A photograph of a bright blue sky filled with numerous small, white, fluffy clouds. The clouds are scattered across the frame, with some larger, more prominent ones in the upper right and lower left corners. The overall scene is bright and clear.

data is macht

Colofon

Redactie en Projectleiding: Dhoya Snijders

Tekst- en taalredactie: Annette Potting, Japke Schreuders

Met medewerking van: diverse auteurs

Vormgeving en illustraties: Ellen Bouma grafisch vormgeven en dtp

Drukwerk: Quantes

Afbeeldingen: De kunstwerken in deze publicatie en op de kaft zijn gemaakt door het zelflerend neurale netwerk *Inceptionism* van Google Research.

ISBN 978-94-91397-15-8

STT-publicatie nr. 86

NUR 950

Trefwoorden: Big Data, digitalisering, kunstmatige intelligentie, toekomstvisie, technologische ontwikkelingen, maatschappij.

© 2017, Stichting Toekomstbeeld der Techniek, Den Haag



Data is Macht. Over Big Data en de toekomst van Stichting Toekomstbeeld der Techniek wordt auteursrechtelijk beschermd zoals vastgelegd onder de Creative Commons Naamsvermelding Niet Commercieel-Geen Afgeleide Werken 3.0 Unported licentie. U kunt dit werk toeschrijven aan Stichting Toekomstbeeld der Techniek / Dhoya Snijders (www.stt.nl), 2017. Bezoek www.creativecommons.org/licenses/by-nc-nd/3.0 voor de volledige tekst van de licentie

Stichting Toekomstbeeld der Techniek

Prinsessegracht 23, 2514 AP Den Haag

Postbus 30424, 2500 GK Den Haag

070-302 98 30

info@stt.nl

www.stt.nl

Data is macht

Over Big Data en de toekomst

Stichting Toekomstbeeld der Techniek

Dhoya Snijders, redactie

INHOUDSOPGAVE

Voorwoord	6
Samenvatting	8
1. Data als DNA	13
2. De staat van data	31
3. Wat telt? Over data en sociale gevolgen	53
4. Tijdsdimensies: willen we meer of minder data?	79
5. Nieuwe verhalen	99
6. Wat nu?	119
Bijlagen	123
Aanpak	123
Dankwoord en betrokkenen	124
Stichting Toekomstbeeld der Techniek	125
Literatuur	128
De auteur	132



Toekomstbeelden

Wat zijn de belangrijkste data-vraagstukken voor de toekomst? Tussen de hoofdstukken door zijn bijdragen van experts opgenomen die met antwoorden en toekomstbeelden komen:

Dirk Helbing over een alternatieve digitale revolutie, pag. 22

Djeevan Schiferli over persoonsgebonden advies, pag. 24

Jeroen Dijkxhoorn over de toekomst van persoonsgegevens, pag. 26

Jeroen van den Hoven over Big Data en ethiek, pag. 43

Maurits Kreijveld over blockchain en prediction markets, pag. 46

Emma AI over de toekomst van AI, pag. 48

Mireille Hildebrandt over persoonlijke digitale assistenten, pag. 70

Eric van Tol over Big Data en de dienstensector, pag. 72

Rogier Creemers over het sociale kredietsysteem in China, pag. 74

Maurits Martijn over de dag dat Google valt, pag. 91

Theo de Vries over de onomkeerbaarheid van ICT, pag. 93

Marijn Janssen over Big Data en de maakbare samenleving, pag. 110

Brenno de Winter over Big Data als privacykans, pag. 112

VOORWOORD

Bij nagenoeg alles wat we doen genereren we data. Of het nu gaat over een bezoek aan de dokter, boodschappen doen of een reis plannen en maken, bij al deze activiteiten worden data gegenereerd die vervolgens ook weer ergens worden opgeslagen. Al deze data brengen ons nieuwe mogelijkheden in de vorm van diensten en services, maar er ontstaan ook ethische dilemma's rond eigenaarschap en gebruiksrecht. Zo kan het vrij beschikbaar maken van medische dossiers helpen om resultaten van klinische studies te vergelijken, maar wanneer verzekeraars hier ook toegang toe hebben kan dat leiden tot selectief uitsluiten van mensen of duur maken van bepaalde verzekeringen. Daarnaast geeft het bezit van data een zekere manipulatieve macht die zich uiteindelijk ook kan uiten in de vorm van geopolitieke spanningen. Er zijn in dit opzicht voldoende voorbeelden van regeringen die inzicht in (internet-)data van burgers gebruiken om bepaalde bevolkingsgroepen te controleren en te onderdrukken.

Op dit moment worden data vooral ingezet om gevraagd en ongevraagd op de persoon toegespitste proposities te doen. Dat kan variëren van slimme weekaanbiedingen van de supermarkt tot reisaanbiedingen die in de vorm van pop-ups als lastige horzels je beeldscherm blijven bevolken. Je kunt dit meestal met enige moeite uitzetten, maar de balans tussen nut en nuisance is niet voor iedereen hetzelfde. Er zijn ook voorbeelden van websites waarop patiënten (maar soms ook artsen) slim doktertje spelen door gegevens te delen en daarmee tot een snelle en betere diagnose te komen. Door gebruik te maken van de kennis en ervaringen van velen biedt Big Data zeker aanvullende mogelijkheden ten opzichte van de huidige medische praktijk. Dit is de situatie van vandaag; de grote vraag is echter waar het in de toekomst naartoe gaat met onze data, ofwel "data, quo vadis?"

In deze toekomstverkenning hadden we te maken met een lastig dilemma: gaat het nu om de data als zodanig of



zijn het juist de handige bewerkingen die maken dat deze data waardevol zijn? Mijn persoonlijke beeld hierbij is toch vooral dat data gezien kunnen worden als de noodzakelijke brandstof voor allerlei slimme bewerkingen. De eerste echte grote stappen die we hierbij zien hebben te maken met *deep learning* en *artificial intelligence*. In het geval van deep learning zijn eerder gegenereerde *real-life data* de toets voor een systeem. Door het kunstmatig gegenereerde antwoord steeds te vergelijken met de werkelijkheid treedt een leereffect op waarmee het vervolgens mogelijk is om steeds betere antwoorden of voorspellingen te geven. Ook kan hiermee steeds beter worden ingespeeld op een situatie-specifieke context. Dit alles maakt het systeem steeds 'intelligenter'.

Waar gaan we uiteindelijk naartoe? Hierover lopen de meningen nogal uiteen. Zelf zie ik data als het vehikel waarmee we uiteindelijk de grens tussen mens en machine zullen slechten. Technieken als deep learning zijn hiervoor cruciaal, maar ook de verbinding tussen neurowetenschappen, high-techsystemen en (nano)materialen zal hiervoor van groot belang zijn. Wanneer we ons brein draadloos zouden kunnen koppelen met datasystemen en informatie uit ons brein zowel kunnen zenden als ontvangen, levert dit ongekende mogelijkheden op. Het zal duidelijk zijn dat de ethische dilemma's die we nu rond data zien alleen maar heftiger naar voren zullen komen. Zo verwacht ik dat een belangrijk dilemma zal ontstaan rond de vraag op welk gebied systemen beslissingen zullen gaan overnemen van de mens; het besturen van een auto of een vliegtuig is nog wel te overzien, maar het bepalen van een medisch behandelplan (of misschien wel juist het stoppen daarvan) door een artificieel systeem zal al minder snel worden geaccepteerd.

In deze toekomstverkenning worden een aantal beelden geschetst over hoe de wereld er ten aanzien van data en daaraan gerelateerde zaken over 25 jaar uit zou kunnen zien. Backcasten van deze toekomstbeelden naar het nu maakt het voor onderzoekers en beleidsmakers mogelijk langetermijnsscenario's te ontwikkelen om ons daarmee tijdig voor te bereiden op een sterk veranderende toekomst.

Ik wens u heel veel leesplezier, waarbij ik ervan overtuigd ben dat het voorliggende boekwerk u zal aansporen tot creatief denkwerk.

Jos Keurentjes

Chief Scientific Officer, en lid van Raad van Bestuur TNO

SAMENVATTING

We leven in tijden van digitale overvloed en de digitalisering van onze wereld lijkt geenszins af te nemen. De hoeveelheid data op aarde verdubbelt zich nu iedere tweeëneenhalf jaar en de snelheid waarmee dit gebeurt neemt toe. Binnen tien jaar zijn er naar verwachting zo'n 150 miljard sensoren via netwerken verbonden en de hoeveelheid data die hiermee geproduceerd wordt komt met nieuwe ambities. Nieuwe technologie brengt de mogelijkheid om alles wat we zien, doen, horen, denken en uitspreken te digitaliseren. Het antwoord op de vraag 'wat kunnen we niet in data vatten?', moet onder voortschrijdend inzicht steeds worden aangepast.

Deze publicatie presenteert de uitkomsten van een toekomstverkenning naar de sociale impact van Big Data. Aan de hand van de resultaten van interdisciplinaire foresight-workshops, tientallen interviews met experts en externe bijdragen van dertien vooraanstaande denkers schetst de publicatie een palet aan toekomstbeelden met belangrijke datavraagstukken voor morgen, maar ook voor vandaag. Uit de toekomstbeelden en scenario's komen een aantal belangrijke lessen naar voren:

Kennisbescheidenheid. Met de opkomst van datatechnologie die rekenkracht, opslag en algoritmische nauwkeurigheid maximaliseert, krijgen we nieuwe gereedschappen om naar de wereld te kijken en daarmee een goudmijn aan kansen om tot meer kennis te komen. Toch is er een

noodzaak voor kennisbescheidenheid. Zelfs als we met duizenden data-elementen – waaronder genetische informatie, locatiedata en een goed gevuld profiel – conclusies willen trekken over mensen moeten we beseffen dat de analyse onjuist kan zijn en niet blind vertrouwen op data-technologie. Big Data verleidt ons om te denken dat we een objectief oordeel over een persoon kunnen vellen. Het is in een tijdperk van Big Data belangrijk de morele autonomie van de mens te beschermen en mensen, ongeacht de data, eigen keuzes te laten maken en hun vrijheid te bewaken.

Mens centraal. "Als je niet betaalt voor een dienst ben jij niet de consument maar het product", stelde Andrew Lewis. Een kritiekpunt in onze relatie met organisaties die data over ons vastleggen is dat men niet goed begrijpt dat onze data verhandeld worden, hoe dit gebeurt en wat de waarde van onze data is. De in deze publicatie gepresenteerde machtsscenario's presenteren alternatieve datamodellen die laten zien dat decentralisatie van datadiensten zowel materiële als immateriële waarde zou kunnen opleveren voor gebruikers en de samenleving.

Nederland, dataland? Datagedreven technologie kan een belangrijke motor zijn voor economische groei en sociale vooruitgang. Hiervoor zullen het bedrijfsleven, het maatschappelijk middenveld, de overheid en de wetenschap moeten

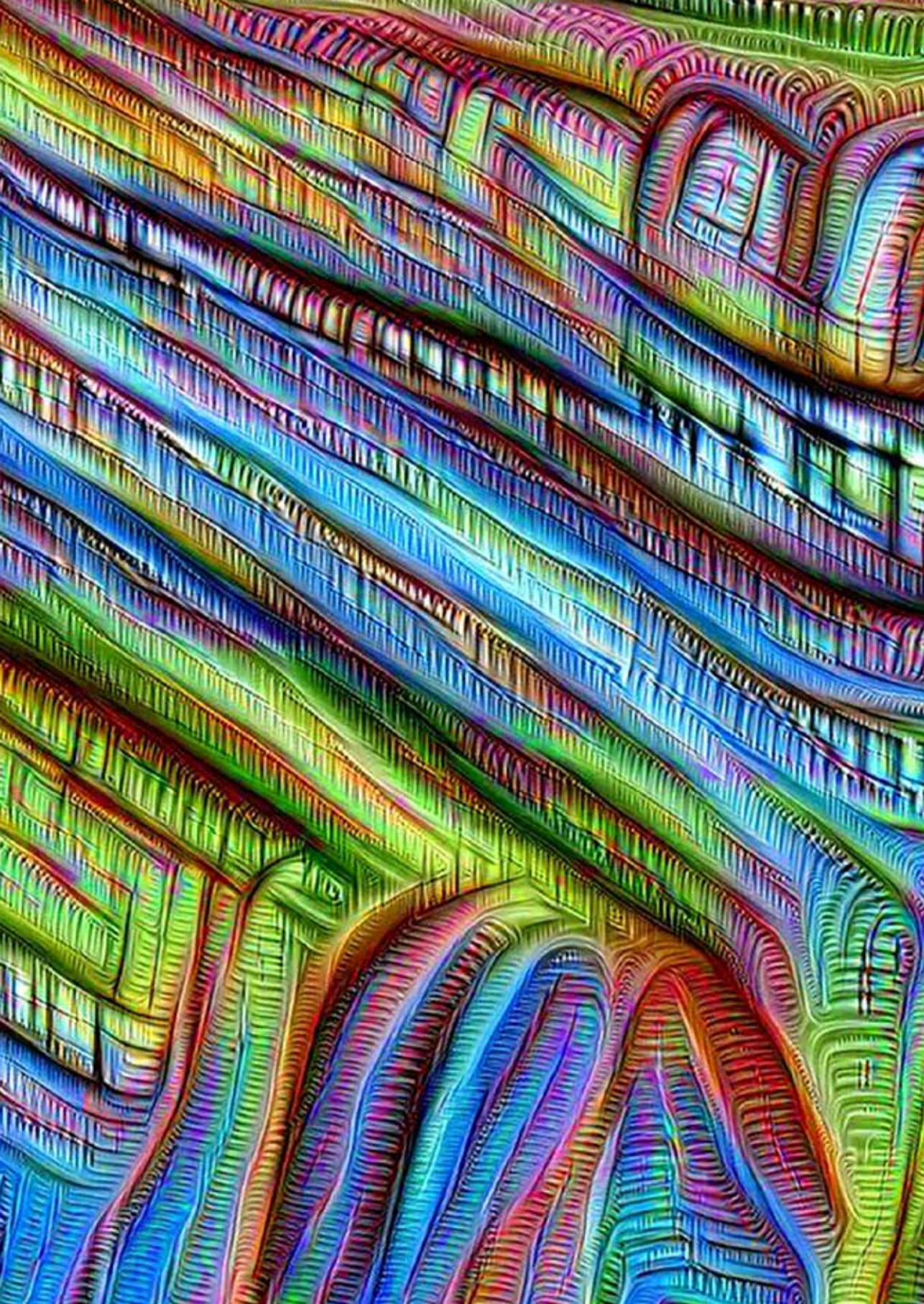
samenwerken om zowel aandacht voor technologie als gepaarde risico's te hebben. De spanning die nu ervaren wordt tussen Big Data à la Silicon Valley en onze eigen grondrechten zou een uitgangspunt kunnen bieden om lokaal tot nieuwe infrastructuur, data-applicaties, en intelligente systemen te komen die lokale waarden en de bescherming van de mens en eigen data centraal stellen.

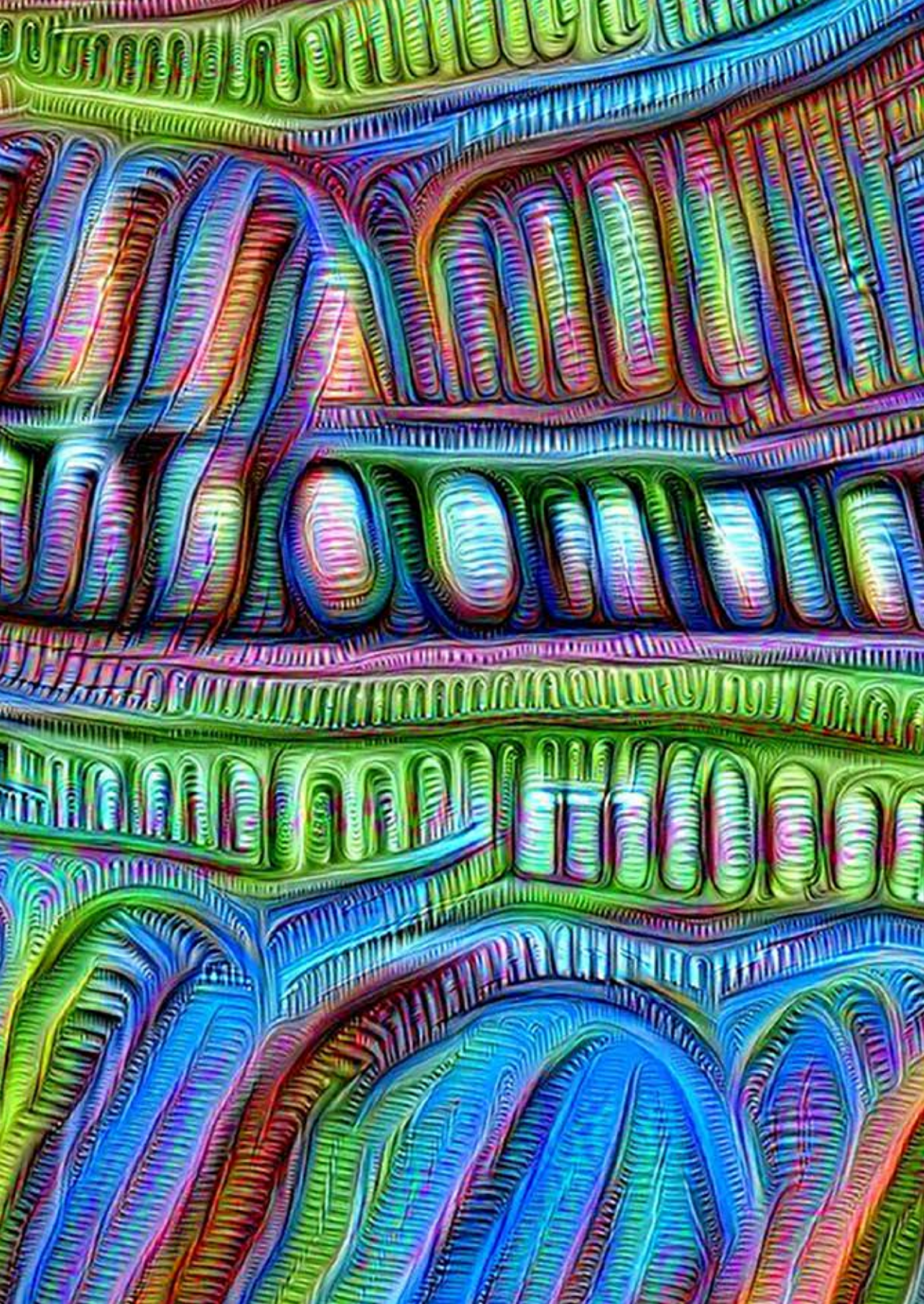
Big Data en pluraliteit. In het Big Data-veld is meer behoefte aan meerstemmigheid. Niet alleen in de discussie over het fenomeen, maar ook in de ontwerpfase, op bestuursniveau, in het aanstellen van ethische commissies is diversiteit wenselijk. Als we dit niet doen riskeren we dat machines, misschien wel heel weldenkende machines, een afspiegeling worden van een klein gedeelte van de samenleving en bij anderen juist weerstand zullen oproepen. We zouden juist het pluralisme, wat we in onze samenleving als belangrijke waarde bestempelen, in onze systemen moeten inbouwen. We zijn gebaat bij de wisdom of crowds, ook als de crowd voor een gedeelte uit slimme systemen bestaat.

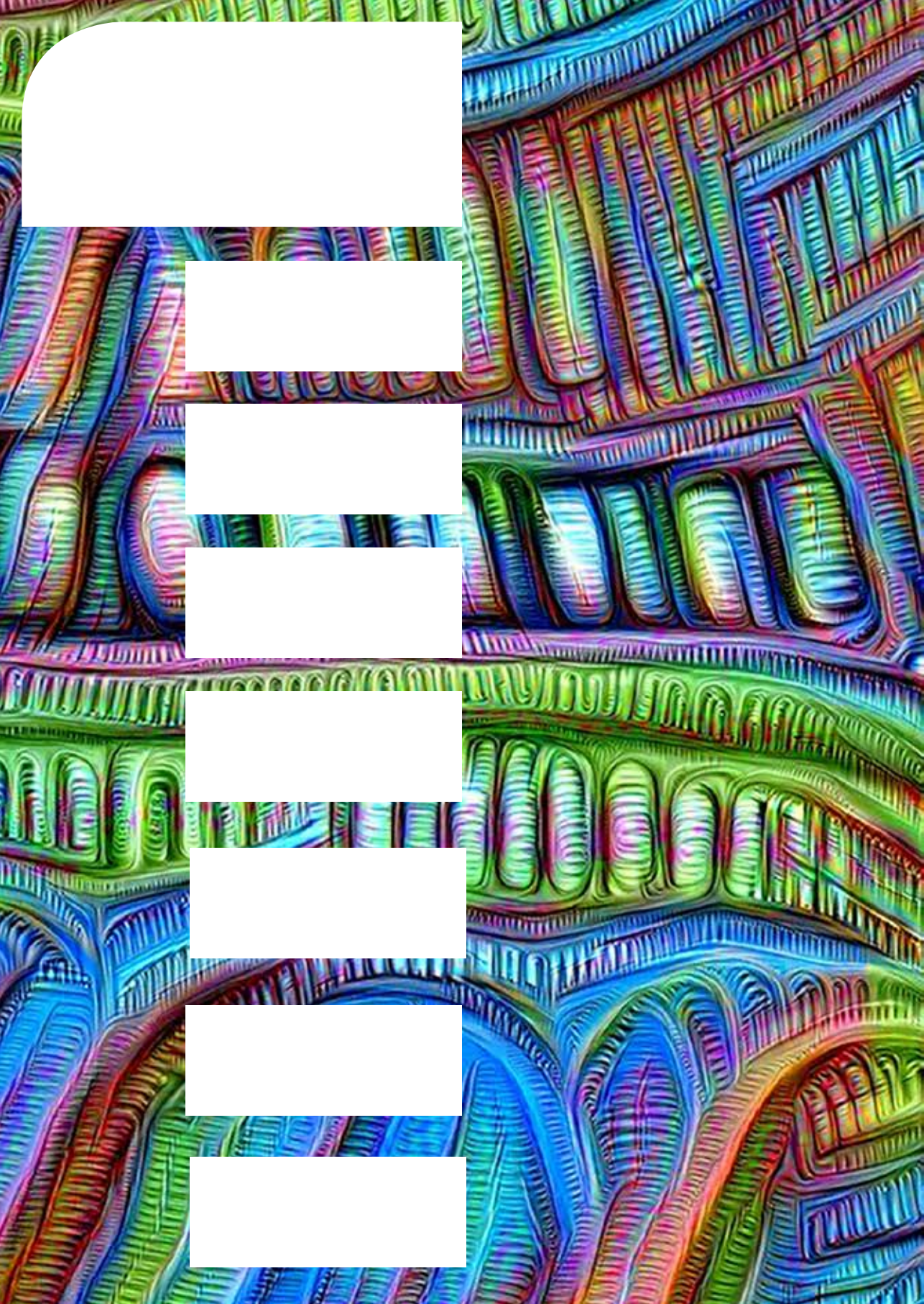
Bewust van onbewuste intelligentie. Hoewel de realisatie van superintelligentie volgens de meeste KI-experts nog toekomstmuziek is, zien we de komende jaren overal om ons heen nieuwe intelligentie opduiken. Sprekende huiskamers, denkende straatverlichting en persoonlijke

digitale assistenten staan in de virtuele rij om u van dienst te zijn. Intelligente zelflerende systemen zijn op komst en gebruikers en zelfs bedenkers van dit soort systemen zullen steeds minder grip hebben op de output hiervan. We hebben te maken met steeds autonomer wordende intelligente machines en moeten leren anticiperen op interactie met deze non-organische intelligentie.

Big Data schept hoge verwachtingen voor vele domeinen; van de zorg tot landbouw, onderwijs tot milieu, van de financiële sector tot transport. Hoewel er veel over Big Data geschreven is en veel organisaties de mogelijkheden en beperkingen verkennen ontbreekt er, vooral in de Nederlandse context, een verkenning naar de impact van deze technologie op de lange termijn. Wat ontbreekt zijn alternatieve beelden van onze samenleving, onze organisaties en onze overheden waarin verkend wordt hoe ze aan de hand van Big Data-technologie op geheel andere wijze zouden kunnen werken of worden ingedeeld. Welke bestuurstaken nemen algoritmes in de toekomst over? Zou persoonlijke informatie persoonlijk eigendom kunnen worden waaraan individuen kunnen verdienen? Hoe definiëren we privacy in 2040? In het beantwoorden van dergelijke vragen en het creëren van toekomstbeelden die ons alternatieve toekomst helpen verbeelden ligt de kracht en meerwaarde van deze publicatie.







1. DATA ALS DNA

We bouwen aan een nieuw universum. Bit voor bit digitaliseren we onze wereld en slaan deze op in virtuele ruimte. Volgens de Italiaanse schrijver Italo Calvino werken we aan 'het geheugen van de wereld', een volledig digitale kopie van onze fysieke wereld. En we gaan verder dan kopiëren. We maken modellen van onze wereld waarin we simuleren hoe het weer is, hoe het was en hoe het zal zijn. We maken bedrijfsmodellen om grip te krijgen op logistiek, orders en productie. Financiële modellen helpen ons bij het maken van investeringen. Misdaadmodellen voorspellen waar een volgende overval zal plaatsvinden. We scheppen digitale werelden met het doel om meer grip te krijgen op onze eigen werkelijkheid. En onze eigen werkelijkheid wordt steeds digitaal. Een steeds groter deel van onze communicatie is digitaal, we spenderen meer tijd online, onze huishoudens en bedrijven zijn virtueel verbonden en ons digitale alterego wordt actief of passief samengesteld.

In honderden databases, waarvan u de meerderheid waarschijnlijk niet kent, is informatie over u opgeslagen. Uw mailbox archiveert een belangrijk gedeelte van uw correspondentie en contacten, veelal met voornaam, achternaam, e-mailadres, telefoonnummer en beroep. Locatiedata toont waar u de afgelopen maanden van dag tot dag bent geweest. Ongemerkt produceert u een lange lijst van zoekresultaten en browsegeschiedenis, vaak tientallen registraties per dag, die inzicht geven in uw voorkeuren, meningen, kennis, uw carrière en sociale contacten. De data is persoonlijk:

Uw volledige naam. Huisadres. Foto's. Video's. Emailadressen. BSN-nummer. Paspoortnummer. IP-adressen. Kentekenbewijs. Gedigitaliseerd profiel van uw gezicht, vingerafdrukken, schrift. Creditcardnummers. Geboorteplaats. Genetische Informatie. Telefoonnummers. Inlognamen. Leeftijd. Geboortedatum. Naam van uw werkgever. Uw professionele geschiedenis. Uw medische dossiers. Geslacht. Etnische achtergrond. Uw strafblad. Cookies. Contacten. Status-updates. Gedachten. Gevoelens.

Big Data-toepassingen zorgen ervoor dat deze lijst zich constant uitbreidt. Aan de hand van deze informatie, krachtige algoritmes en snelle datasystemen worden nieuwe verhalen over u gemaakt. Bent u geneigd een product te kopen, heeft u een groot sociaal netwerk, bent u geschikt voor een bepaalde baan en hoe communiceert u? Data worden

steeds meer gezien als een blauwdruk van ons gedrag, als de behavioristische variant op DNA. Als je genoeg data hebt kun je inzicht krijgen in de manier waarop een organisatie functioneert of een mens handelt, spreekt, slaapt of stemt. Je kunt met data in kaart brengen wat de geschiedenis is van iemands doen en laten en als je voldoende weet kun je een poging doen de toekomst te voorspellen. Deze informatie heeft invloed op onze levens: we worden aan de hand van data-analyses gefouilleerd op luchthavens, komen in aanmerking voor korting op bepaalde producten, krijgen banen aangeboden of worden versierd op het net.

Het proces is in zekere zin onomkeerbaar. Zelfs als u uw e-mailaccounts nu zou sluiten, uw sociale-media profielen zou opzeggen en van het digitale netwerk zou verdwijnen, blijft u nog jaren bestaan in de digitale ruimte.

Meer data

De digitale ruimte dijt uit en doet dit met toenemende snelheid. Er gaat per seconde meer informatie over het internet dan twintig jaar geleden op het hele net te vinden was. Per minuut worden er meer dan 150 miljoen e-mails verzonden en 300 uur aan video-materiaal op YouTube gezet. Naar schatting hebben we meer bits opgeslagen dan er sterren zijn in het fysieke heelal.

Er is zoveel data dat we niet exact weten hoeveel het is. Velen hebben geprobeerd om een getal te geven aan de hoeveelheid informatie die bestaat. Martin Hilbert van de University of South California is hiervan een van de meest vooraanstaande onderzoekers. Hilbert schat dat we gedurende de afgelopen dertig jaar iedere tweeëneenhalf jaar een verdubbeling van digitale informatie hebben gezien. Dit komt neer op een totaal aantal van 10 zettabyte in 2016, waarbij een zettabyte in cijfers wordt uitgedrukt als 10^{21} ofwel 1.000.000.000.000.000.000 bytes (Hilbert 2016). Deze hoeveelheid is moeilijk te bevatten. Als het op dvd's zou worden gezet en we zouden die dvd's stapelen dan zouden we vijf stapels krijgen die allemaal tot de maan reiken.

Megabyte	1.000.000 bytes
Gigabyte	1.000.000.000 bytes
Terabyte	1.000.000.000.000 bytes
Petabyte	1.000.000.000.000.000 bytes
Exabyte	1.000.000.000.000.000.000 bytes
Zettabyte	1.000.000.000.000.000.000.000 bytes
Yottabyte	1.000.000.000.000.000.000.000.000 bytes