Redactioneel

Autonome systemen

‘Computer says no’

Rein Halbersma*

Computer says ‘no’

Met de opkomst van Big Data (zie het themanummer 2018-2/3 van dit tijdschrift) en al maar krachtiger computers en algoritmen om die data te analyseren, worden steeds meer beslissingen genomen door autonome systemen. Er zijn tal van nieuwe maatschappelijke ontwikkelingen: van zelfrijdende auto’s en terugpratende digitale assistenten tot automatische gezichtsherkenning en fraudedetectie. Zulke autonome systemen nemen ons steeds meer beslissingen uit handen, maar stellen ons soms ook voor voldongen feiten.

Deze ontwikkelingen werpen vragen op. Juist ook voor het toezicht. Wat zijn de belangrijkste ontwikkelingen op het terrein van autonoom toezicht? Op welke terreinen van de samenleving spelen deze? Hoe spelen de bestaande toezichtsinstituten hierop in? Wie houdt er toezicht op autonome systemen? Is autonoom toezicht nog verre toekomstmuziek of worden we binnenkort ingehaald door technologie die we ons nu nog niet goed kunnen voorstellen? Op welke wijze worden rollen en verantwoordelijkheden verdeeld tussen mens en machine? Wie doet wat en wie is verantwoordelijk en aansprakelijk als autonome systemen falen? Hoe kunnen besluiten door autonome systemen worden getoetst? Hoe borgen we zorgvuldigheid en uitlegbaarheid? En hoe voorkomen we discriminatie en willekeur?

* Dr. R.S. Halbersma is onderzoekscoördinator bij de Kansspelautoriteit en redactielid van *Tijdschrift voor Toezicht*.

Een nieuwe industriële revolutie

Voor wie denkt dat het allemaal zo’n vaart niet zal lopen, is het goed om een kort historisch uitstapje te maken om de razendsnelle ontwikkelingen te schetsen. Al vanaf eind jaren tachtig van de vorige eeuw zijn computers in staat gebleken om de mens te overtreffen in gespecialiseerde, overzichtelijke domeinen zoals het herkennen van handgeschreven postcodes op poststukken.¹ In de jaren negentig van de vorige eeuw stootten computers de mens van de troon bij klassieke bordspellen als dammen, schaken en backgammon. Bij de twee eerstgenoemde spellen leunde de computer nog sterk op brute rekenkracht in combinatie met een kleine dosis door de mens toegevoegde kennis, maar bij het laatste spel bleken voor het eerst ook zelflerende algoritmen succesvol te kunnen wedijveren met de beste menselijke spelers.² De hooggespannen verwachtingen om deze successen naar complexere domeinen over te kunnen hevelen werden in de jaren erna echter niet bewaarheid. Het wetenschapsveld van artificiële intelligentie (AI) had zich in het verleden al vaker overgeven aan overdreven optimisme, en leek opnieuw af te stevenen op een ‘AI-winter’ van teruglopende onderzoeksgelden.³ De grote doorbraak en ommekeer in het verwachtingspatroon kwam in 2012 toen een zelflerend algoritme met behulp van zogeheten neurale netwerken met overmacht een jaarlijkse wedstrijd in geautomatiseerde beeldher-

1. P. Ahmed en C.Y. Suen, ‘Computer recognition of totally unconstrained handwritten ZIP codes’, *International Journal of Pattern Recognition and Artificial Intelligence* 1987, nr. 1, p. 1-15.
2. G. Tesauro, ‘TD-Gammon, a self-teaching backgammon program, achieves master-level play’, *Neural computation* 1994, nr. 2, p. 215-219.
3. J. Hendler, ‘Avoiding another AI winter’, *IEEE Intelligent Systems* 2008, nr. 2, p. 2-4.

kenning wist te winnen.⁴ Drie jaar later waren verbeterde neurale netwerken de mens definitief voorbijgestreefd in dit domein. In het Japanse bordspel Go werden met vergelijkbare technieken de beste menselijke spelers verslagen, en in 2018 bleek het mogelijk om zowel schaken als Go via een zelflerend algoritme zonder menselijke kennis op bovenmenselijk niveau te laten spelen.⁵ Behalve superioriteit in complexe maar geïsoleerde domeinen zijn computers anno 2020 ook in staat om beeldverwerking, spraakherkenning en vertaling te bundelen in autonome digitale assistenten als Apple Siri, Google Assistant en Amazon Alexa, en staat de auto-industrie onder aanvoering van Tesla aan de vooravond van zelfrijdende auto's.

Transformatie van toezicht

Behalve nieuwe mogelijkheden zijn er ook risico's verbonden aan deze technologische ontwikkelingen. Bij commerciële toepassingen zijn privacy (bij digitale assistenten) en fysieke veiligheid (zoals bij auto's) belangrijke thema's. In het eerder genoemde themanummer Big Data zijn zulke aspecten ook al aan bod gekomen. Deze problemen worden alleen nog maar versterkt wanneer overheden gebruikmaken van autonome systemen, waarbij er sprake is van besluitvorming met voor burgers soms onbegrijpelijke of oneerlijke besluiten. De waarborgen die de Grondwet en het algemeen bestuursrecht bieden tegen willekeur vanuit de overheid, kunnen op termijn worden uitgehold als de overheid zich door onverschilligheid of kennisachterstand verschuilt achter door algoritmen genomen besluiten. Overheden kunnen zelfs nog verder gaan, en de mogelijkheden van Big Data en zelflerende algoritmen benutten voor surveillance en controle van burgers. Het is voor autoritaire regimes zowel financieel als technologisch haalbaar om opnamen van telefoongesprekken en beelden van bewakingscamera's integraal op te slaan en achteraf te analyseren.⁶ In China is zelfs sprake van *real-time* monitoring van de bevolking met het *social credit* systeem.⁷ In westerse landen zijn de instituties hopelijk sterk genoeg om burgerrechten te waarborgen tegen de ongebreidelde technologische mogelijkheden van autonome surveillancesystemen. Toch zijn ook er in Nederland recent incidenten geweest zoals de interne fraudelijsten bij de Belastingdienst en het fraudedetectiesysteem SyRI bij de gemeente Rotterdam. Wie de bewaker

moet bewaken bij deze transformatie naar autonoom toezicht is een belangrijke rode draad in dit themanummer.

Burgers, bedrijven en bestuurders

Het voorliggende nummer is zoals eerder opgemerkt in zijn geheel gewijd aan het thema 'Autonome Systemen'. Het is een gevarieerd nummer geworden met bijdragen vanuit verschillende invalshoeken, veelal van essayerende of opiniërende aard. Daarnaast mocht de redactie zich ook verheugen in twee gedegen artikelen en een interview.

Het artikel van Bart van der Sloot en Sasha van Schendel (beiden van Tilburg Institute for Law, Technology and Society van Tilburg University) is gebaseerd op een in 2019 verschenen lijvig WODC-rapport over procedurele waarborgen en toegang tot het recht in een datagedreven samenleving waarin autonome systemen steeds meer beslissingen nemen. Het artikel komt met tien concrete aanbevelingen om burgers een betere rechtspositie te geven tegenover overheden en grote bedrijven die steeds meer van hen weten en voor hun beslissen.

Het *peer reviewed* artikel van Jan Sviták en Rob van der Noll (beiden van de Autoriteit Consument en Markt) laat aan de hand van computersimulaties zien hoe algoritmen voor prijszetting zonder expliciete afstemming toch tot collusie tussen bedrijven kunnen leiden. Ze nemen de lezer bij de hand en geven daarbij een kijkje in de keuken van de algoritmen. Een noviteit voor het *Tijdschrift voor Toezicht* hierbij is dat de broncode van de computersimulaties door de auteurs voor derden toegankelijk is gemaakt door de software te plaatsen op de populaire website GitHub. Hiermee geven de auteurs concreet invulling aan transparantie en reproduceerbaarheid van hun onderzoek.

De in nummer 2019-2 van dit tijdschrift gestarte reeks van estafette-interviews met recent aangetreden bestuurders van markttoezichthouders wordt voortgezet door het vraaggesprek tussen redactieleden Marije Bating en Martijn Groenleer en bestuursvoorzitter Martijn Snoep van de Autoriteit Consument en Markt. Een opvallende constatering in dat gesprek is dat de ACM naast de energietransitie ook de digitale economie en de verandering in de economie als gevolg van digitalisering als een van de twee belangrijkste maatschappelijke thema's heeft geagendeerd.

Zorgplicht in het analoog en digitaal bestuursrecht

Het essay van Thijs Drouen en Marlies van Eck (beiden Hooghiemstra & Partners) gaat in op de recente jurisprudentie en beleidsvorming rondom het gebruik van

4. A. Krizhevsky, I. Sutskever en G.E. Hinton, 'Imagenet classification with deep convolutional neural networks', in: *Advances in neural information processing systems* 2012, p. 1097-1105.
5. D. Silver e.a., 'A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play', *Science* 2018, 362(6419), p. 1140-1144.
6. J. Villaseñor, *Recording everything: Digital storage as an enabler of authoritarian governments*, Center for Technology Innovation at Brookings 2011.
7. F. Liang, V. Das, N. Kostyuk en M.M. Hussain, 'Constructing a data-driven society: China's social credit system as a state surveillance infrastructure', *Policy & Internet* 2018, nr. 4, p. 415-453.

algoritmen door de overheid, en pleit voor een zorgplicht hierbij. Een belangrijk element om daar daadwerkelijk effectief invulling aan te kunnen geven is om in een vroeg stadium, bij de aanschaf of ontwikkeling van algoritmen, stil te staan bij de uitlegbaarheid van algoritmen, het *accountable by design* zijn.

In een op zijn recente oratie gebaseerd essay bespreekt Johan Wolswinkel (Tilburg University) hoe in het licht van de recente SyRI-uitspraak het bestuursrecht in staat kan worden gesteld om het gebruik van zelflerende systemen adequaat te reguleren. De auteur schetst hoe de rechtsontwikkeling voor autonoom overheids-toezicht vorm kan krijgen als een proces van laveren tussen analoog en digitaal bestuursrecht, waarbij de bestuursrechtelijke verworvenheden van het analoge toezicht worden geconfronteerd met de uitdagingen van het digitale toezicht. Dat vergt de bereidheid van toezichthouders om de normering van digitale toezichthandelingen expliciet vorm te geven.

Kansen en kwaliteitsborging

Daniël Frijters en Jelle Attema (beiden Platform voor de InformatieSamenleving) benadrukken in hun essay dat het gebruik van algoritmen ook kansen biedt voor het werk van toezichthouders. Daarbij dient de toezichthouder zich te verdiepen in de werking van de technologie en de eisen die eraan gesteld moeten worden, welke kansen technologie biedt en wat nodig is om deze te realiseren. De toezichthouder moet oog hebben voor de gevolgen van technologie voor de positie van burgers en andere belanghebbenden, voor de veranderingen in verhouding tussen toezichthouder en burgers, en voor rollen en bijpassende verantwoordelijkheden die op de toezichthouder afkomen als gevolg van het toepassen van technologie.

Jeroen Goudsmit (Vrije Universiteit Amsterdam) en Jonas Teuwen (Nederlands Kanker Instituut) bespreken in hun essay dat kwaliteitstoezicht op door algoritmen genomen beslissingen zich met name zou moeten richten op de stappen om te komen van probleemomschrijving tot een geïmplementeerd algoritme. Door deze keuzes expliciet en met de juiste overwegingen te maken, verkleint het risico op misdragingen. De auteurs bevelen aan dat toezichthouders en ontwikkelaars van algoritmen hier tezamen een handreiking voor kunnen opstellen.

Het essay van Petra Helwig (Inspectie Overheidsinformatie en Erfgoed) besteedt aandacht aan eisen aan onderliggende algoritmen. Eén daarvan is dat algoritmen die de overheid gebruikt bij (deels) geautomatiseerde besluitvorming, transparant en uitlegbaar horen te zijn. Bij het transparant maken van algoritmen kan niet altijd worden volstaan met een beroep op de kennis van direct betrokkenen, zoals programmeurs. De auteur betoogt dat daarom 'iets' van gebruikte algoritmen opgenomen moet worden in het archief van de

organisatie en moet worden beheerd in lijn met de Archiefwet.

Menselijke maat en leren van fouten

De opiniebijdrage van Giel Stoeper (Centrale Raad van Beroep en Tilburg University) en Ivo Stoeper (Eindhoven University) betoogt dat menselijk toezicht van groot belang is wanneer gebruik wordt gemaakt van voorspellende modellen. Vaak leveren de inzichten uit de datawetenschap weliswaar goede voorspellingen op over mogelijke fraudegevallen, maar in de praktijk zijn er toch problemen, bijvoorbeeld met de steekproef en de data op basis waarvan voorspellingen worden gedaan. Dat probleem is sterker aanwezig naarmate de gebruikte modellen complexer zijn. De auteurs zijn van mening dat algoritmen daarom nooit volledig autonoom zouden mogen werken.

Gerben Feddes (Rijksdienst Wegverkeer) gaat in zijn opiniebijdrage in op de onvoorspelbaarheden van menselijk gedrag in het verkeer die de laatste 'drempel' vormen voordat zelfrijdende auto's alle bestuurderstaken veilig kunnen overnemen. De auteur blijft weg van de bekende ethische dilemma's in hypothetische situaties, maar formuleert in plaats daarvan een aantal concrete beleidsvragen rondom het toezicht op zelfrijdende auto's. Vooral hoe de verantwoordelijkheden bij ongevallen te beleggen en hoe een veilige cultuur te creëren rondom het leren van fouten, staan daarbij voorop.

Wisseling van de wacht

Dit themanummer is een momentopname van wat er nu speelt bij toezicht op autonome systemen. Het thema zal de komende jaren nog volop in de belangstelling staan. De redactie kijkt dan ook uit naar nieuwe bijdragen op dit vlak. Zoals in de inleiding reeds geschetst ontwikkelt de techniek zich razendsnel. De thema's van zorgplicht, transparantie en menselijke maat bij autonome systemen zullen echter van belang blijven.

Over het menselijke aspect gesproken: het zal de oplettende lezer niet zijn ontgaan dat de redactie van dit tijdschrift in 2019 gaandeweg is gewijzigd. De redactie nam afscheid van gewaardeerde leden Paul de Bijl, Karin van Wingerde en Margot Aelen. Dit themanummer is tevens mijn laatste redactionele bijdrage aan het *Tijdschrift voor Toezicht*. Ik kijk terug op een mooie tijd waarin ik als redacteur twee mij nauw aan het hart liggende themanummers heb mogen begeleiden (2017-4 over Toezicht en Sport en het voorliggende 2020-1 over Autonome Systemen). Met de komst van vier nieuwe redactieleden, Anna Merz, Frans van Bruggen, Elianne van Steenberghe en Arnt Mein, is de redactie weer op volle sterkte, klaar voor de viering van het tweede lustrium in 2020!