



UvA-DARE (Digital Academic Repository)

De Grondwet en Artificiële Intelligentie

Dommering, E.

Publication date

2024

Document Version

Final published version

Published in

De Grondwet en nieuwe technologie: klaar voor de toekomst?

[Link to publication](#)

Citation for published version (APA):

Dommering, E. (2024). De Grondwet en Artificiële Intelligentie. In B. van Beers, B. Custers, E. Dommering, J. Goossens, A. W. Heringa, W. Hins, B.-J. Koops, S. Ligthart, G. Meynen, A. Meijer, B. van der Sloot, L. Trapman, & J. Wolswinkel (Eds.), *De Grondwet en nieuwe technologie: klaar voor de toekomst? : Twaalf pleidooien voor modernisering van de Grondwet : essaybundel* (pp. 69-83). Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

<https://www.rijksoverheid.nl/documenten/rapporten/2024/08/01/essaybundel-de-grondwet-en-nieuwe-technologie-klaar-voor-de-toekomst>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

De Grondwet en nieuwe technologie: klaar voor de toekomst?

Twaalf pleidooien voor modernisering van de Grondwet

Essaybundel

Britta van Beers
Bart Custers
Egbert Dommering
Jurgen Goossens
Aalt Willem Heringa
Wouter Hins
Bert-Jaap Koops
Sjors Ligthart & Gerben Meynen
Albert Meijer
Bart van der Sloot
Leon Trapman
Johan Wolswinkel

Colofon

De Grondwet en nieuwe technologie: klaar voor de toekomst?

Twaalf pleidooien voor modernisering van de Grondwet

Deze essaybundel is een uitgave van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Tekst: de essays zijn de verdienste en verantwoordelijkheid van de desbetreffende auteurs en reflecteren niet de opvattingen of aanbevelingen van het ministerie.

Concept en project begeleiding: Vormvijf, Den Haag

Grafisch ontwerp en vormgeving: Textcetera, Den Haag

Illustratie omslag en deelpagina's: Kylièn Bergh

Augustus 2024

De Grondwet en Artificiële Intelligentie

Egbert Dommering*

1. Probleemstelling

Er is in 2023 een hype ontstaan over Artificiële Intelligentie, met name rond *Chat GPT*, een taalrobot op internet die alle menselijke vragen kan beantwoorden, lezingen en scripties kan schrijven, en nog veel meer. En het blijft doorgaan met die publiciteit. Op het moment van het schrijven van dit artikel meldt de Volkskrant dat de maker van *Chat GPT*, Open AI, Sora heeft aangekondigd een AI-machine te maken die via een eenvoudige omschrijving in tekst een levensechte video produceert die ‘zeer realistisch’ oogt.¹ Hoe blijven wij deze pseudo-werkelijkheid de baas?

In dit artikel zal worden ingegaan op wat de eventuele implicaties van AI voor de Nederlandse Grondwet zouden kunnen zijn. De onvermijdelijke vraag is ook of de Gw zou moeten worden aangepast aan deze ‘nieuwe tijd’ en hoe: moet een grondwet niet ‘techniek neutraal’ zijn anders kun je wel aan de gang blijven met veranderingen.

Daartoe zal ik eerst ingaan op de vraag wat AI nu eigenlijk is en hoe het – juridisch – gedefinieerd moet worden. Vervolgens behandel ik de vraag of de Gw (telkens) aan veranderde technologieën moet worden aangepast. Ten slotte ga ik dieper in op de implicaties van AI voor de Gw. Ik sluit af met een algemene opmerking over de noodzaak de Gw aan te passen aan de effecten van wereldwijde communicatienetwerken in Nederland en de ontwikkeling van de EU-rechtsorde in dit domein.

* Emeritus-hoogleraar Informatierecht, Universiteit van Amsterdam.
1 *De Volkskrant*, 17 februari 2024, Rubriek *Ten Eerste*: ‘Open AI’s ‘Sora’ kan realistische video’s genereren op basis van eenvoudige tekst’.

2. Wat is AI?

In het laatste beleidsstuk dat er over dit onderwerp in Nederland op het moment van het schrijven van dit stuk voorhanden is, treffen we de herziene definitie van een 'AI-systeem' van de OESO (2023) aan. Het gaat om noot 5 van 'de aanbiedingsbrief (overheidsbrede) visie op generatieve AI'², met de staatsecretaris van 'Digitalisering en Koninkrijksrelaties' als eerste ondertekenaar:

'Een op machines gebaseerd systeem dat, voor expliciete of impliciete doelstellingen, afleidt, uit de input die het ontvangt, hoe de output zoals voorspellingen, inhoud, aanbevelingen of beslissingen moet genereren die fysieke of virtuele omgevingen kunnen beïnvloeden. Verschillende AI-systemen variëren in hun mate van autonomie en aanpassingsvermogen na implementatie/toepassing (deployment) ervan.'

Het vergt close reading om te begrijpen of de grammatica van deze zin klopt en wat deze betekent. Er staat: een systeem dat met behulp van computers overeenkomstig de doelstelling die het heeft meegekregen uit de input die het ontvangt output in de vorm van een beslissing, een aanbeveling, een feitenbestand of een voorspelling genereert die effect heeft op de omgeving waarin zij functioneert. Het ene AI-systeem functioneert zelfstandiger dan het andere en past de output uit zichzelf makkelijker aan dan de ander, voegt de tekst als kwalificatie aan deze definitie toe. Toch is deze definitie nog zo slecht niet, omdat het laat zien dat het zowel om de menselijk geselecteerde (data) input als de machinale bewerking daarvan gaat.

Sinds Descartes' 'cogito ergo sum' is de denkende geest een essentieel begrip om 'de mens' te definiëren. Er is de laatste jaren een groeiende discussie over Artificiële Intelligentie en de vraag of deze 'de geest' gaat overvleugelen of zelfs (op vitale gebieden) gaat vervangen. Naar aanleiding van ChatGPT heb ik mij recentelijk met deze vraag beziggehouden.³ Ik neem een citaat uit dit artikel op, omdat daarin het probleem waar het om draait uiteen wordt gezet:

'ChatGTP is een toepassing van een lerend AI-taalsysteem dat het mogelijk maakt dat een computer op basis van ingevoerde taaldata en een gecontroleerd 'leerproces' (algoritmen) vragen die aan hem

2 Brief Ministerie van Binnenlandse Zaken, 18 januari 2024, kenmerk 2024-0000026119, ondertekend door de Staatsecretaris van Digitalisering en Koninkrijksrelaties.

3 'Artificiële Intelligentie: Waar is de werkelijkheid gebleven', in: Computerrecht 2023 (nr. 258), p. 476-483.

worden gesteld, herkent en kan beantwoorden. Alle AI-paniek richt zich op de 'output' van AI, dat wat uit de computer komt: machtsmisbruik (verborgen camera's die aan 'gezichtsherkenning' doen) en discriminatoire beslissingen. Van even groot, zo niet groter, belang is de 'input', dat wat er in de computer gaat: data en instructies (algoritmen). Ik sta daarom eerst stil bij de input.

De opwinding over AI is in 2023 ontstaan doordat ChatGPT door het bedrijf OpenAI eind 2022 op de markt werd gebracht.⁴ De website van het bedrijf informeert ons dat het een 'AI and Research Company' is en dat het zijn missie is 'to ensure that artificial general intelligence benefits all of humanity.'⁵ Het bedrijf streeft dit doel na met ChatGPT die getraind is om met de gebruiker in gesprek te gaan. Deze dialoog maakt het mogelijk 'to answer follow up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.' In deze omschrijving valt al meteen iets op: In tegenstelling tot de hondse zeden en gewoonten van de gebruikers van de sociale media, presenteert de ChatGPT zich als een bijzonder beleefd wezen, net zoals vroeger de djinn die door op een lamp te wrijven werd opgeroepen om problemen op te lossen in het Duizend-en-één-nacht sprookje 'Aladin en de wonderlamp'. Zo vreemd is die beleefdheid niet. Machines zijn als bedienden: zij volgen de instructies van de bestuurder op. Als een machine deze gehoorzame vriendelijkheid niet vertoont, krijgt zij meteen een monteur op haar dak.

Maar let even op: de Chat-machine beantwoordt vragen; dat is iets anders dan een bewijsvoering of iets waar of onwaar is. "Waarheid is een bijproduct."⁶

De kennis en intelligentie van de ChatGPT is het product van de turbulente geschiedenis van het internet, en dan met name het World Wide Web. Deze toepassing heeft immers het 18^e eeuwse encyclopedie-ideaal verwezenlijkt door alle op de aarde gepubliceerde kennis (webpagina's), uit welke bron dan ook, toegankelijk te maken en met elkaar te verbinden.⁷ De steeds 'intelligenter' zoekmachine (een voorloper van de chatbot) deed de rest.

4 29 november 2022 in San Francisco. Het werd meteen een hype. De winst van de onderneming wordt dit jaar geschat op \$200 miljoen; bron *Financial Times* 17 juni 2023.

5 <https://openai.com/blog/chatgpt>, geraadpleegd juli 2023.

6 Citaat uit de door mij bijgewoonde lezing van 12 juni 2023 van Jelle Zuidema, specialist Neuro linguïstisch programmeren aan de Faculteit Wis- en Natuurkunde van de UvA.

7 Ik verwijs voor de geschiedenis naar Egbert Dommering, *De Europese Informatierechtsorde*, Amsterdam: deLex 2019, Hoofdstuk II, paragrafen 1 en 2.

Om het kenniswonder ChatGPT mogelijk te maken is allereerst nodig dat het ‘taalmodel’ (dat teruggaat op het model uit 1948 van Claude Shannon, de man die de ‘bits’ heeft bedacht) waaruit deze put zo groot mogelijk is. De kans dat hij bij een open vraag het antwoord herkent, wordt daardoor immers groter. De ontwikkeling van ChatGPT is in 2018 begonnen, en in vier jaar tijd is het taalmodel waarmee wordt gewerkt duizenden malen groter geworden.

Deze enorme berg taal (en bijbehorend beeld) moet worden bewerkt, vervolgens in nullen en enen opgeslagen in het computergeheugen, dan afgebroken in kleinere eenheden en dan door de computer zo te trainen dat deze bij een vraag, door vergelijking met de ingevoerde data, vrij nauwkeurige voorspellingen over het juiste antwoord kan geven. Dat bewerkingsproces heeft de laatste twee jaar – door de enorme schaalvergroting – grote stappen gemaakt.

Heeft de Geest nu de Werkelijkheid overwonnen dankzij het internet? Helaas is dat niet zo. We zijn veeleer getuige van een moorddadig datificeringsproces, zoals Kate Crawford dat in haar boek *Atlas of AI* onder woorden brengt.⁸ De werkelijkheid wordt omgezet in een massa neutrale bits:

“The data are not seen as carrying meaning or ethical weight as images of individual people. The personal, the social and the political meanings are all neutralised. It represents the shift from *image* to *infrastructure*, where the meaning or care that might be given to the image of an individual person or context behind a scene, is presumed to be erased at the moment it becomes part of an aggregate mass that will drive a broader system.”

Crawford wijst er terecht op dat AI niet alleen een wereld-omspannende database is, maar ook een verdeling van markten en infrastructuren tussen wereldbedrijven zoals Microsoft (de grootste aandeelhouder van Open AI), Meta, Adobe, Google, en Open AI (allen met eigen generatieve systemen onder andere merknamen). China (en TikTok) rukt op met een eigen data-infrastructuur. Dit is een begin van een nauwelijks gereguleerde innovatie in een bestaande technologie. De sprong die er binnen Google met het vervolmaken van het vertaalsysteem is gemaakt is een essentieel moment in deze ontwikkeling geweest.⁹

8 Kate Crawford, *Atlas of AI*, New Haven and London: Yale University Press 2021, pp. 18-21.
9 ‘Transformers: the Google scientists who pioneered an AI revolution.’, in: *Financial Times* 23 juli 2023.

Tot zover dit citaat. Er zijn natuurlijk ook andere AI-systemen die in de aan het begin genoemde OESO-definitie worden meegenomen. Dit zijn systemen die in tal van wetenschappen worden ingezet om op basis van een enorme hoeveelheid empirische data wetenschappelijke voorspellingen te doen. Denk aan astronomie, medische diagnostiek, weervoorspellingen, marktontwikkelingen, juridische en maatschappelijke risico inschattingen, enzovoort. Het gaat ook om vervanging van menselijke routine arbeid bij het verzamelen en ordenen van informatie, en dit kan dan weer gekoppeld worden aan beslissystemen, die functioneren volgens door mensen aan dit systeem meegegeven beslisregels.

Terug naar de filosofische vraag of AI het ‘cogito ergo sum’ kan verslaan. Dit is in wezen de filosofische vraag of de menselijke geest geheel wordt bepaald door ervaringsdata (empirisme: Locke, Hume) of door een ‘geestelijke’ structuur (rationalisme: Plato, Descartes) of door beide (Kant: empirisme binnen een rationele geestelijk structuur). Ik verwijs de geïnteresseerde lezer naar de bijzonder mooie behandeling van dit onderwerp in het boek van Guido van der Knaap, *Van Aristoteles tot Algoritme*.¹⁰ De ontwikkeling van AI beweegt zich tussen de twee polen ‘rationalisme’ en ‘empirisme’. Enige citaten uit dit boek:

‘Door sterk de nadruk te leggen op de ontwikkeling van het juiste symboolsysteem past het eerdere onderzoek naar AI goed binnen de rationalistische traditie. Het doel was niet om een AI-programma ervaringen te laten opdoen, maar om AI van kennis te voorzien om een eventuele ervaring tot een goed einde te brengen. Bij schaken was het doel om AI een veelvoud van strategieën te geven om het spel te winnen, niet om zoveel mogelijk spelletjes schaak te laten spelen om een eigen strategie te ontwikkelen. Het is als een voetbaltrainer die ellenlange tactische besprekingen houdt in de hoop dat zijn spelers zijn tactiek tijdens de wedstrijd goed uitvoeren, maar zijn spelers bar weinig laat trainen.’

(...)

We zien bij de Go-computer, in 2017 ‘Deep Mind’ en later ‘Alpha Go Zero’, beide ontwikkeld door Google, de geleidelijke overgang naar het empirische model:

‘Twintig jaar later komt ‘Deep Mind’ (...) met het programma *AlphaGo*. Go is een in China populair spel en wordt door vele mogelijke zetten en speculaties gezien als een spel dat voor de computer moeilijker is dan schaken. Toch lukt het *AlphaGo* de wereldkampioen te verslaan. De eerste versies van *AlphaGo* doen dat door een veelvoud van go-wedstrijden van specialisten te bekijken. Op die manier leert het programma wat goede strategieën zijn. Door vervolgens veel oefenspelletjes te spelen wordt *AlphaGo* steeds beter. (...). De succesvollere versie *AlphaGo Zero* gaat nog een stap verder. Het programma krijgt alleen de spelregels van GO mee en is verder een *tabula rasa*, een onbeschreven blad. Kennis van experts ontbreekt, evenals voorbeelden van gespeelde wedstrijden. *AlphaGo Zero* leert het spel door heel vaak tegen zichzelf te spelen, zo’n 29 miljoen keer. Het levert een uiterst succesvolle go-speler op. (...) Go-grootmeester Lee besluit in het najaar van 2019 het bijltje erbij neer te gooien. *AlphaGo Zero* is volgen hem onverslaanbaar.’¹¹

De imput van data is dus niet meer uit een AI-systeem weg te denken. Toch is er een geestelijke component, namelijk de onderscheidende kwaliteit van de ingevoerde data en de ordenende beslisregels (algoritmes) die worden ingezet. Ik citeer nogmaals uit mijn in noot 3 genoemde artikel:

‘Maar dit is ook het gebied dat wordt aangeduid met de ‘algoritmes’ die verkeerd of discriminatoir kunnen uitpakken, met name als het gaat om data met conflicterende waarden in categoriserings-systemen. Waar het bij de eerste input gaat om alle data van hun positieve of negatieve maatschappelijke en culturele waarde te ontdoen, gaat het bij categorisering juist om toekenning van positieve of negatieve waarde aan die data aan de hand van categoriseringssystemen.’

Een scherpe formulering van dit aspect van ordening door algoritmes is die van Louise Amoore:¹²

11 Citaten a.w. noot 10, pp. 36-37. Zie ook de verhelderende studie, Dominique Cardon, *A quoi rêvent les algorithmes?* Parijs: Editions de Seuil et la République des idées 2015.

12 Louise Amoore, *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*, Durnham: Duke University Press 2020, hoofdstuk 1. Oudere bronnen zijn James C. Scott, *Seeing like a State*, Yale: Yale University 1998 en John Cheney-Lippold, *We are Data, Algorithms and the Making of our digital selves*, New York: New York University Press 2017.

“The very essence of algorithms is that they afford greater degrees of recognition and value to some features of a scene than they do to others. In so doing, algorithms generate themselves as ‘ethicopolitical’ beings in the world. (...) ‘I have studied the algorithm not as a finite series of programmable steps but as perennially adjustable and modifiable in relation to a target output.”

En met die laatste zinsnede raken we dus de kern: het algoritme is in deze benadering niet een instructie om de machine beter te laten werken (een betere machinecode), maar een instrument om door middel van een machine een menselijke doelstelling te verwezenlijken (een betere maatschappelijke code).

Om een voorbeeld van een verkeerd algoritme te noemen: Het AI beslisregelsysteem dat mede de ‘Toeslagenaffaire’ heeft veroorzaakt, was te simpel, omdat het iemand die een vergissing maakte bij de terugbetaling van een uitkering, ongeacht de aard en de ernst daarvan, bleef behandelen als een ‘frauderende Bulgaar’ en de daarop van toepassing zijnde draconische sancties genereerde. Er was nog een andere menselijke component: de Afdeling bestuursrechtspraak van de Raad van State achtte zich niet in staat daar iets aan te doen, omdat het oordeelde dat de toepasselijke wettelijke bepaling geen ruimte bood om in het concrete geval af te wijken van de regel.

De slotsom is dat een ‘AI-systeem’ bestaat uit algoritmen plus data. Artikel 1 van de AI-ontwerpverordening (waarvan de tekst op het moment van het schrijven van dit artikel definitief door EU Commissie en Parlement is vastgesteld) definieert alleen op basis van de software (algoritme):

‘Artificiële-intelligentiesysteem (AI-systeem): software die is ontwikkeld aan de hand van een of meer van de technieken en benaderingen die zijn opgenomen in de lijst van bijlage I en die voor een bepaalde reeks door mensen gedefinieerde doelstellingen output kan genereren, zoals inhoud, voorspellingen, aanbevelingen of beslissingen die van invloed zijn op de omgeving waarmee wordt geïnterageerd.’

De verboden systemen worden dan verder gedefinieerd aan de hand van de middelen die worden ingezet: subliminaal, gericht op kwetsbare personen, beoordelingscriteria van mensen gebaseerd op sociaal gedrag en biometrische identificatie van potentiële daders van strafbare feiten. Het ‘hoge risico’ wordt in Bijlage III gedefinieerd

aan de hand van mogelijke schending van grondrechten en dat gebeurt ook aan de hand van de doelstelling van de systemen die weer zijn ondergebracht in categorieën, bijvoorbeeld ‘biometrische identificatie op afstand’ of in de rubriek ‘beheer en exploitatie van kritieke infrastructuur’. De verordening bevat ten aanzien van de aanbieders van ‘hoge risico’-systemen transparantie- en controlevoorschriften. Zij krijgen in hoofdstuk 2 van Titel III specifieke administratieve verplichtingen opgelegd die betrekking hebben op beheer, beperken van risico’s en toegankelijkheid van de gebruikte software. Ook legt zij bijzondere zorgvuldigheidsverplichtingen op aan de gebruikers van AI-systemen met bijzondere risico’s. Artikel 14 schrijft permanent menselijk toezicht voor.

Het bezwaar van deze definitie(s) is dat het de (data) input als onderdeel van een AI-systeem niet duidelijk benoemt en het risico niet in het systeem zelf legt, maar in het gebruiksdoel. Je bent dus een ‘slagersmes’ aan het reguleren, omdat je er zowel vlees mee kunt snijden als een moord plegen. Je kunt al die verboden doelstellingen ook techniekonafhankelijk reguleren.

Toch zijn grote maatschappelijke gevaren aan AI verbonden en dit brengt mij op algemene AI-beginselen.

3. Algemene AI-beginselen

Bij AI gaat het dus niet om een substituuat van de menselijke geest (ik ga hier voorbij aan de vraag hoe je die ‘menselijke geest’ precies filosofisch moet definiëren), maar om een krachtige door de mens te controleren informatiemachine, zowel op het vlak van de data in- en output, als op het vlak van de daarbij gehanteerde automatische ordeningsprincipes en beslisregels. Maar AI drukt ons wel op het ontologische feit dat mens en wereld steeds meer gezien worden als ‘informatie’, of zoals David Chalmers het in zijn boek *Reality +* uitdrukt van ‘it from bit’ of voluit: ‘Every physical quantity, every it, derives its ultimate significance from bits, binary yes-or-no indications, a conclusion we epitomize in the phrase, *it from bit*. (...) According to the it-from-bit consciousness view, *mind* in some form comes first, and physics derives from it.’¹³ We hoeven de ontologische consequenties van deze stelling niet te aanvaarden om te kunnen vaststellen dat een steeds groter deel van de maatschappelijke en

natuurlijke werkelijkheid die wij waarnemen en beïnvloeden, waarin wij leven en waarop wij ons gedrag en beslissingen afstemmen, gegenereerd wordt door digitale (bits) informatiemachines. Het is tegelijkertijd én echte én virtuele realiteit. Anders dan virtuele werkelijkheid-specialisten, meen ik dat er altijd een mengvorm is van de echte en van de virtuele werkelijkheid. Sociale media zijn daarvan een goed voorbeeld: die bestaan zowel in de constante werkelijkheid van personen met maatschappelijke achtergronden als in de constante informatie-uitwisseling tussen hen via digitale media. De deelnemers leven permanent in een mix van die twee. Chalmers: ‘Importantly, both the Apple reality and the Google reality are part of objective reality. (...) Nothing in our discussion of virtual worlds within virtual worlds suggests that there is not a base reality at the top of the chain.’¹⁴

Dit heeft implicaties voor alle rechtsregels op het gebied van dataverwerking (bv auteursrecht, privacy, informatievrijheid, besluitvorming). Daarnaast onderscheid ik in het geciteerde artikel een paar fundamentele algemene AI-regels¹⁵:

Transparantiebeginsel

Ik heb al eens bepleit dat dit een beginsel van behoorlijk bestuur zou moeten zijn om overheidsbesluiten te beoordelen.¹⁶ In dat kader moet ook de ‘algoritmische’ basis van een bestuursbeleid worden beoordeeld, maar helaas is met de aanpassing van de Awb met dit aspect in het geheel geen rekening gehouden.¹⁷ Maar het moet ook gelden voor andere rechtsgebieden, die steeds meer door automatische processen worden beheerst. Dat transparantiebeginsel houdt dus kort gezegd in, dat verantwoord moet (kunnen) worden hoe en met welke technische middelen beslissingen (die feitelijke en rechtsgevolgen voor individuen hebben) tot stand zijn gekomen.

Verificatiebeginsel

Ik bepleit in aanvulling op het transparantiebeginsel een algemeen menselijk verificatiebeginsel. Als zodanig zie ik in artikel 22 van de

14 Chalmers a.w., p. 233.

15 Zie het in noot 3 geciteerde artikel.

16 E.J. Dommering, ‘Naar een algemeen transparantiebeginsel?’, in: *Nederlands Tijdschrift voor Bestuursrecht (ntb)*, 2022, 5, p. 265-271.

17 Zie in de in noot 3 geciteerde aflevering van Computerrecht, het artikel van Johan Wolswinkel ‘De grote onbekende in de vergelijking: geautomatiseerde besluitvorming binnen de Wet versterking waarborgfunctie Awb’. In gelijke zin: Mirte Houweling & Lotte Kuijl, ‘Algoritmische besluitvorming. Een leemte in de bestuursrechtelijke rechtsbescherming’, in: C.J.H. Janssen (red), *Digitalisering en recht*, p. 117-130 en de door hen aangehaalde literatuur.

AVG (verbod op automatische beslissingen ten aanzien van zaken die iemands privacy raken) de uitdrukking van een algemener beginsel dat ook andere grondrechten regardeert, met name de waarheidsvinding die door het recht van vrije meningsuiting wordt beschermd. Een mens moet controleren of de feiten en de rechtsregels waar de beslissing op steunt wel kloppen. Een voorbeeld van een geleidelijke toepassing van het verificatiebeginsel door de rechter, zou je de door het EU Hof ontwikkelde jurisprudentie rond de zoekmachine kunnen noemen, die meer verantwoordelijkheid bij de exploitant van de zoekmachine is gaan leggen voor het zoekresultaat in het kader van bescherming van het privéleven.¹⁸ (...)

Maar dit is mijns inziens niet voldoende, omdat AI-systemen vervangers lijken te worden van menselijke waarheidsvinding, terwijl ze dat niet zijn. (...) Ze zijn ten opzichte van zoekmachines eigenlijk een stap terug. Die presenteren immers een zoekresultaat waarvan de feitelijke en logische uitgangspunten door de menselijke vrager geverifieerd kunnen worden, omdat bronnen worden getoond, althans door de vrager kunnen worden gevonden.¹⁹ (...) Het is geen *zoekmachine* maar een *antwoordmachine*. Om de machine te kunnen controleren moet je dus weten hoe de machine tot een antwoord is gekomen. (...)

En eenzelfde verhaal is bij ‘beeld’ en ‘geluid’ te houden. We kunnen aan met AI gefabriceerd digitaal beeld of geluid niet meer zien of horen of het ‘echt’ is.

Het behoeft geen betoog dat het transparantiebeginsel en het verificatiebeginsel nauw samenhangen, maar ze moeten m.i. toch van elkaar onderscheiden worden, omdat niet alleen de ‘black box’ moet worden geopend, maar ook de middelen aanwezig moeten zijn om de juistheid van de inhoud daarvan langs menselijke waarneming te verifiëren. Dit kan in verschillende vormen gebeuren. De zwaarste vorm is dat er altijd (een weg naar) bronnen moet worden geopend. Een lichtere vorm is een waarschuwingsplicht en adequate

18 HvJEU 13 mei 2014, C-131/12, NJ 2014, 385, (Google I); HvJ EU 24 september 2019, C-136/17 (Google/Frankrijk I), NJ 2019, 434; HvJ EU 24 september 2019, C-507/17 (zaak Google/Frankrijk II), NJ 2019/435, beide met noot E.J. Dommering.

19 Daar is natuurlijk ook nog het nodige over te zeggen, toch blijft dit een belangrijk na te streven uitgangspunt. Er is een interessante analyse van de rol van Wikipedia bij het behoud van ‘feiten’ op het World Wide Web, Heather Ford, *Writing the revolution. Wikipedia and the Survival of Facts in the Digital Age*, Cambridge/Massachusetts: MIT Press 2022.

voorziening op verzoek bronnen te ontsluiten.²⁰ En er is bij een begane fout correctie en rectificatieplicht.

Beginsel van menselijke tussenkomst

Ik gaf hiervoor al aan dat dat de regel over automatische besluitvorming in artikel 22 AVG de uitdrukking is van een veel algemener beginsel dat degene die geraakt wordt in individuele rechten door een automatisch dataverwerkingsproces de mogelijkheid moet hebben beroep te doen op een menselijke (correctieve) beslissing. In de Amerikaanse ‘AI Bill of Rights’ neemt het beginsel aangeduid als ‘Human alternatives, Consideration and Fall Back’ een prominente plaats in.²¹

4. AI-beginselen in de Gw: moet de Grondwet aan de technologie worden aangepast?

Hoe moet de Gw met dit fenomeen omgaan? In dit verband rijst de algemene vraag hoe de Gw zich tot technologische ontwikkelingen moet verhouden. Ook dat is niet geheel nieuw voor mij.²² In de bundel *150 jaar grondwet* uit 1998 tijdens de opkomst van het internet heb ik mij al eens met die vraag beziggehouden in een artikel ‘De Grondwet en de informatiemaatschappij’. Die bundel stond sterk in het teken van de technologie, waarvan in de inleiding van J.H.P. Donner²³ een beeld wordt geschetst alsof het om een sciencefiction verhaal van Jules Verne gaat:

‘Drijvende kracht achter die ontwikkelingen zijn technische veranderingen. Ontwikkelingen op het punt van telecommunicatie, miniaturisering en informatieverwerking breiden het menselijk kunnen in verschillende richtingen fundamenteel uit. Daardoor kunnen wij ons niet alleen fysiek verplaatsen, maar ook in beeld, woord en geschrift. Afstanden zijn ook niet alleen overbrugbaar geworden; ze nemen absoluut in betekenis in het maatschappelijk beeld af. We kunnen niet alleen ver zien en – horen, maar voortaan ook elders actief ingrijpen zonder dat men zich daarvoor hoeft te verplaatsen. Beperkingen van tijd, afstand, materiaal en verwerking

20 Ik wijs op een recente uitspraak van HvJEU over de omvang van het ‘inzagerecht’ volgens de AVG, HvJEU 4 mei 2023 NJ 2024, 1, m.nt E.J. Dommering: omvang inzagerecht persoonsgegevens.

21 <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>

22 Egbert Dommering, ‘De Grondwet in de informatiemaatschappij’ in: M.C. Burkens e.a. (red.), *Gelet op de grondwet, 150 jaar Grondwet*, Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijkszaken 1998, hoofdstuk 7.

23 A.w. noot 22, p. 13.

van gegevens nemen daardoor in betekenis af voor de organisatie van het maatschappelijk verkeer. Onder invloed daarvan veranderen de economische dynamiek en de structuren van het dagelijkse leven, beperkingen die tot dusver impliciet gesteld waren, vallen weg. Nieuwe morele vragen en verantwoordelijkheden ontstaan ook naarmate het Faustiaanse ideaal van het scheppen van leven binnen handbereik komt.’

Teruglezend na zoveel jaar vraag ik mij af of dat de juiste aanpak is. Het artikel dat ik toen vanuit het perspectief van het beginnende internet schreef, vind ik nu sterk gedateerd. Toegegeven moet worden dat het grondrechtenhoofdstuk in de Gw (en daartoe beperk ik mij, omdat de discussie over AI m.i. op dat niveau moet worden gevoerd) een wonderlijk allegaartje is van geschiedenis, techniek, beginselen en het niet regelen van het probleem. De artikelen 1, 3 en 4 bevatten het fundamentele gelijkheidsbeginsel, artikel 2 de rechten verbonden aan de nationaliteit, de artikelen 6 en 23 de godsdienst en onderwijsvrijheid. Artikel 11 de lichamelijke integriteit, artikel 12 de bescherming van de particuliere woning, artikel 14 het eigendomsrecht, artikel 15 het recht van vrijheid van beweging. Dit zijn allemaal artikelen die de inhoud van het recht (min of meer) regelen. Artikel 7 over de vrijheid van meningsuiting bevat geen inhoudelijke regel, alleen een sterk censuurverbod. Hier komt de techniek om de hoek kijken door de drukpers een betere bescherming te geven dan radio, televisie en ‘andere middelen’. Artikel 13 is ook techniek gebonden omdat het spreekt over onschendbaarheid van het briefgeheim en het telegraaf- en telefoongeheim. De artikelen 8, 9 en 10 benoemen de rechten van vereniging, vergadering, betoging en privacy zonder deze rechten inhoudelijk te definiëren; zij zeggen dat de wetgever dat moet doen. Door het toetsingsverbod van artikel 120 Gw is er ook geen update via jurisprudentie van de nationale rechter. De nieuwe ontwikkelingen werden door de nationale rechter via de artikelen 93 en 94 Gw door rechtstreekse toepassing van het EVRM (dat voor fundamentele rechten wel inhoudelijke normen bevat) ingepast.²⁴ Voorts is er de werking van het Europese Handvest via het VwEU.

24 Hiervoor noemde ik de Toeslagenaffaire als een gecombineerd probleem van AI en menselijke toepassing. In dit verband is relevant de beslissing in de zaak LB/Hongarije, EHRM (Grand Chamber) 9 maart 2023 (nr. 36345/16), NJ 2024, 144 m.nt E.J.Dommering, waarin het Hof een fiscale strafmaatregel in strijd acht met artikel 8 EVRM omdat de wet waar deze op was gebaseerd niet voorzag in een proportionaliteitstoetsing. Naar mijn oordeel had de ABRs op grond van artikel 94 Gw in de Toeslagen Affaire de opgelegde sanctiemaatregelen veel eerder kunnen en moeten toetsen aan het EVRM, m.n. artikel 8 EVRM en artikel 1 van het Eerste Protocol bij dit verdrag.

Terugkerend naar de vraag of en hoe AI (Niet Faust maar Dracula!) in de Gw moet worden ingepast, denk ik dat deze niet aan een specifiek vrijheidsartikel moeten worden gekoppeld maar op het niveau van het gelijkheidsbeginsel moeten worden geregeld. Het gaat immers om twee fundamentele beginselen: toepassing en beperking van rechten *door mensen* en bescherming van de *integriteit* van het beeld van de werkelijkheid. Met dit laatste aspect maakt AI wel een blijvend onderdeel van de communicatievrijheden uit.

Ik ben zo vrij om mijn hiervoor geformuleerde beginselen om te zetten in twee korte grondwetsartikelen, die passen in de geserreerde stijl van onze Grondwet.

I De uitoefening en beperking van de bij of krachtens de Grondwet beschermde rechten moet voor mensen doorzichtig zijn en kan slechts plaats vinden door menselijke tussenkomst.

Toelichting: Dit zijn het hiervoor besproken transparantiebeginsel en het beginsel van menselijke tussenkomst. Dit artikel beoogt vast te leggen dat bij alle ‘automatische’ (AI gestuurde) uitoefening of beperking van rechten altijd een menselijk subject als eindverantwoordelijke moet worden aangewezen. Het is dus een eis aan de kant van de uitoefening, bijvoorbeeld Chatbots, verspreiding van informatie, bewerking van persoonsgegevens, AI redactie van informatie enzovoort. Het is ook een eis aan de kant van beperking (inclusief geschilbeslechting waar het laatste woord is aan een mens). Daarbij hoort dus ook de transparantieregel. Het interessante van deze regels is dat zij aansluiten op de regels die de science fiction schrijver Isaac Asimov al in 1940 formuleerde voor robots. De regels luiden, voor zover hier van belang: Een robot (nu dus AI) moet de instructies van mensen opvolgen. Interessant is dat de James Bridle die in zijn boek *New Dark Age* Asimov noemt zijn regels aanvult met een transparantie regel: Een robot moet zijn werking aan mensen kunnen uitleggen.²⁵

II Hij die al of niet door tussenkomst van communicatietechniek informatie openbaar maakt is gehouden de bronnen waarop deze informatie berust al of niet op verzoek openbaar te maken, behoudens beperkingen gesteld bij de wet.

Dit is het hiervoor genoemde transparantie-verificatiebeginsel. Het moet een algemeen artikel zijn omdat het zich niet alleen richt tot de

media, maar ook tot alle burgers en organisaties die informatie bewerken, opslaan en verspreiden. Het vermijdt een inhoudelijke ‘waarheidsnorm’ te formuleren omdat dat strijdig is met de vrijheid van meningsuiting, maar het is dus wel een middel om te achterhalen of het om fake gaat. De opgenomen beperking bij wet ziet op geheimhoudingsplichten, bronbescherming van journalisten, Wet open overheid etc.

De rechtvaardiging voor het opnemen van een dergelijke informatienorm, zou je de consequentie van de vrijheid van meningsuiting kunnen noemen. Die is in de 18^e eeuw ontwikkeld. Zij werd zo wezenlijk geacht voor meningsuitwisseling en bevordering van kennis over mens, samenleving en natuur dat het een grondwettelijke norm in een democratie is geworden. Je zou het daarom de ‘encyclopedieregel’ kunnen noemen.²⁶ Nu wij in een werkelijkheid leven waarin het niet alleen gemakkelijk is door middel van informatiemachines een volstrekt vals beeld van de werkelijkheid te maken, maar daar ook op grote schaal gebruik van wordt gemaakt, is een aanvullende bescherming van die vrijheid op zijn plaats.

5. Een algemene slotopmerking

‘But what if this is not true at all? A close reading of computer history reveals an ever-increasing opacity allied to a concentration of power, and the retreat of power into ever more narrow domains of experience. By reifying the concerns of the present in unquestionable architecture, computation freezes the problems of the immediate moment into abstract, intractable dilemmas; obsessing over a small class of mathematical and material conundrums rather than the broader questions of a truly democratic and egalitarian society.’²⁷

AI staat als uitvinding niet op zichzelf. Sterker nog dan in het boek *Atlas of AI* van Kate Crawford, analyseert de in de aanhef geciteerde Bridle in *New Dark Age* dat AI het product is van de explosieve groei van netwerken en datatechnologie. Dit heeft geleid tot een concentratie van macht in wereldwijde niet aan staatssoevereiniteit gebonden netwerken in handen van een paar internationale spelers (Google,

²⁶ Zie over de geschiedenis van de encyclopedie, Egbert Dommering, *Het Verschil van Mening*, Amsterdam: Prometheus 2017 hoofdstuk 9-11.

²⁷ James Bridle *New Dark Age, Technology and the End of the Future* London: Verso 2023 (updated version), p. 34.

Meta enz.) die daarmee zowel de commerciële en politieke communicatie in handen hebben, ook in Nederland. Bridle vat dat samen in het citaat waar deze paragraaf mee begint. Hij zegt: dat probleem proberen we te verdoezelen door die macht te herleiden tot wiskundige puzzeltjes (gedomineerd door een paar whizzkids zeg ik erbij), maar dan missen we de grotere vragen van een democratische op gelijkheid gebaseerde samenleving die aan de orde zijn. Dat betekent dat AI een onderdeel is van de onderwerpen die door de artikelen 6-10 Gw worden geadresseerd. Wij kunnen wel plechtig in de Gw opschrijven dat de Nederlandse staat moet zorgen voor pluriforme media en privacy van de burgers, maar het is een illusie te denken dat daarmee de kracht van deze internationale netwerken en diensten wordt beperkt. Deze hebben ook belangrijke invloed op het functioneren van het democratisch systeem als geheel (hoofdstuk 3 van de Gw) omdat zij sociale bewegingen buiten de staatkundige democratische structuur mobiliseren. Wij zullen dus moeten proberen deze problemen in de Grondwet aan te pakken om een steviger constitutionele basis te creëren deze netwerken op nationaal niveau te reguleren.²⁸

Het heeft ook een nieuwe Europeesrechtelijke dimensie. Wij hebben geleerd te denken in het monistische schema van artikel 94 Gw-EVRM (de Nederlandse rechter toetst de Nederlandse wet rechtstreeks aan het EVRM), maar deze gehele materie is zo langzamerhand EU recht (en het daarin opgenomen Europese Handvest) geworden. Ook dat vergt een nieuwe plaatsbepaling van de Nederlandse Gw.²⁹

28 Ik noem enige recente studies over het machtsblok van de sociale media: Zeynep Tufekci, *Twitter and Gas. The Power and Fragility of Networked Protest*, New Haven/Londen: Yale University Press 2017; Jonathan Taplin, *Move and Break Things, How Facebook, Google and Amazon cornered Culture and undermined Democracy*, New York: Little, Brown and Company 2017; Philip Napoli, *Social Media and the Public Interest. Media Regulation in the Disinformation Age*, New York: Columbia University Press 2019; Richard Seymour, *The Twittering Machine*, Londen: The Indigo Press 2019; Rana Foroohar, *Don't be Evil*, New York: Currency 2019, Philip N. Howard, *Lie Machines*, New Haven and Londen: Yale University Press 2020; James Ball, *The System, who owns the internet, and how it owns us*, Londen: Bloomsbury Publishing 2020; Max Fisher, *The Chaos Machine The inside story how social media rewired our minds and our world*, Londen: Quercus Editions 2022; Emily Hunt, *The Influencer Industry*, Princeton: Princeton University Press 2023.

29 Zie: Hanco Jürgens, 'Het debat over het Nederlandse staatsrecht kan niet zonder gevolgen blijven', in: *Opiniepagina NRC* 21 maart 2024.