

### **Data-analyse binnen een maatschappelijke normenkader**

*Data-analyse biedt belastingautoriteiten een groot (en deels nieuw) potentieel voor onder meer het inschatten van risico's en het voorspellen van gedrag. Een zorgvuldige aanpak is daarbij echter geboden want er kan ook veel mis gaan, met alle maatschappelijke onrust van dien. Er staat dan ook veel op het spel.*

*Naast een professionele uitvoering is het door de impact van data-analyses op ons leven noodzakelijk geworden om overheidstoezicht in te richten. Daarbij gaat het niet alleen over cybersecurity en data governance, waar de meldplicht datalekken een begin mee heeft gemaakt, maar ook over welke toepassingen we wel en niet maatschappelijk wenselijk achten en welke normen we stellen.*

*De belastingdienst heeft een unieke kans om een voortrekkersrol te spelen in deze ontwikkeling. De tijd is rijp voor een initiatief waarbij academische wereld, bedrijfsleven en toezichthouders de handen in één slaan om het bovengenoemde toezicht in te richten en een stap verder te brengen.*

---

Een groeiende groep wetenschappers maakt zich zorgen over hoe het interpreteren van informatie door algoritmes ongewenste impact kan hebben op ons leven. Dat is logisch, want die algoritmes rukken op in een samenleving met een exploderende hoeveelheid data, ze zijn niet altijd even transparant en soms aantoonbaar vooringenomen. De rode draad in de visies is dat technologie steeds meer vat krijgt op ons menselijk handelen, op een manier die velen decennia geleden niet konden vermoeden. Het motto onder doemdenkers was toen vaak *Big Brother is watching you*. Anno 2017 is de realiteit een stuk verder: *Big Brother is guiding you*.

Een tastbaar voorbeeld is hoe huisartsen zich laten ondersteunen door beslissingsondersteunende systemen die gebruik maken van algoritmes; het is maar te hopen dat deze daadwerkelijk in het belang van je gezondheid handelen; Een ander voorbeeld is hoe recruiters met algoritmes zoeken naar de beste kandidaat; het is maar te hopen dat het algoritme daarbij niet discrimineert; Dergelijke vraagstukken zullen de komende jaren in alle domeinen opkomen, en zullen ook bepalen hoe belastingautoriteiten de mogelijkheden van dergelijke nieuwe vormen van technologie mogen benutten.

Dit onderwerp begint ook op de politieke radar te staan. Het Europese Parlement nam dit voorjaar een resolutie aan waarmee ze de Europese Commissie oproept internationaal het voortouw te nemen.<sup>1</sup> De opkomst van algoritmes heeft volgens de resolutie mooie en minder mooie kanten. Lerende machines hebben “immense economische en innovatieve voordelen voor de samenleving” maar bieden tegelijk ook nieuwe uitdagingen. Het gaat daarbij om het

<sup>1</sup> <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//EN>

waarborgen van onder meer “niet-discriminatie, degelijke verwerking, transparantie en begrijpelijkheid in het besluitvormingsproces.”

De vraag die dus steeds vaker op tafel komt te liggen is of het onder de motorkap – waar de data-analyse plaatsvindt – wel deugt. Burgers kunnen dat in veel gevallen niet zelf vaststellen. En dus is er een nieuwe oplossing nodig om ons vertrouwen daarin te borgen. Want dat vertrouwen is essentieel om met een gerust hart de voordelen van een slimme samenleving te kunnen benutten.

Een eerste stap om dat vertrouwen te borgen is een professionele inrichting van de organisatie. Veel organisaties zitten daarbij nog in een transitiefase waarin ze zichzelf opnieuw aan het uitvinden zijn rondom data. Ze bouwen in hoog tempo kennis op over geavanceerde data-analyse. Daarbij gaat het niet alleen over analysetechnieken, zoals machine learning, maar minstens zo belangrijk is de professionele inrichting van de organisatie er omheen. Vraagstukken die spelen zijn: Hoe richt je een goed peer review proces in om analyses te controleren? Hoe zorg je voor traceerbaarheid en validatie in een snel veranderende omgeving? Kunnen op open source gebaseerde oplossingen helpen bij het vergroten van transparantie aangezien deze inzichtelijk zijn voor de buitenwacht? Hoe bouwen we betrouwbare en vertrouwde ecosystemen waarbij inzichten in data ook beschikbaar zijn buiten de organisatie, bijvoorbeeld voor andere deelnemers in de waardeketen?

Een professionele inrichting van de organisatie alleen is echter niet genoeg. Gezien de enorme impact die data-analyses kunnen hebben op ons leven is overheidstoezicht gewenst. Concreet gaat het daarbij om vijf thema's die het vertrouwen van de samenleving in data-analyses bepalen: Cyber Security, IT Controls, Data Governance, Software Kwaliteit en Ethiek. Alle incidenten die zich de afgelopen jaren hebben voorgedaan op het gebied van data-analyse zijn terug te voeren op één of meer van deze thema's. Voor een groot deel bestaan voor deze thema's al normenkaders en is het 'slechts' een kwestie van een holistische aanpak met de juiste focus: het creëren van vertrouwen in data-analyse. Een aantal andere elementen, zoals de uitlegbaarheid van meer geavanceerde algoritmes en hoe we kunnen toetsen of algoritmes voldoen aan onze normen en waarden, zullen mede op basis van een maatschappelijke dialoog nog verder moeten worden ontwikkeld.

Het inrichten van overheidstoezicht op data-analyses leidt tot een uitdaging op het gebied van complexiteit en schaalbaarheid. De meldplicht datalekken, die vorig jaar is ingesteld, is een goed voorbeeld van hoe we daar in de praktijk mee om kunnen (of zouden moeten) gaan.

Ten eerste de complexiteit. De meldplicht kan het begin zijn van een hanteerbare norm voor data-analyse. Tegelijkertijd is de meldplicht maar één klein onderdeel van wat er nodig is. We verwachten namelijk niet alleen dat organisaties netjes omgaan met data maar ook dat ze analyses uitvoeren die deugen, in de breedste zin van het woord. Volgens de principes van agile ontwikkeling, gericht op het beheersbaar houden van complexiteit, kan de norm het beste iteratief en modulair (i.e. in kleine stapjes) worden uitgebreid.

Ten tweede de schaalbaarheid. De meldplicht draait namelijk de verantwoordelijkheid om. In plaats van het aan de handhaver over te laten om overtredingen te identificeren, wordt die verantwoordelijkheid bij de organisaties zelf gelegd. Dat heeft twee belangrijke voordelen.

Ten eerste is het bijna onmogelijk om zo een massale taak door één (centrale) organisatie uit te laten voeren. Per organisatie is de inspanning daarentegen relatief gering, dus door de activiteiten bij de organisaties in te richten ontstaat de beoogde schaalgrootte. Ten tweede legt het de kosten van deze activiteit op de juiste plek. Als organisaties zelf verantwoordelijk zijn voor het rapporteren van de lekken, zullen ze daar ook zelf in moeten investeren. De kosten vallen dus waar de baten zijn.

Vanaf het moment dat de meldplicht is ingericht is er al discussie over de effectiviteit van de maatregel. Het is niet ondenkbaar dat de handhaver binnenkort de touwtjes wil aantrekken.

Een logische aanvulling zou zijn om rapportages te laten controleren door derde partijen. Vergelijkbaar met het model van de jaarrekeningcontrole. Ook dit model is decentraal en dus schaalbaar en bovendien bestand tegen een toename van complexiteit als de norm wordt uitgebreid. Er ontstaat dan wel behoefte aan onafhankelijke, betrouwbare partijen die de vertrouwensrol op zich kunnen nemen. Zij moeten niet alleen bewaken hoe wordt omgegaan met toegang tot data en daaraan gerelateerde controleerbare processen, maar ook bepalen of de analyses binnen de normen en waarden blijven die wij daar als maatschappij aan stellen.

Accountants weten als geen ander hoe ingewikkeld die rol is. Het vereist een organisatie waarbij het maatschappelijk belang bovenaan staat en de maatschappelijke vertrouwenspositie onder geen enkel beding geschaad mag worden. Dat betekent een vergaande bereidheid om transparant te zijn, nauw samen te werken met nationale en internationale toezichthouders en publiekelijk verantwoording af te leggen. Het vergt daarnaast vertrouwen van zowel de overheid als primaire hoeder van verantwoorde data-analyse als het bedrijfsleven dat bereid moet zijn data en daaraan gerelateerde processen controleerbaar te maken en te delen. En het vergt tenslotte aanzienlijke investeringen en de bereidheid om langjarige verplichtingen aan te gaan in een snel veranderende omgeving.

De 'winst' die gemaakt kan worden is onschatbaar, namelijk het behoud van maatschappelijk vertrouwen in de slimme samenleving.

De belastingdienst heeft een unieke kans om een voortrekkersrol te spelen in deze ontwikkeling. Wat nu noodzakelijk is, is een initiatief waarbij academische wereld, toezichhouders en bedrijfsleven de handen ineenslaan om het overheidstoezicht in te richten en een stap verder te brengen. Daarvoor is nog veel kennisontwikkeling nodig bij alle genoemde partijen. Politiek en accountants zullen hun verantwoordelijkheid moeten onderkennen en het voortouw moeten nemen in deze ontwikkeling. Uiteindelijk hebben zowel de samenleving als de accountants immers het meeste baat bij deze ontwikkeling.

Sander Klous  
Managing Partner  
Lighthouse Data & Analytics  
KPMG Advisory

prof. dr. Sander Klous  
Chair of Big Data Ecosystems  
Informatics Institute  
Faculty of Science  
University of Amsterdam