

MAATSCHAPPELIJKE WAARDEN IN DE DIGITALE INNOVATIE: WIE, WAT EN HOE?

KVAB Denkersprogramma 2019

Jan Rabaey

Rinie van Est

Peter-Paul Verbeek

Joos Vandewalle



KVAB STANDPUNTEN

66

Koninklijke Vlaamse Academie van België
voor Wetenschappen en Kunsten - 2020

MAATSCHAPPELIJKE WAARDEN BIJ DIGITALE INNOVATIE: WIE, WAT EN HOE?

KVAB - DENKERSPROGRAMMA 2019



KVAB Press

KVAB STANDPUNTEN

66

Concept cover: Francis Strauven
Ontwerp cover: Charlotte Dua
Afbeelding: Shutterstock

De tekening van het Paleis der Academiën is een reproductie van het originele perspectief van Charles Vander Straeten in 1823. Jozef Cantré ontwierp het logo van de KVAB in 1947. De KVAB Standpunten worden gepubliceerd door de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Hertogsstraat 1, 1000 Brussel.
Tel. 00 32 2 550 23 23 – info@kvab.be – www.kvab.be

MAATSCHAPPELIJKE WAARDEN BIJ DIGITALE INNOVATIE: WIE, WAT EN HOE?

KVAB - DENKERSPROGRAMMA 2019



Jan Rabaey
Rinie van Est
Peter-Paul Verbeek
Joos Vandewalle

Gedeeltelijke reproductie is toegelaten mits uitdrukkelijke bronvermelding.
Partial reproduction is permitted provided the source is mentioned.
Aanbevolen citeerwijze: Jan Rabaey, Rinie van Est, Peter-Paul Verbeek,
Joos Vandewalle, *Maatschappelijke waarden bij digitale innovatie: wie, wat
en hoe?*, KVAB - Denkersprogramma 2019, KVAB Standpunt 66, 2020.

© Copyright 2020 KVAB
D/2019/0455/05
ISBN 978 90 656 919 72

Drukkerij Universa

MAATSCHAPPELIJKE WAARDEN BIJ DIGITALE INNOVATIE: WIE, WAT EN HOE?

INHOUD

Samenvatting	7
Executive summary	9
Voorwoord	11
1. Situering, doelstellingen en aanpak van het denkersprogramma	13
2. Essay: Zoektocht naar het goede digitale leven. Zet mensen en waarden centraal bij het vormgeven van digitale innovatie	20
2.1. Ontwaken uit onverantwoordelijkheid	20
2.2. Blinde vlekken in het Nederlandse governance-landschap	22
2.3. Denken over digitale innovatie en doen in België en Vlaanderen	26
2.3.1. Nationale veiligheid, economische voorspoed en maatschap- pelijk welzijn	26
2.3.2. Nationale veiligheid: de terugkeer van de drones	28
2.3.3. Economische voorspoed: geloof in technologie én samen- leving	29
2.3.4. Maatschappelijk welzijn: waardevol digitaliseren	32
2.4. Slotgedachte: Het goede digitale leven	34
2.5. Referenties	36
3. Essay: Technologisch burgerschap en het ethisch begeleiden van digi- talisering	38
3.1. Inleiding	38
3.2. De relaties tussen de mens en digitale technologie	39
3.2.1. Het huidige ethische debat	41
3.2.2. Burgerschap in de digitale samenleving	43
3.3. De digitale openbare ruimte	43
3.3.1. De openbare ruimte online	44
3.3.2. Digitale technologie in de fysieke openbare ruimte	45
3.3.3. Digitale geletterdheid	46
3.4. Onderwijs in een digitale samenleving	47
3.4.1. Digitalisering en de onderwijspraktijk	47
3.4.2. De rol van het onderwijs in een digitale samenleving	48
3.4.3. Kritisch denken en transdisciplinariteit	49
3.5. Politiek en digitalisering	50

3.6. Ethiek als begeleiding van technologie: de herkansing van Frankenstein.....	51
3.7. Conclusie	53
3.8. Referenties.....	54
4. Essay: Maatschappelijke waarden in digitale innovatie – Het perspectief van een technoloog	55
4.1. Technologie en disruptie.....	55
4.2. Ontwerpprincipes voor een betere, veiligere en eerlijkere digitale wereld	58
4.2.1. Authenticiteit en vertrouwen creëren	59
4.2.2. Eigenaarschap versterken	60
4.2.3. Openheid en transparantie	61
4.2.4. Traceerbaarheid en responsabilisering	62
4.2.5. Ingebouwd beleid, regelgeving en ethiek.....	62
4.2.6. Opvoeding en opportunititeiten voor allen	64
4.3. De mensheid op een kruispunt	64
4.4. Referenties.....	65
5. Besluiten en aanbevelingen van het Denkersprogramma	67
5.1. Wat?	67
5.2. Wie?.....	67
5.3. Hoe?.....	68
Appendix 1 - CV van de Denkers	71
Appendix 2 - Samenstelling van de stuurgroep	73
Appendix 3 - Relevante literatuur	74

Samenvatting

De digitalisering is een disruptief proces dat veel aspecten van de samenleving grondig beïnvloedt. Dat ervaren we in de covid-19-pandemie aan den lijve. Aan de basis daarvan liggen de toenemende rekenkracht van chips en computers en de veel grotere collectie en opslagcapaciteit van data, wat een veel krachtiger verwerking van gegevens mogelijk maakt. Daardoor worden nieuwe en betere diensten en vormen van communicatie en samenwerking mogelijk. We beleven een digitalisering waarin alle informatie, zoals spraak, muziek, beeld, video, data en ook metadata, in digitale vorm voorgesteld kan worden. De computerkracht en de algoritmes die van die kracht gebruik maken, kunnen deze data snel verwerken en zorgen zo voor nieuwe diensten en voor de automatisering van heel wat processen. Ze nemen veel standaardtaken van mensen over, maar creëren tegelijk ook veel nieuwe taken. De computers kunnen ook leren uit voorbeelden en overtreffen de menselijke prestaties dankzij kunstmatige intelligentie. Vaak levert de combinatie van menselijke en kunstmatige intelligentie momenteel de beste resultaten op.

Deze transformatie is volop aan de gang. Zowel in de politiek, bij de overheid en in het onderwijs als in het bedrijfsleven en de samenleving als geheel krijgt de 'nieuwe digitalisering' recent heel wat aandacht. Het denkersprogramma dat geleid heeft tot dit Standpunt past daarin: *Maatschappelijke waarden bij digitale innovatie: wie, wat en hoe?* Bedrijven en overheidsinstellingen maken gebruik van de opportuniteiten van digitalisering om betere producten en diensten aan te bieden. De digitalisering is in veel sectoren dan ook de drijvende kracht achter innovatie en verandering. Ook klanten maken meer en meer gebruik van digitale diensten van bedrijven, instellingen en overheden: denk onder meer aan onlineshopping en -banking. Met z'n allen laten we in onze normale dagelijkse activiteiten vele digitale sporen na, zoals locaties. Nieuwe bedrijven doen hun intrede in dit marktsegment en jongeren interageren meer en meer via sociale media: om zich te ontspannen, maar ook in de onderwijs sfeer (van het basis- tot het hoger onderwijs en de naschoolse vorming). Het onderwijs is al in grote mate gedigitaliseerd en doet volop ervaring met online leren op. Ook gepensioneerden ervaren de voordelen van de digitalisering. Tegelijk mogen we niet blind zijn voor de nadelen en de gevaren van de digitaliseringsgolf, zoals hacking, het misbruik van gegevens, fake news, virussen, phishing, privacyschendingen, ongewenste profilering, gebrek aan transparantie enz.

Dit alles pleit voor een weloverwogen aanpak van de digitalisering, breed gedragen en met respect voor maatschappelijke waarden. Het gaat om een thematiek met veel facetten, invalshoeken en vormen van expertise, en met tal van belanghebbenden en beleidsniveaus. Een ontwerptraject waarin ICT-experts nieuwe systemen en diensten ontwerpen, los van hun toepassing en zonder interactie met gebruikers en sociale wetenschappers, is niet te verkiezen. Cocreatie en codesign dienen zich aan, in een multidisciplinaire aanpak waarbij de concipiëring

van de functionaliteiten, technische keuzes, gebruikerservaringen en ethische keuzes hand in hand gaan met aandacht voor publieke waarden als vrijheid, onafhankelijkheid, autonomie en gelijkheid.

De stuurgroep van het denkersprogramma maakte een selectie van drie complementaire internationale experts, Jan Rabaey, Peter-Paul Verbeek en Rinie van Est, elk met diepgaande en specifieke ervaring met dit thema: in Nederland, Silicon Valley en internationale organisaties. In samenspraak met de stuurgroep hebben zij gedebatteerd met belanghebbende experts en verantwoordelijke personen uit Vlaanderen. Op basis daarvan hebben de Denkers hun onafhankelijk oordeel gevormd en presenteerden ze hun bevindingen en aanbevelingen op een publiek slotsymposium. Mede aangescherpt door de discussies die daar werden gevoerd, werd dit Standpunt opgesteld, met gefundeerde argumenten en toekomstgerichte boodschappen voor samenleving en beleid. Elk met hun eigen achtergrond, invalshoek en competentie vormen de drie essays complementaire en tegelijk consistente analyses, zijn het waardevolle benaderingen en bieden ze aanzetten voor verdere stappen in de digitale transitie. De omvang en impact daarvan op de samenleving zijn niet te onderschatten en het volle effect is nog niet te overzien.

De Denkers brengen geen doemscenario's of schrikbeelden. Ze bepleiten wel een nieuwe relatie tussen mens en technologie, met onder andere een zich verdiepende bewustwording, een zorgzame houding en verantwoordelijk handelen. De covid-19-uitdaging confronteert ons deze maanden scherp met deze relatie, in een context van thuiswerken, afstandsonderwijs, onlineshopping en social distancing. Meer dan ooit vereist zo'n ongeziene problematiek een coöperatieve houding, interdisciplinaire samenwerking van diverse actoren en een synergetische aanpak. De drie Denkers – nog in tempore non suspecto – reiken hier vanuit hun internationale en (inter)disciplinaire ervaring vele verdiepende inzichten, inspirerende praktische voorbeelden en waardevolle aanbevelingen aan. Dat kan ons leiden naar een algemeen kader voor de ethische begeleiding van de toepassingen van digitale technologie als basis voor technologisch burgerschap.

Gezien de hoge kwaliteit van de gesprekken die de Denkers hebben gevoerd met maatschappelijke partijen en hun constructieve bevindingen is er alle reden om erop te vertrouwen dat er met dit Standpunt in Vlaanderen een stevig fundament wordt aangereikt voor het verder vorm geven van wat burgerschap in een digitale samenleving betekent.

Executive summary

Societal values in digital innovation: who, what and how?

Digitalisation is a disruptive process that has a significant impact on many aspects of society. We have experienced this first-hand during the current pandemic. At the basis lies the increasing processing power of chips and computers and the much greater data collection and storage capacity, enabling much more powerful processing of information. The result is new and better services, communication, and partnerships. We are now living through an age of digitalisation, in which information such as speech, music, images, video, data and even metadata can all be offered in digital forms. The power of the computer, and the algorithms that make use of that power, can rapidly process this digital data, creating new services and a vast array of automated processes. As a consequence they take over many of the standard tasks normally carried out by humans, while at the same time creating a host of new tasks. These computers can therefore learn from examples and, using artificial intelligence, can even outperform human beings in a number of tasks. Often though, it is the combination of human intelligence and artificial intelligence that produces the best results. It is within this current and ongoing transformation that this Thinkers' Programme has evolved, with the principal aim of debating 'Societal values in the face of digital innovation: who, what and how?' In politics, in government, in education and industry, and in society as a whole this digitalisation has recently attracted a great deal of attention. Businesses as well as public institutions want to make the best possible use of the opportunities offered by digitalisation, in order to make better products and to offer better services. In many of these sectors digitalisation is the driving force behind innovation and change. Customers are also making ever more use of digital services offered by businesses or institutions, such as online shopping and banking. Citizens are also making increasingly use of digital channels to interact with government bodies. We are all leaving multiple digital traces of data and metadata such as location while performing our normal everyday activities. Furthermore, new businesses are emerging in this market segment. Young people are interacting on an increasing scale with social media. And not just for leisure. Education, from primary through secondary to higher, and extracurricular training are also now digitalised to a large extent and are building up online experience. Even the retired population in our society has experienced the benefits of digitalisation. But we should not turn a blind eye to the disadvantages and the dangers such as hacking, data misuse, fake news, viruses, phishing, privacy breaches, unwanted profiling, lack of transparency, etc. *All of this calls for a well-balanced approach that respects societal values and enjoys widespread support.* So, this is a multi-faceted theme with many angles and areas of expertise as well as multiple stakeholders and policy levels. After all, a design process in which ICT experts first develop new systems and services, without considering the application, without any interaction with the users and with social scientists, is less preferable to one of co-creation

and co-design. This is a multidisciplinary approach whereby the conception of the functionalities and the technical choices go hand in hand with the user experience and the ethical choices, with a focus on public values such as freedom, independence, autonomy and equality.

A steering group of this Thinkers' Programme selected three complementary international experts, Jan Rabaey, Peter-Paul Verbeek en Rinie van Est, each with their own extensive experience in this area in the Netherlands, Silicon Valley, and in international organisations. In consultation with the steering group, the Thinkers conducted a debate with the interested experts and stakeholders in Flanders. On the basis of this debate, the Thinkers formed their independent assessment. They presented their findings and recommendations at a public final symposium. Refined in part by the discussion that was held there, a Position Paper was drafted with well-founded arguments and forward-looking messages for society and policy. Drawing on their unique backgrounds, perspectives and competencies, the three essays therefore constitute very complimentary but also consistent analyses, valuable approaches and suggested stimuli for further steps in this digital transition. The scope of the transformation and its impact on society cannot be underestimated, nor can the total effect be predicted. However, drawing on their experience and insight, the Thinkers have not come up with doom scenarios or scare stories. On the other hand, the Thinkers call for a new relationship between society and technology that includes, among other things, a deeper awareness, a cautious attitude, and responsible action. The Covid-19 challenge has also confronted us directly with this relationship, in the context of teleworking, distance learning, online shopping and social distancing. More than ever, an unforeseen problem like this demands a cooperative attitude, interdisciplinary partnerships of diverse actors and a synergetic approach. In this context, the three Thinkers have used their international and disciplinary experience *in tempore-non-suspecto* to bring us many in-depth insights, inspiring practical examples and valuable recommendations. That may lead to a general framework for the ethical guidance of the applications of digital technology as a basis for technological citizenship. Given the extremely high quality of all the discussions that the Thinkers have had with societal parties and the constructive findings of the Thinkers, there is every reason to be confident that, with this Position Paper, a firm foundation has been established in Flanders for the ethical guidance of digitalisation, and for the further modelling of what citizenship means in a digital society.

Voorwoord

De Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten (KVAB) organiseert jaarlijks twee Denkersprogramma's over maatschappelijke thema's, die de KVAB autonoom kiest vanwege hun grote relevantie en specifieke uitdagingen voor deze regio. Doel van elk Denkersprogramma is de verdieping van het brede maatschappelijke debat over de gekozen thema's en de ontwikkeling van een visie op de middellange en langere termijn. Zo'n verdieping heeft drie belangrijke facetten: er wordt voor het onderwerp een holistische, multidisciplinaire aanpak nagestreefd; er worden buitenlandse Denkers uitgenodigd die in gesprek gaan met lokale actoren; de bevindingen van de Denkers zijn ook gebaseerd op hun internationale ervaring, inzichten in de feiten en evoluties wereldwijd en wetenschappelijke evidentie. Soms worden specifieke deelthema's uit deze bevindingen en aanbevelingen van de Denkers achteraf concreet uitgewerkt door werkgroepen die worden opgericht door de KVAB. Die leiden tot Standpunten en kunnen ook opgenomen worden in de beleidsadviezen die via de Europese netwerken van academies worden voorbereid in het kader van SAPEA (Science Advice for Policy by the European Academies; <https://www.sapea.info>).

De KVAB heeft in 2018 de permanente reflectiegroep Responsible Research & Innovation (RRI) en Wetenschapsethiek opgericht, met leden uit alle klassen en experts uit de Vlaamse wetenschappelijke wereld. De groep behandelt horizontale thema's in het Europese wetenschaps- en innovatiebeleid en verbindt aspecten van de relatie tussen onderzoek, innovatie en maatschappij met een zestal deelthema's: maatschappelijke betrokkenheid, *open access/open science*, gendergelijkheid, wetenschapsonderwijs, *governance* en ethiek. Doel is te reflecteren over de maatschappelijke impact van onderzoek en innovatie, de ethische aspecten van het onderzoek en wetenschappelijke integriteit. Op die basis organiseert deze reflectiegroep activiteiten over specifieke thema's. Als eerste grote activiteit heeft de groep ervoor gekozen een Denkersprogramma op te zetten over de maatschappelijke aspecten van digitalisering. De digitale transformatie van de samenleving is zo ingrijpend dat de nieuwe Europese Commissie en de nieuwe Europese president dit als een van hun centrale acties hebben gekozen. Door een stuurgroep van dit Denkersprogramma werd een selectie gemaakt van drie internationale experts met diepgaande specifieke ervaring hierover in Nederland, Silicon Valley en internationale organisaties. Zij vullen elkaar ook aan. In samenspraak met de stuurgroep hebben de denkers gedebatteerd met de betrokken experts en verantwoordelijken uit Vlaanderen. Op basis hiervan hebben zij onafhankelijk hun oordeel gevormd. Hun bevindingen en aanbevelingen hebben ze gepresenteerd tijdens een publiek slotsymposium. Aangescherpt door de discussie die daar is gevoerd, hebben ze het voorliggende eindverslag neergelegd met bevindingen en aanbevelingen. We wensen u veel leesgenot, inspiratie én ambitie om ermee aan de slag te gaan.

Nog voor het ter perse gaan van dit rapport sloeg de Covid-19-pandemie toe en werd de hele wereld voor ongeziene uitdagingen geplaatst. Er gaat hierbij veel aandacht naar de gezondheids- en de ouderenzorg, en er wordt hard gezocht naar geneesmiddelen en een vaccin, in een voorbeeldige interdisciplinaire en open internationale samenwerking. De rol van wetenschappers en hun op evidentie gebaseerde adviezen voor de verantwoordelijke politici, zowel in de lockdownfase als tijdens de exitstrategie, is erg belangrijk gebleken. Als wetenschappers goed communiceren en steunen op een multidisciplinaire wetenschappelijke aanpak, krijgen ze ook appreciatie bij het bredere publiek, zo blijkt. Binnen dit geheel komt digitalisering op heel diverse fronten naar voren als een belangrijk instrument, onder meer in het dagelijkse leven: denk aan thuiswerken, afstandsonderwijs, sociale contacten, onlineshopping, corona-apps... We ontdekken hierdoor niet alleen waardevolle elementen die voortvloeien uit digitalisering, maar ook beperkingen en problemen. Dit alles ligt in de lijn van de analyse en aanbevelingen van de Denkers. Daarom werden in dit Standpunt op enkele relevante plaatsen nog toevoegingen gedaan.

Een exhaustieve bespreking van dit thema valt buiten het huidige opzet. Het is zeker wenselijk dat een brede wetenschappelijke studie gerealiseerd wordt. Een aanzet hiertoe werd recent door 120 onderzoekers uit vele disciplines in België gelanceerd.¹

19 april 2020

Dit Standpunt werd goedgekeurd voor publicatie door de Klasse van de Technische Wetenschappen op vrijdag 17 april 2020.

Joos Vandewalle,
coördinator van het denkersprogramma
31 januari 2020

¹ 120 humane wetenschappers willen stem in exitstrategie: 'Mens is meer dan drager van virus', De Standaard, 17 april 2020.

1. Situering, doelstellingen en aanpak van het denkersprogramma

JOOS VANDEWALLE

De digitalisering is een disruptief proces dat vele aspecten van de samenleving grondig beïnvloedt. Aan de basis ervan liggen de toenemende rekenkracht van chips en computers en een veel grotere collectie en opslagcapaciteit van data, waardoor een krachtigere verwerking van gegevens mogelijk wordt. Dat brengt dan weer nieuwe en betere diensten, vormen van communicatie en samenwerking binnen ons bereik. Uit het (verre) verleden kennen we minstens twee enigszins vergelijkbare disruptieve processen. Een eerste is de overgang van gesproken boodschappen naar geschreven taal, waardoor de overdracht van informatie niet langer gebonden was aan plaats of tijd. Deze 'ontkoppeling' was voor de Griekse filosoof Socrates, bij monde van zijn leerling Plato, een voorwerp van zorg omdat de geschreven informatie los komt te staan van de persoon die de woorden uitspreekt of schrijft. Een tweede disruptie kwam er in de 16de eeuw door de verspreiding van de boekdrukkunst, waardoor het geschreven woord sindsdien 'gemakkelijk' vermenigvuldigd en verspreid kan worden. Dit heeft in Europa een grote maatschappelijke impact gehad. Nu beleven we de digitaliseringsgolf, waardoor alle informatie – spraak, muziek, beeld, video, data en zelfs metadata – in digitale vorm voorgesteld kan worden: de digitale transformatie en innovatie. De computerkracht en de algoritmes die daarvan gebruik maken, kunnen deze digitale data snel verwerken en zorgen voor nieuwe diensten en voor de automatisering van heel wat processen. Daarbij nemen ze veel standaardtaken van mensen over, maar tegelijk creëren ze ook veel nieuwe taken. Deze computers kunnen ook leren uit voorbeelden en menselijke prestaties wat een aantal taken betreft overtreffen dankzij kunstmatige intelligentie (AI). Vaak levert de combinatie van de menselijke en de kunstmatige intelligentie de beste resultaten op.

Binnen deze transformatie, die volop aan de gang is, kadert dit denkersprogramma, met als grote ambitie antwoorden te zoeken op de vraag hoe het staat met de maatschappelijke waarden en digitale innovatie: wie, wat en hoe? Zowel in de politiek, bij de overheid en het onderwijs als in het bedrijfsleven krijgt de digitalisering van de samenleving recent heel wat aandacht.^{2,3,4} Bedrijven en overheidsinstellingen willen van de opportuniteiten zo veel mogelijk gebruik maken om betere producten en diensten aan te bieden. Digitalisering is in vele sectoren de

² <https://ec.europa.eu/digital-single-market/en/news/benchmarking-national-innovation-procurement-policy-frameworks-across-europe> last update sept. 2019.

³ Reinhilde Veugelers, Désirée Rückert en Christoph Weiss, 'Bridging the divide: new evidence about rms and digitalization', *Breugel Policy Contribution issue n°17*, december 2019.

⁴ Joint ALLEA/Royal Society 'Flourishing in a data-enabled society', ALLEA Discussion Paper #4 2019 https://allea.org/wp-content/uploads/2019/07/DiscussionPaper_DataGov_Digital.pdf

drijvende kracht van innovatie en verandering. Klanten maken ook meer en meer gebruik van digitale diensten die bedrijven of instellingen aanbieden; denk onder meer aan onlineshopping en onlinebankieren. Burgers gebruiken ook steeds meer digitale kanalen voor hun interactie met de overheden. Iedereen laat door zijn of haar normale dagelijkse activiteit ook vele digitale sporen na (data en metadata), zoals locaties en tal van andere persoonsgegevens. Meer nog, nieuwe bedrijven ontstaan in dit marktsegment. Jongeren interageren meer en meer via sociale media. Niet enkel de ontspanningswereld, maar ook het onderwijs – van basis tot secundair en hoger – en de naschoolse vorming zijn in hoge mate gedigitaliseerd. Ook gepensioneerden ervaren de voordelen. Daarnaast mogen we niet blind zijn voor de nadelen en gevaren. Denk aan hacking, misbruik van gegevens, fake news, virussen, phishing, privacyschendingen, ongewenste profilering, gebrek aan transparantie enz.

Dit alles pleit voor een weloverwogen aanpak, die essentiële maatschappelijke waarden respecteert en breed gedragen wordt. Dit werd recent ook op Europees niveau in de verf gezet. Charles Michel, de president van de Europese Raad, meldde dat de Europese Commissie en de Europese Raad twee topprioriteiten delen: de strijd tegen de klimaatverandering en de digitale agenda. Dit is dan ook een thematiek met veel facetten, invalshoeken en vormen van expertise, en met veel belanghebbenden en beleidsniveaus. Een ontwerptraject waarin ICT-experts eerst nieuwe systemen en diensten ontwerpen, los van hun toepassing, gebruikers en sociale wetenschappers, is daarom niet te verkiezen. Het is wenselijk om te kiezen voor vormen van cocreatie, codesign en een multidisciplinaire aanpak waarbij het concipiëren van functionaliteiten, technische keuzes, gebruikerservaring en ethische keuzes met aandacht voor waarden als vrijheid, onafhankelijkheid, autonomie en gelijkheid hand in hand gaan.

Denkers aan het woord

Begin 2019 hebben de leden van de stuurgroep (voor de samenstelling: zie Appendix 2) drie vooraanstaande internationale experts als Denkers geselecteerd. Zij werden aangezocht om het hele jaar 2019 samen met de stuurgroep te werken rond de gekozen thematiek. De drie complementaire Denkers (voor hun cv: zie Appendix 1) beschikken over een brede kennis en veel ervaring voor een studie van het thema vanuit drie invalshoeken: ethische aspecten, beleid en wetenschap. Peter-Paul Verbeek is hoogleraar in de wijsbegeerte aan de Universiteit van Twente. Rinie van Est is themacoördinator Slimme Samenleving aan het Rathenau Instituut en deeltijds hoogleraar Technology Assessment and Governance aan de Technische Universiteit Eindhoven. Jan Rabaey is professor Electrical Engineering aan UC Berkeley.

Deelthema's die aan bod kwamen zijn onder meer: drijvende krachten in digitalisering; ethisch verantwoord ontwerpen; digitalisering en onderwijs, *learning analytics*, permanente vorming en werkgelegenheid; eigenaarschap van data;

impact op beroepssectoren (zoals juridische diensten), retail, autonome voertuigen, de gezondheidssector..

In een eerste fase werden de Denkers aan het woord gelaten en deelden zij hun visie en ervaring met de stuurgroep. Daarna werden de essentiële elementen van het transformatieproces en zijn maatschappelijke impact in kaart gebracht:

- Belangrijke drijvende krachten in digitalisering (rekencapaciteit, energieverbruik, kunstmatige of artificiële Intelligentie (AI), Internet of Things...) en de relatie met de huidige samenleving en de evoluties daarin. Binnen AI ligt de focus op datagedreven AI en minder op kennisgebaseerde AI.
- Ethisch verantwoord ontwerpen van waardevolle digitale diensten en waardegedreven innovatie. Implementeren van ethische codes voor digitalisering en AI. De relatie tussen standaarden (vrijwillige basis, consensusopbouw tussen belanghebbenden), de ethiek (niet afdwingbaar, flexibel, snel), wetgeving (afdwingbaar, complex, traag).
- Digitalisering, onderwijs, permanente vorming en werkgelegenheid (jobaanbod, jobverandering, jobinhoud en bijscholing), digitale skills, *learning analytics*, *digital humanities* en sociale vaardigheden.
- Democratisering:
 - Digitalisering en de relatie tussen burgers, overheden, lokale besturen en bedrijven: technologisch burgerschap.
 - Digitalisering en concentratie van data (bv. Tesla-wagens, *learning analytics*, Google, Facebook,..). Intelligent en sociaal delen van data.
 - Eigenaarschap van data/toegang tot data.
- Digitalisering in belangrijke Vlaamse bedrijfssectoren, zoals de maakindustrie.
- Digitalisering in onderzoek/open data /eigenaarschap van onderzoeksdata.
- Mondiale evoluties en strategieën, economie, en publieke waarden. AI *made in Europe*, een derde weg voor de EU naast de VS en China.
- Selectie van voorbeelden van beroepssectoren waar digitalisering een grote impact zal hebben: juridische diensten (contracten, notariaten), detailhandel (*blockchain*), autonome voertuigen (robotten, drones, autonome wapens), gezondheidssector, *wearables*.

Ook andere academie, koepels van academies en gremia zetten activiteiten op met interessante discussies die belangrijk zijn voor dit denkersprogramma, zoals ALLEA⁵ met *Putting societal values at the heart of data use and establishing governance and regulation*. Zij buigen zich over de volgende kwesties:

- "What are the roles of different stakeholders in ensuring the use of data enables human flourishing?"
- How can public trust in data use and data governance be established and maintained? How can the public be meaningfully engaged and empowered with the development of technology?

⁵ Joint ALLEA/Royal Society 'Flourishing in a data-enabled society' ALLEA Discussion Paper #4 2019 https://allea.org/wp-content/uploads/2019/07/DiscussionPaper_DataGov_Digital.pdf

- How fit for purpose are existing governance frameworks to adapt to the accelerating pace of technological development?
- What is the potential impact of new regulations or data governance infrastructures on industry and business?
- What is the role of technologies in the enforcement of, and compliance with, regulation? In particular, how might Privacy Enhancing Technologies, Technologies for Governance and Government contribute to good governance?"

Debatten met stakeholders

In deze context en over deze afgebakende thema's zijn de Denkers in debat gegaan met talrijke stakeholders, partners en opiniemakers. Hierdoor kregen zij geleidelijk aan inzicht in de lokale situatie en konden ze hierover reflecteren vanuit hun internationaal perspectief.

Stakeholdersdebat 1 met bedrijven, bedrijvenorganisaties, sociale partners en consultants:

- Saskia Van Uffelen, Digitaliserings Champion België
- Karel Van Eetvelt, Febelfin/Belgische Vereniging van Banken
- Guido Van Humbeeck, VDAB
- Bart Steukers, Agoria
- Hendrik Delagrange, SERV, onderzoeksceel Stichting Innovatie en Arbeid
- Koen Debaus, Digital Innovation Deloitte
- David Bruyneel, Accenture, Digital applied intelligence

Stakeholdersdebat 2 met onderwijskoepels, universiteiten, en strategische onderzoeksinstellingen:

- Peter Goyvaerts, GO! Gemeenschapsonderwijs, Beleid en Strategie
- Greet Vanderbiesen, Katholiek Onderwijs
- Peggy Valcke, KU Leuven, imec
- Lieven De Marez, UGent, imec
- Steven Latré, UAntwerpen, imec

Stakeholdersdebat 3 met overheidsinstellingen en stedelijke initiatieven:

- Leo Van de Loock, VLAIO Vlaams Agentschap Innoveren en Ondernemen
- John Baekelmans, imec, smart cities, digitale stad Antwerpen
- Erik Mannens, iMinds-UGent

Stakeholdersdebat 4 met overheidsinstellingen, onderzoeksinstellingen, burgerverenigingen:

- David Stevens, GBA Gegevensbeschermingsautoriteit, voorzitter Privacycommissie
- Bart Preneel, KU Leuven, adviseur Privacycommissie
- Frank Robben, Kruispuntbank van de Sociale Zekerheid
- Raf Buyle, Agentschap Informatie Vlaanderen
- Eline Kindt, Liga voor Mensenrechten
- Rob Heyman, Kenniscentrum Data & Maatschappij

- Ruben Verborgh, imec IDLab UGent, start-up Solid/Disrupt
- Koen Vanthournout, Energyville

Symposium

De intense gesprekken gaven een realistisch beeld van de activiteiten in Vlaanderen: de aanpak van digitalisering, uitdagingen, problemen en vooruitzichten. Vanuit hun internationale ervaring hebben de Denkers een vergelijkingsbasis voor de Vlaamse context gemaakt, met een visie op sterktes en zwaktes. Hun ideeën en aanbevelingen voor de overheden en diverse andere betrokkenen werden gepresenteerd op een slotsymposium in het Paleis der Academiën, op woensdag 4 december 2019.⁶

Programma van het slotsymposium

Opening door Freddy Dumortier, vast secretaris KVAB

Toelichting bij het denkersprogramma Joos Vandewalle, KVAB, KU Leuven (organisator)

Preliminair rapporten van de Denkers:

- *Mensen en waarden centraal. Reflecties op de situatie in België* - Rinie van Est, Rathenau Instituut, Nederland
- *Digitale geletterdheid. Reflecties op de situatie in België en wereldwijd* - Peter-Paul Verbeek, Universiteit Twente, Nederland
- *Rol van de digitale technologie. Wat kunnen wij als technologen doen om een inclusieve digitale technologie en waardevolle digitalisatie waar te maken en wat is de interactie met overheid, de leidende bedrijven etc.* - Jan Rabaey, UC Berkeley, VS/buitenlands lid KVAB

Digitalisering en internet in een midlifecrisis? - Bart Preneel, COSIC, KU Leuven⁷

Parallele break-outsessies:

1. *Kenniscentrum Data en Maatschappij* - Rob Heyman, imec-SMIT VUB
2. *De digitale stad* - Bieke Zaman
3. *Eigenaarschap van de data, modellen voor het delen van data* - Jan Rabaey
4. *Digitalisering, jobinhoud, opleidingen en bijscholing* - Peter-Paul Verbeek

Bevindingen van de parallele sessies en debat met de Denkers (plenair) en met Jan Hautekiet als moderator

⁶ Programma van het symposium met de presentaties van de sprekers <https://www.kvab.be/nl/activiteiten/maatschappelijke-waarden-bij-digitale-innovatie-wie-wat-en-hoe>

⁷ Bart Preneel, 'The Future of Security and Privacy', *Proceedings Santa's Crypto Get-Together 2019*, Prague, Czech Republic, 5-6 December 2019, Springer-Verlag, Lecture Notes in Computer Science. https://www.kvab.be/sites/default/rest/blobs/2442/preneel_santav2_preprint.pdf

Reactie vanwege de Denkers

Slotwoord Joos Vandewalle, KVAB, KU Leuven

Toelichting van het Kenniscentrum Data en Maatschappij vanuit de presentatie en discussie in de break-outsessie 1.

De krachtlijnen van de visienota van het recent opgerichte Kenniscentrum⁸, die op 9 december 2019 publiek werden gepresenteerd, werden op het symposium toegelicht en besproken. Het Kenniscentrum wil het aanspreekpunt worden voor ethische aspecten van AI en datagedreven toepassingen voor alle belanghebbenden van AI in Vlaanderen. Het wil dit uitvoeren als een samenwerking met alle bestaande initiatieven. Zodra het deze positie verworven heeft, wil het inzetten op het uniform maken van methodes, standaarden en kaderwerken om AI maatschappij-klaar te maken. Indien het bovenstaande is gelukt kan het Kenniscentrum, samen met andere betrokkenen, werken aan *regulatory sandboxes* om AI-innovatie, regulering en beleid te verbeteren als dat nodig mocht blijken.

Uit de bespreking bleek dat het Kenniscentrum in zijn advies wetenschappelijk en onafhankelijk is. Het werkt voor drie doelgroepen: bedrijven, beleidsmakers en het brede publiek. Het vertegenwoordigt niemand en het voorziet deze doelgroepen van inhoud om ethische, juridische en sociaal verantwoorde AI in Vlaanderen mogelijk te maken. Op dit moment bevat de visienota een algemene aanpak die als heel ambitieus ervaren kan worden. Naast deze ambitie zal een meer gedetailleerde jaarplanning gemaakt worden en zullen de aangeboden diensten verder gedefinieerd en uitgebouwd worden.

Drie essays

Dit Standpunt van de KVAB bevat de drie essays van de Denkers en een afsluitend gemeenschappelijk hoofdstuk met besluiten en aanbevelingen. Elke Denker heeft zijn essay autonoom uitgeschreven na het colloquium op 4 december 2019, steunend op de diepgaande gesprekken die in de loop van 2019 in Vlaanderen met de stuurgroep en de stakeholders gemeenschappelijk werden gevoerd. Elk met hun eigen achtergrond, invalshoek en competentie vormen de drie essays dan ook complementaire en tegelijk consistente analyses. Het zijn stuk voor stuk waardevolle benaderingen met aanzetten voor verdere stappen in de digitale transitie. Vanwege de onafhankelijkheid van de verschillende essays volgt de bibliografie ofwel op het einde van de tekst ofwel als voetnoot.

⁸ *Visienota Kenniscentrum Data en Maatschappij*, <https://www.data-en-maatschappij.ai/publicaties/visienota>

Gezien de zeer hoge kwaliteit van alle gesprekken die de Denkers hebben gevoerd met maatschappelijke partijen en gezien de constructieve bevindingen van de Denkers is er alle reden om erop te vertrouwen dat er met dit Standpunt in Vlaanderen een stevig fundament wordt aangereikt voor een ethische begeleiding van digitalisering en voor het verder vormgeven van wat burgerschap in een digitale samenleving betekent.

2. Essay: Zoektocht naar het goede digitale leven. Zet mensen en waarden centraal bij het vormgeven van digitale innovatie

RINIE VAN EST

2.1. Ontwaken uit onverantwoordelijkheid

'It is important not to ignore this history of failure to assign and assume responsibility in the internet age, both by legislators and by Tech corporations, which led to the fiascos of the internet, in the form of spreading of mass surveillance, recruitment to terrorism, incitement to racial and religious hate and violence as well as multiple other catastrophes for democracy.'

Paul Friedrich Nemitz (2018)

Veel te lang is er gezwegen over hoe de informatierevolutie de wereld sluipenderwijs radicaal verandert (Est 2012). Een belangrijke reden voor dat zwijgen is dat vele wetenschappers, ondernemers, politici en journalisten voortdurend de 'technologische droom' verkondigen. Dat is voor de samenleving een hoopvol en verleidelijk verhaal waarin burgers met nieuwe technologie dingen beter kunnen doen, zoals ongestoord in de trein naar tv-series kijken of genezen van kanker. In deze instrumentele visie leidt technologische vooruitgang vanzelf tot maatschappelijke vooruitgang. Voor de aanjagers van innovatie biedt het een zorgeloos bestaan, met heldere rollen en weinig verantwoordelijkheden. Wetenschappers en hightechondernemers worden gezien als helden en hoeven geen verantwoording af te leggen voor de risico's van de technologie. Overheden mogen innovatie richtingloos stimuleren en laten de aanpak van negatieve maatschappelijke 'neveneffecten' over aan de volgende lichter politici en bestuurders. De technologische droom gaat zo hand in hand met de organisatie van onverantwoordelijkheid (Beck 1986).

De technologische droomtoestand kenmerkt zich door idolatie verheerlijking van innovators en aversie tegen critici. Tekenend is het verschil in de wijze waarop het publiek enerzijds reageerde op Steve Jobs toen hij in 2007 de eerste iPhone presenteerde en anderzijds op de Kerstrede van koningin Beatrix eind 2009. Toen Jobs de iPhone introduceerde met de woorden 'Once in a while a revolutionary product comes along that changes everything', werd hij bij iedere zin toegejuicht door een uitbundige menigte. Twee jaar later waarschuwde koningin Beatrix voor informatiestress en stelde ze kritische vragen bij de kwaliteit van het gesprek via sociale media: 'De moderne technische mogelijkheden lijken mensen dichter bij elkaar te brengen, maar ze blijven op "veilige" afstand, schuilgaand achter hun schermen.' Op sociale media werd gezegd dat dit 'oudje' geen verstand had van nieuwe technologie en beter haar mond kon houden. Zo'n publieke atmosfeer

waarin kritiek op technologie haast taboe is, staat een goed debat en dus democratische sturing van technologie in de weg. Bij ICT is dat zeker veel te lang het geval geweest.

Anno 2020 lijkt de wereld wellicht ietwat ontwaakt uit de digitale roes. Moedige klokkenluiders speelden daarbij een cruciale rol. In 2013 bracht Edward Snowden de massa-surveillance van de NSA, de Amerikaanse veiligheidsdienst, onder de aandacht van de wereld. Klokkenluider Christopher Wylie onthulde hoe Cambridge Analytica in 2018, in opdracht van Trump, met data van Facebook het gedrag van kiezers tijdens de Amerikaanse presidentsverkiezingen probeerde te beïnvloeden. Marc Zuckerberg, de CEO van Facebook, moest zich verantwoorden voor het Amerikaanse Congres en gaf toe: 'We didn't take a broad enough view of our responsibility and that was a big mistake.' Zo werd duidelijk dat het privacyparadijs van weleer inmiddels is verworden tot een controlestaat (Vedder et al. 2007) en dat er inmiddels ook sprake is van een mondiale controle-economie, het zogenaamde surveillance-kapitalisme (Zuboff 2019). Dat leidt tot verlies aan controle over persoonlijke data én informatiekeuzevrijheid. Dat wil zeggen: verlies aan grip op de informatie die we ontvangen en op basis waarvan we keuzes maken (Hof et al. 2012).

Door de technologische droom en de bijbehorende 'cultuur van wetteloosheid en onverantwoordelijkheid' (Nemitz 2018) heeft digitale innovatie de afgelopen twee decennia ruim baan gekregen en als gevolg daarvan hebben de bijbehorende risico's welig kunnen tierren, met soms catastrofale gevolgen. Nu het internet gestaag verandert in het Internet of Things, en daarmee werkelijk op alle terreinen van de samenleving een rol gaat spelen, is het belangrijk om deze geschiedenis van onverantwoordelijkheid te erkennen (zie de beginquote van Nemitz 2018). Het roept namelijk de vraag op hoe we vanaf nu digitalisering wél op een verantwoorde wijze kunnen vormgeven.

Dit essay gaat in op verantwoord digitaal innoveren door allereerst te kijken naar de Nederlandse situatie. Hoe is de *governance* van de maatschappelijke en ethische kwesties rondom digitalisering in Nederland georganiseerd? *Governance* betreft 'het vermogen van een samenleving om bepaalde middelen te ontwikkelen voor het maken en uitvoeren van collectieve keuzes' (Peters & Pierre 2009: 91). Vervolgens kijken we naar de situatie in België en Vlaanderen. Op basis van de drie gesprekken tussen de drie Thinkers-in-Residence over digitale innovatie van de KVAB en stakeholders, probeer ik zicht te krijgen op de wijze waarop relevante betrokkenen in België en Vlaanderen spreken over digitale innovatie. Worden het debat en beleid over digitale innovatie bepaald door de technologische droom of door de wens om verantwoord digitaal te innoveren?

2.2. Blinde vlekken in het Nederlandse governance-landschap

In Nederland doet het Rathenau Instituut onderzoek naar de impact van wetenschap, innovatie en technologie op de samenleving. Het informeren van het Nederlandse Parlement (Eerste en Tweede Kamer) staat daarbij centraal. Het instituut probeert ook het debat over digitalisering te verdiepen en te verbreden. Begin 2014 signaleerde het instituut dat er een intiem-technologische revolutie gaande is: digitale technologie nestelt zich in een snel tempo tussen en in ons, verzamelt heel veel informatie over ons en simuleert menselijk handelen (Est 2014). Dat via IT steeds meer in onze persoonlijke levenssfeer ingegrepen kan worden, leidt tot een politieke, economische en sociale strijd om onze intimiteit. Digitale innovatie roept belangrijke ethische vragen op, die raken aan de (grond) rechten en waardigheid van mensen: recht op privacy, verbod op aantasting van de lichamelijke en geestelijke integriteit, recht op een veilige omgeving, recht op eigendom en vrijheid van gedachte en geweten. Het Rathenau Instituut riep destijds politiek en bestuur op om tijdig kaders te ontwikkelen om deze intiem-technologische revolutie maatschappelijk in goede banen te leiden (Est & Rerimassie 2014).

Het werk van het instituut inspireerde Eerste Kamerlid Gerkens tot het indienen van een parlementaire motie op 23 september 2014, waarin zij de opkomst van het Internet of Things signaleerde en zich zorgen maakte over de sociale, juridische en psychologische effecten daarvan. In de motie, die vrijwel unaniem werd aangenomen, verzocht de Eerste Kamer de regering het Rathenau Instituut te vragen 'onderzoek te doen naar de wenselijkheid van een commissie die kan adviseren over de ethische kant van de digitalisering van de samenleving'. De Eerste Kamer vermoedde dat belangrijke waarden in het geding zijn bij digitalisering. De studie *Opwaarderen: borgen van publieke waarden in de digitale samenleving* (Kool et al. 2017) bevestigde dit. Daarin onderzochten we welke maatschappelijke en ethische uitdagingen samenhangen met digitalisering en in hoeverre de omgang daarmee al politiek is geagendeerd en al dan niet institutioneel verankerd is.

De begrippen 'digitalisering' en 'Internet of Things' verwijzen naar een cluster van digitale technologieën, zoals robotica, kunstmatige intelligentie, algoritmen en big data, digitale platformen, biometrie, persuasieve technologie, augmented reality en virtual reality. Samen zorgen zij voor een nieuwe digitaliseringsgolf. Steeds meer onderdelen uit de fysieke wereld krijgen een virtuele representatie. Daardoor ontstaat op steeds meer plekken een continue terugkoppeling tussen de fysieke en de virtuele wereld, waarmee producten of diensten direct of indirect worden aangepast op basis van een analyse van digitale gegevens. De genoemde technologieën spelen zodoende een sleutelrol in datawaardeketens, dat wil zeggen bij het verzamelen, analyseren en toepassen van data. Toegepast op mensen gaat het dan om het meten van mensen, het profileren van mensen en het ingrijpen in het leven van mensen, zoals het sturen van gedrag. Denk aan de nieuwsfeeds

die sociale-mediabedrijven 'op maat' aanpassen, op basis van het surfgedrag van de gebruiker. De datawaardenketen is daarmee een cybernetische feedback-loop geworden. Juist dat is kenmerkend voor de huidige fase van de digitalisering en vergroot de maatschappelijke impact ervan.

Sinds de jaren 1970 houdt de politiek zich bezig met het privacyvraagstuk. Vanuit publiek perspectief gaat het daarbij om het formuleren en implementeren van een eerlijke manier om met persoonlijke data om te gaan en deze automatisch te verwerken. Bekende zogenaamde *fair information*-principes zijn bijvoorbeeld: doelspecificatie, beperking van gegevensverzameling, doelbinding van gegevens, kwaliteit van data, beveiligingswaarborgen en *informed consent*. Waar digitalisering aanvankelijk vooral ging over het verzamelen van data, richt het publieke en politieke debat zich nu ook op de vraag hoe op een eerlijke wijze een grootschalige analyse én toepassing van die gegevens mogelijk is. De afgelopen paar jaar heeft de wereldwijde discussie over artificiële intelligentie (AI) geleid tot een identificatie van allerlei principes voor het toepassen van algoritmische bewerkingen, zoals het idee dat AI-technologie uitlegbaar moet zijn (zogenaamde *explainable AI*) (cf. Jong et al. 2019). Door deze verbreding van het debat is het besef doorgebroken dat digitalisering, naast privacy en veiligheid, ook andere belangrijke publieke waarden onder druk kan zetten, zoals autonomie, controle over technologie, menselijke waardigheid, rechtvaardigheid en eerlijke machtsverhoudingen (zie tabel 1).

Tabel 1. Maatschappelijke en ethische vraagstukken bij digitalisering
(bron: Kool et al. 2017, 75)

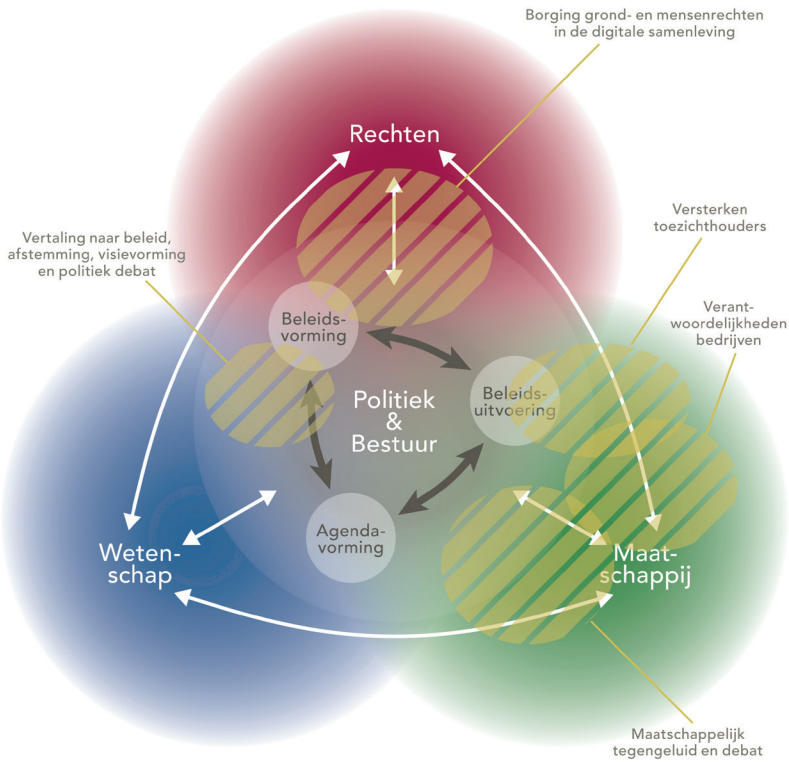
Thema	Maatschappelijke en ethische vraagstukken
Privacy	Gegevensbescherming, privacy, digitaal huisrecht, mentale privacy, surveillance, doelverschuiving
Veiligheid	Informatieveiligheid, identiteitsfraude, fysieke veiligheid
Autonomie	Keuzevrijheid, vrijheid van meningsuiting, manipulatie (verspreiding van desinformatie, <i>microtargeting</i>), bescherming democratie, paternalisme, vaardigheden, grenzen zelfredzaamheid
Controle over technologie	Controle over en inzicht in AI-technologie, verantwoordelijkheid, voorspelbaarheid
Menselijke waardigheid	Dehumanisatie, instrumentalisering, <i>de-skilling</i> , de-socialisatie, werkloosheid
Rechtvaardigheid	Discriminatie, uitsluiting, gelijke behandeling, stigmatisering
Machtsverhoudingen	Ooneerlijke concurrentie, uitbuiting, relatie consument-bedrijf, relatie bedrijf-platform

Drie voorbeelden hiervan. Het Cambridge Analytica-schandaal uit 2018 liet zien hoe de autonomie van de democratische kiezer aangetast kan worden en zorgde voor een groeispurt van het publieke bewustzijn over hoeveel data er verzameld worden, dat aanbieders verdienen aan gebruikersdata en dat bedrijven en overheden via profilering het gedrag van mensen kunnen proberen te sturen. Het tweede voorbeeld betreft menselijke waardigheid, of beter het gebrek aan aandacht voor de mens. Op 18 maart 2018 werd Elaine Herzberg in Tempe, Arizona, gedood door een zelfsturende auto van Uber. Een belangrijke oorzaak van dit ongeval was het feit dat de programmeurs van Uber de software van de robotauto zo hadden ontworpen dat die slechts voetgangers die overstaken op een zebrapad detecteerde (Marshall & Davies 2019). Herzberg deed wat mensen vaker doen: met haar fiets aan de hand oversteken op zo'n 100 meter afstand van een officieel zebrapad. Daarom detecteerde de robotauto haar te laat. De gebrekkige veiligheidscultuur bij Uber was er de oorzaak van dat er te weinig rekening werd gehouden met menselijk gedrag. Het derde voorbeeld betreft eerlijke economische machtsverhoudingen. Inmiddels bestaan er in Europa veel zorgen over de data- en marktmacht van grote Amerikaanse en Chinese platformen, zoals Google, Amazon en Alibaba. Tijdens een rondetafelgesprek in de Tweede Kamer stelde mode-ondernemer Erik van Rosmalen bijvoorbeeld dat de opkomst van platformen tot flink wat omzetverlies bij het midden- en kleinbedrijf heeft geleid: 'Veel detaillisten hebben in de afgelopen 8 tot 10 jaar de omzet en daarmee het inkomen fors zien afnemen, waarbij percentages als 30-40 procent eerder regel dan uitzondering zijn.' (Tweede Kamer 2018).

Het Rathenau Instituut onderzocht tevens in hoeverre de wettelijke kaders en toezichtarrangementen voldoende toegerust waren op de genoemde kwesties, die veelal raken aan essentiële publieke waarden en fundamentele rechten. Het gaat hier dus om de *governance* van de maatschappelijke en ethische kwesties rondom digitalisering. Daarbij is gekeken naar het gehele *governance*-landschap: de rol van de wetenschap, de rol van grond- en mensenrechten, de rol van het maatschappelijk middenveld en de rol van politiek en bestuur bij agendavorming, beleidsbepaling en -uitvoering. De analyse liet zien dat de bescherming van publieke waarden die door digitalisering onder druk komen te staan, op dat moment in Nederland sterk tekortschoot. Vijf blinde vlekken of *governance*-uitdagingen werden gesignaleerd (zie figuur 1):

1. Vertaling van nieuwe maatschappelijke en ethische vraagstukken naar beleid, interdepartementale afstemming en coördinatie rond digitalisering, en het politieke debat over de nieuwe vraagstukken.
2. Borging van grond- en mensenrechten in de digitale samenleving.
3. Versterken van toezichthouders en zorgen voor onderlinge afstemming tussen hen.
4. Nieuwe verantwoordelijkheden van ontwikkelaars van digitale diensten en producten.

6. Organiseren van het maatschappelijk tegengeluid: versterken maatschappelijk middenveld, kennis en vaardigheden van burgers en maatschappelijk debat over digitalisering.



Figuur 1. Blinde vlekken in het Nederlandse *governance*-landschap rondom digitalisering, ethiek en maatschappij (bron: Kool et al. 2017, 124).

In Nederland heeft de afgelopen paar jaar het publieke en het politieke debat over de *governance* van de maatschappelijke inbedding van digitalisering een flinke stimulans gekregen. In juni 2018 bracht het kabinet een integrale digitaliseringsstrategie uit (EZK 2018). Om meer grip te krijgen op digitalisering stelde de Tweede Kamer in juli 2019 de tijdelijke commissie Digitale Toekomst in, die onderzoekt hoe de kamer beter sturing kan geven aan digitale ontwikkelingen. Daarnaast zijn diverse politieke partijen, zoals D'66 (Verhoeven et al. 2019), CDA (Havelaar & Dijkman 2019) en GroenLinks (Wouters 2019) digitalisering vanuit hun ideologische grondslagen gaan doordenken. Het politieke zwijgen over hoe de informatierevolutie onze samenleving verandert, waarover ik aan het begin van dit essay sprak, is in 2019 in Nederland doorbroken. Hierdoor wordt het mogelijk om een democratisch en pluralistisch politiek debat over de toekomst van de digitale

samenleving te houden. Door het Cambridge Analytica-schandaal en de grote belangstelling voor AI en ethiek bij ontwikkelaars en media is de afgelopen jaren ook de publieke aandacht voor tal van maatschappelijke en ethische kwesties rondom digitalisering sterk gegroeid (Jong et al. 2019).

2.3. Denken over digitale innovatie en doen in België en Vlaanderen

In het kader van het Denkersprogramma van de KVAB over digitale innovatie roept de bespreking hierboven de vraag op: zijn Vlaanderen en België klaar voor de digitale samenleving? Met andere woorden: hoe is het in Vlaanderen en België gesteld met de *governance* van maatschappelijke en ethische kwesties die door digitalisering zijn ontstaan? Het gaat daarbij om vragen als: welke publieke problemen zijn gesignaleerd en politiek geagendeerd? Hoe discussiëren diverse actoren in de samenleving over deze problemen? Worden debat en beleid over digitale innovatie bepaald door de technologische droom of de wens om verantwoord digitaal te innoveren? Wie is daarbij wel of niet betrokken? Op welke wijze worden publieke waarden institutioneel geborgd? Welke instituties zijn er door de jaren heen opgebouwd om dat te doen?

Het goed adresseren van dat soort vragen vergt veel en diepgaand onderzoek. Noodzakelijkerwijs reikt mijn ambitie hier dan ook minder ver. Op basis van de drie gesprekken tussen de drie Thinkers-in-Residence over digitale innovatie, de KVAB en stakeholders probeer ik zicht te krijgen op de wijze waarop men in België en Vlaanderen spreekt over digitale innovatie. Die gesprekken tonen aan dat stakeholders zich bewust zijn van het feit dat digitalisering een breed scala aan kwesties met zich meebrengt (zie tabel 2) en tal van *governance*-uitdagingen. De Sociaal-Economische Raad van Vlaanderen (SERV 2018) pleit bijvoorbeeld voor experimenteerruimtes en proeftuinen, een voortrekkersrol van de overheid en het stimuleren van de dialoog over de impact van digitalisering, en in het bijzonder over de integratie van ethische en maatschappelijke kwesties bij de technologieontwikkeling.

2.3.1. Nationale veiligheid, economische voorspoed en maatschappelijk welzijn

In het vervolg reflecteer ik op drie belangrijke innovatiedoelen – nationale veiligheid, economische voorspoed en maatschappelijk welzijn – en gerelateerde innovatieprocessen. Tussen deze doelen kan spanning bestaan. De huidige economie kan op allerlei manieren op gespannen voet staan met sociale en ecologische grenzen (Raworth 2017). Het besef dat de mondiale uitstoot van broeikasgassen de planetaire ecologische grenzen, waarvan ons leven afhankelijk is, in ruime mate overschrijdt (Rockström et al. 2009), vormt bijvoorbeeld de morele grondslag voor de duurzame energietransitie. Een andere spanning in het debat is de angst dat te veel aandacht voor maatschappelijk welzijn economische voorspoed ondermijnt. Deze angst komt goed naar voren in de volgende quote van ALLEA (2019, 7): 'De

Tabel 2. Overzicht van maatschappelijke en ethische vraagstukken die tijdens de drie stakeholdergesprekken over digitale innovatie naar voren werden gebracht.

Thema	Maatschappelijke en ethische vraagstukken
Privacy	Grenzen van het meten en delen van intieme informatie, bijvoorbeeld bij leerlingen en studenten in het geval van <i>learning analytics</i>
Veiligheid	Informatieveiligheid
Autonomie	Fake news, verslaving aan games en sociale media
Controle over technologie	Gevaar voor bias in computeralgoritmes, het belang van uitlegbaarheid van AI-technologie om betekenisvol beslissen mogelijk te maken
Menselijke waardigheid	Kortere aandachtsspanne van leerlingen (<i>deskilling</i>). Terwijl het publiek vreest voor banenverlies door automatisering, stelt de technologiefederatie Agoria (2018) dat voor elke baan die in de toekomst verloren gaat vanwege digitalisering, er 3,7 nieuwe voor in de plaats zullen komen.
Rechtvaardigheid	Een grote digitale kloof dreigt in de samenleving. Er is een digitale kloof tussen leerkrachten en er bestaat een digitale kloof tussen leerlingen. Houden we genoeg rekening met mensen die niet meekunnen? Online als standaard zorgt voor uitsluiting
Machtsverhoudingen	Rol van (grote) bedrijven bij onderwijs, die op termijn mogelijk de democratische functie van onderwijs onder druk kunnen zetten

spanning hier is dat Europa zijn technologische voorsprong verliest en dat dit kan leiden tot de achteruitgang van de levensstandaard en de politieke macht die het beschermt.' ALLEA, de Europese Federatie van Academies voor Wetenschappen en Kunsten, vraagt zich af of Europa de bescherming van de menselijke waardigheid niet te veel boven economische groei waardeert en of het verkeerde gebruik van data niet opweegt tegen de kansen die dat biedt? (ibid.) De genoemde spanningen roepen de vraag op in hoeverre nationale veiligheid, economische voorspoed en maatschappelijk welzijn hand in hand kunnen gaan.

Hieronder reflecteer ik, schetsmatig en essayistisch, over hoe België en Vlaanderen vormgeven aan digitale innovatie op het terrein van nationale veiligheid, economische voorspoed en maatschappelijk welzijn, en in hoeverre er rekening gehouden wordt met de maatschappelijke inbedding daarvan. De wijze waarop digitale innovatie wordt vormgegeven vergelijk ik met een brede visie op innovatie, die geïnspireerd is door de notie 'maatschappelijk verantwoord innoveren' (cf. Schomberg & Hankins 2019). Bij verantwoord (digitaal) innoveren is er aandacht voor de maatschappelijke impact van innovatie en nemen de betrokken actoren

hun verantwoordelijkheid om innovatie in een vanuit publiek perspectief gewenste richting te sturen (cf. Est et al. 2017).

2.3.2. Nationale veiligheid: de terugkeer van de drones

Op het terrein van de nationale veiligheid is het streefdoel vaak om technologisch superieur te blijven. Dat wordt helder onder woorden gebracht door Yoge Patel, hoofontwikkelaar Blue Bear Systems Research: 'De enige manier om gevaren voor te blijven, is voorsprong te behouden. (...) Ik zeg altijd: trek die hardloopschoenen aan, en ren! Je moet op kop blijven lopen.' (geciteerd in Hofman 2019) Dat leidt op het gebied van militaire drones en cyberwapens tot een AI-wapenwedloop. Aan de nieuwe robotica ligt de ingenieurswens ten grondslag om machines te bouwen die autonoom kunnen bewegen en handelen in complexe omgevingen (Royakkers et al. 2012). De wil om technologisch superieur te zijn zorgt er zodoende voor dat (militaire en commerciële) drones steeds intelligenter en autonomer zullen worden. Dergelijke technologie kan ook in verkeerde handen komen en op een kwaadaardige manier worden gebruikt. In de documentaire *Drones: The next air disaster* wordt Jason Derick, de hoofdontwikkelaar van Exyn Technologies, een Amerikaans bedrijf uit Philadelphia dat autonome drones ontwikkelt, gevraagd of een dergelijk scenario mogelijk is. Hij antwoordt: 'Ik denk dat we deze technologie moeten scheiden van de manier waarop mensen het gebruiken. Hoe mensen het gebruiken, is in wezen hun zaak.' Deze uitspraak laat zien dat hij geen enkele verantwoordelijkheid wil nemen voor het mogelijke misbruik van de drones die zijn bedrijf ontwikkelt. Hoe kan vanuit het perspectief van nationale veiligheid wel op een verantwoordelijke wijze omgegaan worden met deze ontwikkeling?

De ontwikkeling van autonome militaire drones roept de vraag op of bewapende onbemande systemen zelfstandig (dus autonoom) mogen beslissen over dodelijk geweld. Deze automatisering van de dood roept uiteraard ethische vragen op. Kunnen autonome militaire robots voldoen aan belangrijke beginselen van het humanitaire oorlogsrecht, zoals proportionaliteit (het uitgeoefende geweld moet in verhouding staan tot het bereikte voordeel) en discriminatie (bij doelwitkeuze moet onderscheid worden gemaakt tussen strijders en burgers en tussen militaire en civiele objecten)? Het Federale Parlement van België heeft in 2018 besloten dat dergelijke *killer robots* nooit door België gebruikt mogen worden. België is het eerste land ter wereld dat een dergelijk preventief verbod vanuit ethische overwegingen heeft ingesteld. Onderzoek naar automatische wapens is in België nog wel mogelijk. België doet dus nog wel mee met de wapenwedloop.

De ontwikkeling van militaire drones roept ook de vraag op in hoeverre de proliferatie van bewapende drones een onverantwoord gevaar vormt. In Syrië zette Islamitische Staat sinds 2017 commerciële drones in om bommen te werpen. De proliferatie van militaire drones en het misbruik van commerciële drones behoeven dan ook serieuze aandacht. Dit laat ook zien dat het idee dat technologische

superioriteit op het gebied van militaire techniek de nationale veiligheid dient, in ieder geval te simplistisch is. Ook simpele en goedkope dronetechologie kan veel schade veroorzaken. Daarnaast kan men zich afvragen in hoeverre België defensief is voorbereid op een droneaanval. Hoewel het schrikscenario van de bewapende drones die terugkeren en terugbijten (cf. Tenner 1997) militair-strategisch zo voor de hand ligt, is het lange tijd genegeerd. Er lijken eerst concrete incidenten nodig – denk aan de tijdelijke sluiting begin 2019 van het Londense vliegveld Heathrow omdat een persoon met behulp van een drone het luchtverkeer in gevaar bracht – om aandacht te creëren voor dat soort maatschappelijke impact.

Kortom, vanuit een brede visie op digitale innovatie dient wapenontwikkeling steeds gelijk op te gaan met wapenbeheersing. België heeft met het besluit om nooit *killer robots* in te zullen zetten een krachtig moreel signaal afgegeven. Vanuit die positie kan meegewerkt worden aan mondiale verdragen voor wapenbeheersing, die er zorg voor kunnen dragen dat de proliferatie van wapens tegengaan kan worden en militaire robots op een verantwoorde wijze kunnen worden ontwikkeld en ingezet bij militaire operaties. Ten slotte dient België zichzelf voor te bereiden op een mogelijke (terroristische) droneaanval.

Tabel 3. Brede visie op digitale innovatie op het terrein van nationale veiligheid, economische voorspoed en maatschappelijk welzijn.

Beleidsdoel	Nationale veiligheid	Economische voorspoed	Maatschappelijk welzijn
Doel innovatie	Technologische superioriteit	Ontwikkeling en gebruik van sleuteltechnologie	Maatschappelijke uitdagingen
Gevaar van innovatie	Strategische kortzichtigheid	Georganiseerde onverantwoordelijkheid	Technologie-gedreven; technologisch verouderd
Brede visie op innovatie	Wapenontwikkeling en -beheersing	Gelijktijdige ontwikkeling van technologie, economie, debat en regulering	Waardevol digitaliseren: mens en publieke waarden centraal

2.3.3. Economische voorspoed: geloof in technologie én samenleving

Op het gebied van economische voorspoed is de grote angst dat een land technologisch achterop gaat lopen en als gevolg daarvan economisch in verval raakt. Zo vindt Vande Reyde, een Vlaams parlementslid voor de partij Open Vlaamse Liberalen en Democraten (Open VLD), dat er in het politieke debat veel te weinig aandacht is gegaan naar een passage in het Vlaamse regeerakkoord: 'Vlaanderen wil werk maken van de uitrol van 5G.' (Reyde 2019) Volgens Vande Reyde viel daarover tijdens de debatten in het Vlaams Parlement geen woord,

terwijl de uitrol van 5G het beleidsvoornemen is met wellicht de grootste impact op de Vlaamse samenleving. Hij waarschuwt dat Europa, België en Vlaanderen in vergelijking met China en Amerika op het gebied van 5G de trein volledig dreigen te missen. Vande Reyde heeft een punt: een belangrijke rol voor de overheid is inderdaad het stimuleren van innovatie, zodat een land internationaal competitief kan blijven. Vanuit het oogpunt van verantwoorde innovatie is het daarbij van belang oog te hebben voor de maatschappelijke inbedding van technologie. In hoeverre is er bij innovatiebeleid aandacht voor de impact voor innovatie en wat wordt er gedaan om innovatie in een gewenste richting te sturen?

Laat ik inzoomen op de rol die de Vlaamse overheid voor zichzelf ziet op het terrein van artificiële intelligentie (AI). Tijdens de plenaire vergadering van het Vlaams Parlement op 22 mei 2019 vond een interessant onderonsje plaats tussen twee leden van de Nieuw-Vlaamse Alliantie (N-VA), een Vlaams-nationalistische en liberaal-conservatieve politieke partij. Parlements lid Gryffroy stelde vast: 'Ik ben ingenieur. Ik geloof in technologie.' Daarna vroeg hij aan zijn partijgenoot Muyters, de toenmalige minister van Werk, Economie, Innovatie en Sport: 'Wat is de rol van de Vlaamse overheid om enerzijds AI te stimuleren, maar anderzijds ook een antwoord te bieden op die uitdagingen en die negatievere connotaties?' De minister zag drie rollen: 1) investeerder in wetenschappelijk onderzoek, 2) stimulator en facilitator en 3) bewustmaker en vormer. De eerste rol spreekt voor zichzelf. De tweede rol betreft het informeren van het brede publieke en het sensibiliseren van het midden- en kleinbedrijf opdat zij de trein niet zullen missen. De derde rol betreft scholing van jonge mensen en herscholing van werknemers. Om de dialoog over ethische vraagstukken te stimuleren en de overheid daarover advies te geven werd het Kenniscentrum Data en Samenleving opgericht (zie ook het *Vlaams Beleidsplan Artificiële Intelligentie* uit 2019).

Deze blik op innovatie gaat duidelijk voorbij aan de technologische droom. Toch is ze nog sterk techno-economisch gedetermineerd. Om vrije baan te maken voor technologie en economische activiteiten dienen diverse maatschappelijke randvoorwaarden te worden gecreëerd, zoals het informeren van ondernemers en burgers, het scholen van (toekomstige) werknemers en het adequaat adresseren van ethische kwesties. Toch zit de angst dat de ethische dialoog de 'voortgang' van technologie en economie in de weg zal zitten er nog sterk in. Het *Vlaams Beleidsplan Artificiële Intelligentie* (2019, 2) stelt dat er vaak terecht zorgen zijn over ethische, legale en andere vraagstukken, maar dat discussies daarover 'vaak verzanden op mogelijke ongewenste of zelfs negatieve gevolgen, zonder te zien welke nieuwe opportuniteiten voor bestaande maatschappelijke uitdagingen AI met zich kan meebrengen'.

Bij verantwoord innoveren is de maatschappij geen randvoorwaarde maar het uitgangspunt, en vormen technologie en economie de middelen. Bij het vormgeven van innovatie vanuit een publiek perspectief staan maatschappelijke

wenselijkheid en maatschappelijke uitdagingen centraal. Er is geen angst voor sociale kwesties, omdat die leidend dienen te zijn voor innovatie. Er is dus sprake van een verandering van perspectief. Het gaat niet langer om digitale technologie die ingebed dient te worden in de samenleving, maar om het vormgeven van de digitale samenleving. Dat vereist de kunst om vanuit sociale ambities en publieke waarden richting te geven aan de energie en vitaliteit van innovatie (Est et al. 2018).

Innovatie dient daarbij breed – technologisch en sociaal – te worden opgevat. In aanvulling op de uitspraak van Gryffroy is het motto: 'Ik ben burger. Ik geloof in technologie én samenleving.' Sociale innovatie betreft onder meer economische innovatie (nieuwe verdien- of organisatorische modellen), politieke en maatschappelijke innovatie (ontwikkeling van een nieuwe taal, publiek debat en politieke denkkaders) en ethische en juridische innovatie (nieuwe morele kaders en/of wetgeving). Het vormgeven van de digitale transitie vraagt om aandacht voor en gelijktijdige investering in al deze types van innovaties. De geschiedenis leert namelijk dat digitale ontwikkelingen zeer snel kunnen gaan en disruptieve, soms catastrofale, gevolgen kunnen hebben. Wanneer er niet gelijktijdig energie gestoken wordt in de *governance* van digitalisering, dan lopen het publieke debat, ons ethisch denken en wet- en regelgeving haast spreekwoordelijk achter op de technologische ontwikkeling. Maar maatschappelijk achter de technologie aanlopen is een politieke keuze. Scherper gezegd: het belijden van de technologische droom is een politieke keuze geweest die de afgelopen twintig jaar ruim baan heeft gegeven aan een 'cultuur van wetteloosheid en onverantwoordelijkheid' (Nemitz 2018) op het gebied van digitale innovatie.

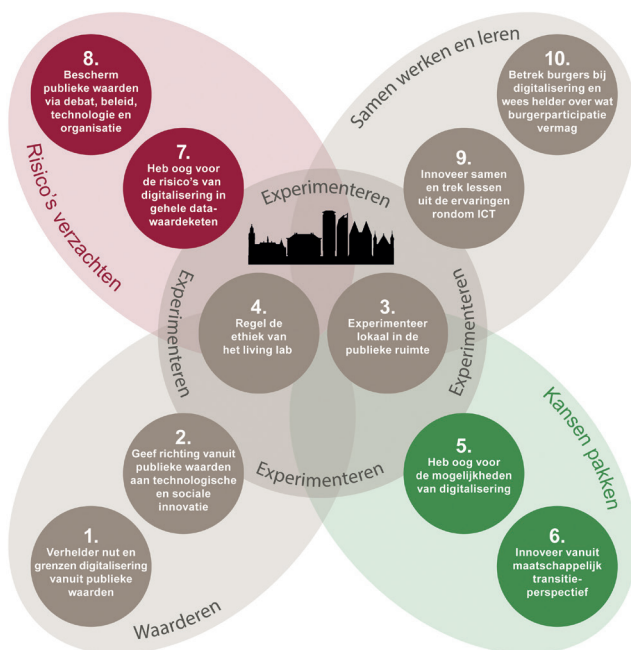
Ontwikkelingen op het gebied van bijvoorbeeld gezichts- en emotieherkenning en van virtual reality (VR) – allebei toepassingen van AI – vragen om gelijktijdige innovatie en soms zelfs om het radicaal herdenken van de manier waarop we naar digitale innovatie kijken en deze reguleren. Op het gebied van VR is er een groeiende kloof tussen enerzijds het gebrek aan politieke aandacht voor de vele maatschappelijke en ethische vraagstukken die samenhangen met VR, en anderzijds de noodzakelijke tijdige ontwikkeling van kaders om deze technologie in te bedden (Snijders et al. 2019). Volgens het Rathenau Instituut is de meest fundamentele vraag hierbij in hoeverre VR als een biomedische technologie moet worden gezien en gereguleerd (ibid.). De redenering hierachter is dat de versmelting van de computer, de camera, biometrische sensoren, de VR-bril en het lichaam leidt tot steeds meer mogelijkheden om individuen real-time en ongemerkt te beïnvloeden. Kortom, een VR-bril kan niet gezien en gereguleerd worden als een simpel gadget, maar is een intieme technologie die politiek en bestuur uitdaagt om (gelijk)tijdig kaders te ontwikkelen en de inbedding van VR in goede banen te leiden.

2.3.4. Maatschappelijk welzijn: waardevol digitaliseren

Bij innovatie op het gebied van scholing, overheidsdiensten, zorg en leefomgeving gaat het niet primair om technologische superioriteit of economische competitie, maar om maatschappelijke doelen, zoals publieke gezondheid en kwaliteit van leven. Tijdens de stakeholdergesprekken kwamen een aantal mooie Vlaamse voorbeelden naar voren. John Baekelmans, vicepresident van imec's IoT and Connected Health Solutions Group, vertelde over zijn persoonlijke leertraject. (imec is het grootste onafhankelijke Europese onderzoekscentrum op het gebied van digitale technologie). Vooraleer hij bij imec kwam, heeft Baekelmans 21 jaar voor Cisco gewerkt, de laatste jaren als chief technology officer voor de Internet of Everything (IoE) Solutions group. Bij Cisco verkocht men de technologische droom en volgde men een *technology push*-strategie. Baekelmans heeft ervaren dat er veel fouten worden gemaakt door technologie in te zetten 'for the sake of technology', zonder rekening te houden met de burger. Volgens hem dient de technologie noch de overheid leidend te zijn bij de vormgeving van bijvoorbeeld de stad, maar de burger, omdat het in essentie gaat om de kwaliteit van leven en de vraag of burgers hun leefomgeving prettig vinden om in te leven. Voor de democratische invulling van de openbare ruimte is participatie nodig op alle niveaus: buurt, regionaal, bestuurlijk en nationaal. Terwijl Nederland worstelt met de vraag hoe lessen en innovaties op lokaal niveau opgeschaald kunnen worden, is het slimme-stedenbeleid in Vlaanderen een regioverhaal. Antwerpen is een test- en experimenterstad. Dat wat goed gaat kan in andere steden worden toegepast.

Frank Robben is betrokken geweest bij de oprichting van de Kruispuntbank van de Sociale Zekerheid – de coördinator van e-government in de sociale zekerheid – en het eHealth-platform, dat de elektronische dienstverlening en informatie-uitwisseling in de gezondheidszorg wil verbeteren. Tijdens een stakeholderbijeenkomst beschreef Robben vier elementen voor het succesvol digitaliseren van dergelijke publieke diensten. Ten eerste is het van belang om alle belangrijke stakeholders bij het proces van innovatie te betrekken (en 'het bataljon juristen dat belangen komt verdedigen buiten de deur te houden'). Ten tweede: met behulp van die stakeholders dienen gezamenlijk principes en doelstellingen vastgesteld te worden. Deels liggen die al vast in wetgeving, zoals de GDPR (*General Data Protection Regulation*) op het gebied van de bescherming van persoonlijke data. Ten derde is het van belang te bepalen wat wenselijk en onwenselijk is en daartussen een goed evenwicht te bepalen. Zo is er bij de Kruispuntbank gekozen om de identificatie te laten beheeren door de deelnemende partners – omdat men dit niet wilde overdragen naar een ICT-gigant als Apple, bestaat er geen app voor de Kruispuntbank. Ten slotte dienen de waarden en normen die van belang zijn in het ontwerp van het digitale systeem hun beslag te vinden. Het gaat hier om architectuurpolitiek: politiek en ethiek door ontwerp. Het ontwerp dient in de praktijk bijvoorbeeld te zorgen voor een veilig systeem dat de persoonlijke levenssfeer van patiënten in acht neemt.

De twee voorbeelden uit de twee voorgaande paragrafen passen goed in het perspectief dat het Rathenau Instituut 'waardevol innoveren' noemt (Est et al. 2018) en dat draait om een constructieve interactie tussen technische en sociale innovatie. Daarbij spelen vijf processen een rol: waarderen, experimenteren, kansen grijpen, risico's verzachten en samen werken en leren (zie figuur 2). 'Waarderen' gaat om het verhelderen van de publieke waarden en doelen die men wil nastreven. 'Experimenteren' betekent ruimte maken om nieuwe zaken uit te proberen. 'Kansen grijpen' betekent een open oog houden voor de mogelijkheden die digitale technologieën bieden. De overheid behoort ernaar te streven dat onze samenleving zo veel mogelijk kan profiteren van de zegeningen van digitalisering. 'Risico's verzachten' houdt in dat burgers zo veel mogelijk worden beschermd tegen de risico's van digitalisering. Het in goede banen leiden van digitalisering is een gezamenlijke uitdaging. Het vraagt om afstemming en samenwerking tussen verschillende bestuurslagen en tussen diverse overheden en kennisinstellingen, bedrijven en, niet te vergeten, burgers. 'Samen werken en samen leren' is dus het vijfde devies. Het betrekken van relevante betrokkenen is cruciaal, maar niet evident. Er kunnen bijvoorbeeld stakeholders zijn die zich niet graag laten betrekken of die 'ongrijpbaar' zijn. Hoe zorg je er bijvoorbeeld voor dat kinderen uit gezinnen zonder internetverbinding op een of andere manier toch als stakeholder aanwezig kunnen zijn in de discussie over de digitalisering van onderwijs?



Figuur 2. Waardevol innoveren door middel van vijf centrale innovatieprocessen

2.4. Slotgedachte: Het goede digitale leven

In dit essay keek ik hoe men in Nederland en België en Vlaanderen praat over en omgaat met digitale innovatie. In beide landen is het zwijgen over hoe de informatierevolutie de wereld sluipenderwijs radicaal verandert voorbij. Het getuigt niet langer van ruimdenkendheid om alleen de zegeningen van digitalisering te bezingen. Het debat is voorbij de technologische droom. Zowel in Nederland als België is het besef doorgedrongen hoe groot en breed de impact is van digitalisering op de samenleving. Digitalisering verwijst niet langer vooral naar een verzameling gadgets, maar wordt inmiddels gezien als een transitie met kansen en risico's en vol onzekerheden. Het transitieperspectief legt de vraag naar het digitale goede leven op tafel en daarmee ook de democratische kernvraag: in welke digitale samenleving willen we leven? (cf. Kool et al. 2018)

In beide landen wordt ingezien dat met digitalisering tal van ethische, juridische en maatschappelijke vraagstukken samenhangen. In Vlaanderen is vanuit dat inzicht het Kenniscentrum Data en Samenleving opgericht. Het is echter de vraag of aandacht voor maatschappelijke en ethische kwesties niet te veel als een verplicht nummer gezien wordt, een appendix om digitale innovatiepolitiek te legitimeren. Aan het begin van mijn essay stelde ik dat de technologische droomtoestand zich toont in de idolate verheerlijking van innovators en de aversie tegen critici. Het *Vlaams Beleidsplan Artificiële Intelligentie* (2019) hinkt wat betreft de houding ten opzichte van kritiek nog duidelijk op één been. Enerzijds ziet men dat men niet zonder aandacht voor lastige vraagstukken kan, maar anderzijds is men bang dat die kritiek zand in de technologische en economische motor zal strooien.

Die angst is zowel gegrond als ongegrond. Gegrond, omdat ecologische, sociale en ethische aspecten inderdaad grenzen op kunnen werpen voor technologie en economie. Ongegrond, omdat de afweging tussen waarden als *uitgangspunt* voor innovatie gezien dient te worden. Zie hierboven het advies van Frank Robben en het perspectief van waardevol innoveren. Met het achter ons laten van de technologische droom verlaten we de mythe dat technologische vooruitgang vanzelf leidt tot maatschappelijke vooruitgang en dat economische innovatie vanzelfsprekend ook positief is vanuit publiek perspectief. Een breed perspectief op innovatie betekent dat allerlei relevante waarden en normen een rol mogen hebben in het richting geven van innovatie. De vraag die ALLEA stelt over de relatie tussen economische groei en menselijke waardigheid is dus in principe een goede vraag. Ik zeg 'in principe', omdat een goede omgang met die vraag leidt tot een constructieve en meer integrale discussie over digitale innovatie. In het slechtste geval kan zo'n vraag juist angst oproepen en leiden tot een terugval naar de technologische droom.

Hierboven heb ik laten zien wat een dergelijk breed waardegedreven perspectief op digitale innovatie betekent op het terrein van nationale veiligheid, economische

voorspoed en maatschappelijk welzijn (zie tabel 1). Nationale veiligheid is niet voldoende gebaat bij wapenontwikkeling, wel bij zo min mogelijk wapenontwikkeling – gelet op de toestand van de wereld – en zo veel mogelijk wapenbeheersing. Het blind stimuleren van digitale innovatie leidt tot veel techno-economische innovatie en is tevens een formule voor georganiseerde onverantwoordelijkheid en sociale catastrofes. Het is de kunst om technische en sociale innovatie hand in hand te laten gaan. Dat vraagt om het gelijktijdig stimuleren van beide vormen van innovatie. Dat is precies hoe Vlaanderen digitale innovatie vormgaf en -geeft op het gebied van e-government en de 'slimme stad'. Het Rathenau Instituut pleit in Nederland voor deze vorm van waardevol digitaliseren. Bij maatschappelijk gedreven innovatie weet men in België en Vlaanderen blijkbaar al hoe dat moet. Laat dat een zet in de rug zijn bij het vormgeven van het gehele pallet aan digitale innovatie, zodat die op een positieve manier bij kan dragen aan de democratische zoektocht naar het goede digitale leven.

Naschrift: 'Digitale vaardigheden zijn ook voor kappers van belang'

Het lijkt me voor de discussie over digitalisering goed wanneer helder in kaart gebracht wordt hoe het in Vlaanderen en België gesteld is met de *governance* van maatschappelijke en ethische kwesties die door digitalisering zijn ontstaan. In de Rathenau studie *Opwaarderen* (Kool et al. 2017) en aan het begin van paragraaf 3 staat uitgebreider om welke vragen het daarbij gaat.

Tijdens de stakeholderbijeenkomsten werd regelmatig de zorg geuit dat er een grote digitale kloof dreigt te ontstaan in de Belgische samenleving: er is een digitale kloof tussen leerkrachten en er bestaat ook een digitale kloof tussen leerlingen. Die laatste wordt bestendigd door de mythe van de *digital native*, het idee dat als kinderen opgroeien in een digitale wereld zij automatisch mediawijs worden. De uitspraak 'Digitale vaardigheden zijn ook voor kappers van belang' van Greet Vanderbiessen (Katholiek Onderwijs Vlaanderen) maakte diepe indruk op me. Deze zin raakt namelijk op drie manieren de kern van de digitaliseringsopgave. Hij laat ten eerste zien dat digitalisering iedereen in de samenleving aangaat en stelt tevens dat digitalisering voor iedereen een positief perspectief dient te hebben, en dus eerlijk en inclusief dient te zijn. Ten slotte geeft Vanderbiessen aan dat iedere burger recht heeft op digitale scholing en vaardigheden. Dit sluit direct aan bij mijn mening dat technologisch burgerschap dé democratische uitdaging is van deze eeuw (Est 2016). Politiek-bestuurlijke instituties dienen een drietal vaardigheden mogelijk te maken en te stimuleren. Burgers dienen ten eerste optimaal gebruik te kunnen maken van de zegeningen van technologie. Daarnaast dienen ze weerbaar te zijn tegen de risico's van technologie en daartegen beschermd te worden. Ten slotte dienen burgers in staat te zijn om deel te nemen aan het publieke en politieke debat over de toekomst van onze digitale samenleving. Slechts op die manier kunnen burgers hun rechten en plichten in het digitale tijdperk op een democratische wijze opeisen. En daarom zijn digitale vaardigheden dus zeker ook voor kappers van belang.

2.5. Referenties

Agoria (2018) *Shaping the future of work – Digitalisering en de Belgische arbeidsmarkt*. Brussel: Agoria.

ALLEA (2019) *Flourishing in a data-enabled society*. ALLEA discussion paper #4. All European Societies (ALLEA) and Royal Society.

Beck, U. (1996 [1986]) *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Mainz: Suhrkamp.

Est, R. van (2012) 'De ideologische leegte van het techniekdebat: Het grote zwijgen over hoe de informatierevolutie ons verandert', in: C. van Monfort, A. Michels & W. Van Dooren (red.) *Stille ideologie: Onderstromen in beleid en bestuur*. Den Haag: Boom Lemma, pp. 85-102.

Est, R. van, met medewerking van V. Rerimassie, I. van Keulen & G. Dorren (2014) *Intieme technologie: De slag om ons lichaam en gedrag*. Den Haag: Rathenau Instituut.

Est, R. van & V. Rerimassie (red.)(2014) *Strijd om onze intimiteit. Bericht aan het Parlement*. Den Haag: Rathenau Instituut.

Est, R. van (2016) 'Technologisch burgerschap als dé democratische uitdaging van de eenentwintigste eeuw'. *Christen Democratische Verkenningen* 3: 108-115.

Est, R. van, E. de Bakker, J. van den Broek, J. Deuten, P. Diederer, I. van Keulen, I. Korthagen & H. Voncken (2018) *Waardevol digitaliseren: Hoe lokale bestuurders vanuit publiek perspectief mee kunnen doen aan het 'technologiespel'*. Den Haag: Rathenau Instituut.

EZK (juni 2018) *Nederlandse digitaliseringsstrategie: Nederland Digitaal – Hier kan het. Hier gebeurt het*. Den Haag: Ministerie van Economische Zaken en Klimaat.

Havelaar, R. & P.J. Dijkman (2019) *Mens-zijn in de digitale samenleving: Perspectief op technologisch burgerschap*. Den Haag: Wetenschappelijk Instituut voor het CDA.

Hof, C. van 't, J. Timmer & R. van Est (red.)(2012) *Voorgeprogrammeerd: Hoe internet ons leven leidt*. Den Haag: Boom Lemma.

Hofman, L. (2019) 'Op de grootste wapenbeurs ter wereld zag ik de oorlog van de toekomst'. *De Correspondent* 2 oktober.

Jong, R. de, L. Kool & R. van Est (maart 2019) *Zo brengen we AI in de praktijk vanuit Europese waarden*. Den Haag: Rathenau Instituut.

Kool, L., J. Timmer, L. Royakkers & R. van Est (2017) *Opwaarderen: Borgen van publieke waarden in de digitale samenleving*. Den Haag: Rathenau Instituut.

Kool, L., E. Dujso & R. van Est (2018) *Doelgericht digitaliseren: Hoe Nederland werkt aan een digitale transitie waarin mensen en waarden centraal staan*. Den Haag: Rathenau Instituut

- Marshall, A. & A. Davies (2019) 'Uber's self-driving car didn't know pedestrians could jaywalk'. *Wired* May 11.
- Nemitz, P.F. (2018) *Constitutional democracy and technology in the age of Artificial Intelligence*. Royal Society Philosophical Transactions A. August.
- Peters, B. G. & J. Pierre (2009) 'Governance approaches', in: A. Wiener & T. Diez (eds.) *European integration theory* (2nd edition). Oxford: Oxford University Press. pp. 91-104.
- Raworth, K. (2017) *Doughnut economics: Seven ways to think like a 21st-century economist*. London: Penguin.
- Reyde, M. vande (2019) 'Opinie: 5G in communautaire frigo laten komt ons duur te staan', in: *De Tijd*, 29 november.
- Rockström, J., et al. (2009) 'A safe operating space for humanity', in: *Nature* 461: 472-475.
- Royakkers, L., F. Daemen & R. van Est (2012) *Overal robots: Automatisering van de liefde tot de dood*. Den Haag: Boom Lemma.
- Schomberg, R. von & J. Hankins (eds.)(2019) *International Handbook on Responsible Innovation: A global resource*. Cheltenham: Edward Elgar.
- SERV (juli 2018) *De transitie naar een digitale samenleving: Aanbevelingen en actie*. Brussel: Sociaal-Economische Raad van Vlaanderen.
- Snijders, D., S. Horsman, L. Kool & R. van Est (2019) *Verantwoord virtueel: Bescherm consumenten in virtual reality*. Den Haag: Rathenau Instituut.
- Tenner, E. (1997) *Why things bite back: Technology and the revenge of unintended consequences*. Vintage.
- Tweede Kamer (13 september 2018) *Rondetafelgesprek Digitaliseringsstrategie in de praktijk*. Den Haag: Tweede Kamer der Staten Generaal.
- Vedder, A., L. van der Wees, B.-J. Koops & P. de Hert (2007) *Van privacyparadijs tot controlestaat? Misdaad- en terreurbestrijding in Nederland aan het begin van de 21ste eeuw*. Den Haag: Rathenau Instituut.
- Verhoeven, K., M. van Vliet & J. Gerritsen (2019) *Digitale revolutie: Een politiek antwoord op de datamacht van techreuzen en controlestaten*. Den Haag: D'66.
- Wouters, R. (red.)(2019) *Handvest voor de slimme stad: Technologie sturen op basis van waarden*. Utrecht: Wetenschappelijk Bureau GroenLinks.
- Zuboff, S. (2019) *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. London: Profile Books.

3. Essay: Technologisch burgerschap en het ethisch begeleiden van digitalisering

PETER-PAUL VERBEEK

3.1. Inleiding

Het jaar 2019 zal de geschiedenis ingaan als het jaar van de ethiek van de digitale technologie – met een hoofdrol voor kunstmatige intelligentie. Bedrijven, overheden, maatschappelijke organisaties: er is haast geen speler in onze samenleving te noemen die niet op de een of andere manier heeft gewerkt aan een ethische code.⁹ De reden hiervoor is ongetwijfeld dat de ontwikkelingen in digitale technologie momenteel zo snel gaan en maatschappelijk zo goed voelbaar worden dat er steeds meer zorgen ontstaan over de maatschappelijke gevolgen ervan: zijn we wel in staat om in een digitale samenleving kernwaarden als vrijheid, rechtvaardigheid en solidariteit overeind te houden? Vragen als deze lagen ook mede ten grondslag aan het Denkersprogramma van de KVAB rondom digitale technologie, waar dit essay onderdeel van uitmaakt.¹⁰

Het ontstaan van de digitale samenleving wordt wel aangeduid als 'de vierde industriële revolutie'. Na de uitvinding van de stoommachine in de eerste revolutie bracht de tweede revolutie de massaproductie en de derde de informatisering. In de vierde revolutie wordt die informatietechnologie *fysiek*: in het tijdperk van de digitalisering wordt het internet een 'Internet der Dingen' en krijgt de computer een lichaam als *robot*. Daardoor raakt digitale technologie verweven met alle facetten van de samenleving en wordt steeds duidelijker zichtbaar dat er een nieuwe maatschappij aan het ontstaan is. Van de politiek tot de zorg en van liefdesrelaties tot de openbare ruimte: overal verandert digitale technologie de infrastructuur van ons dagelijks leven.

In Japan wordt de maatschappij die ontstaat uit deze vierde revolutie inmiddels al aangeduid als '*Society 5.0*': de vijfde samenleving.¹¹ Vanuit dit gezichtspunt bracht de ploeg de overgang van een samenleving van jagers-verzamelaars naar de agrarische samenleving, waarna de stoommachine de industriële samenleving mogelijk maakte en de computer tot de informatiesamenleving leidde. De vierde revolutie leidt nu tot het vijfde samenlevingstype: de digitale samenleving, in Japan ook wel aangeduid als de '*super smart society*'.

⁹ UNESCO heeft een aantal van deze initiatieven verzameld in dit overzicht: <https://unesdoc.unesco.org/ark:/48223/pf0000369756>

¹⁰ Een gedeelte van dit essay is toegestuurd aan de Vaste Kamercommissie Economische Zaken en Klimaat van het Nederlandse parlement, ter voorbereiding van een rondetafelgesprek over kunstmatige intelligentie op 20 februari 2020.

¹¹ Zie: https://www8.cao.go.jp/cstp/english/society5_0/index.html

Wat betekent deze digitale revolutie voor de maatschappij waarin wij leven, en in het bijzonder voor de Vlaamse maatschappij? De gesprekken die in 2019 werden georganiseerd rondom het thema 'digitalisering' in het kader van het Denkersprogramma van de KVAB hebben vele waardevolle inzichten opgeleverd om deze vraag te beantwoorden. In dit essay wil ik mij op grond van die inzichten primair richten op de vraag naar het verantwoord vormgeven van *burgerschap* in de digitale samenleving. Veel van de uitdagingen waarvoor de digitale revolutie ons stelt, bleken namelijk juist in de thematiek van technologisch burgerschap bij elkaar te komen: die thematiek maakt het nodig om de vraag te beantwoorden hoe mensen volwaardig deel uit kunnen maken van een technologische samenleving, en daarmee om de impact van digitale technologie op mens en maatschappij beter te doorgronden, en de knelpunten voor burgerschap in kaart te brengen. Drie domeinen bleken in deze context van bijzonder belang: de openbare ruimte, het onderwijs, en de politiek.

In dit essay zal ik allereerst in algemene zin ingaan op de nieuwe relaties tussen mens en technologie die de digitale revolutie met zich meebrengt. Wat is het karakter van deze relaties, en hoe kunnen we de nieuwe impact van technologie op mens en maatschappij duiden? Daarna zal ik in drie afzonderlijke paragrafen ingaan op digitalisering in relatie tot de openbare ruimte, het onderwijs en de politiek. Welke invloed heeft technologie in deze domeinen, welke waarden staan erin centraal, en hoe kunnen we die waarden optimaal de ruimte geven? Tot slot zal ik de contouren schetsen van een algemeen kader voor het ethisch begeleiden van digitale technologie als basis voor technologisch burgerschap.

3.2. De relaties tussen de mens en digitale technologie

Digitale technologieën vormen geheel nieuwe configuraties in de relaties tussen mens en techniek. Het klassieke beeld dat technologieën worden 'gebruikt' als middel voor een doel gaat niet meer op voor de meeste digitale technologieën. Het onderscheid tussen een 'actieve gebruiker' en een 'passieve technologie' maakt plaats voor een veel nauwere verwevenheid van mens en techniek. In veel gevallen is er bijvoorbeeld eerder sprake van 'immersie' dan van 'gebruik'. Het Internet der Dingen rust onze materiële omgeving uit met sensoren die mensen waarnemen, onderling communiceren en menselijk gedrag beïnvloeden. Zulke technologieën gebruiken we niet, we zijn er eerder in ondergedompeld. In andere gevallen treedt er eerder een 'fusie' op, zoals wanneer diepe hersenstimulatie de stemmingen van mensen beïnvloedt of insulinepompjes op eigen kracht de suikerspiegel van mensen met diabetes reguleren. Augmented reality-brillen leiden tot een configuratie van 'augmentatie', waarin technologie een tweede laag van interpretatie over de werkelijkheid heen legt, die simultaan verwerkt kan worden. En kunstmatig intelligente systemen maken configuraties mogelijk die voorheen alleen tussen mensen onderling bestonden, zoals interactie (robot en mens reageren op elkaar, zijn erop gericht elkaar te begrijpen en worden door

elkaar beïnvloed) of coöperatie (mens en algoritme zijn op dezelfde werkelijkheid gericht en ontwikkelen zichzelf op grond van hun onderlinge interactie, zoals medische expertsystemen die advies geven aan een arts vanuit een gezamenlijke gerichtheid op dezelfde patiënt).

Van de technologieën uit deze vierde revolutie is kunstmatige intelligentie (verder ook AI: artificiële intelligentie) zonder enige twijfel de meest in het oog springende – en tevens zowel de meest gevreesde als de meest bejubelde. Terwijl de VS, China en de EU in een digitale concurrentiestrijd verwickeld zijn geraakt, waarschuwen invloedrijke techneuten als Elon Musk en Bill Gates voor het gevaar dat deze kunstmatige intelligentie de mens gaat overheersen, en schetsen bestsellerauteurs als Yuval Noah Harari angstaanjagende toekomstbeelden. Enkele jaren geleden richtte de ethiek zich vooral op 'mensenverbetering': zijn we niet onderweg naar een 'supermens', die nieuwe ongelijkheid, onrechtvaardigheid en onwaardigheid met zich meebrengt? Inmiddels gaat de discussie eerder over 'mensvervanging': zijn we onszelf niet kwetsbaar aan het maken voor overheersing door onze eigen technologie? Van angst voor 'te veel mens' zijn we opgeschoven naar angst voor 'te weinig mens'.

AI is in deze zin exemplarisch voor de digitale revolutie waar onze samenleving zich in bevindt. Deze technologie stelt computers met behulp van algoritmes in staat om dingen te doen die voorheen voorbehouden waren aan mensen: van diagnoses stellen en beslissingen nemen tot vonnissen vellen en robot-lichamen besturen, van het activeren van de stormvloedkering en het analyseren van het studiegedrag van leerlingen tot het genereren van zoekresultaten op het internet en het live vertalen van gesproken woord. Het is zeer waarschijnlijk dat kunstmatige intelligentie een grote invloed op onze samenleving zal hebben. En het is niet altijd eenvoudig om met die impact om te gaan, omdat ze door bestaande kaders breekt. Tot nu toe zagen we technologie namelijk vaak als 'passief' en de mens als 'actief', maar kunstmatige intelligentie bezit een technologische vorm van 'autonomie'. Computersystemen kunnen 'leren' en daardoor ontwikkelen ze zich op manieren die hun programmeurs niet altijd hebben kunnen voorzien, en zijn ze in staat om zelfstandig beslissingen te nemen en mensen te adviseren over zaken die het menselijk denken minder goed kan overzien dan een computer met kunstmatige intelligentie.

De kern van de maatschappelijke impact van AI ligt in haar invloed op het *cognitieve functioneren* van de mens: AI verandert het menselijke denken. Historisch gezien is dat niet voor het eerst. Het schrift heeft bijvoorbeeld grote invloed gehad op hoe ons geheugen werkt: Plato was al ongerust dat we niets meer zouden kunnen onthouden als we alles zouden gaan opschrijven. Ook de boekdrukkunst is zo'n cognitieve technologie: opeens was kennis niet meer opgesloten achter kloostermuren, wat leidde tot een democratisering ervan en uiteindelijk de basis legde voor de wetenschappelijke revolutie. Wat is de nieuwe stap in ons denken die

AI teweeg gaat brengen? Wat gaat er gebeuren met ons begrip van de wereld en van onszelf, onze morele en politieke opvattingen en besluiten, onze existentiële vragen en kwetsbaarheden? En hoe zal AI invloed hebben op het functioneren van onze democratie, op de relatie tussen burgers en overheid, en op mogelijkheden tot participatie aan de samenleving?

Zoals ook bij veel andere technologieën het geval is, is het zinvol om in de impact van digitale technologie drie niveaus te onderscheiden. Op *individueel* niveau beïnvloedt ze de keuzes, gedragingen en interpretatiekaders van individuele gebruikers: zoekmachines kleuren met welke informatie mensen hun wereldbeeld opbouwen, expertsystemen geven artsen en rechters een kader voor een diagnose of een vonnis, slimme steden sturen automobilisten naar specifieke parkeerplaatsen en regelen dynamisch de maximumsnelheid. Op *sociaal* niveau werkt deze invloed door in sociale domeinen zoals de zorg en de rechtspraak (via expertsystemen), verkeer en vervoer (via zelfrijdende voertuigen), de journalistiek (in de vorm van *automated journalism*). Op *maatschappelijk* niveau ten slotte beïnvloedt digitale technologie de organisatie van de samenleving: *predictive policing* verandert de aard van de handhavende taak van de politie, zoekmachines beïnvloeden politiek engagement, algoritmes zijn betrokken in militaire besluitvorming.¹²

3.2.1. Het huidige ethische debat

Om de maatschappelijke implicaties van digitale technologie, en met name kunstmatige intelligentie, goed tegemoet te kunnen treden, is er de afgelopen tijd een veelheid aan ethische analyses en kaders ontwikkeld.¹³ In die analyses komen stevast een aantal kernthema's terug. Allereerst draait de discussie vaak om *bias*: AI-systemen kunnen bevooroordeeld zijn, door hun eigen ontwerp of door de data waarmee ze getraind zijn. Een tweede centraal thema is *accountability*: in hoeverre is het mogelijk aansprakelijkheid te organiseren voor het functioneren en de impact van kunstmatig intelligente systemen, bijvoorbeeld voor besluiten die mede genomen zijn op basis van AI? Daarnaast zijn *transparantie* en *uitlegbaarheid* belangrijke ethische issues: in welke mate kan een systeem uitleggen hoe het tot zijn conclusies is gekomen, zodat mensen kunnen overzien op grond waarvan ze het systeem kunnen vertrouwen? Ten slotte wordt ook de *veiligheid* van de systemen vaak benoemd: in hoeverre zijn ze betrouwbaar en bestand tegen vijandig gebruik (*dual use*)?

Al deze issues gaan impliciet uit van een strijd tussen mens en technologie. Computers kunnen dingen die vroeger alleen mensen konden, maar omdat ze het

¹² Een korte video met daarin een kernachtig beeld van de ethische uitdagingen van AI is te vinden op: <https://www.youtube.com/watch?v=TlihPx4z0Nw&t=14s>

¹³ Zie: Roos de Jong, *Het debat over de betekenis van AI begint pas net*: Rathenau Instituut, 2019 (<https://www.rathenau.nl/nl/maakbare-levens/het-debat-over-de-betekenis-van-ai-begint-pas-net>).

beter en sneller kunnen, en op een manier die mensen niet altijd helemaal kunnen overzien, worden ze bedreigend. Als computers sommige vormen van kanker al beter kunnen herkennen dan pathologen, wat gaat dat dan betekenen voor de rol van de arts? En wat is de toekomst van de journalistiek als nieuwssites niet alleen gevuld worden door journalisten maar ook door kunstmatig intelligente systemen? Hoe kunnen we controleren of het systeem de juiste analyse maakt en niet bevooroordeeld is, en kunnen we überhaupt wel verantwoordelijkheid nemen voor beslissingen die tot stand komen op basis van AI?

Vanuit dit tegenover elkaar plaatsen van mens en technologie richt de ethiek zich als vanzelf op het beschermen van de menselijke autonomie tegen de nieuwe vorm van 'autonomie' die de techniek ontwikkelt. De centrale vraag is dan of er nog voldoende menselijke controle mogelijk is, en of het systeem de mens of groepen mensen niet te veel buitensluit. Niet voor niets richt de High Level Expert Group for the Ethics of AI van de EU zich in haar recente rapport¹⁴ expliciet op *trustworthy AI*: de ethiek moet het vertrouwen tussen mens en techniek waarborgen.

Er is echter ook een hele categorie ethische vragen die hierbuiten valt, en die juist van belang zijn voor de thematiek van technologisch burgerschap. Die vragen richten zich minder op het individu, en meer op de gemeenschap. *Inclusiviteit* is zo'n thema: hoe waarborgen we diversiteit in onze samenleving? *Welzijn* is een ander voorbeeld: hoe ontwikkelen we kunstmatige intelligentie die een positieve bijdrage levert aan het floreren van de gehele samenleving?

In dit licht is het interessant te vermelden dat de World Commission for the Ethics of Science and Technology (COMEST) van de UNESCO een heel andere invalshoek gekozen in deze discussie. Kunstmatige intelligentie, zo is de gedachte, heeft invloed op de menselijke geest. En dat is precies het domein van de UNESCO, met haar focus op onderwijs, wetenschap, cultuur en communicatie. De *Preliminary Study on the Ethics of Artificial Intelligence* die COMEST onlangs publiceerde,¹⁵ richt zich dan ook precies op deze domeinen. Hoe kan het *onderwijs* mensen opleiden voor een wereld waarin technologie sommige banen overbodig maakt? Hoe kan het kritisch denken stimuleren en ingenieurs leren om AI verantwoord te ontwerpen? Wat gaat AI betekenen voor de *wetenschap*: hoe verandert ons begrip van wat een 'wetenschappelijke verklaring' inhoudt, en op welke 'feiten' baseren we straks beslissingen? Wat doet AI met *cultuur* en culturele diversiteit? Hoe kunnen we ook de kleine talen een toekomst geven als automatische vertaling steeds normaler wordt? Hoe kunnen de suggesties die platformen als Netflix en Spotify met behulp van kunstmatig intelligente systemen doen aan hun gebruikers, ook ruimte maken voor diversiteit? En tot slot: hoe kan de kwaliteit van menselijke

¹⁴ Zie: <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

¹⁵ Zie: <https://unesdoc.unesco.org/ark:/48223/pf0000367823>

communicatie worden bewaakt, als AI nieuwe vormen van desinformatie mogelijk maakt en steeds nauwer verweven raakt met de journalistiek?

In deze analyses neemt COMEST expliciet een mondiaal perspectief in, dat niet alleen is gericht op het waarborgen van individuele vrijheid maar ook op het floreren van de gemeenschap. Daarom vraagt COMEST ook aandacht voor een aantal bredere *global ethics*-thema's, die in het werk van UNESCO altijd terugkeren: Afrika, gender, duurzaamheid en vrede. Hoe zorgen we ervoor dat Afrika volop betrokken blijft in de ontwikkeling en implementatie van AI? Hoe kunnen we voorkomen dat AI nieuwe vormen van gender-bias creëert? Hoe realiseren we een duurzame technologische infrastructuur voor AI, in termen van zeldzame aardmetalen en de energieconsumptie van servers? En hoe zorgen we ervoor dat AI vreedzaam en niet gewelddadig wordt gebruikt?

3.2.2. *Burgerschap in de digitale samenleving*

Het bovenstaande maakt duidelijk dat de nieuwe configuraties van mens en techniek die digitale technologie met zich meebrengt grote implicaties hebben voor mens en maatschappij. Technologie geeft op een nieuwe manier vorm aan ons doen en laten en zelfs aan ons denken, wat ons voor de uitdaging plaatst om die invloed nader te analyseren en er een kritische verhouding toe in te nemen. Dat is nu precies wat 'digitaal burgerschap' vereist: om volwaardig deel uit te maken van de digitale samenleving is het nodig om de invloed van digitale technologie op mens en maatschappij te doorgronden en kritisch te evalueren, en om deze technologie op een verantwoorde manier te ontwerpen, te gebruiken en in te bedden in de samenleving.

Om deze notie van technologisch burgerschap nader uit te werken zal ik in wat volgt onderzoeken welke implicaties digitale technologie heeft voor drie domeinen die een centrale rol hadden in de gesprekken met stakeholders binnen het Denkersprogramma en die ook een belangrijke relatie hebben met burgerschap: de openbare ruimte (paragraaf 3), het onderwijs (paragraaf 4) en de politiek (paragraaf 5).

3.3. De digitale openbare ruimte

In de openbare ruimte speelt digitale technologie een steeds grotere rol, omdat digitale technologie zelf een nieuwe openbare ruimte is gaan vormen, en omdat digitale technologie zich ook vermengt met de fysieke openbare ruimte, bijvoorbeeld in het ontwerp van *smart cities*. Een aantal kernthema's keerde terug in de gesprekken met stakeholders, zowel ten aanzien van de onlinewereld als ten aanzien van de gebouwde omgeving.

3.3.1. De openbare ruimte online

Een van de meest centrale ethische thema's ten aanzien van de digitale openbare ruimte is *desinformatie* – ook wel aangeduid als 'fake news'. Het onderscheid tussen echt en onecht wordt in onze digitaliserende wereld steeds minder evident, en is ook steeds moeilijker vast te stellen. Dat heeft belangrijke politieke consequenties, omdat dit het maatschappelijk debat sterk bemoeilijkt. Hannah Arendt waarschuwde er in *The Orgins of Totalitarianism* al voor: 'The ideal subject of totalitarian rule is not the convinced Nazi or the convinced Communist, but people for whom the distinction between fact and fiction (i.e., the reality of experience) and the distinction between true and false (i.e., the standards of thought) no longer exist.' Als meningen en besluiten niet meer kunnen rusten op een gedeelde basis van feiten, vervalt de mogelijkheid om een samenleving bijeen te houden door het bereiken van overeenstemming en gedeeld inzicht.

Een tweede terugkerend thema is dat van *internetverslaving*. De tweede werkelijkheid die de onlinewereld creëert, raakt in toenemende mate in concurrentie met de echte, fysieke werkelijkheid. De eigen aard van de sociale interactie die sociale media met zich meebrengen heeft een aanzuigende en soms ook verslavende werking, door de permanente mogelijkheid om zaken te delen, daar reactie op te ontvangen en op de hoogte te blijven van wat anderen delen. De digitale openbare ruimte is daardoor voor sommigen niet zozeer een ontmoetingsplaats in aanvulling op de fysieke wereld, maar eerder een ruimte die hen uit de fysieke wereld weghoudt. De vele notificaties die mensen ontvangen via hun mobiele telefoon leiden hen permanent af, en trekken hen uit de fysieke wereld in de digitale wereld, waarbij de behoefte aan erkenning (*likes* op berichten) en de 'beloning' van het steeds weer vinden van nieuwe informatie een belangrijke rol spelen.

Hier nauw mee verbonden is een derde thema dat terugkeerde in de gesprekken: dat van *digitale uitsluiting*. Nu digitale technologie een vitaal onderdeel is geworden van de infrastructuur van onze samenleving, is het van wezenlijk belang dat iedereen ook via deze infrastructuur volwaardig kan meedoen. Dat blijkt niet altijd het geval te zijn. Met name ouderen blijken een kwetsbare groep te zijn. Omdat veel ouderen onvoldoende vertrouwd zijn met digitale technologie, en ze die technologie ook niet altijd vertrouwen, ontstaat hier een nieuwe generatiekloof die aandacht verdient. Maar deze kloof, die overigens al ruim aandacht krijgt in het publieke debat, bleek niet de enige. Ook één generatie daaronder bestaat er een tweede kloof, die met name zichtbaar wordt in het onderwijs: de kloof tussen de *digital natives*, die geboren zijn in een digitale wereld, en de generatie boven hen, die wel vertrouwd is met digitale techniek maar deze op een andere manier gebruikt, en die vaak ook andere technologie gebruikt, met name op het gebied van sociale media.

Tot slot – en weinig verrassend – maakt de digitalisering van de openbare ruimte onze *privacy* tot iets kwetsbaars, het vierde thema dat een centrale rol speelde in de gesprekken met stakeholders. De *General Data Protection Regulation* (GDPR) heeft op Europees niveau weliswaar tot een betere privacybescherming geleid, maar de problematiek blijkt veel dieper te liggen dan het beschermen van data. Naast een waakhond die reeds bestaande technologie beoordeelt, blijkt er ook behoefte te zijn aan een systeemaanpak, waarbij ethische vragen over privacy ook in de ontwerp- en de inbeddingsfase van technologie worden meegenomen. Alleen op die manier kunnen we de waarde van privacy in balans brengen met andere waarden, die ook van cruciaal belang zijn voor onze samenleving, zoals veiligheid, solidariteit en duurzaamheid. Wanneer we te voorzichtig worden met het delen van data, en wanneer er niet voldoende wordt gezocht naar een balans tussen individuele rechten en collectieve belangen, verliezen we in onze strijd om de privacy bijvoorbeeld belangrijke mogelijkheden om op een meer duurzame manier met energie om te gaan, criminaliteit te bestrijden of onze medische data uit solidariteit met anderen te delen, zodat er medisch onderzoek mee gedaan kan worden. De GDPR is kortom onderdeel van een leerproces, dat nog veel verder moet gaan. Een integrale aanpak, waarbij regulering niet achteraf komt maar onderdeel uitmaakt van het innovatieproces, en waarbij een diversiteit aan waarden een rol speelt, lijkt belangrijk te zijn.

3.3.2. *Digitale technologie in de fysieke openbare ruimte*

Naast de openbare onlineruimte speelde ook de fysieke openbare ruimte een rol in de discussies met stakeholders. De toenemende digitalisering van de materiële omgeving verandert het karakter van de publieke infrastructuur. In *smart cities* worden wifi-signalen geanalyseerd om te kunnen meten hoe vaak mensen op bepaalde plekken zijn, worden gegevens verzameld over zaken als verkeersstromen en luchtkwaliteit, en wordt het gedrag van mensen beïnvloed met dynamische verkeersborden en interactieve verlichting. Deze digitalisering van de openbare ruimte brengt een nieuwe verhouding tussen overheid en burger met zich mee, en daarmee nieuwe uitdagingen voor de democratie.

Om te voorkomen dat de overheid een te grote macht zou ontwikkelen en er te veel ruimte voor surveillance zou ontstaan, wordt er geëxperimenteerd met cocreatie. Burgers worden dan actief betrokken in de keuzes die op politiek niveau gemaakt moeten worden, zodat de inrichting van de *smart city* een gezamenlijke verantwoordelijkheid wordt. Het is zaak deze vormen van cocreatie zo democratisch mogelijk in te richten. Om burgers daadwerkelijk een stem te geven zou hun participatie verder moeten gaan dan het kiezen uit een aantal voorgegeven opties. In samenspraak met burgers zouden bijvoorbeeld de controversen kunnen worden geïdentificeerd die een rol spelen in de besluitvorming rondom digitale technologie

in de openbare ruimte.¹⁶ Wat staat er op het spel voor burgers, overheid en bedrijven? Welke waarden zijn in het geding? En hoe kunnen we deze waarden zo goed mogelijk representeren in het besluitvormingsproces? Bovendien is het van groot belang om een optimale inspanning te leveren en zo *alle* relevante stakeholders te betrekken dan wel hun belang te vertegenwoordigen: ook de stakeholders die juist buitengesloten raken of dreigen te raken in een digitale samenleving. 'Slimme steden' maken niet alleen nieuwe vormen van surveillance en gedragsbeïnvloeding mogelijk – die vooral ook een nieuwe macht van de overheid over de burger belichamen – maar tevens nieuwe vormen van politieke invloed en engagement. Sensoren kunnen burgers ook in staat stellen om zaken op de politieke agenda te zetten die anders maar moeilijk aandacht krijgen: overlast door vliegtuiglawaai bijvoorbeeld, of de luchtkwaliteit. Het *citizen science*-project Curieuze Neuzen is hiervan een goed voorbeeld, geïnitieerd door de Universiteit Antwerpen, de krant De Standaard en de Vlaamse Milieumaatschappij (<https://curieuzeneuzen.be>). Hier bewerkt digitale technologie het tegenovergestelde van 'fake news' en desinformatie: met behulp van sensoren wordt betrouwbare informatie beschikbaar gemaakt, die vervolgens een rol kan spelen in politieke processen. Technologie geeft hier mede vorm aan de totstandkoming van politieke issues: *sensing* speelt zo een belangrijke rol in *sense making*. In plaats van overheersing door de overheid te faciliteren geeft deze technologie juist nieuwe invloed aan de burger. En naarmate datastromen breder toegankelijk zijn, neemt ook zijn potentiële invloed toe.

3.3.3. Digitale geletterdheid

Een belangrijke sleutel om deze uitdagingen het hoofd te kunnen bieden lijkt te bestaan in het verder ontwikkelen van digitale geletterdheid. Nog te vaak worden technologie en samenleving als twee afzonderlijke domeinen benaderd: in opleidingen worden de sociale vakgebieden gescheiden van de technische. Daardoor komen maatschappelijke en technologische perspectieven ook in de praktijk te weinig bij elkaar. Het is van cruciaal belang dat we technologie leren 'lezen' als iets maatschappelijks en zo de kloof tussen alfa en bèta dichten. Digitale geletterdheid wordt nog te vaak ingevuld als 'computervaardigheden' of 'leren programmeren'. Het zou moeten worden uitgebreid tot het leren begrijpen van de interactie tussen digitale technologie, de mens en de maatschappij. Deze 'geletterdheid' maakt onderdeel uit van 'technologisch burgerschap': het vermogen om volwaardig deel uit te kunnen maken van de digitale samenleving.

Digitale geletterdheid maakt een kritische houding ten opzichte van digitale technologie mogelijk en biedt op die manier tegenwicht tegen desinformatie, verslaving, uitsluiting en privacybedreigingen. Alleen door de nauwe verbindingen

¹⁶ Zie: <https://www.utwente.nl/en/project-portal/!/project/586053/responsible-smart-city-design>

tussen mens en technologie te begrijpen wordt het mogelijk om desinformatie beter te doorzien, verantwoord om te gaan met de impact van digitale technologie op je eigen leven en op de samenleving, en voluit deel uit te maken van de digitale samenleving.

3.4. Onderwijs in een digitale samenleving

De noodzaak om de digitale geletterdheid te vergroten is nauw verbonden met een tweede thematiek die een grote rol speelde in de gesprekken met stakeholders: de relatie tussen digitalisering en onderwijs. De huidige mondiale crisis vanwege de coronaviruspandemie heeft deze discussie nog urgenter gemaakt dan ze al was: vanwege de noodzaak zo veel mogelijk vanuit huis te werken is de digitalisering van het onderwijs in een stroomversnelling geraakt. Dit uit noodzaak geboren experiment met een ver doorgevoerde digitalisering van het onderwijs biedt vele mogelijkheden om er lessen uit te leren.

De relatie tussen digitalisering en onderwijs heeft niet alleen betrekking op de *inhoud* van het onderwijs (waaronder de nodige aandacht voor digitale geletterdheid) maar ook op het *proces* (de digitalisering van de onderwijsomgeving en de digitale kloof tussen generaties) en de *maatschappelijke rol* ervan, in een wereld waarin 'technologische werkeloosheid' ontstaat en steeds meer nieuwe competenties en vaardigheden nodig zijn. Ook hier is het van belang op te merken dat het in deze discussie van groot belang bleek om aandacht te hebben voor het perspectief van stakeholders die niet aan tafel zitten: juist de groep mensen die buitengesloten zou kunnen raken in een digitale samenleving dient een stem te hebben.

3.4.1. Digitalisering en de onderwijspraktijk

De praktijk van het geven en ontvangen van onderwijs is de afgelopen decennia ingrijpend veranderd door de komst van digitale technologie. Nadat digitale leeromgevingen de interactie tussen leerlingen en docenten, en ook tussen leerlingen onderling, ingrijpend veranderd hebben, doet nu ook het fenomeen *learning analytics* zijn intrede. Door zo veel mogelijk data te verzamelen over het leergedrag van leerlingen en hun interactie met elkaar en hun docenten wordt het mogelijk om correlaties op te sporen die uiteindelijk wellicht tot verbetering van leerprestaties kunnen leiden. Deze data kunnen afkomstig zijn uit het onlinegedrag van leerlingen in digitale leeromgevingen, maar ook van sensoren in onderwijsgebouwen.

Zulke technologieën brengen ethische en maatschappelijke vragen met zich mee. Het monitoren van leer- en studiegedrag heeft uiteraard een dimensie van privacy en surveillance: hoe ver willen en mogen docenten en onderwijsinstellingen gaan in het volgen van de activiteiten van leerlingen? Minstens zo belangrijk is de impact op de leerprocessen zelf. Door bepaalde aspecten van dat leerproces

meetbaar te maken, zoals patronen van aan- en afwezigheid en van interactie met het studiemateriaal, kunnen die aspecten een doel op zich worden. Een leerproces omvat meer dan wat meetbaar gemaakt kan worden met digitale technologie. Om die reden is het van groot belang die technologie te ontwerpen, implementeren en gebruiken vanuit een expliciete visie op wat goed onderwijs is, en op de waarden die daarin centraal staan.

De digitale revolutie raakt de onderwijspraktijk ook op een tweede manier: door de al besproken digitale kloof. Uit de gesprekken met stakeholders kwam naar voren dat die kloof niet alleen bestaat tussen leerlingen en docenten, maar ook tussen docenten onderling. Oudere docenten staan vaak verder af van de digitale wereld van hun leerlingen dan jongere collega's. De technologische ontwikkelingen voltrekken zich in zo'n hoog tempo dat er een noodzaak tot permanente bijscholing lijkt te ontstaan. Om voldoende aan te sluiten bij de belevingswereld en denkwijze van jonge mensen zouden docenten zich vertrouwd moeten blijven maken met de technologie en applicaties die het dagelijks leven van jonge mensen mede bepalen. Een nascholingsprogramma, zoals dat ook bestaat voor medici, zou hier een goede vorm voor kunnen zijn. Door de lesbevoegdheid te koppelen aan een doorlopend opleidingstraject – eventueel gecompenseerd door een reductie van de administratieve taken – kan voorkomen worden dat de digitale kloof te groot wordt. De eerder genoemde thematiek van digitale geletterdheid zou in een dergelijk nascholingsprogramma een centrale rol kunnen hebben.

3.4.2. De rol van het onderwijs in een digitale samenleving

De digitale revolutie leidt niet alleen tot een andere onderwijspraktijk, maar ook tot een andere arbeidsmarkt waar het onderwijs leerlingen voor moet toerusten. Ten eerste verandert de aard van vele beroepen door de komst van digitale technologie, en daarmee ook de eisen die aan opleidingen moeten worden gesteld. Maar vooral ook *verdwijnen* er functies en beroepen, omdat bepaalde werkzaamheden kunnen worden overgenomen door technologie. Dit stelt nieuwe eisen aan het onderwijs.

De thematiek van 'technologische werkeloosheid' is al op vele manieren onderzocht en de conclusie is doorgaans dat er een 'polarisering' van de arbeidsmarkt optreedt.¹⁷ Het soms geschetste beeld dat met name de 'onderkant' van de arbeidsmarkt geraakt zou worden en dat vooral laagopgeleide arbeid zou verdwijnen, blijkt onjuist: het is juist het 'middensegment' dat de gevolgen van de digitale revolutie het sterkste voelt. In een gedigitaliseerde samenleving blijven mensen nodig die de technologische infrastructuur kunnen bouwen, onderhouden en repareren, en er blijft behoefte aan leidinggevenden en creatieve mensen: hun werk is niet zo gemakkelijk over te nemen door robots en algoritmes. Het zijn

¹⁷ Zie: <https://www.utwente.nl/en/project-portal/!/project/586053/responsible-smart-city-design>

vooral routinematige taken, bijvoorbeeld in de administratie en in de secretariële ondersteuning, die het gemakkelijkst door technologie uit te voeren blijken te zijn.

Tegelijkertijd blijken er niet alleen banen te *verdwijnen* onder invloed van technologische ontwikkelingen: er komen ook nieuwe voor in de plaats. Dat betekent dat de tweedeling in de arbeidsmarkt die door digitalisering dreigt te ontstaan niet zozeer loopt tussen degenen van wie het werk wel en niet door technologie overgenomen kan worden, maar eerder tussen degenen die wel en die niet de veerkracht hebben om zichzelf om te scholen en van beroep te veranderen. Juist hier kan het onderwijs een belangrijke rol spelen. Naast kennis en vakinhoud worden studievaardigheden en 'leren leren' steeds belangrijker. 'Levenslang leren' lijkt zo een nieuwe invulling te krijgen: technologische ontwikkelingen scheppen een zeer dynamische context voor de arbeidsmarkt, die vraagt om nascholing en bijscholing, maar soms ook om omscholing. In alle gevallen moeten mensen toegerust zijn met een basis om deze veerkracht te bezitten.

3.4.3. *Kritisch denken en transdisciplinariteit*

De digitalisering van de samenleving stelt tevens nieuwe eisen aan de *inhoud* van het onderwijs. Een centraal onderdeel daarvan kwam hierboven al aan de orde: digitale geletterdheid als het vermogen om de grondbeginselen van digitale technologie te begrijpen op zowel technologisch als sociaal gebied. Daar nauw aan verbonden zijn nog twee elementen die niet mogen ontbreken in het onderwijs in een digitale samenleving.

Allereerst is dat *kritisch denkvermogen*. In een wereld waarin kunstmatige intelligentie steeds belangrijker wordt, is het zaak om kritische distantie te kunnen innemen ten aanzien van de analyses van expertsystemen, de resultaten van zoekmachines, de antwoorden van *voice assistants*. Het onderwijs moet leerlingen leren om informatie niet voor lief nemen, maar altijd terug te vragen naar de herkomst ervan en de condities waaronder die informatie tot stand is gekomen. Waarop baseert een kunstmatig intelligent systeem zijn conclusies, met welke datasets is het getraind, in welke omgeving heeft het verbanden leren leggen?

Ten tweede gaat het om een *transdisciplinaire houding*. Het klassieke onderscheid tussen alfa en bèta, sociaal en technisch, mensgericht en exact, volstaat niet meer om de maatschappij van de 21ste eeuw te begrijpen. Digitale technologie is niet alleen iets technisch, maar ook iets sociaals. Sociale problemen laten zich niet alleen langs sociale maar ook langs technologische wegen oplossen. Technologie heeft een fundamentele invloed op mens en maatschappij en moet ook als zodanig benaderd worden bij het ontwerpen, implementeren en gebruiken ervan, net zoals de huidige samenleving alleen goed begrepen kan worden als we daarbij ook de invloed van technologie een plek geven.

3.5. Politiek en digitalisering

De digitalisering van de samenleving heeft ook grote gevolgen voor het functioneren van onze democratie. Voorbeelden daarvan zagen we al in de context van de digitalisering van de openbare ruimte. Door de enorme macht van grote technologiebedrijven blijkt het voor overheden zeer ingewikkeld te zijn om verantwoordelijkheid te nemen voor de ontwikkeling en inbedding van digitale technologie. En dat betekent dat het niet altijd goed mogelijk is om die digitale technologie onder democratische controle te krijgen, terwijl ze intussen wel diep ingrijpt in het functioneren van de samenleving. Het belang van digitale technologie is inmiddels dermate groot dat ze tot de primaire infrastructuur is gaan behoren. Zeker in het Europese model van democratie draagt normaliter de overheid daar zorg voor.

In gesprekken met stakeholders kwam deze democratische uitdaging helder naar voren. Zo blijkt het moeilijk om *privacy by design* af te dwingen bij ontwikkelaars van digitale technologie, en zo blijkt de Franse politie haar toevlucht te nemen tot methodes van hacking omdat de versleuteling van WhatsApp-berichten het opsporingswerk aanzienlijk bemoeilijkt. Tegelijkertijd wordt er gepleit voor softwareontwikkeling door de overheid zelf, die het mogelijk zou moeten maken data van burgers op een centraal, beveiligd punt te beheren en toegankelijk te maken, zonder daarbij apps van de grote tech-giganten te gebruiken.

Deze thematiek sluit aan bij de mondiale discussie van dit moment rondom ethische AI in relatie tot democratie. In deze discussie wordt de EU vaak gepositioneerd als alternatief voor het Amerikaanse en het Chinese model. In een soort *revival* van de 20ste-eeuwse zoektocht naar een Derde Weg tussen het kapitalistische Westen en het communistische Oosten zoekt de EU nu naar een kader voor AI waarin niet alleen commerciële belangen van bedrijven de hoofdrol spelen, en evenmin de behoefte van de staat om de samenleving te sturen. Deze zoektocht heeft onder andere geleid tot het ethische kader voor AI van de al genoemde High Level Expert Group on the Ethics of AI van de EU.

Het lijkt van groot belang om deze Derde Weg niet zozeer te positioneren als een *alternatief* voor het Chinese en Amerikaanse model, maar eerder als een *verbinding*. Vanuit het mondiale kader dat de UNESCO World Commission for the Ethics of Science and Technology heeft neergelegd, is juist een goede verbinding van individuele vrijheid en verantwoordelijkheid voor het collectief essentieel. Wanneer de EU zich presenteert als 'het ethische alternatief' of zoekt naar ruimte om de derde 'grote speler op de markt' te zijn, wordt die kans op verbinding onnodig beperkt. Gezien de te verwachten mondiale impact van AI is het juist zaak om niet verder te polariseren, maar op mondiale fora de verbinding op te zoeken.

De EU is voor deze verbinding van cruciaal belang, niet alleen als medeontwikkelaar van technologie, maar ook als proeftuin voor een succesvol beleid ten aanzien van technologie. Hoeveel kritiek er soms ook mag worden geuit op de effectiviteit en zinvolheid van Europees beleid en Europese wetgeving: als er één domein is waarin de EU uitermate effectief en belangrijk werk heeft gedaan, dan is het wel de informatietechnologie. De al besproken *General Data Protection Regulation* is daar een levend voorbeeld van. De EU slaagt er daadwerkelijk in om individuele en collectieve belangen met elkaar in balans te brengen en staatsinterventie samen te laten gaan met vrijheid en democratie. Naar dergelijke vormen van balans moeten we blijven zoeken: de mogelijkheid om verantwoord om te gaan met de impact van AI op de mensheid is hiervan afhankelijk. De voorhoedepositie die de EU heeft op het gebied van Responsible Research and Innovation vormt hiervoor een goede basis: er is ruime ervaring met het integreren van ethische en maatschappelijke reflectie in technologische ontwikkeling. Die ervaring moet ten volle worden benut en uitgebouwd.

3.6. Ethiek als begeleiding van technologie: de herkansing van Frankenstein

Om deze thema's goed te kunnen agenderen is het belangrijk om digitale technologie in de ethische en maatschappelijke discussie niet *tegenover* de mens te plaatsen maar te beschouwen als *verbonden* met de mens. Feitelijk is de huidige discussie over digitalisering een nieuwe versie van het Frankenstein-verhaal. In de vierde revolutie komt techniek alleen niet tot leven in vlees en bloed, maar via algoritmes en robots: of het nu gaat om technologische werkeloosheid, algoritmische besluitvorming of geautomatiseerde journalistiek, steeds is er de vrees dat technologie op de stoel van de mens gaat zitten, en schade aanricht. En hoewel de vrees voor negatieve implicaties van digitale technologie zeker niet ongegrond is, is het tegenover elkaar plaatsen van mens en technologie dat wel. Om de implicaties ervan helder te kunnen begrijpen is het immers juist nodig om te zien hoe digitale technologie *onderdeel* wordt van menselijke praktijken en vormgeeft aan interpretatiekaders en beslissingen – net zoals het schrift en de boekdrukkunst dat eerder deden.

Dit is precies wat er mis ging bij Frankenstein. Zijn monster werd alleen maar een monster omdat hij het in de steek liet. Victor Frankenstein raakte na de euforie over het tot leven komen van zijn schepsel namelijk zo in paniek dat hij wegvlochtte en het monster aan zijn lot overliet. Dat monster moest daardoor helemaal alleen zijn weg zien te vinden in de wereld. Vrij vertaald: als we willen voorkomen dat kunstmatig intelligente systemen en robots ontsporen, dan moeten we ze niet vrezen, maar opvoeden. Als we ons uitsluitend laten regeren door de angst voor wat er allemaal kan gebeuren, laten we ze in de steek in plaats van ze een goede plek te geven in de samenleving. Wie blijft hangen in de keuze tussen 'toestaan' of 'verbieden', slaagt er niet in door te dringen tot de werkelijke impact van digitale technologie.

Dat 'opvoeden' heeft in het geval van AI bovendien een heel letterlijke betekenis. De crux van kunstmatig intelligente systemen is immers dat ze *leren*. Ze ontwikkelen zich zelfstandig in interactie met hun omgeving en gaan dingen doen die hun ontwerpers er nooit zelf in hebben gestopt. Uiteindelijk is daardoor niet meer de ontwerper of de omgeving verantwoordelijk, maar het systeem zelf. En dus moet ethiek zich ook op die systemen zelf richten en op de manier waarop ze leren. Welke datasets gebruiken we om ze te trainen? En hoe bewerken de algoritmes deze datasets?

Ethiek van digitale techniek moet dan ook veel meer zijn dan het stellen van de vraag welke techniek we wel en niet willen. In plaats van digitale technologie alleen *van buitenaf te beoordelen* moet de ethiek haar *van binnenuit begeleiden*. Door centrale ethische waarden een plek te geven in het verantwoord ontwerpen ervan, en daarnaast in haar inbedding en gebruik. Algoritmes en systemen belichamen waarden, ontwikkelen zich in de richting van bepaalde waarden door de manier waarop ze getraind en ingebed worden, en worden vanuit waarden geïmplementeerd en gebruikt. Het selectie-algoritme waarmee Amazon nieuw personeel zocht is hiervan een iconisch voorbeeld geworden: door het profiel van de zittende succesvolle medewerkers als model te nemen bleek het algoritme uiteindelijk alleen nog blanke mannen tussen de veertig en de vijftig te selecteren en werd het personeelsbestand wel erg eenzijdig en discriminatoir. Doordat alleen de waarde van 'succesvol zijn' een plaats had, en 'inclusiviteit' en 'diversiteit' niet expliciet werden meegenomen in de ontwikkeling en training van het algoritme, had deze ontwikkeling negatieve gevolgen die voorkomen hadden kunnen worden.

Responsible Design en *Responsible Guidance* van technologie zijn daarom minstens zo belangrijk voor de ethiek van AI als het opstellen van codes en kaders. Op het gebied van *Responsible Design* is veel expertise in de EU, en ook in Nederland is hier een aantal groepen zeer actief;¹⁸ *Responsible Guidance*, dat zich sterker richt op implementatie en gebruik, is een thema dat op landelijk niveau is uitgewerkt vanuit de werkgroep Ethiek en Digitalisering van ECP, waarbij niet alleen kernspelers uit de Nederlandse overheid en techniekethiek aangesloten zijn, maar ook grote bedrijven als IBM, Microsoft en Facebook. Dit heeft geleid tot een *Aanpak begeleidingsethiek* die de enorme veelheid aan ethische codes en uitgangspunten wil vertalen naar concrete handelingsperspectieven.

Deze aanpak, zoals beschreven in Verbeek en Tijink 2019,¹⁹ kan van waarde zijn bij het ontwikkelen, inbedden en gebruiken van digitale technologie vanuit de

¹⁸ Deze expertise bevindt zich onder andere in het Delft design for values <https://www.delftdesignforvalues.nl> en het designlab van de Universiteit Twente. Vanuit de Universiteit Twente wordt ook een aangeboden op het gebied van de filosofie en ethiek van techniek en design. Zie ook: Verbeek, P.P. (2011), *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago and London: University of Chicago Press.

¹⁹ Verbeek, P-P., & Tijink, D. (2019). *Aanpak begeleidingsethiek: een dialoog over technologie met handelingsperspectief*. Leidschendam: ECP | Platform voor de InformatieSamenleving



waarden die centraal staan in de gebruikspraktijk. De aanpak begint bottom-up in plaats van top-down: de eerste stap is niet het formuleren van algemene principes, maar juist het in detail beschrijven van de technologie in haar concrete context van gebruik en implementatie. Op grond daarvan volgt in stap twee een analyse van de impact die de technologie heeft op alle betrokken actoren, om van daaruit de waarden te inventariseren die op het spel staan. Op grond hiervan wordt in de derde en laatste stap uitgewerkt welke handelingsopties er zijn om de geïdentificeerde waarden te bevorderen. Deze handelingsopties vallen uiteen in het ontwerp en herontwerp van de technologie, de aanpassing van de context waarin de technologie functioneert, en het kritisch toerusten van de gebruiker van de technologie.

3.7. Conclusie

Een concrete aanpak als deze kan een belangrijke rol spelen in technologisch burgerschap. Ze belichaamt en stimuleert de digitale geletterdheid die nodig is om ons kritisch te verhouden tot digitale technologie, ze verbindt een technologisch en een sociaal perspectief, en ze richt zich op de kwaliteit van digitale technologie in de context van het persoonlijke bestaan, maatschappelijke praktijken en de politiek. Ethiek van digitalisering is niet alleen een kwestie van abstracte principes en grote verhalen, maar vooral ook van concrete analyses en actiepunten. Juist op het gebied van verantwoord innoveren en ethiek van technologie ligt de kracht van het Europese perspectief, als verbinder tussen West en Oost. Gegeven de zeer hoge kwaliteit van alle gesprekken die wij hebben gevoerd met maatschappelijke partijen is er alle reden om erop te vertrouwen dat er in Vlaanderen een stevig fundament is voor een ethische begeleiding van digitalisering, en voor het verder vormgeven aan burgerschap in de digitale samenleving.

3.8. Referenties

Arendt, H. (1951). *The Origins of Totalitarianism*. New York: Harcourt, Brace and Co

COMEST (2019). Preliminary Study on the Ethics of Artificial Intelligence. (UNESCO World Commission for the Ethics of Science and Technology: <https://unesdoc.unesco.org/ark:/48223/pf0000367823>)

Jong, R. de (2019). Het debat over de betekenis van AI begint pas net. Rathenau Instituut website (<https://www.rathenau.nl/nl/maakbare-levens/het-debat-over-de-betekenis-van-ai-begint-pas-net>)

Smits, W. en J. De Vries (2015). 'Toenemende polarisatie op de Nederlandse arbeidsmarkt': ESB Jaargang 100 (4701), 8 januari 2015

UNESCO (2019). Report on the work of other international organizations and conventions on different aspects of artificial intelligence (AI). Item for 207th Executive Board session, 207 EX/5.I.A (<https://unesdoc.unesco.org/ark:/48223/pf0000369756>)

Verbeek, P.P. (2011), *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago and London: University of Chicago Press

Verbeek, P.P. (2019). De Ethiek van Kunstmatige Intelligentie: Verantwoord Ontwerp en de Herkansing van Frankenstein. Position paper voor hoorzitting Vaste Kamercommissie EZK over Strategisch Actieplan voor Kunstmatige Intelligentie, 20 februari 2020 (https://www.tweedekamer.nl/debat_en_vergadering/commissievergaderingen/details?id=2019A05632)

Verbeek, P.P. (2020). 'Politicizing Postphenomenology'. In: Miller G., Shew A. (eds) Reimagining Philosophy and Technology, Reinventing Ihde. *Philosophy of Engineering and Technology*, vol 33. Springer, Cham

Verbeek, P.P. en D. Tijink (2019). Aanpak Begeleidingsethiek. Den Haag: ECP (<https://ecp.nl/wp-content/uploads/2019/11/060-001-Boek-Aanpak-begeleidingsethiek-240165-binnenwerk-digitaal.pdf>)

4. Essay: Maatschappelijke waarden in digitale innovatie – Het perspectief van een technoloog

JAN RABAEY

4.1. Technologie en disruptie

Het lijkt geen twijfel dat nieuwe technologieën tot verstoringen leiden. Dat is in de hele geschiedenis van de mensheid het geval geweest. Denk aan hoe de introductie van georganiseerde landbouw de manier waarop mensen leefden en met elkaar omgingen fundamenteel heeft veranderd. Of hoe het schrift een collectief geheugen creëerde dat de tijd kon overleven. Deze 'revoluties' hebben ons gevormd tot wie we zijn als mens en hebben het functioneren van de samenleving bepaald. Het opmerkelijke is dat deze veranderingen de neiging hadden zich zeer snel te verspreiden, ook in een wereld waarin communicatiekanalen schaars en ad hoc waren. Denk bijvoorbeeld aan de innovatie waardoor sindsdien informatie door het schrift in het collectieve geheugen vastgelegd kon worden. Nadat het schrift rond 3000 voor Christus in Sumer werd 'uitgevonden', verspreidde het zich in nauwelijks een paar duizend jaar over de hele beschaafde wereld.

Al deze technologische innovaties leidden tot diepgaande veranderingen in de manier waarop mensen leefden, werkten en communiceerden. De Industriële Revolutie van rond 1800 veranderde fundamenteel de aard van het werk. De introductie van de auto in de vroege jaren 1900 zorgde voor een revolutie in transport en mobiliteit. Enzovoort. Altijd werd onze manier van leven beïnvloed en moesten we ons aanpassen. De mensheid is daar heel bedreven in gebleken. Men kan het oneens zijn over de waarde die wordt gecreëerd door technologische disrupties en over de voor- en nadelen, maar de realiteit is: zodra een technologie nuttig wordt geacht en wordt geaccepteerd, is er vrijwel geen manier om ze terug 'in de doos te stoppen'.

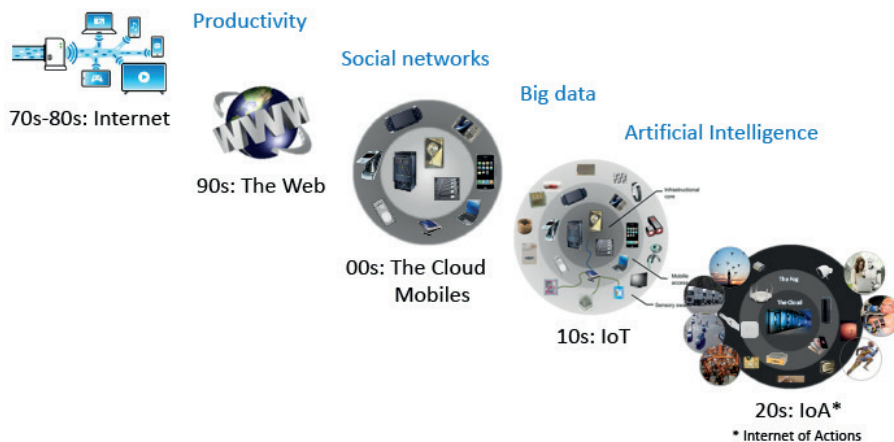


Fig. 4: De evoluerende voetafdruk van de digitale technologie

De digitale revolutie begon eigenlijk in de jaren 1950 en is daarna op verbazing-wekkende wijze veranderd en uitgebreid (figuur 1). Met elke innovatiefase kwamen nieuwe kansen, uitdagingen en verstoringen. Banen en industrieën werden aangetast en traditionele levenspatronen veranderden. De introductie van het internet in de jaren 1970 veranderde oorspronkelijk de aard van het wetenschappelijk onderzoek en verkortte daarna snel de tijdsspanne van informatie-uitwisseling, van weken tot seconden. De (personal) computer veranderde de werkplek en creëerde een heel ander perspectief op het concept 'productiviteit'. Maar het is misschien wel de gelijktijdige introductie van het wereldwijde web en de smartphone die de grootste effecten op de samenleving had, omdat ze traditionele modellen van communicatie, interactie en dagelijks leven radicaal omverwierp en een gevoel van instabiliteit veroorzaakte, in één of twee decennia tijd. Nog verontrustender is dat deze toestand van verandering en onevenwicht nog onophoudelijk aan de gang is en geen tekenen van verslapping vertoont. De combinatie van de introductie van het Internet of Things (en zijn vele varianten) en artificiële intelligentie (AI) verandert effectief ons perspectief van de fysieke wereld om ons heen en verbindt het rechtstreeks met een virtuele wereld, gecreëerd door onze eigen verbeelding. Binnenkort zullen we die wereld delen met autonome entiteiten, die ons zullen helpen om taken onafhankelijk of in nauwe samenwerking uit te voeren: het Internet of Actions (Institute of the Future 2016).

Het komt erop neer dat er nog meer snelle veranderingen zullen komen, of je dat nu wilt of niet. Dat zal impact hebben op ons mens-zijn en op hoe we als mensheid evolueren. De opportuniteit is nu dat we mee vorm kunnen geven aan die toekomst en dat we er tevens voor kunnen zorgen dat de opkomende digitale wereld beter, veiliger en eerlijker wordt.

Aan elke technologie die de mensheid heeft uitgevonden, ontwikkeld of geconcipieerd, zijn er goede en slechte kanten (bv. Anderson 2018).

- In de omgeving geplaatste sensoren helpen ons de luchtkwaliteit te monitoren, de verkeersstroom te bewaken, energie efficiënter te gebruiken, ons huis te beschermen of een inbreker op te sporen. Dezelfde technologie kan worden gebruikt door overheden om afwijkende meningen te onderdrukken of door bedrijven om ons gedrag in kaart te brengen en te manipuleren, en onze privacy te schenden.
- Draagbare apparaten kunnen onze gezondheid volgen en mogelijke problemen opsporen of aanpakken, of ons helpen om onze levensstijl te verbeteren. Dezelfde informatie kan, wanneer ze op grote schaal wordt verspreid, worden benut om de kosten van onze ziektekostenverzekering te verhogen (of zelfs om ons de kostendekking te ontfemen), of kan de kansen op een baan beïnvloeden.
- Autonome of semiautonome systemen, zoals zelfrijdende auto's, kunnen het aantal verkeersdoden verminderen en hersenmachine-interfaces en exoskeletten kunnen de mobiliteit van verlamde patiënten herstellen. Maar zelfrijdende vrachtwagens kunnen voor honderdduizenden vrachtwagenchauffeurs tot banenverlies leiden en autonome wapens kunnen een grote populatie bedreigen.
- Enz. enz.

Samenvattend: deze technologieën kunnen enorm veel goed doen en de mensheid helpen om een betere wereld te creëren, maar de mogelijk negatieve effecten zijn ernstig en kunnen de goede overtreffen. Het is ook terecht te stellen dat veel van deze innovaties slechts een fractie van de bevolking bedienen, wat leidt tot toenemende ongelijkheid en een steeds groter wordende digitale kloof.

In het licht van dit alles zou een plausibel antwoord kunnen zijn dat we als samenleving al onze inspanningen erop moeten richten om het goede te maximaliseren en de kansen voor het slechte te minimaliseren, met behoud van de doelstellingen die centraal stonden in het oorspronkelijke internet: open, onbevooroordeeld en schaalbaar. De geschiedenis heeft ons geleerd dat dit alleen kan worden gedaan door een veelzijdige aanpak, waarbij alle belanghebbenden worden betrokken: de vernieuwers / makers / technologen; de bedrijven; de regeringen; de burgers / gebruikers. Te vaak wordt technologie de wereld ingegooid met een ongebreideld optimisme, zonder een diepgaande reflectie en een gesprek over mogelijke negatieve of neveneffecten. Tegelijk helpt technologiebashing alleen ook niet: meestal is het een combinatie van enkele of alle belanghebbenden die ervoor zorgen dat het misgaat. Toen Facebook werd gelanceerd, had niemand het vermoeden dat sociale netwerken – 'een verzamelpaats voor (oude) vrienden' – konden of zouden worden gebruikt om verkiezingen te ondermijnen. Veel hiervan kwam later tot stand door de introductie van twijfelachtige bedrijfsmodellen en de (on)vindingrijkheid van gebruikers en adoptanten. De overheid kwam pas tussenbeide toen het al veel te laat was.

Daarom is een initiatief als het Denkersprogramma zo belangrijk. Het helpt als alle partijen nadenken over hoe ze kunnen bijdragen aan een betere en eerlijkere digitale wereld. Een uitstekend voorbeeld hiervan werd gegeven in de loop van dit onderzoek met de bespreking van het *Antwerpen - Stad der Dingen*-initiatief, waar veel van de ideeën voortkwamen uit kleine stakeholdersbijeenkomsten die vervolgens over een bepaalde tijd werden doorgelicht (imec 2016).

In het vervolg van dit essay zal ik een technologisch perspectief presenteren en de volgende vraag beantwoorden: wat kunnen we vanuit het perspectief van een maker/innovator doen om op zijn minst enkele zorgen over de opkomende digitale wereld van tevoren aan te pakken en de waarschijnlijkheid te maximaliseren dat technologie ten goede wordt gebruikt? Het antwoord is vrij eenvoudig: we moeten de basis herbekijken.

4.2. Ontwerpprincipes voor een betere, veiligere en eerlijkere digitale wereld

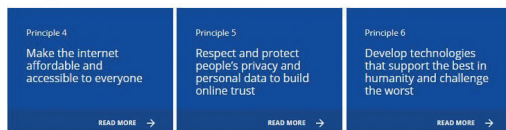
In het afgelopen decennium hebben heel wat denkers en visionairen het gehad over hoe de digitale technologie ingrijpt en evolueert in de samenleving en wat er kan worden gedaan om bepaalde fenomenen weer onder controle te krijgen. Toen J.C.R. Licklider van MIT in de vroege jaren 1960 zijn concept van *Galactic or Global Network* zag, merkte hij op: 'Over een paar jaar zullen mensen effectiever kunnen communiceren via een machine dan face to face.' (Leiner 1997) Hij en zijn collega's, die aan de oorsprong stonden van wat nu 'het internet' wordt genoemd, hadden een lappendeken van gedecentraliseerde netwerken voor ogen, zonder enig idee hoe het zou werken als een handvol bedrijven de meeste software zouden schrijven en het meeste verkeer zouden beheren (Brandon 2017). Nu is die visie uiteen aan het rafelen: 'Het sociale basispact van internet heeft zijn limiet bereikt – en begon te breken.' Het hoeft dan ook niet te verbazen dat sommige ontwerpers van het internet en het wereldwijde web aan de alarmbel trekken en onderzoeken hoe een herzien internet zijn oorspronkelijke doelstellingen – open en gratis – kan herstellen.

In november 2019 onthulde Tim Berners-Lee, de uitvinder van het web, een wereldwijd plan om het web te redden. Dat wordt ondersteund door meer dan 150 organisaties: van Microsoft, Twitter, Google en Facebook tot de digitale rechtengroep The Electronic Frontier. Het voor-gestelde *Contract voor het web* vermeldt de volgende missie: 'Het web is ontworpen om mensen samen te brengen en kennis vrij beschikbaar te maken. Het heeft de wereld voorgoed veranderd en het leven van miljarden mensen verbeterd. Toch hebben veel mensen nog steeds geen toegang tot de voordelen en voor anderen brengt het web te veel onaanvaardbare kosten met zich mee. Iedereen heeft een rol te spelen bij het veiligstellen van de toekomst van het web.'

Governments



Companies



Citizens



Fig. 5: Principles van de belanghebbenden voor het Contract for the Web [<https://contractfortheweb.org>]

Om de doelstellingen van het contract te bereiken moeten overheden, bedrijven, het maatschappelijk middenveld en individuen zich inzetten voor duurzame beleidsontwikkeling, belangen behartiging en de implementatie van decontract-tekst. (Contract2019). Vervolgens identificeert het contract een reeks principes, met de basisregels voor elk van de belanghebbenden (overheden, bedrijven en burgers) (figuur 2). De gestelde doelen zijn zeker zeer fundamenteel en nobel, en ze stellen duidelijke verwachtingen. Toch zijn er geen echte richtlijnen over hoe ze kunnen worden gerealiseerd en, meer nog, hoe ze moeten worden afgedwongen. De enige 'stok' in het contract luidt als volgt: 'Degegenen die het contract steunen, moeten aantonen dat ze de principes implementeren en werken aan oplossingen voor de moeilijkere problemen, anders kunnen ze worden verwijderd van de lijst met goedkeurders.'

Daarom loont het de moeite om nog een stap verder te gaan en enkele ontwerp-principes uit te werken voor een betere, veiligere en eerlijkere digitale wereld. Hoewel wat volgt zeker niet compleet is (en waarschijnlijk niet zonder controversie), helpt het om te bepalen hoe de verschillende belanghebbenden zouden kunnen bijdragen.

4.2.1. Authenticiteit en vertrouwen creëren

Het uitbouwen van een eerlijk en rechtvaardig digitaal ecosysteem is buitengewoon moeilijk als er weinig of geen vertrouwen is tussen de partijen, of het nu mensen, machines of een combinatie van beide betreft. Hoe kan men op informatie vertrouwen als men de bron of herkomst niet kent? En hoe vertrouw je sensorgegevens als de bron een Trojaans Paard is? Deze vragen leggen enkele van de meest uitdagende en ingewikkelde problemen bloot waarmee de digitale wereld wordt geconfronteerd. Een antwoord kan bestaan in een stevige aanpak van de authenticatie en het vestigen van een eenduidige identiteit op het netwerk. Dat zou voorkomen dat malafide instanties verkeerde informatie verspreiden en dat nepsensoren het energiedistributiesysteem verstoren. Dit botst echter met het fundamentele internetprincipe van vrijheid. De vrijheid van meningsuiting zou in gevaar kunnen worden gebracht als autoritaire regimes informatie kunnen opsporen en wreken, bij bronnen die anoniem zouden moeten zijn. Hier kunnen

innovatieve benaderingen geconcipieerd worden. Technisch gezien is het mogelijk om een enkelvoudige en echte authenticatiebenadering te bieden die:

- multi-variaat is (wat wil je controleren?) en veilig;
- meerdere profielen ondersteunt, afhankelijk van de geïnteresseerde partijen (sociaal, financieel, gezondheid, overheid...) die sterk en juridisch orthogonaal zijn. Dit bouwt voort op het concept van 'bevoegdheidsverdeling' dat de afgelopen eeuwen als leidraad heeft gediend voor democratische regimes;
- kan worden afgeschermd voor anonimiteitsdoeleinden, maar toch als betrouwbaar kan worden verklaard.

Verskillende benaderingen voor het opbouwen van vertrouwen (vertegenwoordigd door agenten) die aan de bovenstaande vereisten voldoen, worden momenteel overwogen. Dat vereist wel dat de belanghebbenden tot overeenstemming komen ondanks hun concurrerende doelen. Hoewel het redelijk is dat een regering mogelijke veiligheidsdreigingen bekampt, is het onaanvaardbaar dat zij dezelfde macht gebruikt om haar tegenstanders te onderdrukken. Hetzelfde geldt voor burgers en bedrijven. Dit is duidelijk een plaats waar derden en onafhankelijke organisaties, bestuurd en goedgekeurd door alle democratische instellingen, een rol kunnen spelen. Bedenk hoe de naamgevingsruimte voor het internet lange tijd door een onafhankelijke entiteit werd beheerd.

4.2.2. Eigenaarschap versterken

Data zijn de jongste tijd een kostbaar goed geworden en worden dat elke dag meer. De grote internetbedrijven, financiële systemen, verzekeringsmaatschappijen en ook overheden rekenen op de informatie die gratis ter beschikking wordt gesteld en vervolgens wordt geëxploiteerd in ruil voor gratis of zelfs betaalde diensten. Helaas is dit model volledig eenzijdig en verkeerd.

Dit moet de stelregel zijn: 'Ik ben de producent van mijn data. Dus moeten die mijn bezit zijn en blijven.' Te lang hebben we onze gegevens gratis weggegeven. Meestal betalen er we zelfs voor: denk aan de jaarlijkse bloedonderzoeken als onderdeel van een gezondheidsonderzoek. Vaak wordt ons verhinderd dat we de gegevens zelfs maar zien of openen. Pas onlangs is een deel van ons kredietrapport – dat bepaalt of we in aanmerking komen voor leningen – of zijn onze medische rapporten beschikbaar gekomen voor inzage. Toch hebben we geen controle over die gegevens en hoe ze worden verspreid, uitgewisseld en verhandeld voor commerciële of overheidsdoeleinden. Het operationeel basismodel is dat, zodra we onze gegevens hebben 'gegeven', de houder vrij is/was om ze naar eigen inzicht te gebruiken en verspreiden.

Een voor de hand liggende manier om dit aan te pakken is door middel van regelgeving en gedragsregels - zoals blijkt uit de vele wetten inzake gegevensbescherming die door regeringen over de hele wereld zijn of worden ingevoerd. Dit model heeft

echter veel valkuilen en weinig ingebouwde bescherming tegen niet-naleving en misbruik van mazen in de wet. In feite waren regeringen zélf vaak de eersten die hun eigen regels omzeilden.

Er is dus behoefte aan een heel ander operationeel model, waarin mijn gegevens exclusief eigendom zijn van mij en nooit worden prijsgegeven. Om bij mijn gegevens te komen moeten bedrijven en overheden naar mij komen. Ze krijgen een tijdelijke en beperkte toegang, vaak in ruil voor een dienst of betaling. De 'gerechtigde' krijgt de gegevens nooit zelf in handen (tenzij de eigenaar ze expliciet vrijgeeft) en heeft dus geen middelen om ze door te geven aan derden. Met andere woorden: de eigenaar van de gegevens behoudt de controle. Dit model is gebaseerd op sterke encryptietechnieken, authenticatie en microtransacties. Het idee is de afgelopen jaren in een aantal voorstellen uitgewerkt en krijgt nu serieuze aandacht. Een artikel in de NY Times van 19 november 2019 beschrijft een dergelijke aanpak, die werd ontwikkeld door de startup Oasis Labs in San Francisco, opgericht door professor Dawn Song van de EECS-afdeling aan UC Berkeley (Smith 2019): daar is zij 'een platform aan het bouwen dat mensen de mogelijkheid kan geven om hun gegevens te beheren en te controleren hoe deze worden gebruikt. Ze is van mening dat, zodra gegevens als eigendom worden beschouwd, ze de wereldeconomie op ongeziene manieren kunnen voortstuwten. Hierop kunnen nieuwe bedrijfsmodellen worden gebouwd.'

Om zo'n model mogelijk te maken zijn tijd, doorzettingsvermogen en een mentaliteitsverandering nodig. De toenemende aandacht van het publiek voor onderwerpen als privacy kan een doorbraak forceren.

4.2.3. Openheid en transparantie

Zwarte-doosystemen, een naam die voor veel van de momenteel gebruikte digitale systemen geldt, zijn standaard onverklaarbaar, ondoorzichtig en kwetsbaar. Fouten die worden veroorzaakt door onvolledige of ongedefinieerde interfaces kunnen tot uitbuiting leiden. En de resulterende acties zijn moeilijk uit te leggen of te motiveren.

Dit staat in schril contrast met opensourceprojecten, -producten of -initiatieven. Opensource.com zegt het zo: '(Zij) omarmen en viere principes van open uitwisseling, gezamenlijke deelname, snelle prototyping, transparantie, meritocratie en gemeenschapsgerichte ontwikkeling.' Dat leidt doorgaans tot robuustere en schaalbare systemen, in de loop van de tijd. Bart Preneel, professor aan de KU Leuven, stelt: 'Een belangrijk onderdeel van de oplossing is een radicale keuze voor open oplossingen: open hardware en open software zijn de enige manier om onafhankelijkheid te creëren en achterdeuren te vermijden.' (Preneel 2019). Opensourcesoftware wordt allang geaccepteerd en vormt de hoeksteen van veel besturingssystemen (UNIX, LINUX) en meer recent ook

voor netwerksoftware (Open Network Foundation). Hardwaremodules (vaak 'IP' genoemd voor intellectuele eigendom) zoals processors zijn bijna altijd als black box gekomen. Dit model staat zwaar onder druk door de introductie van de publiek domein RISC-V instructieset (en de ondersteunende RISC-V Foundation).

Een vergelijkbare trend doet zich voor in de gegevenswereld, waar veel overheidsinstanties en uitgevers eisen dat de datasets die aan een studie, onderzoeksproject of publicatie ten grondslag liggen, openbaar worden gemaakt: 'Open data is het idee dat sommige data voor iedereen vrij beschikbaar moeten zijn om te gebruiken en opnieuw te publiceren, zonder beperkingen van copyright, patenten, abonnementen of andere controlemechanismen.' (Wikipedia). Hoewel dit inderdaad de creatie van enorme datasets mogelijk heeft gemaakt die ontwikkelingen in de gezondheidszorg, wellness, rampenbeheersing enz. hebben gestimuleerd, hebben ze ook onverwachte neveneffecten en zorgen gecreëerd inzake privacy en betrouwbaarheid.

4.2.4. *Traceerbaarheid en responsabilisering*

Deze eigenschappen hangen nauw samen met de eerder uitgewerkte principes van authenticatie, vertrouwen en openheid. De vele lagen van obfuscatie in de huidige complexe digitale systemen maken het vaak buitengewoon moeilijk om de eenvoudige maar belangrijke vragen te beantwoorden die in dit Standpunt worden gesteld: Wat? Hoe? Wie? Het lijkt logisch dat elk verantwoordelijk digitaal systeem doorzoekbaar moet zijn en in staat moet zijn inzicht te geven in ten minste enkele aspecten van deze vragen. Twee eenvoudige voorbeelden illustreren de bijbehorende uitdagingen:

- AI-systemen moeten kunnen uitleggen waarom beslissingen zijn genomen en wat de redenering erachter is ('uitlegbaarheid'). Dat is zeker niet het geval met de huidige *deep-learning* systemen. Het belang van deze vragen neemt echter dramatisch toe wanneer AI wordt gebruikt in autonome systemen en impliciet verantwoordelijk is voor de resultaten van acties. Hetzelfde geldt voor elke situatie waarin aansprakelijkheid een rol speelt.
- Als een afbeelding of video wordt gewijzigd en online wordt geplaatst, moet de informatie over wanneer, waar en door wie expliciet worden vermeld, op een manier die niet kan worden verwijderd of gewijzigd. Dit zou bijvoorbeeld de verspreiding van *deep-fake* video's kunnen voorkomen.

Deze onderwerpen zijn brandend actueel in de wereld van artificiële intelligentie. Er wordt vooruitgang geboekt, maar verdere innovatie is zeker nodig. Wat ontbreekt, zijn een mondiaal beleid en het kader om de invoering ervan af te dwingen.

4.2.5. *Ingebouwd beleid, regelgeving en ethiek*

De handhaving van beleid en regels vereist dat ze eerst worden geformuleerd of zelfs maar bestaan. Vaak zijn ze impliciet en zijn ze ingebakken in de realisatie van het systeem, waardoor ze moeilijk te veranderen of aan te passen zijn. Zo zijn

de operationele regels en beslissingspunten van een autonoom systeem meestal vast en hard gecodeerd, waardoor ze moeilijk aan te passen zijn bij veranderende omstandigheden. Hetzelfde geldt voor de werking van draadloze netwerken en verkeersmanagementsystemen. Gezien de uiteenlopende omstandigheden en toestanden waarin een systeem kan werken, moeten de operationele regels aanpasbaar zijn. Denk aan een natuurramp: het zou logisch zijn in zo'n omstandigheid alleen voorrang te verlenen aan het draadloze netwerk voor EHBO'ers en de normale inzetregels te veranderen. Hetzelfde geldt voor het verkeersmanagementsysteem onder deze voorwaarden.



Fig. 6: Beleid en regelgeving.

Het expliciet maken van het beleid, als een reeks regels die worden beheerd door een 'beleidsbeheerder', helpt om deze bezorgdheid weg te nemen. Dit maakt het mogelijk dat veranderingen in beleid, regels van opdrachten en regelgeving dynamisch en adaptief plaatsvinden, op basis van de evaluatie van operationele omstandigheden en/of door menselijk ingrijpen. Nieuwere generaties van enkele van de bovengenoemde gedistribueerde systemen omarmen dit idee al.

Ethiek, over het algemeen gedefinieerd als de tak van de filosofie die betrekking heeft op het systematiseren, verdedigen en aanbevelen van concepten van goed of fout gedrag (Wikipedia), vormt een consistent geheel van dergelijke regels en kan daarom op een vergelijkbare manier worden uitgebouwd. De uitdaging is om te definiëren, evalueren en meten wat goed of fout is. De realiteit is vaak ingewikkeld en dus niet zwartwit. Dit is een domein dat interactie vereist tussen alle belanghebbenden en waarin filosofische, morele, maatschappelijke en juridische argumenten moeten worden gecombineerd.

Een klassiek voorbeeld zijn de drie wetten van robots, opgesteld door Isaac Asimov in 1942, toen praktische robots nog een verre toekomst waren (Asimov 1942-1950):

- Een robot mag een mens niet verwonden of, door niets te doen een mens schade toebrengen.
- Een robot moet de bevelen opvolgen die de mens hem heeft gegeven, behalve wanneer dergelijke bevelen in strijd zijn met de eerste wet.
- Een robot moet zijn eigen bestaan beschermen, zolang deze bescherming niet in strijd is met de eerste of tweede wet.

Op het eerste gezicht lijken deze wetten volkomen logisch en geven ze een gevoel van geordende beslissingsprincipes. Maar Asimov zelf publiceerde kort daarna een aantal verhalen waarin hij mazen in de wet uitbuitte of situaties verkende die inherente tegenstellingen veroorzaakten.

Een ander klassiek voorbeeld wordt vaak aangehaald om aan te tonen hoe moeilijk het is onfeilbare ethische regels te definiëren: 'Als een zelfrijdende auto

geconfronteerd wordt met een catastrofale botsing die de passagier zou doden, moet hij een van de volgende acties uitvoeren: (a) uitwijken naar links en een spelend kind doden; (b) rechts afslaan en een oude dame doden; (c) de aanrijding gewoon laten gebeuren.' Ik laat het aan de lezer over om hierover te reflecteren.

Deze reeks vragen heeft de aandacht getrokken van een aantal onderzoekers op alle betrokken gebieden. Het is een zeer rijk veld geworden met veel aspecten, zoals vooroordelen of bias in AI-systemen, door AI gecontroleerde autonome wapens, onbedoelde gevolgen enz. Een goed compendium vind je in White (2019). Deze discussies kunnen inzichtelijke aanbevelingen opleveren. De uitdaging blijft echter om een digitaal systeem te realiseren dat zich aan deze regels houdt en dat niet kan worden ondermijnd.

4.2.6. Opvoeding en opportuniteiten voor allen

Het lijkt geen twijfel dat alle genoemde problemen complex zijn en voor de algemene bevolking overweldigend kunnen lijken. De enige manier om dit aan te pakken is te beginnen aan een missie om voor iedereen geletterdheid in digitale innovatie binnen bereik te brengen. Gelukkig neemt dit langzaam toe op universitair niveau, waar veel instellingen zijn begonnen met het aanbieden van cursussen en programma's in *data science*. Deze programma's, die bedoeld zijn voor de hele campusbevolking, bieden de studenten niet alleen basisinzichten in datastatistiek en kunstmatige intelligentie (en de ondersteunende werktuigen), maar passen ze ook toe op voorbeelden uit specifieke gebieden (gezondheid, kunst, recht, enz.) en plaatsen die in een menselijke context en ethiek. Een goed voorbeeld is de Berkeley Data Science Major, een Bachelor of Arts degree (Berkeley 2019).

Dit is nog maar een begin. Digitale innovatie moet een onderdeel worden van het curriculum in de onderwijsketen, te beginnen bij het secundair onderwijs. Gezien de snelle evolutie en het potentieel voor dramatische veranderingen in de manier waarop we werken, sociaal leven en spelen, mag onderwijs nooit stoppen, moet het levenslang zijn en moet het ernaar streven alle segmenten van de samenleving te bereiken.

4.3. De mensheid op een kruispunt

Samenvattend: het is duidelijk dat de mensheid opnieuw op een kruispunt staat. Tijdens haar evolutie zijn dergelijke kruispunten al een aantal keer voorgekomen, zoals we in de inleiding aangaven. Deze keer kan echter 'anders' zijn, aangezien het voor het eerst is dat de mensheid zelf bepaalt waar ze heen gaat en dat de manier waarop we onszelf en onze lotgenoten waarnemen, kan veranderen. Daarom kunnen we een grote invloed uitoefenen op hoe onze toekomst eruit kan zien.

De digitalisering van de samenleving, die een aantal decennia geleden is begonnen, heeft het inherente potentieel al bewezen om veel goeds te kunnen doen en zal dat ongetwijfeld blijven doen. We leven langer en gezonder, en het percentage van de wereldbevolking dat uit de spiraal van armoede komt, neemt toe, ondanks de voortdurende groei van die bevolking. Tegelijkertijd is ook het potentieel voor grote schade en voor onrust zeker aanwezig. Niets doen is geen optie, evenmin als de natuurlijke menselijke reflex om de tijd terug te draaien en zo de vaak verontrustende bijwerkingen van veranderingen te vermijden. Zoals we in dit essay hebben uitgewerkt, kunnen oplossingen, als ze eenmaal zijn geïdentificeerd en begrepen, worden overwogen – zolang ze de steun van alle betrokken partijen krijgen. Ze zullen zeker fundamentele veranderingen vereisen in de manier waarop individuen, de samenleving, de overheid en bedrijven met elkaar omgaan. Een ander element is ook duidelijk: wat we ook doen, we zullen allemaal nodig zijn en het zal een aantal gedurfde stappen vergen. In een aantal opzichten zijn we al te laat... Laten we daarom niet nog meer tijd verliezen.

4.4. Referenties

J. Anderson and L. Rainie (2018), *Fifty-fifty anecdotes: How digital life has been both positive and negative*. Pew Research Center, <https://www.pewresearch.org/internet/2018/07/03/fifty-fifty-anecdotes-how-digital-life-has-been-both-positive-and-negative/>

I. Asimov (1942 en 1950), "Runaround". I, *Robot* (The Isaac Asimov Collection ed.). New York City: Doubleday, p. 40.

Berkeley (2019), Berkeley Data Science Major, University of California at Berkeley, <http://guide.berkeley.edu/undergraduate/degree-programs/data-science/>

R. Brandon (2017), *The internet we left behind*. The Verge, <https://www.theverge.com/2017/12/19/16792306/fcc-net-neutrality-open-internet-history-free-speech-anonymity>.

Contract (2019), *Contract for the Web, A global plan of action to make our online world safe and empowering for everyone*, <https://contractfortheweb.org>.

Imec (2019), *City of Things*, <https://www.imec-int.com/en/cityofthings>.

Institute of the future (2016), *Towards an Internet of Actions*, <http://www.iftf.org/future-now/article-detail/searchable-matter-from-an-internet-of-information-to-an-internet-of-action/>

B. Preneel (2019), *The future of security and privacy*. KVAB, https://www.kvab.be/sites/default/rest/blobs/2442/preneel_santav2_preprint.pdf.

C. Smith (Nov. 2019), *Building a World Where Data Privacy Exists Online*, NY Times, <https://www.nytimes.com/2019/11/19/technology/artificial-intelligence-dawn-song.html>

J. White (2019), *Ethics of Artificial Intelligence*. Philpapers, <https://philpapers.org/browse/ethics-of-artificial-intelligence/>

5. Besluiten en aanbevelingen van het Denkersprogramma

JAN RABAËY, RINIE VAN EST, PETER-PAUL VERBEEK EN JOOS VANDEWALLE

5.1. Wat?

Het thema van dit Denkersprogramma – de ingrijpende veranderingen in de samenleving door de digitalisatie – werd eind 2018 gekozen door de reflectiegroep Responsible Research & Innovation en Wetenschapsethiek van de KVAB. Sindsdien is dit thema in de media, de politiek, het bedrijfsleven en in het algemeen in de maatschappelijke discussie in Vlaanderen, de EU en de wereld alleen maar prominenter geworden. Een bewijs daarvan is onder andere de enorme toeloop eind januari 2020 in Antwerpen en Kortrijk voor de lezingen van Yuval Harari over de AI-revolutie.²⁰ De Denkers hebben dit ook uit de eerste hand kunnen vaststellen in de diverse gesprekken die ze gehad hebben in de loop van 2019.²¹ De omvang van de transformatie en de impact ervan op de samenleving zijn niet te onderschatten en de volle ontplooiing en de impact zijn nog niet in zicht. Vanuit hun ervaring en inzichten hebben de Denkers geen doemscenario's of schrikbeelden geprojecteerd. Er is bij hen en evenzeer bij Harari (*Lesson 18: the future is not what you see in the movies*) een distantiëring van de angstbeelden in sciencefictionachtige scenario's. Anderzijds pleiten de Denkers voor een nieuwe relatie tussen mens en technologie, met onder andere een zich verdiepende bewustwording, een zorgzame houding en verantwoordelijk handelen. De Covid-19-uitdaging confronteert ons ook scherp met deze relatie, in een context van thuiswerken, afstandsonderwijs, online-shopping en social distancing. Meer dan ooit vereist een dergelijke ongeziene problematiek een coöperatieve houding, interdisciplinaire samenwerking van diverse actoren en een synergetische aanpak. In deze context hebben de drie Denkers in tempore non suspecto en vanuit hun internationale en disciplinaire ervaring vele verdiepende inzichten, inspirerende praktische voorbeelden en waardevolle aanbevelingen aangebracht. Dat kan ons brengen tot een algemeen kader voor het ethisch begeleiden van de toepassingen van digitale technologie als basis voor technologisch burgerschap. De recente opstart van het Vlaams Kenniscentrum voor Data en Maatschappij is in deze ontwikkeling een zeer waardevol element.

5.2. Wie?

Heel de samenleving – jongeren en ouderen, de overheden van steden, regio's, landen, de EU en wereldwijd, het onderwijs en onderzoek, de bedrijfswereld,

²⁰ Yuval Harari (2018), *21 Lessons for the 21st century*, Pinguin Random House.

²¹ Academieberichten KVAB, 15 oktober 2019: interview met Peter-Paul Verbeek, Jan Rabaey en Rinie van Est.

sociale organisaties – moet zich aangesproken voelen om de transitie te proberen te begrijpen en haar mee in goede banen te leiden, elk vanuit zijn/haar rol in de samenleving. Er wordt duidelijk gesteld dat digitale vaardigheden relevant zijn voor álle burgers, beroepssectoren, organisaties en bedrijven. Het gaat daarbij zeker niet alleen om digitale beroepsvaardigheden. Burgers dienen inzicht te hebben in de mogelijkheden en de risico's van technologie om zo hun eigen leven vorm te kunnen geven, weerbaar te zijn tegen risico's en mee te doen aan het publieke debat en de democratische besluitvorming over de toekomst van de digitale samenleving. Digitale kloven – tussen leerlingen, leraren, burgers, *digital natives* en digitale immigranten, tussen bedrijven onderling en tussen bedrijven en consumenten enz. – dienen voorkomen en gedicht te worden.

5.3. Hoe?

De drie Denkers zijn het erover eens dat er een enorm potentieel is voor digitalisering en AI. Maar net zo goed kan dit ook leiden tot veel schade en onrust in de maatschappij, als deze technieken ongebreideld worden toegepast. De samenleving staat dan ook op een kruispunt. Niets doen of de tijd terugdraaien is nog nooit een zinvolle of werkbare optie geweest. Er moet een traject gevolgd worden, waarin de problemen bestudeerd worden en mogelijke scenario's voor oplossingen geïdentificeerd, geschetst en begrepen worden, met inbreng van alle betrokken partijen. Daarbij moet men openstaan voor fundamenteel andere, creatieve manieren waarop individuen, de maatschappij, de overheid en bedrijven met elkaar omgaan en communiceren. We hebben het allemaal nodig en het zal een aantal moedige stappen vergen. Aangezien het al laat is, mogen we niet nog meer tijd verliezen.

De overheid stimuleert techno-economische innovatie. De drie Denkers brengen elk vanuit hun ervaring de boodschap dat het bij waardevol innoveren gaat om een constructieve interactie tussen technische en sociale innovatie. Dat betekent dat zowel techno-economische als sociale innovatie gestimuleerd moet worden, hand in hand. Bij waardevolle innovatie spelen vijf processen een rol: waarderen, experimenteren, kansen grijpen, risico's verzachten en samen werken en leren (cfr. Rathenau). 'Waarderen' gaat om het verhelderden van de publieke waarden en doelen die men wil nastreven. 'Experimenteren' betekent ruimte maken om nieuwe zaken uit te proberen. 'Kansen grijpen' betekent een open oog houden voor de mogelijkheden die digitale technologieën bieden. De overheid behoort ernaar te streven dat onze samenleving zo veel mogelijk kan profiteren van de zegeningen van digitalisering. 'Risico's verzachten' houdt in dat burgers zo veel mogelijk worden beschermd tegen de risico's van digitalisering. Het in goede banen leiden van digitalisering is een gezamenlijke uitdaging. Het vraagt om afstemming en samenwerking tussen bestuurslagen, overheden en kennisinstellingen, bedrijven en, niet te vergeten, burgers. 'Samen werken en samen leren' is dan ook het vijfde devies. Zo wordt er gepleit voor 'cocreatie': in een samenwerking tussen

burgers en de overheid of bedrijven worden samen experimenten met nieuwe diensten of systemen opgezet. Men voorkomt zo dat de overheid een te grote macht ontwikkelt en dat er te veel ruimte is voor surveillance. Burgers worden dan actief betrokken bij de keuzes die op politiek niveau gemaakt worden, zodat de inrichting van bijvoorbeeld de *smart city* een gezamenlijke verantwoordelijkheid wordt.

De drie Denkers benadrukken ook de belangrijke rol van het onderwijs, vanaf het lager tot en met het hoger onderwijs, en ook van bijscholing en herscholing tijdens de professionele levensfase. Een gedragen en gedeelde open visie op de toekomst kan jongeren in hun leerproces stimuleren om hun maatschappelijke verantwoordelijkheden op te nemen, en andere mensen om de nodige flexibiliteit in de uitbouw en evolutie van hun carrière in te bouwen. Bovendien zorgen digitalisering en AI voor heel wat ingrijpende veranderingen in het onderwijsproces zelf. We vermelden hier de digitale leeromgevingen, die de interactie tussen leerlingen en docenten, en ook tussen leerlingen onder elkaar, ingrijpend veranderen. De onlineleerervaringen die tijdens de coronacrisis in het kleuter-, basis-, secundair en hoger onderwijs worden opgedaan, zullen op evidentie gebaseerd inzichten aanleveren in de waarde, de mogelijkheden en de beperkingen van digitale leeromgevingen. Ook tijdens leeractiviteiten kunnen heel wat data verzameld worden over het leergedrag van leerlingen en hun interactie met elkaar en hun docenten, en wordt het mogelijk om correlaties op te sporen die tot een verbetering van leerprestaties kunnen leiden (*learning analytics*). Hiervoor verwijzen we naar een recent Standpunt van de KVAB.²² Ook aan de inhoud van het onderwijs stelt de digitalisering van de samenleving nieuwe eisen, zoals kritisch denkvermogen en een transdisciplinaire houding. Een van de aandachtspunten is zeker ook de aanpak van de ongelijkheden in het onderwijs en in de samenleving in het algemeen.

Voor de ontwerpers van nieuwe diensten en producten is er nood aan ontwerp-principes die een meer rechtvaardige, transparante en veilige digitale wereld realiseren. De ethiek van digitalisering is niet alleen een kwestie van abstracte principes en grote verhalen, maar vooral ook van concrete analyses en actiepunten. Op het gebied van verantwoord innoveren en van de ethiek van technologie ligt de kracht van het Europese perspectief, als verbinder tussen West en Oost.²³ De invoering van corona-apps bij de exit uit de lockdown is een voorbeeld van de zorgvuldige toepassing van digitale technieken, waarin afwegingen worden gemaakt inzake precisie, veiligheid, respect voor de privacy, het gevaar voor de

²² T. De Laet e.a. (2018), "*Learning Analytics*" in het Vlaams hoger onderwijs, Standpunt KVAB.

²³ M. Ressler-Lavigne, 'Putting Ethics at the Heart of Innovation Stanford centers the humanistic and ethical considerations of new technology', *Stanford Magazine*, maart 2020. <https://stanfordmag.org/contents/putting-ethics-at-the-heart-of-innovation-marc-tessier-lavigne>

heropflakking van de epidemie en gedragenheid en vertrouwen bij het bredere publiek.²⁴

²⁴ Communication from the Commission, *Guidance on Apps supporting the fight against COVID 19 pandemic in relation to data protection*, Brussel, 16 april 2020.

Appendix 1 - CV van de Denkers

Jan Rabaey bekleedt de Donald O. Pederson Distinguished Professorship aan de University of California at Berkeley. Hij behaalde het diploma van burgerlijk ingenieur en doctorerde aan de KU Leuven. Hij verrichtte onderzoek aan imec, is de stichter van het Berkeley Wireless Research Center (BWRC) en de Berkeley Ubiquitous SwarmLab, en was tot tweemaal toe voorzitter van de afdeling Electrical Engineering aan Berkeley.

Professor Rabaey leverde belangrijke bijdragen in diverse domeinen, waaronder *advanced wireless systems, low power integrated circuits, sensor networks, and ubiquitous computing*. Zijn huidige onderzoek concentreert zich op de volgende generatie geïntegreerde draadloze systemen met haar brede gamma aan toepassingen en op onderzoek naar de interactie tussen de virtuele en de biologische wereld.

Hij ontving talrijke onderscheidingen, zoals de IEEE Mac Van Valkenburg Award, de European Design Automation Association (EDAA) Lifetime Achievement Award en de Semiconductor Industry Association (SIA) University Researcher Award. Hij is IEEE Fellow en buitenlands lid van de Koninklijke Vlaamse Academie van België, en ontving eredoctoraten van de universiteiten van Lund (Zweden), Antwerpen en Tampere (Finland). Hij is tevens betrokken bij een groot aantal start-ups.

Peter-Paul Verbeek is universiteitshoogleraar Filosofie van Mens en Techniek en wetenschappelijk codirecteur van het DesignLab van de Universiteit Twente. Hij is tevens honorary professor aan Aalborg University, Denemarken. Zijn onderzoek richt zich op de relaties tussen mens, technologie en samenleving, en op filosofische analyse, ethische reflectie, en praktijken van ontwerp en innovatie. In het Nederlands publiceerde hij onder andere *Op de Vleugels van Icarus* (2013), *De Grens van de Mens* (2011) en *De Daadkracht der Dingen* (2000); in het Engels onder andere *Moralizing Technology: Understanding and Designing the Morality of Things* (2011) en *What Things Do: Philosophical Reflections on Technology, Agency, and Design* (2005).

Professor Verbeek is voorzitter van de World Commission for the Ethics of Science and Technology (COMEST) van UNESCO, vice-voorzitter van de UNESCO Ad Hoc Expert Group for the Ethics of Artificial Intelligence en lid van de Programmaraad Maatschappelijk Verantwoord Innoveren van NWO, het bestuur van het Rathenau Instituut, de Raad voor Geesteswetenschappen (KNAW), de Commissie voor de Vrijheid van Wetenschapsbeoefening (KNAW) en de Raad van Toezicht van TNO.

Hij is ook lid van de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW), de Academie voor Technologie en Innovatie (ActI) en de Koninklijke Hollandse Maatschappij van Wetenschappen (KHMW). In 2012 ontving hij de Prof. Borghgraef prijs voor biomedische ethiek (KU Leuven) en in 2016 de World Technology Award in Ethics (World Technology Network). Meer informatie: www.ppverbeek.nl.

Rinie van Est is themacoördinator aan het Rathenau Instituut. Het Rathenau Instituut stimuleert in Nederland het publieke en politieke debat over maatschappelijke en ethische kwesties in verband met wetenschap en technologie. Van Est coördineert het thema Slimme Samenleving. Als natuurkundige en politicoloog is hij gespecialiseerd in de politiek van waardevolle innovatie. Hij is een wereldwijd erkend expert op het gebied van *technology assessment*, *governance* en publieksparticipatie. Bij het Rathenau Instituut houdt hij zich al meer dan twintig jaar bezig met ontwikkelingen op het gebied van de energietransitie en digitale transitie, en met name met de rol van opkomende technologieën daarin, zoals robotica en AI. Duurzaamheid van de natuurlijke en menselijke omgeving – respectievelijke ecologische en menselijke duurzaamheid – is een kernbegrip in zijn werk. Tevens werkt hij sinds 2000 een dag per week aan de Faculteit Industrial Engineering & Innovation Sciences van de Technische Universiteit Eindhoven, momenteel als hoogleraar Technology Assessment and Governance. Enkele publicaties waaraan hij heeft meegewerkt zijn: *Waardevol digitaliseren* (2018), *Human rights in the robot age* (2017), *Opwaarderen: Borgen van publieke waarden in de digitale samenleving* (2017), *Just ordinary robots: Automation from love to war* (2016), *De kracht van platformen* (2014) *Intieme technologie: De strijd om ons lichaam en gedrag* (2014).

Appendix 2 - Samenstelling van de stuurgroep

Luc Bonte (KTW)
Geert Bouckaert (KMW)
Hugo De Man (KTW)
Inez Dua (KVAB staf)
Freddy Dumortier (KVAB vast secretaris)
Vincent Ginis (Jonge Academie)
Johan Hanssens (EWI)
Liesbet Lagae (KTW)
Lode Lauwaert (KU Leuven)
Egbert Lox (KTW)
Jan Rabaey (KTW, UC Berkeley, denker)
Rinie van Est Rathenau, (TU Eindhoven, denker)
Ine Van Hoyweghen (Jonge Academie)
Marc Van Hulle (KAGB)
Joos Vandewalle (KTW, coordinator KVAB reflectiegroep RRI & onderzoeksethiek)
Peter-Paul Verbeek (U Twente, denker)
Pascal Verdonck (KTW)
Simon Verschaeren (EWI)
Paul Verstraeten (KTW)
Bieke Zaman (KU Leuven)

KTW = Klasse van de Technische Wetenschappen

KMW = Klasse van de Menswetenschappen

KAGB = Koninklijke Academie voor Geneeskunde van België

Appendix 3 - Relevante literatuur

Rathenau Instituut en Wetenschappelijke Raad voor het Regeringsbeleid Nederland WRR

Rathenau Instituut, *Overzicht van ethische codes en principes voor AI*, 19 maart 2019. <https://www.rathenau.nl/nl/digitale-samenleving/overzicht-van-ethische-codes-en-principes-voor-ai>

Rathenau Instituut, *Zo brengen we AI in de praktijk vanuit Europese waarden*, 19 maart 2018. <https://www.rathenau.nl/nl/digitale-samenleving/zo-brengen-we-ai-de-praktijk-vanuit-europese-waarden>

Rathenau Instituut, *Jaarverslag 2018*. https://www.rathenau.nl/sites/default/files/inline-files/Jaarverslag_Rathenau_2018.pdf

Rathenau Instituut, *Zo staat AI op de agenda van de EU*, maart 2019. <https://www.rathenau.nl/nl/digitale-samenleving/zo-staat-ai-op-de-agenda-van-de-eu>

Rathenau Instituut, *Naar een verantwoorde digitale samenleving Van kwesties naar acties*, 2018, Verslag conferenties van de SER en het Rathenau Instituut op 06-12-17 en 25-1-2018.

Rathenau Instituut, *Doelgericht digitaliseren Hoe Nederland werkt aan een digitale transitie waarin mensen en waarden centraal staan*, Linda Kool, Elma Dujso en Rinie van Est, 2018.

Rathenau Instituut, *Desinformatie in Nederland*, 13 februari 2019.

Scientific Council for Government Policy, WRR, *iGovernment*, Corien Prins, Dennis Broeders, Henk Griffioen, Anne-Greet Keizer & Esther Keymolen.

Ministerie van Economische Zaken en Klimaat, *Nederlandse digitaliseringsstrategie, Hier kan het. Hier gebeurt het*, 2018.pdf.

Rathenau Instituut, *Nieuwe regels voor kunstmatige intelligentie?*, 29 november 2018.

Corien Prins, Dennis Broeders, Henk Griffioen, Anne-Greet Keizer, Esher Keymolen, *iGovernment*, p.p., R086e-iGovernmentWRR-NL.pdf.

Rathenau Instituut, *Sensordata leefbaarheid veiligheid*, 2019.

Rathenau Instituut, *Zo brengen we AI in de praktijk vanuit Europese waarden*, maart 2020. <https://www.rathenau.nl/nl/digitale-samenleving/zo-brengen-we-ai-de-praktijk-vanuit-europese-waarden>

Academies internationaal

National Academies Press, *Data Matters: Ethics, Data, and International Research Collaboration in a Changing World*, Proceedings of a Workshop, 2018.

Österreichische Akademie der Wissenschaften, *Stellungnahme_Digitaler-Stillstand, die Verletzlichkeit der digital vernetzten Gesellschaft*, mei 2017.

José van Dijck (2018). *Vertrouwen in expertise in een digitale samenleving. Jaarrede President KNAW 2018*. Amsterdam, KNAW

Ethische codes/wetgeving

High-Level Expert Group on Artificial Intelligence, *Draft Ethics Guidelines for Trustworthy AI*, European Commission, Directorate-General for Communication, 18 December 2018. E-mail CNECT-HLG-AI@ec.europa.eu

Lambèr Royakkers, Jelte Timmer, Linda Kool, Rinie van Est, 'Societal and ethical issues of digitization' *Ethics and Information Technology*, 2018 pp. 127-142.

Vlaanderen

Benchmark studie Artificiële Intelligentie, 3 juli 2018, PWC.

Benchmark studie EWI, juli 2018. https://www.ewi-vlaanderen.be/sites/default/files/bestanden/departement_economie_wetenschap_en_innovatie_-_benchmark_studie_over_artificiele_intelligentie_-_webvriendelijk.pdf

Vlaamse Beleidsnota Artificiële Intelligentie, november 2018. VARIO_ADVIES_NR 5_VLAAMSE_BELEIDSAGENDA_AI_.pdf.

Derek Jan Fikkers, Elma Dujso, en Robert van der Vooren, *Flemish Open Science Board De governance structuur voor Open Data en Research Data Management in Vlaanderen*, Technopolis group. Januari 2019.

Voka Paper-Verkiezingsmemorandum_Vlaams_HR_p.p..pdf

RAPPORT Digimatuurtest, De digitale transformatie van HR & Well-being. www.attentia.be

Europa

Automating Society: Taking stock of automated decision-making in the EU, Bertelsmann Stiftung (België pp. 40-44). https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/001-148_AW_EU-ADMreport_2801_2.pdf

Robert Atkinson and Stephen Ezell, *Promoting European growth, productivity and competitiveness by taking advantage of the next digital technology wave*, ITIF Information Technology & Innovation Foundation, maart 2019. <http://www2.itif.org/2019-europe-digital-age-a4.pdf>

Coordinated Plan on Artificial Intelligence (EC Communication - 7 december 2018). <https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-795-F1-EN-MAIN-PART-1.PDF>

<https://www.kowi.de/de/Portaldata/2/Resources/fp/2018-COM-CP-Artificial-Intelligence-Annex.pdf>

Craglia M. (ed.), *Artificial Intelligence – A European Perspective*. JRC, 2018, 140pp. https://www.researchgate.net/profile/Stefano_Nativi/publication/329449889_Artificial_Intelligence_A_European_Perspective/links/5c091cf9299bf139c7432a66/Artificial-Intelligence-A-European-Perspective.pdf?origin=publication_detail

EU Science Hub, European Commission, *Digital Living*, 13 december 2018

The Age of Artificial Intelligence – Towards a European Strategy for Human-Centric Machines (EPSC, maart 2018).

Gabriel Rissola, Jens Sörvik, *Digital Innovation Hubs in Smart Specialisation Strategies, Early lessons from European regions*, JRC Technical Reports, European Commission, 2018.

Raad van de Europese Unie van het Comité van permanente vertegenwoordigers (1e deel) aan: Raad nr. vorig doc.: 5808/19 COMPET 75 MI 70 IND 22 DIGIT 15 JUSTCIV 26 RECH 56 EDUC 33 Betreft: Kunstmatige intelligentie b) Conclusies over het gecoördineerd plan inzake kunstmatige intelligentie, 11 februari 2019.

Innovation in Europe, A Deloitte survey on European companies and how digital technologies can strategically enhance innovation. Deloitte, 2019, 24 p.p.

Jess Whittlestone, Rune Nyruup, Anna Alexandrova, Kanta Dihal en Stephen Cave, *Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research*. Nuffield Foundation UK, 2019, 59 p.p.

C. Cath et al., *Artificial Intelligence and the 'Good Society': the US, EU, and UK approach* (2017). https://ec.europa.eu/epsc/sites/epsc/files/epsc_strategicnote_ai.pdf

Notes from the AI Frontier (McKinsey) – Modeling the Impact of AI on the World Economy (september 2018) <https://www.mckinsey.com/~media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Notes%20from%20the%20frontier%20Modeling%20the%20impact%20of%20AI%20on%20the%20world%20economy/MGI-Notes-from-the-AI-frontier-Modeling-the-impact-of-AI-on-the-world-economy-September-2018.ashx>

Tackling Europe's Gap in Digital and AI (McKinsey), februari 2019. <https://www.mckinsey.com/~media/mckinsey/featured%20insights/artificial%20intelligence/tackling%20europes%20gap%20in%20digital%20and%20ai/mgi-tackling-europes-gap-in-digital-and-ai-feb-2019-vf.ashx>

J. Scott Marcus, Georgios Petropoulos en Timothy Yeung, *Contribution to Growth: The European Digital Single Market Delivering economic benefits for citizens and businesses*, European Parliament, Directorate-General for Internal Policies PE 631.044 - januari 2019.

OECD Publishing, *Digital dividend: policies to harness the productivity potential of digital technologies*, OECD Policy paper, februari 2019, N°26.

OECD Publishing, *How's Life in the Digital Age? Opportunities and Risks of the Digital Transformation for People's Well-being*, 2019. DOI: <https://doi.org/10.1787/9789264311800-en>

OECD Publishing, *Going Digital: Shaping Policies, Improving Lives*, 2019. DOI: <https://doi.org/10.1787/9789264312012-en>

Ajuntament de Barcelona, *White Paper on the Future of Work and Jobs Reflections for a new local economic policy*, maart 2018, 228 p.p.

Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence*, 23 augustus 2017, Random House, 384 p.

UNESCO-conferentie over AI, 4 maart 2019.

RECENTE STANDPUNTEN (vanaf 2016)

41. Anne-Mie Van Kerckhoven, Francis Strauven – *Een bloementapijt voor Antwerpen*, KVAB/Klasse Kunsten, 2016.
42. Erik Mathijs, Willy Verstraete (e.a.), *Vlaanderen wijs met water: waterbeleid in transitie*, KVAB/Klasse Technische wetenschappen, 2016.
43. Erik Schokkaert – *De gezondheidszorg in evolutie: uitdagingen en keuzes*, KVAB/Klasse Menswetenschappen, 2016.
44. Ronnie Belmans, Pieter Vingerhoets, Ivo Van Vaerenbergh e.a. – *De eindgebruiker centraal in de energietransitie*, KVAB/Klasse Technische Wetenschappen, 2016.
45. Willem Elias, Tom De Mette – *Doctoraat in de kunsten*, KVAB/Klasse Kunsten, 2016.
46. Hendrik Van Brussel, Joris De Schutter e.a., *Naar een inclusieve robotsamenleving*, KVAB/Klasse Technische Wetenschappen, 2016.
47. Bart Verschaffel, Marc Ruyters e.a., *Elementen van een duurzaam kunstenbeleid*, KVAB/Klasse Kunsten, 2016.
48. Pascal Verdonck, Marc Van Hulle (e.a.) – *Datawetenschappen en gezondheidszorg*, KVAB/Klasse Technische wetenschappen, 2017.
49. Yolande Berbers, Mireille Hildebrandt, Joos Vandewalle (e.a.) – *Privacy in tijden van internet, sociale netwerken en big data*, KVAB/Klasse Technische wetenschappen, 2017.
50. Barbara Baert (e.a.), *Iconologie of 'La science sans nom'*, KVAB/Klasse Kunsten, 2017.
51. Tariq Modood, Frank Bovenkerk – *Multiculturalism. How can Society deal with it?* KVAB/Klasse Menswetenschappen, 2017.
52. Mark Eyskens – *Europa in de problemen*. KVAB/Klasse Menswetenschappen, 2017.
53. Luc Steels – *Artificiële intelligentie. Naar een vierde industriële revolutie?*. KVAB/Klasse Natuurwetenschappen, 2017.
54. Godelieve Gheysen, René Custers, Dominique Van Der Straeten, Dirk Inzé, *Ggo's anno 2018. Tijd voor een grondige herziening*. KVAB/Klasse Natuurwetenschappen, 2017.
55. Christoffel Waelkens (e.a.) – *Deelname van Vlaanderen aan grote internationale onderzoeksinfrastructuren: uitdagingen en aanbevelingen*, KVAB/Klasse Natuurwetenschappen, 2017.
55. Addendum. Jean-Pierre Henriët. – *Mijlpalen in internationale wetenschappelijke samenwerking*, KVAB/Klassen Natuurwetenschappen, 2017.
56. Piet Swerts, Piet Chielens, Lucien Posman – *A Symphony of Trees. Wereldcreatie naar aanleiding van de herdenking van de Derde Slag bij Ieper, 1917*, KVAB/Klasse Kunsten, 2017.
57. Willy Van Overschée e.a. – *De mobiliteit van morgen: zijn we klaar voor een paradigmawissel?*, KVAB/Klasse Technische Wetenschappen, 2018.
58. Tinne De Laet e.a. – *"Learning Analytics" in het Vlaams hoger onderwijs*, KVAB/Klasse Technische Wetenschappen, 2018.
59. Dirk Van Dyck, Elisabeth Monard, Sylvia Wenmackers e.a. – *Onderzoeker-gedreven wetenschap. Analyse van de situatie in Vlaanderen*, KVAB/Klasse Natuurwetenschappen, 2018.
60. Liliane Schoofs – *Doctoraathouders geven het Vlaanderen van morgen vorm*, KVAB/Klasse Natuurwetenschappen, 2018.
61. Luc Bonte, Aimé Heene, Paul Verstraeten e.a. – *Verantwoordelijk omgaan met digitalisering. Een oproep naar overheden en bedrijfsleven, waar ook de burger toe kan/moet bijdragen*, KVAB/Klasse Technische Wetenschappen, 2018.
62. Jaak Billiet, Michaël Opgenhaffen, Bart Pattyn, Peter Van Aelst – *De strijd om de waarheid. Over nepnieuws en desinformatie in de digitale mediawereld*, KVAB/Klasse Menswetenschappen, 2018.
63. Christoffels Waelkens. – *De Vlaamse Wetenschapsagenda en interdisciplinariteit. Leren leven met interdisciplinaire problemen en oplossingen*, KVAB/Klasse Natuurwetenschappen, 2020.
64. Patrick Onghena – *Repliceerbaarheid in de empirische menswetenschappen*, KVAB/Klasse Menswetenschappen, 2020.
65. Mark Eyskens – *Als een virus de mensheid gijzelt. Oorzaken en gevolgen van de Coronacrisis*, KVAB/Klasse Menswetenschappen, 2020.

De volledige lijst met standpunten en alle pdf's kunnen worden geraadpleegd op www.kvab.be/standpunten



Onze samenleving staat op een kruispunt: de digitale transformatie is volop aan de gang en de totale impact ervan is nog niet te overzien. In vele sectoren is de digitalisering een drijvende kracht van innovatie en verandering. Consumenten maken gebruik van talrijke digitale diensten en laten vele digitale sporen na. In de huidige coronacrisis ervaren we meer dan ooit de voordelen van de digitalisering. Tegelijk mogen we niet blind zijn voor de nadelen en gevaren. Een goed overwogen en breed gedragen aanpak, met respect voor onze maatschappelijke waarden, dringt zich op.

De thematiek heeft vele facetten, invalshoeken en expertises. Dit Standpunt brengt drie complementaire essays van internationale experts. Zij brengen een toekomstgerichte boodschap voor de samenleving en het beleid. Aan de hand van verdiepende inzichten, inspirerende praktijkvoorbeelden en concrete aanbevelingen krijgt Vlaanderen hier een stevig fundament aangereikt voor een ethische begeleiding van digitalisering en voor het verder vorm geven van wat burgerschap in een digitale samenleving betekent.

Dit Standpunt brengt geen doemscenario's. Het is een pleidooi voor een nieuwe relatie tussen mens en technologie.

De reeks Standpunten van de Academie is een bijdrage tot het wetenschappelijk onderbouwd debat over actuele maatschappelijke en artistieke thema's. De auteurs, leden en werkgroepen van de Academie schrijven in eigen naam, onafhankelijk en met volledige intellectuele vrijheid. De goedkeuring voor publicatie door een of meerdere Klassen van de Academie waarborgt de kwaliteit van de gepubliceerde studies.