeken blijkt dat het menselijk gedrag akelig nauwkeurig te voorspellen is, veel beter dan menigeen denkt. Voor wiskundigen is dat nauwelijks een verrassing. Stephen Wolfram, een bekende wis- en natuurkundige die onder andere een volgende generatie zoekmachine ontwikkelde (Wolfram Alpha), stelde al eens treffend vast: *'People are more predictable than particles.'*

Die vaststelling is de basis voor veel mogelijkheden om met data-analyse een bijdrage te leveren aan meer veiligheid in de samenleving. Daarbij moeten we niet alleen de exotische voorbeelden uit de film Minority Report op het netvlies houden, maar (vooral) ook denken aan meer *down to earth* mogelijkheden om bijvoorbeeld grote menigten in goede banen te leiden

en een betere risico-analyse door de brandweer. Een kort overzicht van dergelijke toepassingsgebieden volgt later. Eerst even een korte blik op waarom Big Data relevant is.

Explosieve groei van data

De relevantie van Big Data valt niet los te zien van de explosie van de hoeveelheid data, veroorzaakt doordat mensen en vrijwel alle denkbare apparaten 24/7 via internet met elkaar verbonden zijn. Dat netwerk van verbindingen en sensoren (Internet of Things) levert een fenomenale hoeveelheid data op. Een veelzeggende indicator voor die ontwikkeling is het feit dat we de afgelopen twee jaar al meer data hebben geproduceerd dan in alle eeuwen daarvoor. Dat biedt fascinerende nieuwe mogelijkheden, die vaak onder de noemer Big Data worden samengevat, zoals ook in het boek Wij zijn Big Data van Sander Klous en Nart Wielaard[1]. Overigens is het woord 'big' in Big Data eigenlijk wat verwarrend. Want allerlei nieuwe toepassingen gaan helemaal niet om het bewerken of interpreteren van enorme hoeveelheden data maar veel meer om het *slim combineren* van data.

Het potentieel

Hoe ziet het potentieel van Big Data eruit en wordt het al benut? De OESO becijferde in 2014 dat de toepassing van Big Data kan zorgen voor een stijging in productiviteit van 5 tot 10 procent.[2] Recent onderzoek van KPMG laat zien dat het merendeel van grotere bedrijven data-analyse dan ook al toepast binnen (onderdelen van) hun organisaties. Hetzelfde onderzoek laat echter ook zien dat er nog een wereld te winnen is.[3] Meer specifiek voor de overheid kunnen we verwijzen naar een rapport van McKinsey uit 2011 dat liet zien dat de potentiële toegevoegde waarde voor het toepassen van Big Data in de overheid zeer groot is, maar dat het verkrijgen van deze toegevoegde waarde niet zo gemakkelijk is.[4] De (Nederlandse) overheid lijkt zich in elk geval bewust van de potentie en besteedt er op verschillende wijzen aandacht en geld aan. Zo presenteerde de Europese Commissie in juli 2014 haar mededeling 'Naar een bloeiende data-economie' over onder andere het opzetten van een Europees netwerk van kenniscentra om het aantal data-specialisten met de nodige vaardigheden te vergroten. Ook de Nederlandse overheid richt zich op het potentieel van Big Data en stelt bijvoorbeeld de komende twee jaar EUR 40 miljoen beschikbaar voor ICT- onderzoek en innovatie. Het op te richten consortium van bedrijven, kennisinstellingen en overheden gaat zich richten op de toepassing van Big Data op het gebied van energie, zorg, veiligheid en smart industry. Tevens wil het consortium meer ICT-specialisten opleiden op het gebied van cybersecurity en Big Data. Hiernaast is in de innovatieagenda van het ministerie van Veiligheid en Justitie een prominente plek ingeruimd voor Big Data. Ook regionaal worden er diverse initiatieven gestart om met data te innoveren, zoals de smart cities waar Big Data en met name Open Data een belangrijke rol in spelen. De Veiligheidsregio Groningen neemt deel aan een initiatief op het gebied van The Internet of Things door al de (brandweer)locaties beschikbaar te stellen voor plaatsing van een Gateway zodat er een verbinding kan worden gemaakt tussen het netwerk van de brandweer en het internet.[5]

Verkenning van toepassingen

Op basis van gesprekken met vertegenwoordigers van verschillende uitvoeringsorganisaties[6], beschrijven we hierna de toepassingen en mogelijkheden van Big Data in het veiligheidsdomein in Nederland. Het gaat hierbij om een eerste verkenning.

We stellen daarbij vast dat de toepassingen in de praktijk nog beperkt zijn, een constatering die aansluit bij een achtergrondstudie van het Rathenau Instituut die stelde dat "de verwachtingen over wat er mogelijk is vooruitlopen op wat er daadwerkelijk gebeurt. Ondanks de torenhoge verwachtingen is het toepassen van Big Data in de praktijk dus niet vanzelfsprekend."[7]

De gesprekken laten zien dat Big Data op zeer verschillende wijzen wordt ingezet: van real-time inzichten tijdens incidenten tot het voorspellen van risico's. Grofweg is er een tweedeling zichtbaar: enerzijds gaat het om het analyseren en verwerken van data en anderzijds het komen tot voorspellingen. Uiteraard staan deze toepassingen niet los van elkaar: wanneer data goed verwerkt en geanalyseerd wordt is het mogelijk om op basis daarvan te komen tot voorspellingen over gedrag, risico's en misdaad.

1 Klous & Wielaard, Wij zijn Big Data – De toekomst van de informatiesamenleving, 2015
2 OESO, Data-driven innovation for growth and wellbeing, 2014
3 KPMG, Going beyond the data: turning data from insights into value, 2015
4 McKinsey, Big data: The next frontier for innovation, competition, and productivity, 2011
5 http://www.rbenw.nl/veiligheidsregio-groningen-neemt-deel-aan-ontwikkeling-van-internet-of-things/
6 Er is gesproken met vertegenwoordigers van de Politie, het Nationaal Cyber Security Centrum (NCSC), het Nederlands Forensisch Instituut (NFI), de Militaire Inlichtingen- en Veiligheidsdienst (MIVD) en de Veiligheidsregio Haaglanden.
7 Rathenau Instituut, De datagedreven samenleving – Achtergrondstudie, 2015

Analyseren en verwerken van data

Zoals eerder beschreven gaat het bij Big Data vooral om het slim combineren van bestaande data. De meeste organisaties, zeker in het veiligheidsdomein, hebben beschikking over grote hoeveelheden data. Wanneer deze data (aan elkaar gekoppeld) geanalyseerd wordt, kan dit leiden tot een optimalisatie van de werkprocessen. We onderscheiden daarbij drie toepassingen:

- Verwerken van grote hoeveelheden data
- Verbeteren van inzicht door koppelen van (open) data
- Optimalisatie inzet personeel en materieel

- Verwerken van grote hoeveelheden data

  Het NFI ontwikkelde in nauwe samenwerking met de Politie het Hansken-systeem om de alsmaar groeiende hoeveelheid digitale sporen bij strafzaken snel en makkelijk doorzoekbaar te blijven maken. Het systeem ordent digitale informatie zodat rechercheurs sneller relevante informatie uit de stroom van digitale data kunnen halen, waardoor veel tijd wordt bespaard. De Politie is de eerste partner in de veiligheidsketen die met Hansken gaat werken. De verwachting is dat ook andere partners volgen.[8]

- Optimalisatie inzet materieel en personeel

  Data-analyse draagt bij aan het optimaliseren van de inzet van personeel en materieel. Hierbij gaat het bijvoorbeeld om de planning van onderhoud van materieel en de planning van opleidingen voor personeel in relatie tot de inzetbaarheid. Hier kan bijvoorbeeld ook de spreiding van personeel en voertuigen op worden afgestemd. Als deze data vervolgens gecombineerd wordt met data over incidenten zal dit naar verwachting leiden tot een nog hogere effectiviteit omdat er genoeg, maar niet te veel personeel en materieel beschikbaar zijn om het aantal incidenten aan te kunnen.

> **ZENO GERADTS,** NFI:
>
> Een van de volgende uitdagingen is het daadwerkelijk combineren van beelddata en digitale data

- Verbeteren van inzicht door koppelen van (open) data

  Real-time inzicht op het juiste moment op de juiste plaats heeft veel waarde bij de bestrijding van incidenten. Het draagt bij aan een betere inschatting van situaties en leidt daarmee tot hogere veiligheid en effectievere en efficiëntere bestrijding. De inzet van real-time intelligence centers (RTIC) in de meldkamers zijn een start van deze ontwikkeling. In het proces van opsporing verschaffen 'open source' bronnen additioneel inzicht.

Naar het volgende niveau: komen tot voorspellingen

Op basis van goed verwerkte en geanalyseerde data is het mogelijk

> **RUTGER VAN DEN HURK,** Eenheid Noord-Holland, Politie:
>
> Data is van belang op zaaksniveau en bij het analyseren van fenomenen en trends. Iedereen wil uiteindelijk die voorspellende glazen bol

8 https://www.forensischinstituut.nl/over_het_nfi/nieuws/2015/nfi-bouwt-krachtige-forensische-zoekmachine-voor-digitaal-onderzoek.aspx?cp=34&cs=578

om voorspellingen te doen over gedrag, risico's en misdaad. Dit leidt weer tot betere inschatting van de benodigde (veiligheids)maatregelen, betere planning van personeel en materieel en uiteindelijk effectiever beleid. Hieronder noemen we drie mogelijkheden:

- Gedrag voorspellen
  Gedrag kan voorspeld worden op verschillende niveaus, op het niveau van de groep (crowd control), of op het niveau van het individu (opsporing) en begint met het aan elkaar koppelen van verschillende databronnen. Op dit moment gebeurt dit vooral in gebouwen of op afgebakende plekken, zoals de Amsterdam Arena. Een voorbeeld vormde ook de inhuldiging van koning Willem-Alexander: de Politie deed aan crowd control op basis van druktemeters in de vorm van wifi-spots, waarmee ze een nauwkeurig 'reallife'-beeld kregen van mensenmassa's en -stromen. Zo bepaalde men de politie-inzet en bleef men op de hoogte van drukke en minder drukke plekken in de stad.

- Risico's voorspellen
  De brandweer koppelt verschillende databronnen en ontdekt daarmee verbanden tussen preventie, incidenten en (bijvoorbeeld) omgeving. Daardoor wordt betere risicobeheersing mogelijk waardoor incidenten kunnen worden voorkomen. Dat levert tal van vragen op: Waarom zijn er in gemeente A meer incidenten dan in gemeente B? Bestaan er correlaties tussen de uitgifte van vergunningen en incidenten? Kan er op basis van het koppelen van beide soorten data een betere risico-inschatting gedaan worden? De verwachting is dat de koppeling van databronnen (een deel van) de antwoorden gaat opleveren. In internationaal verband kunnen we hier wijzen op FireCast, een algoritme dat gebruikt wordt door de New York City Fire Department (FDNY) om bestrijding en preventie effectiever en efficiënter te maken. De analyse wordt gedaan op basis van drie jaar aan data van verschillende instanties en voor ieder gebouw. Op basis van 7.500 risicofactoren waaronder specificaties, klachten en overtredingen wordt de kans op een brand voorspeld. Op basis van een risicoscore worden de gebouwen gerangschikt voor inspecties en toegewezen aan een kazerne.

- Misdaad voorspellen
  Een veelgenoemde toepassing van Big Data in het veiligheidsdomein is het zogenaamde predictive policing, waarmee we deze publicatie ook begonnen. Ook in Nederland is er een initiatief op dit gebied gestart: het Criminaliteits Anticipatie Systeem (CAS), ontwikkeld binnen het voormalige korps Amsterdam-Amstelland.[9] De kern is dat informatie over criminaliteit, zoals woninginbraken, wordt gekoppeld aan ongeveer 250 kenmerken als sociale samenstelling, aantal gokhallen en historische data over inbraken en op een digitale kaart wordt gezet. Met een algoritme is het mogelijk om voorspellingen te doen en de inzet van de Politie daarop af te stemmen. 36 procent van de woninginbraken bleek voorspeld te kunnen worden, reden om in het najaar 2015 een pilot testarten in vier basisteams buiten Amsterdam.

> **RUTGER RIENKS,** Landelijke Eenheid, Politie:
>
> Nu is het vooral handhaven in het nu en opsporing in het verleden, we moeten doorstappen naar de toekomst

Bezint eer ge begint

De opmars van Big Data biedt niet alleen nieuwe kansen maar gaat ook gepaard met risico's. De toepassing ervan roept vragen op over de privacy van burgers en de sociale consequenties die dit met zich mee kan brengen (hoe gaat onze samenleving eruitzien, wat betekent dit voor onze rechtsstaat?). Een publicatie van KPMG over The Internet of Things biedt een overzicht met verschillende (soms tegenstrijdige) opinies.[10] Waar het in het geval van Big Data om gaat is dat we de kansen weten te benutten en tegelijkertijd de (mogelijke) gevaren adequaat beheersen.

Voorbeelden uit andere sectoren laten zien dat zorgvuldigheid geen overbodige luxe is, zeker gezien de maatschappelijke gevoeligheid ten aanzien van het gebruik van (persoonlijke) data. Dat blijkt onder meer uit de kritiek die opsteekt zodra een financiële instelling data-analyse wil doen op betaalgegevens van klanten. Die gevoeligheid is waarschijnlijk nog hoger zodra het gaat om data die in het veiligheidsdomein wordt gebruikt. Dat noopt dan ook tot grote zorgvuldigheid, en het aloude mantra bezint eer ge begint is dan ook van toepassing. In hoofdlijnen gaat het onder meer om de volgende aspecten:

Het waarborgen van privacy

Indien bij Big Data persoonsgegevens worden verwerkt is het recht op bescherming van persoonsgegevens van

9 Het Tijdschrift voor de Politie editie – 76/nr.4/5/14 – Predictive Policing - wens of werkelijkheid?
10 KPMG, The Internet of Things – should we embrace its full potential, 2015

kracht, zoals geregeld in de Wet bescherming persoonsgegevens (Wbp). Deze implementeert de Europese Privacyrichtlijn uit 1995. In de Wbp zijn twee principes het meest relevant voor het gebruik van Big Data, namelijk rechtmatigheid van de verwerking en informatieverstrekking aan degenen van wie de persoonsgegevens worden verwerkt.[11] Op dit moment wordt de Europese Privacyrichtlijn herzien, een van de belangrijkste doelen is burgers zelf de mogelijkheid te geven hun persoonlijke data beter te beheren. Om een goed gebruik van Big Data met stevige waarborgen voor privacy en andere fundamentele waarden vorm te geven binnen Nederland, heeft de regering de WRR op 26 mei 2014 een advies gevraagd over het thema 'Big Data, privacy en veiligheid'.[12] Dit advies wordt aan het einde van 2015 verwacht. Het advies van de WRR moet leiden tot een normatief kader voor het gebruik van Big Data door de overheid en het identificeren van aanknopingspunten voor normering en regulering in brede zin. Voor het veiligheidsdomein geldt met betrekking tot het experimenteren en toepassen van Big Data dat maatschappelijk draagvlak extra relevant is. In dat kader is het ook duidelijk dat er in de media kritische kanttekeningen worden geplaatst over onder andere het koppelen van databronnen.[13] Er is wel sprake van iets merkwaardigs in dit verband. Veel burgers kennen nauwelijks terughoudendheid op sociale media en plaatsen onophoudelijk informatie over zichzelf online. Van de overheid verwachten ze iets heel anders.

Transparantie: wat gebeurt er met mijn data?

Organisaties zijn niet altijd transparant over wat ze met data doen, nemen de privacy van de burger niet serieus of maken gebruik van slecht doordachte toepassingen voor data-analyse. Dat is

> **ANONIEM,** MIVD:
>
> Onze kinderen zullen deze tijd zien als de '60 van de vorige eeuw; vrijheid en blijheid op het internet, zonder na te denken over eventueel negatieve consequenties

een serieus risico. Voor iedere organisatie, maar voor de overheid in het bijzonder, is het noodzakelijk om transparant te zijn waar persoonlijke data voor wordt gebruikt en om burgers een keuze te geven in het wel/niet beschikbaar stellen van hun persoonlijke data. Vanaf 1 januari aanstaande wordt het zelfs verplicht voor organisaties om een ernstig datalek te melden bij het College bescherming persoonsgegevens en soms ook aan de betrokken personen. Er is ook bijzondere aandacht nodig voor open communicatie en een zeer actieve houding in de vorming van de publieke opinie, zeker nu data-analyse een grotere rol gaat spelen. Denk bijvoorbeeld aan de opkomst van machine learning, die over tien jaar waarschijnlijk alomtegenwoordig zal zijn. Veel apparaten kunnen allerlei taken dan minstens even goed als de mens en hebben daarmee een grote invloed op ons leven. De vraag is of we wel in staat zijn om te laten zien hoe die apparaten werken en hoe ze tot beslissingen komen.

Bij dit alles spelen moeilijke afwegingen die rekening houden met een heterogene groep stakeholders, onder meer omdat het vanuit het typische karakter van het veiligheidsdomein niet altijd mogelijk is om geheel transparant te zijn. Een mooi voorbeeld van hoe dat spanningsveld ook in andere overheidssectoren speelt – en hoe ermee om te gaan – is te lezen in een interview met de Belastingdienst en hoe deze tot risico-inschattingen over burgers komt. Zij gebruiken voor het opstellen van risicoprofielen alleen gegevens die zij al in huis hebben.[14]

Cruciale kwaliteitsaspecten

Met het oprukken van data-analyse in het veiligheidsdomein – en het complexer worden van de data-analyses – nemen niet alleen de mogelijkheden toe maar ook het risico en de impact van foutieve data-analyses. Dat stelt

> **ELIN BERFELO,** KPMG:
>
> Zorg voor veilige structuren om data te verwerken en te analyseren en stel burgers op de hoogte wat je met hun persoonlijke data doet

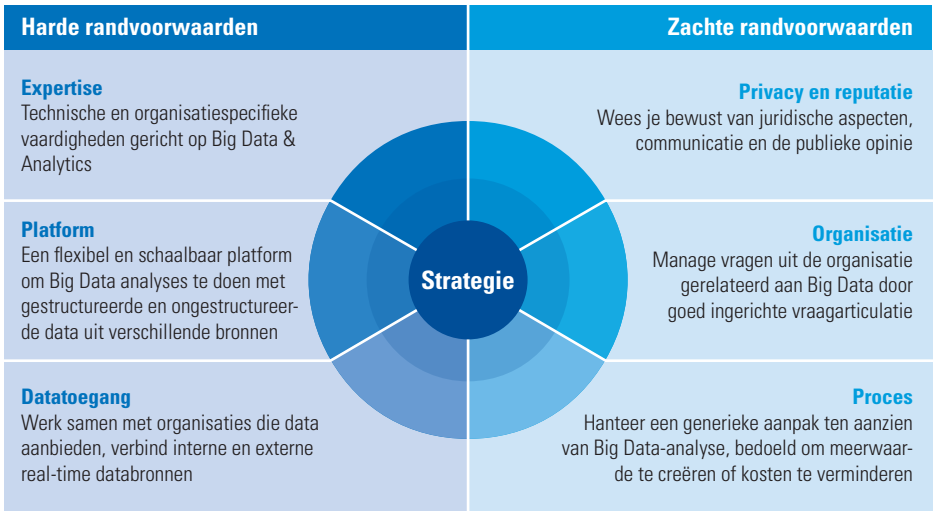
11 Big Data en privacy, kamerbrief kenmerk: DGETM-TM / 14179608, 19 november 2014

12 http://www.wrr.nl/projecten/project/article/big-data-privacy-en-veiligheid/

13 https://decorrespondent.nl/3044/De-politie-van-de-toekomst-houdt-iedere-burger-non-stop-in-de-gaten/343168743192-63143445 en https://news.vice.com/article/predictive-policing-is-not-like-minority-report-its-worse

14 https://decorrespondent.nl/2720/Baas-Belastingdienst-over-Big-Data-Mijn-missie-is-gedragsverandering/30632155840-0b4112e6

dus hoge eisen aan de professionals die de data-analyse toepassen. Het zou een misvatting zijn om te veronderstellen dat data-analyse weinig meer is dan number crunching, wat met de aankoop van wat tools eenvoudig kan worden geïmplementeerd in de werkwijze van organisaties. Immers: *a fool with a tool is still a fool.* Waar het om gaat is dan ook dat goed geschoolde data-scientists aan de knoppen zitten van de tools en de resultaten in de juiste context kunnen interpreteren. Tot slot moeten ook andere kwaliteitsaspecten – zoals het afschermen van data voor ongeoorloofd gebruik en/of de beveiliging tegen cybercrime – buiten kijf staan.



Tot slot: Handelingsperspectief

Wat is er nodig om op een verantwoorde wijze aan de slag te gaan met Big Data? Er zijn zes met elkaar samenhangende terreinen die aandacht behoeven, zoals blijkt uit de figuur hieronder. Allereerst zijn er drie harde eisen: organisaties hebben (1) kennis van zaken nodig, (2) een platform en de tools om goede data-analyses uit te voeren en (3) data om op dat platform te gebruiken (zowel intern als extern). De zachtere factoren zijn minstens even relevant. Het gaat achtereenvolgens om (1) aandacht voor privacy en wettelijke eisen (onder meer om reputatieschade te voorkomen), (2) het effectief omgaan met de vragen vanuit de werkvloer en (3) een consistente aanpak waarin alle Big Data-projecten een duidelijke gemeenschappelijke deler hebben die is gebaseerd op de strategische uitgangspunten.

CAMILLE MICHEL, VEILIGHEIDSREGIO HAAGLANDEN:

Een breed draagvlak om aan de slag te kunnen is essentieel

Dat klinkt overzichtelijk, maar kan in de praktijk weerbarstig zijn. Elk van de zes factoren kent specifieke aandachtspunten en het voert te ver om daarop in deze publicatie nader in te gaan.

Een slotopmerking is wel op zijn plaats. Vrijwel geen enkele organisatie is in staat om op alle genoemde terreinen vanaf het begin te excelleren, al was het maar omdat Big Data-initiatieven vaak een heel andere manier van werken vragen. Er is dan ook ruimte voor experiment nodig. Tegelijkertijd is het zaak dat het experimenteren niet vrijblijvend is omdat het er uiteindelijk om gaat dat goede projecten ook daadwerkelijk worden verankerd in de organisatie. De uitdaging voor het management: ruim baan geven aan creativiteit en tegelijkertijd de randvoorwaarden zeer strak formuleren. Gas geven en remmen tegelijkertijd dus.

ERIK VAN DEN BROM, KPMG:

Maak ruimte voor experimenten binnen organisaties

Sander Klous is managing director Big Data Analytics bij KPMG, hij is hoogleraar Big Data Ecosystems for Business and Society bij de Universiteit van Amsterdam en heeft samen met Nart Wielaard het boek 'Wij zijn Big Data' geschreven.

Erik van den Brom is binnen KPMG verantwoordelijk voor het ontwikkelen van systemen en toepassingen voor organisaties om data in de volle potentie te gebruiken. Zijn specialisatie is het ondersteunen van organisaties om datagedreven te kunnen werken.

Huub Wilbrink maakt deel uit van het veiligheidsteam binnen KPMG en heeft jarenlange ervaring als adviseur in verschillende organisaties in het veiligheidsdomein. Hij richt zich met name op het herontwerpen en inrichten van organisaties en dient als katalysator voor verandering.

Elin Berfelo heeft vanuit verschillende functies jarenlange ervaring in het veiligheidsdomein. Zij ziet in het gebruik van Big Data grote meerwaarde voor organisaties in het publieke (veiligheids) domein. Haar specialisaties zijn innovatie en organisatieverandering.

Contact

KPMG
Sander Klous
Laan van Langerhuize 1
1186 DS Amstelveen

T: 020 656 7186
E: klous.sander@kpmg.nl

kpmg.nl