

SOCIETAL VALUES IN DIGITAL INNOVATION: WHO, WHAT AND HOW?

KVAB thinkers' programme 2019

Jan Rabaey

Rinie van Est

Peter-Paul Verbeek

Joos Vandewalle



KVAB POSITION PAPERS

66 b

Royal Flemish Academy of Belgium
for Science and the Arts - 2020

SOCIETAL VALUES IN DIGITAL INNOVATION: WHO, WHAT AND HOW?

KVAB THINKERS' PROGRAMME 2019



KVAB Press

KVAB STANDPUNTEN

66 b

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

The drawing of the Palace of the Academies is a reproduction of the original perspective made by Charles Vander Straeten in 1823. The logo of the KVAB was designed in 1947 by Jozef Cantré.

The KVAB Position Papers (Standpunten) are published by the Royal Flemish Academy of Belgium for Science and the Arts, Hertogsstraat 1, 1000 Brussel.
Tel. 00 32 2 550 23 23 – info@kvab.be – www.kvab.be

SOCIETAL VALUES IN DIGITAL INNOVATION: WHO, WHAT AND HOW?

KVAB THINKERS' PROGRAMME 2019



**Jan Rabaey
Rinie van Est
Peter-Paul Verbeek
Joos Vandewalle**

Partial reproduction is permitted provided the source is mentioned.
Recommended citation: Jan Rabaey, Rinie van Est, Peter-Paul Verbeek,
Joos Vandewalle, *Societal values in digital innovation: who, what and how?*
- *KVAB Thinkers' Programme 2019*, KVAB Position paper 66 b, 2020. (Original
text: Dutch). Translation Dutch manuscript and revision English parts: Sandra
McElroy
Printing office Universa

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

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

TABLE OF CONTENTS

| | |
|---|----|
| Executive summary | 7 |
| Foreword | 9 |
| 1. Positioning, aims and approach of the Thinkers' Programme | 11 |
| 2. Essay of a Thinker: 'The search for a good digital life. Put people and values at the heart of digital innovation' | 17 |
| 2.1. Waking up from irresponsibility | 17 |
| 2.2. Blind spots in the Netherlands' governance landscape | 18 |
| 2.3. Thinking about digital innovation and acting in Belgium and Flanders | 22 |
| 2.3.1. National security, economic prosperity and social welfare ... | 24 |
| 2.3.2. National security: The return of the drones | 24 |
| 2.3.3. Economic prosperity: Belief in technology and society | 26 |
| 2.3.4. Social welfare: Value-driven digitalisation | 28 |
| 2.4. Final thoughts: The good digital life | 30 |
| 2.5. References | 32 |
| 3. Essay of a Thinker: 'Technological Citizenship and the Ethical Guidance of Digitalisation' | 35 |
| 3.1. Introduction | 35 |
| 3.2. The relationships between humans and digital technology | 36 |
| 3.2.1. The current ethical debate | 38 |
| 3.2.2. Citizenship in the digital society | 39 |
| 3.3. The digital public space | 40 |
| 3.3.1. The online public space | 40 |
| 3.3.2. Digital technology in the physical public space | 41 |
| 3.3.3. Digital literacy | 42 |
| 3.4. Education in a digital society | 42 |
| 3.4.1. Digitalisation and educational practice | 43 |
| 3.4.2. The role of education in a digital society | 44 |
| 3.4.3. Critical thinking and transdisciplinarity | 44 |
| 3.5. Politics and digitalisation | 45 |
| 3.6. Ethics as guidance for technology: Frankenstein's second chance | 46 |

| | |
|--|----|
| 3.7. Conclusion..... | 48 |
| 3.8. References | 49 |
| 4. Essay from the Thinker: 'Societal Values in Digital Innovation – A Technologist's Perspective'..... | 50 |
| 4.1. Technology and disruption | 50 |
| 4.2. Design principles for a better, safer and fairer digital world..... | 52 |
| 4.2.1. Establishing authenticity and trust..... | 54 |
| 4.2.2. Empowering ownership..... | 54 |
| 4.2.3. Openness and transparency | 55 |
| 4.2.4. Traceability and responsibility | 56 |
| 4.2.5. Embedded policies and ethics | 57 |
| 4.2.6. Education and opportunities for all | 58 |
| 4.3. Humanity at a crossroads | 58 |
| 4.4. References | 59 |
| 5. Conclusions and recommendations of this Thinkers' Programme | 60 |
| 5.1. What?..... | 60 |
| 5.2. Who?..... | 60 |
| 5.3. How? | 61 |
| Appendix 1 - CVs of the Thinkers..... | 63 |
| Appendix 2 - Members of the Steering Group | 65 |
| Appendix 3 - Relevant literature | 66 |

Executive summary

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.

Foreword

Every year the Royal Flemish Academy of Belgium for Science and the Arts (KVAB) organises two Thinkers' Programmes on societal themes that KVAB independently selects for their particular relevance and the specific challenges for the region. The aim of every Thinkers' Programme is to enrich the broad societal debate on these themes with a vision of the medium and long term. This enrichment comprises three important elements: instead of looking at the topic from the perspective of one single discipline, a holistic and multidisciplinary approach is taken; Thinkers from outside Belgium are invited, who then put questions to those in charge of relevant local institutions; finally, the findings of the Thinkers are based on their more extensive international experience, understanding of the facts and evolutions around the world, and scientific evidence. Sometimes specific subtopics from the Thinkers' findings and recommendations are further explored at a later date by work groups set up by KVAB. These result in Position Papers and can also be included in policy recommendations that are prepared via the European Federation of Academies in the context of SAPEA Science Advice for Policy by the European Academies <https://www.sapea.info>.

In 2018 KVAB set up a permanent Reflection Group on Responsible Research & Innovation (RRI) and Science Ethics with members from all classes and experts from the Flemish scientific world. This Reflection Group deals with horizontal themes within European science and innovation policy, and links the aspects of the relationship between research, innovation and society with six themes: societal engagement, open access/open science, gender equality, science education, governance and ethics. The aim is to reflect on the societal impact of research and innovation, the ethical aspects of research, and scientific integrity. On this basis, the Reflection Group then organises activities on specific themes. Its first major activity was to set up a Thinkers' Programme on the theme of the societal aspects of digitalisation. The digital transformation taking place in our society is so far-reaching that the new European Commission and the new European President have already chosen this as one of their central areas for action. A steering group of this Thinkers' Programme then selected three complementary international experts, 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 representatives 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. Further honed by the discussion that took place there, they then submitted a final report with their findings and recommendations, which you will find below. We wish you an enjoyable read, and hope that it provides you with inspiration and ambition to further the cause.

Joos Vandewalle
coordinator of the Thinkers' Programme
31 January 2020

Before this report went to press, the entire world was struck by the Covid-19 pandemic, confronting us all with a series of unprecedented challenges. In addition to the focus on health care and elderly care, there is a concerted effort underway to find medicines and a vaccine in a model of interdisciplinary, open, and international cooperation. In addition to that, the role of scientists and evidence-based advice to the politicians in charge, both in the lockdown phase and during the exit strategy, has been extremely important. When scientists communicate well and take a multidisciplinary, scientific approach, they are also appreciated by the general public. In amongst all this, digitalisation has emerged on various different fronts as an important tool in daily life for working from home, distance learning, social contacts, online shopping, corona apps, etc. This is making us aware not only of a number of valuable elements but also of a number of limitations and problems. This corresponds to the analysis and recommendations of the Thinkers. As a result, a few additions have been included in the relevant places in this Position Paper. An exhaustive discussion of this topic falls outside the scope of this particular project. A broad scientific study in this area is certainly desirable. And the first move in this direction has already been made by 120 researchers from multiple disciplines in Belgium.¹

19 April 2020

The Dutch version of this position paper was accepted for publication by the Class of Technical Sciences of KVAB on 17 april 2020.

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

1. Positioning, aims and approach of the Thinkers' Programme

JOOS VANDEWALLE

Digitalisation is a disruptive process that has a significant impact on many aspects of society. Underlying this is the increasing processing power of chips and computers and the much greater collection and storage capacity of data, enabling much more powerful processing of information. The result is new and better services, communication and partnerships. At least two similar disruptive processes have taken place in history. The first was the transition from spoken messages to written language for transmitting information, so releasing it from the shackles of time or place. This separation troubled Socrates, because the written information was being disconnected from the person who spoke or wrote the words. The second disruption occurred in the 16th century with the arrival of the printing press, making it 'easy' to reproduce and spread the written word. This had a huge impact on society in Europe. Now we are living through an age of digitalisation, in which information such as speech, music, images, video, data and even metadata can all be represented in digital form, hence the digital transformation and innovation. The power of the computer, and the algorithms that make use of that power, can rapidly process this digital data, creating new services and the automation of a vast array of 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, the combination of human intelligence and artificial intelligence 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 recent times, this digitalisation has attracted a great deal of attention in politics, government, education and industry^{2,3,4}. Businesses as well as government institutions want to make maximum use of the opportunities offered by digitalisation, in order to offer better products and 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 online

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

banking. Citizens are also making increasing use of digital channels to interact with government bodies. We are all leaving multiple digital traces of data and metadata such as location in 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 relaxation. Education, from primary through secondary to higher, and extracurricular training is now also digitalised to a large extent. 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 is widely supported. This has recently been highlighted at the European level. Charles Michel, the new President of the European Council, reported that the new European Commission and the European Council share the same two top priorities, namely the fight against climate change, and the digital agenda. 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 develop new systems and services without considering the application, the users and the 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.

The thinkers involved

In early 2019, the Steering group (see Appendix 2 for members) selected three leading international experts as Thinkers. They were invited to work with the Steering group on the chosen theme throughout the year 2019. These three complimentary Thinkers (see Appendix 1 for their CVs) have wide-ranging knowledge and extensive experience, enabling them to study the theme from three perspectives: **ethical aspects, policy, and science**. Peter-Paul Verbeek is a professor of Philosophy at Universiteit van Twente; Rinie van Est is a theme coordinator of *Slimme Samenleving* (Smart Society) at the Rathenau Institute and a part-time professor in Technology Assessment and Governance at Technische Universiteit Eindhoven; and Jan Rabaey is a professor in Electrical Engineering at UC Berkeley.

Some of the subthemes that were also discussed include: major driving forces in digitalisation; ethically responsible design; digitalisation and education, learning analytics, continuous training and employment; data ownership; impact on professional sectors such as legal services, retail, self-driving vehicles, health care sector ...

In the first phase, the Thinkers were allowed free rein, and they shared their vision and experience with the Steering group. After that, the essential elements of this transformation process and the societal impact are scrutinised.

- Major driving forces in digitalisation (processing capacity, energy consumption, artificial intelligence (AI), internet of things, ...) and the relationship with today's society and the associated evolutions. Within AI, the focus is on data-driven AI rather than knowledge-based AI.
- Ethically responsible design of valuable digital services, value-driven innovation. Implementation of ethical codes for digitalisation and AI. The relationship between standards (voluntary basis, consensus building between stakeholders), ethics (not enforceable, flexible, fast), legislation (enforceable, complex, slow).
- Digitalisation, education, continuous training and employment (job supply, job change, job content and training), digital skills, learning analytics, digital humanities, and social skills.
- Democratisation:
 - Digitalisation and the relationship between citizens, governments, local government and businesses, technological citizenship.
 - Digitalisation and concentration of data (e.g. Tesla cars, learning analytics, Google, Facebook,...). Intelligent and social data-sharing.
 - Data ownership/access to data.
- Digitalisation in major Flemish industry sectors, e.g. manufacturing.
- Digitalisation in research/open data/ownership of research data.
- Global evolutions and strategies, economy, and public values. AI made in Europe, a third way for the EU alongside US, China.
- Selection of examples of professional sectors where digitalisation will have a serious impact: legal services (contracts, notary sector), retail (block chain), self-driving vehicles (robots, drones, autonomous weapons), health care sector, wearables.

Other Academies and umbrella Academies are developing activities with interesting discussions that are also important for this Thinkers' Programme such as ALLEA⁵: "Putting societal values at the heart of data use and establishing governance and regulation:

- 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?
- How fit for purpose are existing governance frameworks to adapt to the accelerating pace of technological development?

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

- 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?"

Debates with stakeholders

On the basis of this context and delineation of the theme, the Thinkers entered a debate with multiple stakeholders, partners and opinion-makers. As a result the Thinkers were gradually able to obtain insight into the local situation and could reflect on this from their international perspective.

Stakeholders debate 1 with businesses, business associations, social partners and 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, research unit Stichting Innovatie en Arbeid
- Koen Debaus, Digital Innovation Deloitte
- David Bruyneel, Accenture, Digital applied intelligence

Stakeholders debate 2 with school network organisations, universities and strategic research institutes:

- Peter Goyvaerts, GO! Community education, Policy and Strategy
- Greet Vanderbiesen, Catholic Education
- Peggy Valcke, KU Leuven, imec
- Lieven De Marez, UGent, imec
- Steven Latré, UAntwerpen, imec

Stakeholders debate 3 with government institutions and urban initiatives:

- Leo Van de Loock, VLAIO Vlaams Agentschap Innoveren en Ondernemen
- John Baekelmans, imec, smart cities, digital city of Antwerp
- Erik Mannens, iMinds-UGent

Stakeholders debate 4 with government institutions, research institutions, citizen associations:

- David Stevens, GBA Gegevensbeschermingsautoriteit, chairman of Privacy Commission
- Bart Preneel, KU Leuven, advisor to Privacy Commission
- 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

These intense debates provided a realistic picture of the activities taking place in Flanders together with the approach, challenges, problems and prospects. From the perspective of their international experience, the Thinkers then put together a basis of comparison for the Flemish context with a perception of the strengths and weaknesses. These ideas and recommendations of the three Thinkers for the authorities and various other interested parties were presented at a **final symposium in the Paleis der Academiën on Wednesday 4 December 2019**⁶.

Programme of Final Symposium:

Opening Freddy Dumortier, permanent secretary KVAB

Explanation of the Thinkers' Programme Joos Vandewalle, KVAB, KU Leuven (Organiser)

Preliminary reports of the Thinkers:

- *People and values central, reflections on the situation in Belgium* – Rinie van Est, Rathenau Institute, the Netherlands
- *Digital literacy, reflections on the situation in Belgium and worldwide* - Peter-Paul Verbeek, Universiteit Twente, the Netherlands
- *The role of digital technology, What can we technologists do to bring about an inclusive digital technology and value-driven digitalisation and what is the interaction with government, leading companies, etc.* - Jan Rabaey, UC Berkeley, US/international member KVAB

Digitalisation and internet in a midlife crisis? - Bart Preneel, COSIC, KU Leuven⁷

Parallel break-out sessions:

1. *Knowledge centre for Data and Society* - Rob Heyman, imec-SMIT VUB
2. *The digital city* - Bieke Zaman
3. *Ownership of data, models for sharing data* - Jan Rabaey
4. *Digitalisation, job content, training and continuing education* - Peter-Paul Verbeek

Findings of the parallel sessions and debate with the Thinkers (plenary) Jan Hautekiet, (moderator)

Reaction of the Thinkers

⁶ Programme of the Symposium with the speakers' presentations <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

Explanation about the Knowledge Centre for Data and Society based on the presentation and discussion in break-out session 1.

The main lines of the position paper compiled by the recently established Knowledge Centre⁸, which was presented formally in public on 9 December, were explained and discussed. The Knowledge Centre aims to be the contact point for ethical aspects of AI and data-driven applications for all AI stakeholders in Flanders. The plan is to achieve this aim in partnership with all existing initiatives. Once this position has been achieved, it would like to start standardising methods, norms and frameworks to make AI ready for society. If this succeeds, the Knowledge Centre can work with others involved on 'regulatory sandboxes' to improve AI innovation, regulation and policy where necessary.

It was clear from the discussion that the Knowledge Centre delivers scientific and independent advice. It works on behalf of three target groups: businesses, policy makers and the general public. It represents nobody; it provides these target groups with content to enable ethical, legal and socially responsible AI in Flanders. At the moment, the position paper consists of a general approach that could be regarded as extremely ambitious. Alongside this ambition there will be a more detailed annual plan and the services offered will be further defined and developed.

Three essays

The following parts of this KVAB Position Paper are made up of the three essays by the three Thinkers and a closing joint chapter with conclusions and recommendations. Each Thinker wrote his essay independently after the colloquium on 4 December 2019, aided by the far-reaching discussions held jointly with the steering group and the stakeholders in the course of 2019 in Flanders. Drawing on their specific backgrounds, perspectives and competencies, the three essays therefore constitute very complimentary but also consistent analyses, valuable approaches and proposed stimuli for further steps in this digital transition. Due to the independent nature of the various chapters, the bibliography for each chapter is published either at the end of the chapter or as a footnote. Given the extremely high quality of all the discussions conducted by the Thinkers with social parties and the constructive findings of the Thinkers, there is every reason to be confident that Flanders enjoys a firm foundation for the ethical guidance of digitalisation, and for the further modelling of citizenship in the digital society.

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

2. Essay of a Thinker: 'The search for a good digital life. Put people and values at the heart of digital innovation'

RINIE VAN EST

2.1. Waking up from irresponsibility

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)

For far too long there has been silence about how the information revolution sneaked up and radically changed the world (Est 2012). An important reason for that silence is that many scientists, entrepreneurs, politicians and journalists are continually proclaiming the 'technological dream'. For society, this is a hopeful and seductive narrative in which citizens can do things better with new technology, for example watch TV series uninterrupted on the train or during cancer cure. In this instrumental vision, technological progress leads inexorably to societal progress. For the advocates of innovation, it offers a carefree existence with clear roles and few responsibilities. Scientists and high-tech entrepreneurs are seen as heroes and don't need to take responsibility for the risks associated with the technology. Governments are allowed to aimlessly stimulate innovation and leave future generations of politicians and managers to tackle the negative social 'side effects'. The result is that the technological dream goes hand in hand with organised irresponsibility (Beck 1986).

The technological dream state is characterised by idolatrous worship of innovators and aversion to critics. What is significant is the difference between the way the public reacted to Steve Jobs when he presented the first iPhone in 2007, and the reaction to the Christmas speech given by Queen Beatrix of the Netherlands at the end of 2009. When Jobs introduced the iPhone with the words 'Once in a while a revolutionary product comes along that changes everything', his every sentence was cheered by an ecstatic audience. Two years later Queen Beatrix issued a warning about information stress and asked some critical questions about the quality of debate on social media: 'The modern technical possibilities appear to bring people closer together, but they remain at a 'safe' distance, hiding behind their screens.' Social media exploded with comments that this 'oldie' had no understanding of new technology and would be better to keep her mouth shut. This kind of public atmosphere in which criticism of technology is almost taboo, is

an obstacle to a proper debate and also therefore to the democratic guidance of technology. This has certainly been the case for far too long in ICT.

In the year 2020, the world seems to have partially woken from its digital trance. Courageous whistle-blowers have played a crucial role here. In 2013 Edward Snowden brought the mass surveillance activities of the NSA, the American security services, to the attention of the entire world. Whistleblower Christopher Wylie revealed how Cambridge Analytica, at Trump's request, used data from Facebook to try and influence the behaviour of voters during the American presidential elections in 2018. Zuckerberg, the CEO of Facebook, was forced to give an account of himself before the American Congress and admitted: "We didn't take a broad enough view of our responsibility and that was a big mistake". So, it became clear that the privacy paradise of yesteryear has turned into a control state (Vedder et al. 2007) and that now there is also a global control economy, the so-called surveillance capitalism (Zuboff 2019). The result is a loss of control over personal data *and* freedom of information, i.e. a loss of control over the information we receive and on the basis of which we make choices (Hof et al. 2012).

Because of the technological dream and accompanying 'culture of lawlessness and irresponsibility' (Nemitz 2018), digital innovation has enjoyed a free-for-all in the last two decades, as a result of which the accompanying risks have been allowed to run rampant, sometimes with catastrophic results. Now that the internet is slowly morphing into the Internet of Things and will effectively play a role in all areas of society, it is important to acknowledge this history of irresponsibility (see opening quote by Nemitz 2018). Basically, the question is how we can ensure that digitalisation is modelled responsibly from now on.

This essay addresses the issue of responsible digital innovation by first looking at the situation in the Netherlands. How is governance of the societal and ethical issues around digitalisation organised in the Netherlands? Governance is 'the capacity of a society to develop some means of making and implementing collective choices' (Peters & Pierre 2009: 91). Then we take a look at the situation in Belgium and Flanders. On the basis of the three KVAB debates about digital innovation between the three Thinkers-in-Residence and stakeholders, I have tried to obtain some insight into the way in which the relevant stakeholders in Belgium and Flanders talk about digital innovation. Are debate and policy on digital innovation defined by the technological dream or the desire for responsible digital innovation?

2.2. Blind spots in the Netherlands' governance landscape

The Rathenau Instituut in the Netherlands conducts research into the impact of science, innovation and technology on society. Providing information to the Dutch Parliament (Senate and House of Representatives) plays a major role in this process. The institute also tries to deepen and broaden the debate on

digitalisation. At the beginning of 2014 the institute identified the occurrence of an intimate-technological revolution: digital technology was nestling in and among us at lightning speed, gathering masses of data about us and simulating human behaviour (Est 2014). The fact that there is increasing interference in our personal lives via IT is leading to a political, economic and social fight for our intimacy. Digital innovation raises important ethical questions that touch on the (fundamental) rights and dignity of human beings: the right to privacy, a ban on the violation of physical and mental integrity, the right to a safe environment, the right to property and freedom of thought and conscience. At the time, the Rathenau Instituut called on politics and government to promptly develop frameworks to oversee this intimate-technological revolution on a societal level (Est & Rerimassie 2014).

The work of the Rathenau Instituut inspired Senate member Gerkens to submit a parliamentary motion on 23 September 2014 in which she signalled the advent of the Internet of Things and expressed concern about the social, legal and psychological effects thereof. In the motion, which was almost unanimously accepted, the Senate requested that the government ask the Rathenau Instituut to 'conduct research into the desirability of a committee that could advise on the ethical elements of the digitalisation of society.' The Senate suspected that important values were at stake as a result of digitalisation. The study *Urgent upgrade: Protect public values in our digitized society* (Kool et al. 2017) confirmed this. In this study we investigated which societal and ethical challenges were connected to digitalisation and the extent to which dealing with them is already politically placed on the agenda or institutionally embedded.

The terms digitalisation and Internet of Things refer to a cluster of digital technologies, such as robotics, artificial intelligence and algorithms and big data, digital platforms, biometry, persuasive technology, augmented reality and virtual reality. Together they are responsible for a new wave of digitalisation. An increasing number of components of the physical world are represented virtually. As a result, there are more and more places in which there is continuous feedback between the physical and virtual world, whereby products or services are directly or indirectly adapted on the basis of an analysis of digital data. The abovementioned technologies therefore play an increasingly important role in data value chains, i.e. for collecting, analysing and applying data. Applied to people, this means the measurement of people, the profiling of people and interference in the lives of people, such as controlling behaviour. Take for example the news feeds that social media companies 'personalise' on the basis of the user's surfing behaviour. The data value chain has thus become a cybernetic feedback loop, and as such is the main feature of the current phase of digitalisation, increasing its social impact.

Politics has been concerned with the issue of privacy since the 1970s. From a public perspective this means formulating and implementing a fair method of handling

and automatically processing personal data. Notable so-called 'fair information' principles include: purpose specification, restriction on data collection, purpose limitation of data, quality of data, security guarantees and informed consent. Whereas digitalisation was initially about collecting data, now the public and political debate is about the large-scale analysis and use of those data in a fair way. In recent years, the global debate on artificial intelligence (AI) has led to an identification of all kinds of principles for the use of algorithmic processing, such as the idea that AI technology has to be explainable (cf. Jong et al. 2019). This broadening of the debate has led to the realisation that digitalisation is putting pressure on other important public values besides privacy and security, such as autonomy, control over technology, human dignity, justice and equitable power structures (see Table 1).

Table 1. Societal and ethical issues relating to digitalisation
(source: Kool et al. 2017, 72)

| Topic | Societal and ethical issues |
|-------------------------|--|
| Privacy | Data protection, privacy, digital inviolability, mental privacy, surveillance, function creep |
| Security | Information security, identity fraud, physical security |
| Autonomy | Freedom of choice, freedom of expression, manipulation (dissemination of disinformation, microtargeting), protection of democracy, paternalism, skills, limits of self-sufficiency |
| Control over technology | Control over and understanding of AI technology, responsibility, predictability |
| Human dignity | Dehumanisation, instrumentalisation, de-skilling, de-socialisation, unemployment |
| Justice | Discrimination, exclusion, equal treatment, stigmatisation |
| Power structures | Unfair competition, exploitation, consumer-business relations, business-platform relations |

Below are three examples. The Cambridge Analytica scandal of 2018 showed how the autonomy of the democratic voter could be violated and resulted in a spurt of public awareness about how much data is collected, about the fact that providers earn money from user data and that companies and governments are in a position to try and manipulate human behaviour via profiling. The second example concerns human dignity or rather the lack of attention to human beings. On 18 March 2018 Elaine Herzberg was killed in Tempe, Arizona by an Uber self-driving car. An important cause of this accident was the fact that the Uber programmers had designed the software for the robot car such that it only detected pedestrians using a zebra crossing (Marshall & Davies 2019). Herzberg did what people often do: she walked her bike across the road. Because she did this at a distance of

approx. 100 metres from an official zebra crossing, the robot car detected her too late. The lack of a safety culture at Uber was the reason why insufficient attention was paid to human behaviour in their calculations. A third example concerns fair economic power structures. There is now a great deal of concern in Europe about the data power and market power of large American and Chinese platforms, such as Google, Amazon and Alibaba. During a Round Table discussion in the Dutch House of Representatives, fashion entrepreneur Erik van Rosmalen declared that the advent of platforms had caused a dramatic reduction in revenues among small and medium-sized businesses: 'In the last 8 to 10 years many retailers have seen a catastrophic fall in their sales and consequently their income, where percentages such as 30-40 are the rule rather than the exception' (House of Representatives 2018).

The Rathenau Institute also investigated the extent to which the legal framework and monitoring arrangements were sufficiently equipped for the above issues, which often touch on essential public values and fundamental rights. The issue at hand is therefore governance of the societal and ethical issues around digitalisation. The whole governance landscape was also examined: the role of science, the role of fundamental and human rights, the role of civic society and the role of politics and government in setting the agenda and determining and implementing policy. The analysis showed that the protection of public values under pressure from digitalisation was failing miserably in the Netherlands at that point. Five blind spots or governance challenges were identified (see Figure 1):

1. Translation of new societal and ethical issues into policy, interdepartmental harmonisation and coordination on digitalisation, and the political debate about these new issues.
2. Guarantee of fundamental and human rights in the digital society.
3. Empowering regulators and ensuring mutual coordination between regulators.
4. New responsibilities for developers of digital services and products.
5. Organising societal dissent: reinforcing civic society, knowledge and skills of citizens and societal debate about digitalisation.

In the Netherlands in recent years the public and political debate about governance of the societal embedding of digitalisation have received a real boost. In June 2018 the cabinet rolled out an integrated digitalisation strategy (EZK 2018). In order to get a better grasp on digitalisation, the House of Representatives set up the provisional committee Digital Future in July 2019. Its role was to investigate how the House can better oversee digital developments (Tijdelijke Commissie Digitale Toekomst 2020). In addition, various political parties, such as D'66 (Verhoeven et al. 2019) CDA (Havelaar & Dijkman 2019) and GroenLinks (Wouters 2019) have started thinking about digitalisation from the perspective of their ideological principles. The political silence about how the information revolution is changing

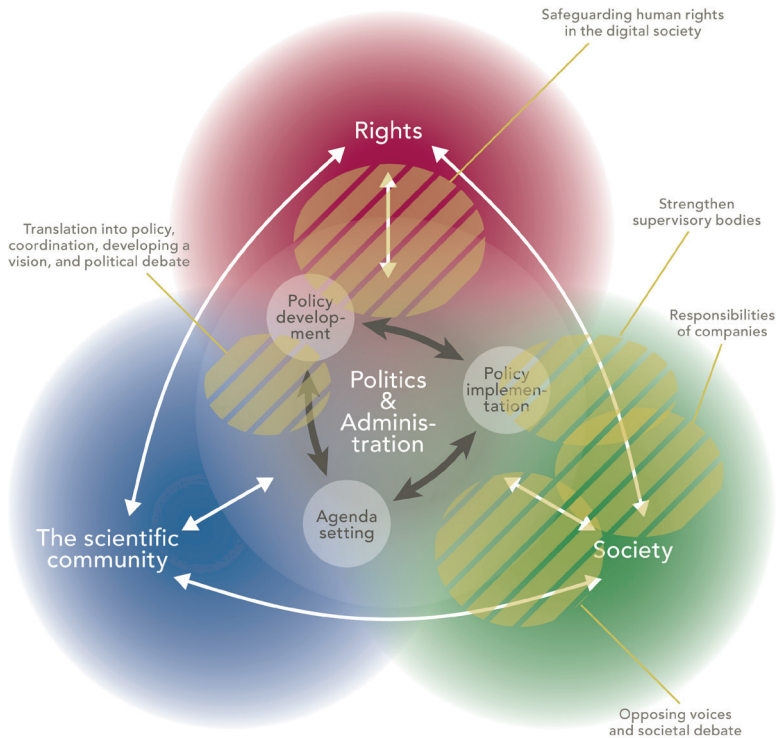


Figure 1. Blind spots in the Dutch governance landscape around digitalisation, ethics and society (source: Kool et al. 2017, 124).

society, which I talked about at the beginning of this essay, was broken in the Netherlands in 2019. As a result it is now possible to have a democratic pluralistic political debate about the future of the digital society. Following the Cambridge Analytica scandal and the enormous interest in AI and ethics among developers and media, the public focus on a number of social and ethical issues surrounding digitalisation has grown dramatically over recent years (Jong et al. 2019).

2.3. Thinking about digital innovation and acting in Belgium and Flanders

In the context of the KVAB Thinkers' Programme on digital innovation the above discussion raised the following question: Are Flanders and Belgium ready for the digital society? In other words: How do things stand in Flanders and Belgium in the matter of the governance of societal and ethical issues that have emerged as a result of digitalisation? The following issues also come up: Which public problems have been identified and put on the political agenda? How do various actors in

society discuss these problems? Are debate and policy on digital innovation defined by the technological dream or by the desire for responsible digital innovation? Who is involved and who is not? How are public values safeguarded at the institutional level? Which institutions have been set up over the years for that purpose?

Proper responses to these sorts of questions require far-reaching and in-depth research. As a matter of necessity, my ambition in this respect is far more limited. On the basis of the three KVAB discussions about digital innovation between the three Thinkers-in-Residence and stakeholders, I have tried to obtain some insight into the way in which people in Belgium and Flanders talk about digital innovation. These conversations show that stakeholders are aware of the fact that digitalisation comes with a wide range of issues (see table 2) and numerous governance challenges. For example, the Social and Economic Council of Flanders (SERV 2018) calls for experimentation spaces and testing grounds, a pioneering role for the government, and the stimulation of a dialogue on the impact of digitalisation and in particular the integration of ethical and social issues in technological development.

Table 2. Overview of societal and ethical issues highlighted during the three stakeholder debates about digital innovation.

| Topic | Societal and ethical issues |
|-------------------------|--|
| Privacy | Limits on measuring and sharing intimate information For example, among pupils and students in the case of learning analytics. |
| Security | Information security |
| Autonomy | Fake news, addiction to 'games' and social media |
| Control over technology | Risk of bias in computer algorithms, the importance of explicability of AI technology to enable meaningful decision-making |
| Human dignity | Shorter attention span among pupils (de-skilling). While the public fears job losses due to automation, the technology federation Agoria (2018) claims that for every job lost in the future because of digitalisation, 3.7 new jobs will be created. |
| Justice | A huge digital divide threatens to emerge in society. There is a digital divide between teachers. And there is a digital divide between pupils. Are we taking sufficient account of those people who cannot participate? Online as standard causes exclusion |
| Power structures | Role of (big) businesses in education that may in time jeopardise the democratic function of education |

2.3.1. National security, economic prosperity and social welfare

Hereafter, I will reflect on three key innovation goals – national security, economic prosperity and social welfare – and related innovation processes. There can be tension between these three goals. The current economy can be at odds with social and ecological limits in all kinds of ways (Raworth 2017). The realisation that the current global emissions of greenhouse gases is more than exceeding the planetary ecological limits upon which our lives depend (Rockström et al. 2009), is for example the moral basis for the sustainable energy transition. Another tension in the debate is the fear that too much focus on social welfare is undermining economic prosperity. This fear is very evident in the following quote by ALLEA (2019, 7): ‘The tension here is that Europe is losing its technological advantage and that this may lead to a fall in living standards and a decline in the political power that protects these standards.’ ALLEA, the European Federation of Academies of Sciences and Humanities, wonders whether Europe is not putting the protection of human dignity too far above economic growth and whether the incorrect use of data shouldn’t be balanced against the opportunities that this offers? (ibid.) The abovementioned tensions raise the question of how far national security, economic prosperity and social welfare can be combined.

In my ‘essay’ below I sketch out some reflections on the way in which Belgium and Flanders are modelling digital innovation in the area of national security, economic prosperity and social welfare and the extent to which they are taking into account its embedding in society. I compare the way in which digital innovation is modelled with a broad vision of innovation, inspired by the notion of socially responsible innovation (cf. Schomberg & Hankins 2019). In responsible (digital) innovation, the focus is on the societal impact of innovation, and those involved take responsibility for steering innovation a desired direction from a public perspective (cf. Est et al. 2017).

2.3.2. National security: The return of the drones

In the area of national security, the objective is often to retain the technological upper hand. This is nicely expressed by Yoge Patel (chief designer at Blue Bear Systems Research): ‘The only way to anticipate danger is to retain the advantage. ... I always say: put on your sneakers and run! You have to stay ahead’ (quoted in Hofman 2019). In the area of military drones and cyber weapons that has led to an AI arms race. The engineering aim underlying the new robotics is to build machines that can move independently and work in complex environments (Royackers et al. 2016). This desire to have the technological upper hand means that (military and commercial) drones will become increasingly intelligent and autonomous. Such technology can also fall into the wrong hands and be used maliciously. In the documentary *Drones: The next air disaster*, Jason Derick, the lead developer of Exyn Technologies, an American company from Philadelphia that

develops autonomous drones, is asked if such a scenario is possible. He replies, 'I think we should separate this technology from the way people use it. How people use it is essentially their business.' This statement shows that he refuses to take any responsibility for the potential abuse of the drones that his company designs. For the sake of national security, how can this development be handled responsibly?

The development of autonomous military drones raises the question of whether armed unmanned systems can be allowed to make independent (i.e. autonomous) decisions about lethal force. This automation of death obviously invokes ethical questions. Can autonomous military robots conform to the main principles of the humanitarian law of war, such as proportionality (the use of violence must be proportionate to the value of the military objective being sought) and discrimination (when making a choice of target, a distinction must be made between combatants and civilians, and between military and civilian objects)? In 2018, the Federal Parliament of Belgium ruled that such 'killer robots' must never be used by Belgium. Belgium is the first country in the world to have introduced such a preventive ban from ethical considerations. Research into automatic weapons is however still possible in Belgium. So, Belgium is still a participant in the arms race.

The development of military drones raises the question of the extent to which proliferation of armed drones forms an irresponsible risk. In Syria the Islamic State has been using commercial drones to drop bombs since 2017. The proliferation of military drones and abuse of commercial drones thus requires serious attention. In any case this also shows that the idea that technological superiority in military technology serves the interests of national security is too simplistic. Cheap and simple drone technology can also cause a lot of damage. In addition, there is a question as to how well Belgium is prepared to defend itself against a drone attack. Although the doom scenario of armed drones that return and take revenge (cf. Tenner 1997) is obvious from a military strategic viewpoint, it has long been ignored. It seems that real-life incidents are needed – e.g. the temporary closure in early 2019 of Heathrow airport because one individual with a drone jeopardised air traffic – before attention is focused on that sort of societal impact.

In short, from a broad vision of digital innovation, arms development should always take place in parallel with arms control. With its decision never to employ 'killer robots', Belgium has sent a powerful moral signal. From that position, the country can cooperate on global treaties for arms control, to ensure that arms proliferation can be stopped and military robots can be responsibly developed and employed in military operations. Finally, Belgium needs to prepare itself for a possible (terrorist) drone attack.

Table 3. Broad vision on digital innovation in the area of national security, economic prosperity and social welfare.

| Policy objective | National security | Economic prosperity | Social welfare |
|-----------------------------------|------------------------------|--|---|
| Innovation objective | Technological superiority | Development and use of key technology | Societal challenges |
| Risks of innovation | Strategic short-sightedness | Organised irresponsibility | Technology-driven; technologically obsolete |
| Broad vision of innovation | Arms development and control | Simultaneous development of technology, economy, debate and regulation | Value-driven digitalisation: people and public values central |

2.3.3. Economic prosperity: Belief in technology and society

I

n terms of economic prosperity there is a very real fear that a country will fall behind technologically and therefore economically. On this point, Vande Reyde, a Flemish member of parliament for Open VLD (liberal democratic party), is of the opinion that there has been too little focus on the passage in the Flemish government agreement: 'Flanders wants to work on the roll-out of 5G' (Reyde 2019). According to Vande Reyde, this got far too little attention during the debates held in the Flemish Parliament, and yet the roll-out of 5G is the policy resolution that will probably have the biggest impact on Flemish society. He warns that Europe, Belgium and Flanders are going to miss out to China and America in the area of 5G. Vande Reyde has a point: one of the important roles of government is to stimulate innovation so that a country can remain competitive. From the perspective of responsible innovation it is important here to address the societal embedding of technology. How much attention is paid to the impact of innovation in innovation policy, and what is being done to steer innovation in the desired direction?

Let me zoom in on how the Flemish government sees its role in the area of artificial intelligence (AI). During the plenary meeting of the Flemish Parliament on 22 May 2019 an interesting get-together took place between two members of the N-VA (New Flemish Alliance), a Flemish nationalist and liberal conservative political party. Member of parliament Gryffroy made the statement: 'I am an engineer. I believe in technology.' He then asked his party colleague Muylers, the Minister for Work, Economy, Innovation and Sport: 'What is the role of the Flemish government when it comes to stimulating AI on the one hand but also responding to those challenges and those more negative connotations on the other?' The minister outlined three roles: 1) investor in scientific research, 2)

stimulator and facilitator, and 3) awareness-raiser and trainer. The first role is self-explanatory. The second role is about informing the general public and raising awareness among small and medium-sized companies so that they don't miss the boat. The third role concerns training young people and re-training employees. In order to stimulate the dialogue on ethical issues and to give the government advice on this subject, the Knowledge Centre for Data and Society was set up (see also Flemish Policy Plan on Artificial Intelligence of 2019).

This view of innovation clearly ignores the technological dream. And yet it is still largely defined on a techno-economical basis. In order to clear the way for technology and economic activities, various social preconditions need to be created, such as providing information to businesses and citizens, training (future) employees and adequately addressing ethical issues. Nonetheless, there is still a real fear that the ethical dialogue will get in the way of the 'progress' of technology and economy. The Flemish Policy Plan on Artificial Intelligence (2019, 2) states that there are often legitimate concerns about ethical, legal and other issues, but that discussions about them 'frequently get stranded on possible undesirable or even negative consequences, without looking at what kind of new opportunities AI can bring about for existing societal challenges.'

In responsible innovation, society is not a precondition but the starting point, and technology and economy are the means. When modelling innovation from a public perspective, social desirability and societal challenges are central. Social issues are not feared, because they must guide innovation. So there is a change of perspective. The issue is no longer about digital technology that has to be embedded in society, but about shaping the digital society. This requires a capacity to give direction to the energy and vitality of innovation from the perspective of social ambitions and public values (Est et al. 2019).

Innovation also needs to be interpreted broadly – i.e. from a technological and social angle. To complement Gryffroy's pronouncement, here is another motto: 'I am a citizen. I believe in technology *and* society'. Social innovation is concerned with economic innovation (new earning or organisational models), political and social innovation (development of new language, public debate and political paradigms) and ethical and legal innovation (new moral frameworks and/or legislation). Shaping the digital transition requires attention and concomitant investment in all these types of innovation. After all, history teaches us that digital developments can happen very quickly and can have disruptive, sometimes catastrophic effects. If there is no simultaneous injection of energy into governance of digitalisation, then the public debate, our ethical thinking and laws and legislation, lags almost proverbially behind the technological development. But if society lags behind technology, this is a political choice. To put it more bluntly, following the technological dream was a political choice that has given free rein in the last twenty years to a 'culture of lawlessness and irresponsibility' (Nemitz 2018) in the area of digital innovation.

Developments in the area of, for example, face and emotion recognition and in the field of virtual reality (VR) – both applications of AI – require concomitant innovation and sometimes even a radical re-thinking of the way in which we look at and regulate digital innovation. In the area of VR there is a growing divide between the lack of political interest in the many social and ethical issues associated with VR on the one hand, and the need for timely development of frameworks in which to embed this technology on the other (Snijders et al. 2020). According to the Rathenau Instituut, the most fundamental question here is the extent to which VR should be regarded and regulated as biomedical technology (ibid.). The reasoning behind this is that the merging of the computer, the camera, biometric sensors, the VR glasses and the body is giving rise to an increasing number of possibilities for influencing individuals in real time and unobserved. In short, VR glasses cannot be regarded and regulated as a simple gadget; instead it is an intimate technology that is challenging politics and government to simultaneously develop timely frameworks in order to oversee the embedding of VR.

2.3.4. Social welfare: Value-driven digitalisation

Innovation in the area of education, government services, care and the living environment is not primarily about technological superiority or economic competition, but social objectives, like public health and quality of life. During the stakeholder discussions, a number of lovely Flemish examples came up. John Baekelmans, vice-president of imec's IoT and Connected Health Solutions Group, told us about his own personal learning curve. (imec is the largest independent European research centre in the area of digital technology). Before he came to work at imec, Baekelmans worked for Cisco for 21 years, the last years as Chief Technology Officer for the Internet of Everything (IoE) Solutions group. At Cisco they sold the technology dream and they adopted a 'technology push' strategy. Baekelmans witnessed many mistakes being made when technology was used 'for the sake of technology' without taking into account the citizen. In his view, neither technology nor the government should take the lead in shaping the city. Instead it should be the citizen, because essentially it is about quality of life and whether citizens consider the living environment a good place in which to live. If public space is to be purposed democratically, participation on all levels is necessary: from local through regional, to the governmental and national level. While the Netherlands is wrestling with the question of how lessons and innovations can be upscaled on a local level, the smart city policy in Flanders is a regional story. Antwerp is a test and experimentation city, a so-called living lab, and anything that works well there can be rolled out to other cities.

Frank Robben was involved in setting up the Crossroads Bank for Social Security – the coordinator of e-government in social security – and the eHealth platform, whose aim is to improve the electronic service and information exchange in health care. During a stakeholders' meeting, Robben described four elements in the

successful digitalisation of these types of public services. Firstly, it is important to involve all key stakeholders in the innovation process (and ‘to keep out the battalion of lawyers who come to defend interests’). Secondly, with the help of those stakeholders joint principles and goals must be identified. Some of them are already laid down in legislation, such as the GDPR (General Data Protection Regulation) concerning data protection and privacy. Thirdly, it is important to decide what is desirable and what is undesirable, and to look for a balance between the two. So, for instance, at the Crossroads Bank it was decided that identification would be managed by the participating partners – since nobody wanted to outsource that task to an ICT giant like Apple, there is no app for the Crossroads Bank. Finally, the values and norms that are important need to be reflected in the design of the digital system. This relates to architecture politics: politics and ethics by design. In practice the design must, for example, ensure a secure system that respects the personal privacy of patients.

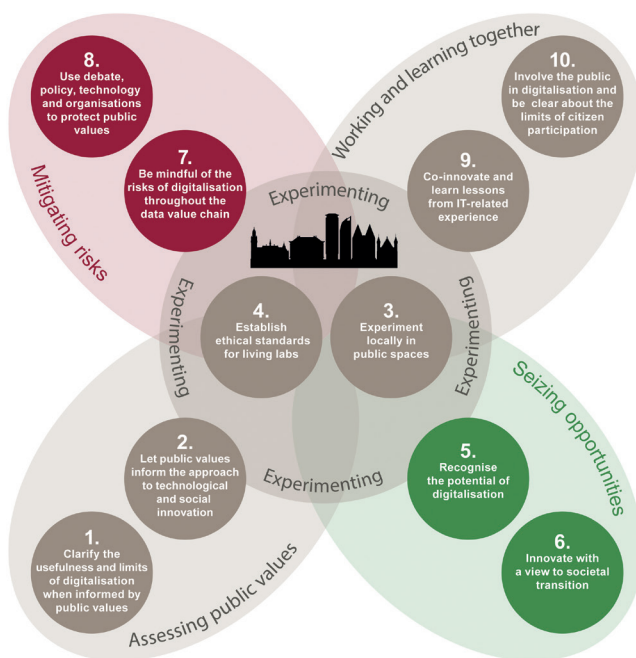


Figure 2. Value-driven innovation by means of five key innovation processes (source: Est et al. 2019).

The two examples from the previous two paragraphs are consistent with the perspective that the Rathenau Instituut refers to as value-driven innovation (Est et al. 2019). Value-driven innovation is about a constructive interaction between technical and social innovation. Five processes play a role here: assessing public

values, experimenting, seizing opportunities, mitigating risks and working and learning together (see Figure 2). 'Assessing public values' is about elucidating public values and objectives that people want to strive for. 'Experimenting' means making space for trying out new things. 'Seizing opportunities' means being open to the possibilities that digital technologies offer. The government undertakes to strive to ensure that our society can profit as much as possible from the blessings of digitalisation. 'Mitigating risks' involves protecting citizens as far as possible from the risks of digitalisation. Steering digitalisation in the right direction is a joint challenge. It requires coordination and cooperation between various levels of government and between diverse authorities and knowledge institutions, companies and of course citizens. 'Working and learning together' is the fifth ideal. The involvement of relevant stakeholders is crucial, but not always straightforward. There may, for example, be 'stakeholders' who do not wish to be involved or who are 'elusive'. How do you ensure, for instance, that children from families without an internet connection can still take part as a stakeholder in whatever way possible in the discussion about the digitalisation of education?

2.4. Final thoughts: The good digital life

In this essay I looked at how people in the Netherlands and Belgium and Flanders talk about and approach digital innovation. In both countries, there is no longer silence about how the information revolution has gradually and radically changed the world. It is no longer seen as open minded purely to sing the praises of digitalisation. The debate has transcended the technological dream. Both in the Netherlands and Belgium, the realisation has sunk in about how big and wide the impact of digitalisation is on society. Digitalisation no longer refers to a collection of gadgets but is now seen as a transition with opportunities and risks and a whole host of uncertainties. The transition perspective brings the question of the digital good life to the table and with it the key democratic question: What kind of digital society do we want to live in? (cf. Kool et al. 2018)

In both countries there is an awareness that digitalisation comes with numerous ethical, legal and societal challenges. In Flanders, the Knowledge Centre Data and Society was set up with that in mind. However, the question is whether attention to social and ethical issues is not too much seen as a compulsory activity; an appendix or necessary evil to legitimise digital innovation in political spheres. At the start of my essay I stated that the technological dream state manifests in the idolatrous worship of innovators and an aversion to critics. The Flemish Policy Plan for Artificial Intelligence (2019) still has an ambivalent attitude towards social criticism. On the one hand, it is clearly not possible to avoid difficult questions; on the other hand there is a fear that this criticism will throw a spanner in the technological and economic wheel.

This fear is both justified and unjustified. Justified because ecological, social and ethical aspects can indeed place limits on technology and economy. Unjustified because this choice between values should be viewed as a starting point for innovation. See above the opinion of Frank Robben and the perspective of value-driven innovation. By leaving behind the technological dream, we leave behind the myth that technological progress by itself leads to social progress and that economic innovation is obviously also positive from a public perspective. A broad perspective on innovation means that all kinds of relevant values and norms can play a role in guiding innovation. The question posed by ALLEA, the European umbrella organisation of scientific academies, about the relationship between economic growth and human dignity is therefore in principle a good question. I say 'in principle' because the proper handling of this question gives rise to a constructively more integrated discussion about digital innovation. In the worst-case scenario such a question can stoke fear and lead to a return to the technological dream.

Above I have shown what such a broad value-driven perspective on digital innovation means in terms of national security, economic prosperity and social welfare (see Table 3). National security is not sufficiently served by arms development. It is served by the least possible – depending on the state of the world – arms development and as much arms control as possible. The blind stimulation of digital innovation is giving rise to a great deal of techno-economic innovation and is also a formula for organised irresponsibility and social disasters. The trick is to make sure that technical and social innovation go hand in hand. That requires the simultaneous stimulation of both forms of innovation. That is precisely how Flanders modelled and continues to model digital innovation in the area of e-government and the smart city. The Rathenau Instituut advocates this kind of value-driven digitalisation in the Netherlands. In the case of socially driven innovation, Belgium and Flanders already seem to know how this can be done. Let that prove to be a boost for the modelling of the entire palette of digital innovation, so that it can make a positive contribution to the democratic search for the good digital life.

Postscript: 'Digital skills are important for hairdressers too'

As concerns the discussion on digitalisation, it seems to me a good idea to check how things are going in Flanders and Belgium with the governance of social and ethical issues that have arisen because of digitalisation. In the Rathenau study *Urgent upgrade* (Kool et al. 2017) and at the beginning of paragraph 3, there are more details on the sorts of questions involved here.

During the stakeholders' meetings, concern was often expressed about the huge digital divide that is threatening to emerge in Belgian society. There is a digital divide between teachers. And there is a digital divide between pupils. This latter

is perpetuated by the myth of the digital native: the idea that if kids grow up in a digital world, they will automatically become media-wise. The statement 'Digital skills are important for hairdressers too' by Greet Vanderbiessen (Katholiek Onderwijs Vlaanderen) made a deep impression on me. This sentence goes to the heart of the digitalisation challenge in three ways. Firstly, it shows that digitalisation affects everyone in society; secondly, it argues that digitalisation must have a positive prospect for everyone and so should be fair and inclusive. Finally, Vanderbiessen points out that every citizen has the right to digital training and skills. This is totally consistent with my opinion that technological citizenship is the democratic challenge of this century (Est 2016). Political-administrative institutions need to enable and stimulate three skills. Citizens must first be able to make optimal use of the benefits of technology. In addition, citizens must be resilient in the face of the risks of technology and protected against them. Thirdly, citizens must be in a position to take part in the public and political debate on the future of our digital society. Only then can citizens democratically claim their rights and responsibilities in the digital era. And that is why digital skills are important for hairdressers too.

2.5. References

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. Diederens, 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 of a Thinker: 'Technological Citizenship and the Ethical Guidance of Digitalisation'

PETER-PAUL VERBEEK

3.1. Introduction

The year 2019 will go down in history as the year of ethics in digital technology – with Artificial Intelligence in the leading role. Businesses, governments, social organisations: there are few parties in our society who have not worked on an ethical code in one way or other. The reason for this is undoubtedly that developments in digital technology are currently moving at such a pace and are so tangible on a societal level, that there are increasing concerns about the societal consequences: are we really in a position to sustain core values like freedom, justice and solidarity in a digital society? Questions like these also lie at the heart of the KVAB Thinkers' Programme on digital technology, of which this essay is a part.

The emergence of the digital society is often referred to as 'the fourth industrial revolution'. Following the invention of the steam engine in the first revolution, the second revolution brought mass production and the third revolution offered computerisation. In the current, fourth revolution, that information technology is *physical*: in the age of digitalisation, the internet has become an internet of things and the computer can have the body of a *robot*. As a consequence, digital technology is intertwined with all aspects of society and it is becoming increasingly obvious that a new society is emerging. From politics to health care and from personal relationships to public space, digital technology is changing the infrastructure of our daily life.

In Japan, the society that has arisen from this fourth revolution is now called 'Society 5.0'.

From this perspective, the plough caused the transition from a hunter-gatherer society to an agricultural society, after which the steam engine enabled the industrial society and the computer led to the information society. The fourth revolution is now leading to the fifth society type: the digital society, also called the 'super smart society' in Japan.

What is the significance of this digital revolution for the society in which we live, and in particular for Flemish society? The discussions organised in 2019 around the theme of 'digitalisation' in the context of the KVAB Thinkers' Programme produced numerous valuable insights into this question. On the basis of these insights, my primary aim in this essay is to address the question of the responsible modelling of *citizenship* in the digital society. Many of the challenges confronting us in the digital revolution seem in fact to convene in precisely this theme of technological

citizenship: this theme necessitates a response to the question how people can play a valuable role in a technological society, and therefore requires a better understanding of the impact of digital technology on people and society, and an analysis of the obstacles for citizenship. Three areas were of particular relevance in this context: public space, education, and politics.

In this essay I will first take a general look at the new relationships between humans and technology arising from the digital revolution. What is the nature of these relationships, and how can we determine the new impact of technology on people and society? In three separate paragraphs, I shall then look at digitalisation in relation to public space, education and politics. What is the effect of technology in these areas, what are the core values, and how can we provide those values with maximum space? Finally, I shall sketch the lines of a general framework for the ethical guidance of digital technology as a basis for technological citizenship.

3.2. The relationships between humans and digital technology

Digital technologies are shaping completely new configurations in the relationships between humans and technology. The traditional image of technologies being 'used' as a means to an end no longer applies to most digital technologies. The difference between an 'active' user and a 'passive technology' is making way for a much closer interconnectedness between humans and technology. In many cases, for example, it is more a case of 'immersion' than 'use'. The 'internet of things' equips our material environment with sensors that observe people, communicate with each other, and influence human behaviour. We don't use such technologies; we are immersed in them. In other cases there is more of a 'fusion', such as when deep brain stimulation influences people's moods, or insulin pumps independently regulate the blood sugar levels of diabetics. Augmented Reality glasses result in a configuration of 'augmentation', in which technology lays a second layer of interpretation over reality, which can be processed simultaneously. And artificial intelligence systems enable configurations that previously only existed between humans, such as interaction (robot and human respond to each other, are configured to understand one another, and are influenced by each other) or cooperation (human and algorithm are focused on the same reality and develop on the basis of their mutual interaction, such as medical expert systems that give a doctor advice on the basis of a common focus on the same patient).

Of all the technologies from this fourth revolution, Artificial Intelligence is without a doubt the most striking – and also the most feared and the most celebrated. While the US, China and the EU become entangled in a digital race, influential techies like Elon Musk and Bill Gates warn of the risk of this artificial intelligence overpowering humankind, and bestselling authors like Yuval Noah Harari write terrifying futuristic stories. A few years ago, ethics focused mainly on 'human improvement': are we heading towards the creation of a 'super human', and will

that be accompanied by a new inequality, injustice and indignity? Meanwhile, the discussion focuses more on 'human substitution': are we laying ourselves open to domination by our own technology? From a fear of 'too much of the human' we have shifted to a fear of 'too little of the human'.

In this sense, AI is exemplary for the digital revolution in which our society now finds itself. With the help of algorithms this technology enables computers to do things that previously could only be done by people: from making diagnoses and taking decisions to delivering judgements and controlling robot-bodies, from activating storm surge barriers and analysing the studying behaviour of schoolchildren to generating search results on the internet and live translations of the spoken word. It is entirely probable that artificial intelligence will have a huge impact on our society. And it is not always easy to cope with that impact, because it smashes through existing frameworks. Until now we have seen technology as 'passive' and humans as 'active', but artificial intelligence possesses a technological form of 'autonomy'. Computer systems can 'learn' and consequently develop in ways that their programmers could not always have predicted; and they are in a position to take decisions independently and to advise people on things that human thinking is less well equipped to oversee than a computer with artificial intelligence.

The core of the societal impact of AI lies in its effect on the *cognitive functioning* of humans: AI changes human thinking. From a historical perspective, this is not something new. Writing, for example, had a huge effect on how our memory worked: Plato worried that we would no longer be able to remember anything if we could write it all down. Book printing is a similar form of cognitive technology: all of a sudden, knowledge was no longer hidden away behind monastery walls. This led to its democratisation and ultimately laid the foundations for the scientific revolution. What is the new step in our thinking that AI will bring about? What will happen to our understanding of the world and ourselves, our moral and political opinions and conclusions, our existential questions and vulnerabilities? And how will AI impact the way our democracy functions, the relationship between citizens and government, and the possibilities to participate in society?

As with many other technologies, it is a good idea to distinguish three levels of the impact of digital technology. On an *individual* level this technology affects the choices, behaviours, and frames of interpretation of individual users: search machines determine which information people use to shape their image of the world, expert systems provide doctors and judges with a context for a diagnosis or a judgement, smart cities guide car drivers to specific parking places and dynamically regulate the maximum speed. On a *social* level this influence feeds through into social domains such as health care and jurisdiction (via expert systems), traffic and transport (via self-driving vehicles), journalism (in the form of 'automated journalism'). Finally, on a *societal* level, digital technology has an effect on the organisation of society: '*predictive policing*' is changing the nature

of the enforcement role of police, search engines influence political engagement, algorithms are involved in military decision-making.

3.2.1. *The current ethical debate*

In order to be able to properly address the societal implications of digital technology, in particular artificial intelligence, a great many ethical analyses and frameworks have been developed in recent times. A number of core themes pop up time and again in these analyses. First and foremost, the discussion often revolves around *bias*: AI systems can be biased, by their own design or by the data with which they are trained. A second central theme is *accountability*: to what extent is it possible to assign liability for the functioning and impact of artificial intelligence systems, for example for decisions that are taken on the basis of AI? *Transparency* and *explicability* are also important ethical issues: to what degree can a system explain how it reached its conclusions, so that humans can assess the basis on which they can trust the system? And finally, the *security* of the systems is often mentioned: to what extent are they trustworthy and resistant to 'dual use'?

All of these issues implicitly assume a battle between humans and technology. Computers can do things that previously only humans could do, but because they do them better and faster, and in such a way that humans can't always make an assessment, they are threatening. If computers can already recognise some forms of cancer better than pathologists, where does that leave doctors? And what is the future of journalism if news sites are provided with copy not only from journalists but also from artificial intelligent systems? How can we monitor whether the system is making the right analysis and is not biased, and can we even take responsibility for decisions made on the basis of AI?

From this perspective of humans pitted against technology, ethics naturally focuses on defending human autonomy against the new form of 'autonomy' that technology is developing. The key question is therefore whether there is sufficient human control, and whether or not the system excludes humankind or groups of people too much. Not surprisingly in its recent report the EU *High Level Expert Group for the Ethics of AI* focuses explicitly on 'trustworthy AI': ethics must guarantee trust between humans and technology.

Nevertheless, there is also an entire category of ethical questions that are not covered here, and which are important for the theme of technological citizenship. Those questions focus less on the individual and more on the community. *Inclusivity* is one such theme: how do we guarantee diversity in our society? *Well-being* is another example: how do we design artificial intelligence that makes a positive contribution to the flourishing of the whole of society?

In this light, it is interesting to report that the UNESCO *World Commission for the Ethics of Science and Technology* (COMEST) has chosen a completely different angle in this discussion. Artificial Intelligence, it is thought, affects the human mind. And that is precisely the domain of UNESCO, with its focus on education, science, culture and communication. The *Preliminary Study on the Ethics of Artificial Intelligence* that COMEST recently published therefore focuses on precisely these areas. How can *education* prepare people for a world in which technology makes some jobs superfluous, and how can it stimulate critical thinking and teach engineers to design responsible AI? What will AI mean for *science*: how is our understanding changing of what comprises a 'scientific explanation', and which 'facts' will we soon base our decisions on? What does AI mean for *culture* and cultural diversity? What sort of future do the minority languages have if automated translation increasingly becomes the norm? How can the suggestions for users made by platforms such as Netflix and Spotify using artificial intelligence systems also make room for diversity? And, finally: how can the quality of human *communication* be safeguarded, if AI enables new forms of disinformation, and becomes even more closely intertwined with journalism?

In these analyses COMEST explicitly takes a global view, that focuses not only on guaranteeing individual freedom but also on the flourishing of communities. As a result, COMEST also demands attention for a number of wider 'global ethics' themes, that come up time and again in UNESCO's work: Africa, gender, sustainability and peace. How do we ensure that Africa remains actively involved in the development and implementation of AI? How can we prevent AI from creating new forms of gender bias? How do we create a sustainable technological infrastructure for AI, in terms of rare-earth metals and the energy consumption of servers? And how do we ensure that AI is used for peaceful and not violent ends?

3.2.2. Citizenship in the digital society

The above makes clear that the new configurations of humans and technology caused by digital technology have huge implications for humans and society. Technology is shaping everything we do and even what we think in a new way, which challenges us to analyse this influence more closely and to take a more critical attitude towards it. And that is precisely what 'digital citizenship' requires: to play a full role in the digital society, it is necessary to scrutinise the effect of digital technology on humans and society and to evaluate it critically, and to design, use and embed this technology in our society in a responsible way.

With a view to further developing this notion of technological citizenship, I will now examine the implications of digital technology in three areas that played a central role in the discussions with stakeholders within the Thinkers' Programme, and that have an important relationship with citizenship: public space (paragraph 3), education (paragraph 4) and politics (paragraph 5).

3.3. The digital public space

Digital technology is playing an increasingly important role in public space, because it has shaped a new public space itself and because digital technology is also intertwined with the physical public space, for example in the design of 'smart cities'. A number of core themes returned in discussions with stakeholders, both in relation to the online world and as concerns the built environment.

3.3.1. The online public space

One of the most central ethical themes in relation to the digital public space is *disinformation* – also called 'fake news'. The difference between real and unreal is becoming increasingly blurred in our digital world, and also increasingly difficult to determine. That has major political consequences because it seriously hampers the societal debate. Hannah Arendt warned us about it in '*The Origins of Totalitarianism*': "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". If opinions and decisions can no longer rely on a shared basis of facts, it is no longer possible to keep a society together with consensus and shared understanding.

A second recurring theme was that of *internet addiction*. The second reality that the online world creates is increasingly competing with the real, physical reality. The very nature of the social interaction that comes with social media acts as a pull-factor and sometimes has an additive effect, due to the permanent possibility of sharing things, receiving reactions to this, and knowing what others are sharing. So, for some people the digital public space is not so much a meeting place to supplement the physical world as a space that keeps them out of the physical world. All the notifications that people receive via their mobile phones are constantly distracting them, and pulling them away from the physical world into the digital world, whereby the need for recognition ('likes' on messages) and the 'reward' of constantly finding new information play an important role.

Closely related to this is a third theme that cropped up again and again in our discussions: that of *digital exclusion*. Now that digital technology has become a vital part of the infrastructure of our lives, it is really important that everyone can partake of it. And that is not always the case. Elderly people are a particularly vulnerable group. Because many elderly people are not sufficiently familiar with digital technology, and don't always trust it either, a new generation gap is emerging that requires our attention. But this gap, which has incidentally already plenty of attention in the public debate, appears not to be the only one. There is also a second gap one generation below, which is especially visible in education: the gap between the 'digital natives', born in a digital world, and the generation

above them, which is familiar with digital technology but is using it in a different way and also often uses other technology, especially in the area of social media. Finally – and not surprisingly – the digitalisation of public space threatens our *privacy*: the fourth theme that played a central role in the conversations with stakeholders. The General Data Protection Regulation has certainly resulted in better privacy protection on a European level, but the problem seems to be much bigger than just protecting data. In addition to a watchdog to evaluate existing technology, there also seems to be a need for a system approach, whereby ethical issues about privacy are also addressed in the design and embedding phase of technology. This is the only way in which we can balance the value of privacy with other values that are also crucial for our society, such as security, solidarity and sustainability. In our battle for privacy, if we are too careful about sharing data, and if we fail to find a balance between individual rights and collective interests, we will, for example, miss out on important opportunities to forge sustainable energy policies, to fight crime, or to share our medical data out of solidarity with others so that it can be used for medical research. In short, GDPR is part of a learning process that still has a long way to go. It is important to take an integrated approach, in which regulation doesn't come as an afterthought but is part of the innovation process and in which a wide array of values play a role.

3.3.2. Digital technology in the physical public space

In addition to the online public space, the physical public space also played a role in the discussions with stakeholders. The increasing digitalisation of the material environment is changing the nature of the public infrastructure. In 'smart cities' Wi-Fi signals are analysed to measure how often people occupy certain places; data is gathered about things such as traffic flows and air quality, and people's behaviour is influenced by dynamic traffic signs and interactive lighting. This digitalisation of the public space introduces a new relationship between government and citizen, and with that new challenges for democracy.

To stop the government from acquiring too much power and to prevent the risk of too much surveillance, an experiment with 'co-creation' is being carried out. In this set-up, citizens are actively involved in the choices that have to be made on the political level, and the design of the 'smart city' then becomes a joint responsibility. The whole point is to make these forms of co-creation as democratic as possible. If citizens are really to be given a voice, their participation must go beyond just choosing from a number of predefined options. So, for example, a consultation with citizens may highlight the controversies that play a role in the decision-making on digital technology in the public space. What is at stake for citizens, government and business? Which values are at issue? And how can we best represent these values in the decision-making process? Moreover, it is extremely important to make the maximum effort to involve *all* relevant stakeholders and to represent their interests, including the stakeholders who are excluded or are at threat of exclusion in a digital society.

'Smart cities' not only enable new forms of surveillance and behaviour-influencing – which are first and foremost the embodiment of a new form of power of government over its people – but also new forms of political influence and engagement. Sensors can also enable citizens to put things on the political agenda that would otherwise not receive attention: noise pollution from air traffic, or air quality. A good example of this is the citizen science project 'Curieuze Neuzen', initiated by Universiteit Antwerpen, De Standaard and the <https://curieuzeneuzen.be>. Here, digital technology processes the opposite of 'fake news' and disinformation: sensors collect reliable information that can then play a role in political processes. Technology contributes to the formation of political issues: thus, 'sensing' plays a key role in 'sense making'. Rather than facilitating government domination, this technology gives the citizen a new channel of influence. And the broader the data streams, the greater this potential influence of the citizen.

3.3.3. Digital literacy

A key element when addressing these challenges is the further development of digital literacy. Technology and society are still approached far too often as two separate areas: in education, social disciplines are separate from technical disciplines, as a result of which social and technological perspectives also overlap far too infrequently in practice. It is vital that we learn to 'read' technology as something social, and so close the gap between alpha and beta. Digital literacy is interpreted far too often as computer skills or learning to programme, and should be extended to learning to understand the interaction between digital technology, humans, and society. This 'literacy' is part of 'technological citizenship': the ability to be a fully-fledged participant in the digital society.

Digital literacy enables citizens to adopt a critical attitude towards digital technology, and as such offers a counterbalance to disinformation, addiction, exclusion and privacy threats. Only by understanding the close connections between humans and technology is it possible to see through disinformation, to handle the impact of digital technology responsibly in your own life and in society, and to participate fully in the digital society.

3.4. Education in a digital society

This need to increase digital literacy is closely related to a second theme that played a major role in the discussions with stakeholders: namely, the relationship between digitalisation and education. The current global crisis caused by the coronavirus pandemic has served to make this discussion more pressing than ever: because of the need to work from home where possible, the digitalisation of education has been precipitated, and this necessary experiment with advanced digitalisation in education offers many opportunities to learn from.

The relationship between digitalisation and education concerns not only the *content* of the education (including the necessary focus on digital literacy) but also the *process* (the digitalisation of the educational environment and the digital divide between generations) and the *societal role* thereof, in a world in which 'technological unemployment' is on the rise and an increasing number of new competences and skills are required. Here too, it is important to note that in this discussion it appeared to be vital to remember the perspective of stakeholders who are not at the table: it is precisely those that may end up being excluded in a digital society who need to have a voice.

3.4.1. Digitalisation and educational practice

The practice of giving and receiving education has changed dramatically in recent decades with the arrival of digital technology. After digital teaching environments had drastically changed the interaction between pupils and teachers and also between pupils themselves, then came the arrival of the phenomenon 'learning analytics'. By collecting as much data as possible about the learning behaviour of pupils and their interaction with each other and with teachers, it is possible to find correlations that may ultimately lead to an improvement in learning performance. This data can come from the online behaviour of pupils in digital learning environments, but also from sensors in school buildings.

There are ethical and social questions connected with such technologies. Obviously the monitoring of learning and studying behaviour has a dimension of privacy and surveillance: to what extent are teachers and educational institutions willing and permitted to follow the activities of pupils? But equally important is the impact on the learning processes themselves. By making certain aspects of that learning process measurable, e.g. patterns of attendance and absence and of interaction with the study material, those aspects can become a goal in themselves. A learning process is about more than what can be measured using digital technology, which is why it is vital that the technology is designed, implemented and used from an explicit vision of what good education is, and of the values central to it.

The digital revolution affects education in another way, namely via the already mentioned 'digital divide'. The discussions with stakeholders revealed that this divide exists not just between pupils and teachers but also between teaching staff themselves. Older teachers are often more removed from the digital world of pupils than younger teachers. Technological developments are unfolding at such high speed that there is a need for lifelong learning. If teachers are to stay sufficiently attuned to the world and mindset of young people, they will have to become familiar with the technology and applications that help determine the everyday life of young people. A continuous learning programme, such as the one medics undertake, would be a good way to achieve this. By coupling the teaching qualification to an ongoing training trajectory, compensated where necessary by

a reduction in administrative tasks, there would be less risk of the digital divide becoming too big. The aforementioned theme of 'digital literacy' could play a central part in such a continuous training programme.

3.4.2. The role of education in a digital society

The digital revolution is leading not only to a different system of education, but also to a different labour market, for which education must prepare its students. Firstly, the nature of many jobs is changing as a result of digital technology, and with it the demands placed on education courses. But most importantly, job functions and professions are *disappearing*, because certain activities can be done by technology. And this puts new demands on education.

This theme of 'technological unemployment' has already been studied in a multitude of ways, and the conclusion is usually that there is a 'polarisation' of the labour market. It seems that the picture sometimes painted of the 'underbelly' of the labour market being affected and the disappearance of 'low-skilled' work is unjust: it is actually the 'middle segment' that will feel the effects of the digital revolution most acutely. In a digital society there will always be a need for people to build, maintain and repair the technological infrastructure, and there will still be a need for leaders and creative people: their work is more difficult to replace with robots and algorithms. It is the routine jobs, for example in administration and in secretarial support, that seem to be the easiest to carry out with technology.

At the same time, technological developments are not just erasing jobs, they are also creating jobs. This means that the division in the labour market that threatens to emerge because of digitalisation will not be so much between those whose work is or is not taken over by technology, but rather between those who do or do not have the resilience to re-train and change job. This is exactly where education can play an important role. Alongside knowledge and subject content, study skills and 'learning to learn' are becoming increasingly important. 'Lifelong learning' seems to have acquired a new purpose: technological developments create a very dynamic context for the labour market, which demands continuous education and further training but also re-training. And in all cases people must be equipped with basic skills in order to show this resilience.

3.4.3. Critical thinking and transdisciplinarity

The digitalisation of society also puts new demands on the *content* of education. A central part of this has already been mentioned above: digital literacy as the ability to understand the basics of digital technology both on a technological and social level. Closely linked to that are two other elements that must be present in education in a digital society.

The first is *critical thinking*. In a world in which artificial intelligence is becoming increasingly important, it is vital to be able to take a critical distance from the analyses of expert systems, the results of search engines, and the responses of 'voice assistants'. Education must teach students not to take information for granted, but always to question its origin and the conditions under which that information came about. What does an artificial intelligence system base its conclusions on, what data sets is it trained with, in what sort of environment has it learned to make connections?

The second is a transdisciplinary attitude. The traditional distinction between alpha and beta, social and technical, human-related and exact, is no longer sufficient for an understanding of society in the 21st century. Digital technology is not just a technical thing, it is also social. Societal problems can be resolved not just with social but also with technological solutions. Technology has a fundamental effect on humans and society and must be approached as such in its design, implementation and use, just as society can only be properly understood if we also attribute a place to the influence of technology.

3.5. Politics and digitalisation

The digitalisation of society also has huge consequences for how our democracy functions. We have already seen examples of this in the context of the digitalisation of public space. It seems that the tremendous power of large technology companies has made it very difficult for governments to take responsibility for the development and embedding of digital technology. And that means it is not always possible to get digital technology under democratic control, even though it is now seriously interfering in our daily lives. The importance of digital technology is now so great that it belongs to the primary infrastructure, for which the government is normally responsible, at least in the European model of democracy.

In discussions with stakeholders this democratic challenge was clearly identified. It seems, for example, to be difficult to enforce 'privacy by design' among developers of digital technology, and as a result the French police have resorted to hacking methods because the encryption of WhatsApp messages seriously hampers detective work. At the same time, there is a plea for the government itself to develop software, which would make it possible to control data on citizens in a central, secure place and to make it accessible, without using apps created by the big tech giants.

This theme ties in with the current global discussion on ethical AI in relation to democracy. In this discussion the EU is often positioned as an alternative for the American and Chinese model. As a sort of revival of the 20th century search for a 'third way' between capitalist West and communist East, the EU is now looking for a framework for AI in which it is not just the commercial interests of companies

that play the main role, and even less the need of the state to be able to control society. This search has led, among other things, to the ethical framework for AI set out by the abovementioned EU *High Level Expert Group on the Ethics of AI*.

It is very important to position this 'third way' not so much as an *alternative* for China and the US, but rather as a *connection*. From the perspective of the global framework laid down by the UNESCO *World Commission for the Ethics of Science and Technology*, it is precisely this good connection between individual freedom and responsibility for the collective that is essential. If the EU presents itself as 'the ethical alternative' or looks for space to be the third 'major player on the market', this opportunity for connection will be unnecessarily limited. Given the expected global impact of AI, it is more vital than ever to stop polarising and look for connections on the global fora.

The EU is crucial for this connection, not just as a joint-developer of technology, but also as a test ground for successful policy in relation to technology. Despite the criticism of the effectiveness and relevance of European policy and European legislation, if there is one area in which the EU has done extremely effective and important work, it is information technology. The abovementioned General Data Protection Regulation is a living example of that. The EU has actually succeeded in balancing individual and collective interests and pairing state intervention with freedom and democracy. We need to keep looking for these kinds of balance: the opportunity to deal responsibly with the impact of AI on humankind depends on it. The vanguard position of the EU in the area of *Responsible Research and Innovation* is a good foundation for this: there is ample experience with integrating ethical and societal reflection in technological development, and that experience should be maximised and developed.

3.6. Ethics as guidance for technology: Frankenstein's second chance

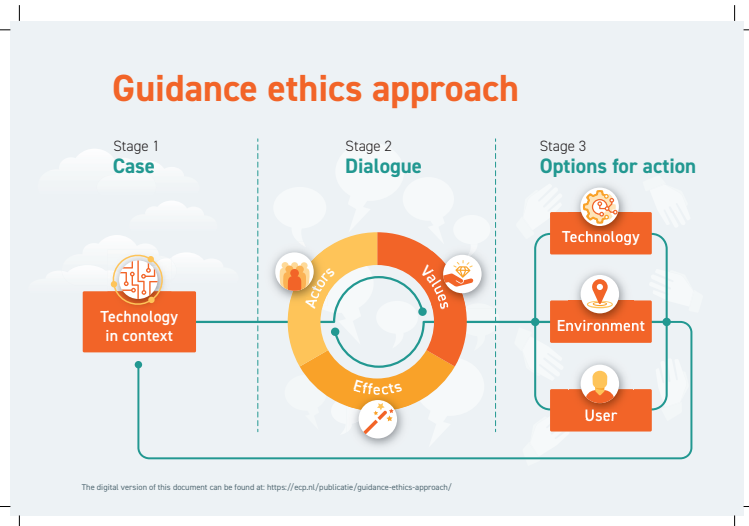
If these themes are to end up on the agenda, it is important not to pit digital technology *against* humans in the ethical and social discussion, but to regard it as *connected* to people. The current discussion about digitalisation is in fact a new version of the Frankenstein story. Except that in the fourth revolution technology doesn't come to life as flesh and blood, but via algorithms and robots: whether we are talking about technological unemployment, algorithmic decision-making or automated journalism, there is still a fear that technology is going to take humankind's place and inflict damage. And although the fear of the negative implications of digital technology is definitely not unjustified, the pitting of man against technology certainly is. In order to be able to understand the implications of this more clearly, it is necessary to see how digital technology is a *part* of human activities and shapes interpretation frameworks and decisions – just as writing and printing did previously.

This is precisely what went wrong with Frankenstein, whose monster only became a monster because Frankenstein abandoned him. After the initial euphoria of having created a living being, Victor Frankenstein became so frightened that he ran away and left the 'monster' to his lot. That monster then had to find its own way in the world. Roughly translated: if we want to prevent artificial intelligence systems and robots from derailing, we must educate them, not fear them. If we only allow ourselves to be ruled by fear of what might happen, we will let them down instead of finding the right place for them in society. Those who remain suspended in the choice between 'permit' or 'prohibit' will not discover the real impact of digital technology.

In the case of AI, moreover, that 'education' also has a literal significance. The crux of artificial intelligence systems is after all that they *learn*. They develop in interaction with their environment and do things that they were never programmed to do by their inventors. In the end, it is no longer the designer or the environment that is responsible, but the system itself. And so, ethics must focus on these systems and the way in which they learn. Which data sets are we using to train them? And how do the algorithms process these data sets?

The ethics of digital technology must therefore be far more than just asking what technology do we want and what technology do we not want. Instead of *judging* digital technology exclusively from the *outside*, ethics must guide it from the inside: by giving central ethical values a place in the responsible design, and also in its embedding and use. Algorithms and systems embody values, develop in the direction of certain values by the way in which they are trained and embedded and are implemented and used from the perspective of values. The selection algorithm that Amazon uses to find new personnel has become an iconic example of this: by using the profile of the existing successful employees as a model, the algorithm only selected white men between 40 and 50 and the workforce became really very homologous and discriminative. Because only the 'successful' value was used, and 'inclusivity' and 'diversity' were not explicitly included in the development and training of the algorithm, this development had negative consequences that could have been avoided.

'Responsible Design' and 'Responsible Guidance' of technology are therefore at least as important for AI ethics as the design of codes and frameworks. There is a lot of experience in the EU in the area of Responsible Design, and there are also a number of very active groups in the Netherlands. Responsible Guidance that focuses more on implementation and use is a theme that has been developed on a national level in the ECP workgroup Ethics and Digitalisation, and involves not only core players from the Dutch government and technical ethics but also big businesses like IBM, Microsoft and Facebook. This has resulted in an '*Aanpak begeleidingsethiek*' (approach to ethical guidance) that aims to translate the huge number of ethical codes and principles into concrete action points.



This approach, as described in Verbeek and Tijink 2019, could be useful in the development, embedding and applications of digital technology on the basis of the values that are central to user practice. The approach is 'bottom-up' rather than 'top-down': the first step is not the formulation of general principles but the detailed description of the technology in its real-life context of use and implementation. On these grounds, step 2 follows with an analysis of the impact of that technology on all the actors involved, in order to make an inventory of the values that are at stake. On the basis of this, the third and final step looks at developing action options for promoting the identified values. These action options can be divided into designing and re-designing the technology, adapting the context in which the technology functions, and critically equipping the user of the technology.

3.7 Conclusion

A concrete approach like this can play an important role in 'technological citizenship'. It embodies and stimulates digital literacy, which is necessary if we are to remain critical about digital technology; it connects a technological and social perspective and it focuses on the quality of digital technology in the context of personal existence, social practices and politics. The ethics of digitalisation is not just a question of abstract principles and grand narratives, but also and mainly concrete analyses and action points. The power of the European perspective as a connector between East and West lies precisely in the area of responsible innovation and ethics of technology. Given the extremely high quality of all the discussions that we conducted with groups from society, there is every reason to be confident that Flanders possesses a firm foundation for the ethical guidance of digitalisation, and for the further modelling of citizenship in the digital society.

3.8. References

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. and J. De Vries (2015). 'Toenemende polarisatie op de Nederlandse arbeidsmarkt': ESB Jaargang 100 (4701), 8 January 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 February 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 from the Thinker: 'Societal Values in Digital Innovation – A Technologist's Perspective'

JAN RABAEY

4.1. Technology and disruption

There is no question that new technologies lead to disruption. This has been the case throughout the history of mankind. Think about how the introduction of organised agriculture fundamentally changed how humans lived and interacted. Or how writing created a collective memory that could survive time. These "revolutions" happened time after time, and have shaped who we are as humans and how we operate as a society. The remarkable factor is that these changes, once induced, tended to spread very quickly, even in a world where communication channels were sparse and ad hoc. Consider, for instance, that the concept of committing information to memory by writing, after being "invented" in Sumer just before 3000 BC, spread all over the civilised world in the space of just a few thousand years. Yet, each of these technological innovations led to profound changes in how humans lived, worked and communicated. The industrial revolutions of the 1800's fundamentally changed the nature of work. The introduction of the automobile in the early 1900's revolutionised transportation and mobility, and so on. In each case, our way of living was affected and required us to adapt. Humanity has proven to be quite adept at that. One can argue about the value created by technology disruption, and the pro's and the con's, but the reality is that once a technology is deemed to be useful and is adopted, there is virtually no way to put it back in the box.

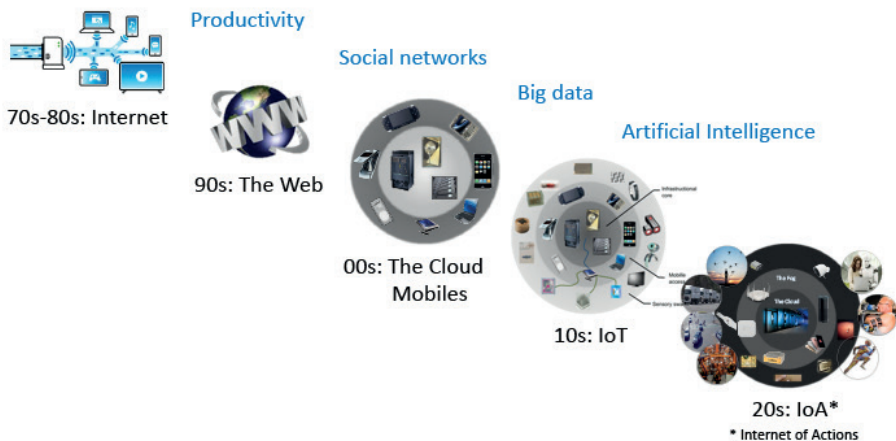


Figure 4.

The digital revolution started actually in the 1950's, and has morphed and expanded in an astonishing way (Fig. 4). With each phase of innovation came new opportunities, challenges, and disruptions. Jobs and industries were affected, and traditional patterns of life upended. The introduction of the Internet in the 1970's originally changed the nature of scientific research, but rapidly thereafter shortened time of information exchange from weeks to seconds. The (personal) computer modified the workplace, and helped to create a completely different perspective on productivity. It is arguably the simultaneous introduction of the World Wide Web and smartphones that had the biggest impact on society, because they radically overturned traditional models of interaction, daily life and business and caused a feeling of instability, all in the space of one or two decades. Even more disturbing is that this state of change and unbalance is continuing at a relentless pace, and is showing no signs of abatement. The combination of the introduction of the Internet of Things (and its many variants) and Artificial Intelligence (AI) effectively changes our perspective of the physical world around us and meshes it directly to a virtual world, created by our own imagination. Pretty soon we will be sharing that world with autonomous entities that help us to perform tasks independently or in close collaboration – creating the “Internet of Actions” (Institute of the Future 2016). The bottom line is that rapid change will continue, whether you want it or not, and it will drive who we are as humans and how humanity will evolve. The opportunity we have now is to help shape how that future may look, and ensure that the emerging digital world is better, safer and fairer.

For every technology that humanity has invented, developed or engineered, there are *good and bad sides* (e.g. Anderson 2018).

- Sensors placed in the environment help us track air quality, monitor traffic flow, use energy more efficiently, protect our homes, or track a burglar. The same technology however can be used by abusive governments to suppress dissent, or by companies to map and manipulate behaviour and invade privacy.
- Wearable devices can monitor your health and spot or address potential problems or help you to improve your lifestyle. That same information, when spread widely, can be exploited to drive up the cost of your health insurance (or even to deny you coverage), or could impact your chances of landing a job.
- Autonomous or semi-autonomous systems such as self-driving cars could reduce the number of road fatalities, and brain-machine interfaces and exoskeletons can bring back mobility to paralysed patients. Yet self-driving trucks can cost hundreds of thousands truck drivers their job, and autonomous weapons can threaten large populations.
- Etc, etc.

In summary, these technologies can do tremendous good and help to make this a better world, but the negative effects are severe and can outdo the good. It

also may be true that many of these goods are only serving a fraction of the population, leading to increasing inequality and a widening digital gap.

In light of all this, one plausible answer could be that as a society we should put all our efforts into maximising the upsides, and minimising the opportunities for the downsides, while preserving the tenets that were central to the original internet (open, unbiased, scalable). History has taught us that this can only be done through a multi-pronged approach that involves all stakeholders: the innovators/creators/technologists; the corporations; the governments; and the citizens/adopters. Too often technology is launched into the world with an unbridled optimism, and without any deep reflection or conversation about the possible negative or side-effects. Yet, at the same time, technology-bashing alone does little to help, as it is most often a co-play of all or some of the stakeholders that causes things to go awry. When Facebook was launched, no one had an inkling that social networks (a gathering place for friends) could or would be used to subvert elections. Much of this emerged later through the introduction of questionable business models and the (dis)ingenuity of users and adopters. Government only stepped in when it was way too late ...

That is why this study is so important. It helps if all parties think about how they can contribute to make the digital world better and fairer. One excellent example of this was given during the course of this study with the discussion about the "Antwerp - City of Things" initiative, where many of the ideas arose from small stakeholder meetings and were then vetted over time (imec 2016).

In the rest of this document, I will present a technological perspective, answering the following question: "What can we do from a creator/innovator perspective to address at least some of the concerns about the emerging digital world upfront, and maximise the probability that technology is used for the good?" The answer is quite simple: we should revisit the basics.

4.2. Design principles for a better, safer and fairer digital world

Over the past decade, quite a number of thinkers and visionaries have been mulling over what is going on with digital technology and what can be done to address those causes. When J.C.R. Licklider of MIT envisioned his "Galactic or Global Network" concept in the early '60s, he observed: "In a few years, men will be able to communicate more effectively through a machine than face to face" (Leiner 1997). He and his colleagues, who were at the origin of what is now called the Internet, envisioned a patchwork of decentralised networks, with no sense of how it would work when a handful of companies would write most of its software and managed most of its traffic (Brandon 2017). Now that vision is fraying at the edges. To quote: "The basic social compact of the Internet has reached its limit - and begun to break". It should therefore come as no surprise that some of the

creators of the Internet and the World Wide Web are ringing the alarm bells, and examining ways in which a revised Internet could restore its original tenets of open and free access.

In November 2019, Tim Berners-Lee, the inventor of the web, unveiled a global plan to save his brainchild. His plan is backed by more than 150 organisations, from Microsoft, Twitter, Google and Facebook to the digital rights group “The Electronic Frontier”. The proposed “Contract for the Web” states the following mission: *“The Web was designed to bring people together and make knowledge freely available. It has changed the world for good and improved the lives of billions. Yet, many people are still unable to access its benefits and for others, the Web comes with too many unacceptable costs. Everyone has a role to play in safeguarding the future of the Web. To achieve the Contract’s goals, governments, companies, civil society and individuals must commit to sustained policy development, advocacy, and implementation of the Contract text.”* (Contract 2019). It goes on to identify a set of “principles”, laying out the basic rules for each of the stakeholders (governments, companies and citizens) (Fig 5.). The stated goals are surely very fundamental and noble, and set out a clear set of expectations. Yet, there are no true guidelines on how to make them happen let alone how to enforce them. The only stick in the contract reads as follows: *“Those who back the contract must show they are implementing the principles and working on solutions to the tougher problems, **or face being removed from the list of endorsers ...”**. (My emphasis)*

Governments



Companies



Citizens



This is why it is worthwhile going one step further, and laying out some design principles for a better, safer and fairer digital world. While the list presented below is by no means exhaustive (and probably not without controversy), it helps define how the various stakeholders could contribute.

Fig.5: Stakeholders Principles for the “Contract for the Web” [<https://contractfortheweb.org>]

4.2.1. *Establishing authenticity and trust*

Building a fair and equitable digital ecosystem is extremely hard if there is little or no trust between the different parties, be they humans, machines, or a mix of both. How can one rely on information if one does not know its source or provenance? Similarly, how do you trust sensor data if the source could be a trojan horse? These questions expose some of the most challenging and gut-wrenching problems the digital world is facing today. One answer may be a strong authentication approach and the establishment of an unambiguous identity on the network. It would prevent rogue agencies spreading misinformation, and fake sensors disrupting the energy distribution system. This, however, collides with the core Internet principle of freedom. Freedom of speech could be threatened if agencies of authoritarian regimes can trace information that was supposed to be anonymous and take revenge. This is where innovative approaches can come in. Technically, it is not impossible to provide a single and true authentication approach that:

- Is multi-variate (what do you want to check?) and safe.
- Supports multiple profiles depending upon the interested parties (social, financial, health, government, ...), which are strictly and legally orthogonal. This expands on the "division of powers" concept that has served as guidance for democratic regimes for the past few centuries.
- Can be shielded for anonymity purposes, yet still be declared trustworthy.

Various trust establishment approaches (represented by agents) that meet the above requirements are under consideration today. The various stakeholders are, however, required to come to a consensus in light of competing goals and targets. While it is fair for a government to pursue potential security threats, it is unacceptable for it to use that same power to suppress its opponents. The same holds for citizens and corporations. This is clearly a place where third parties and independent organisations, governed and approved by all constituencies, can play a role. Think about how the internet name space has been governed for a long time by an impartial entity.

4.2.2. *Empowering ownership*

Data has been a precious commodity as of late and is becoming even more so with every passing day. The large internet corporations, financial systems, insurance companies and governments are banking on information that is made freely available to them and subsequently exploited in exchange for some free or even paid services. Unfortunately, this model is totally lopsided and wrong.

My data is and should be owned by me, the producer. For too long, we have been giving away our data for free. Usually, we even pay for that data – think about the blood tests you do every year as part of your annual health check. Often we have been blocked from even seeing or accessing the data. Only

recently has a fraction of our credit report – which determines if you are eligible for loans – or our medical reports become available for us to check. And yet you still have no control over that data and how it is spread, exchanged and traded for commercial or governmental purposes. The basic operational model is that, once you have relinquished your data, the holder is/was free to use and distribute it as deemed fit. One obvious way to address this is through rules and regulations – as witnessed by the many data protection laws that have been or are being put into place by governments all over the world. But this model has many pitfalls, and there is little built-in protection against non-compliance, abuse and exploitation of loopholes. In fact, governments themselves have often been the first ones to skirt around their own rules.

Hence the need for a totally different operational model, in which my data is owned by me exclusively and is never relinquished. To get to my data, companies and governments have to come to me. Temporal and restricted access is granted, often in exchange for a service or payment. The grantee never gets hold of the data itself (unless the owner explicitly releases it) and hence has no means of forwarding it to third parties. In other words, the owner keeps control over the destiny and distribution. This model is based on strong encryption techniques, authentication and micro-transactions. This idea has been pursued in a number of proposals over the past years, but is now getting some serious traction. An article in the NY Times of 19 November 2019 describes such an approach, under development by a start-up called Oasis Labs in San Francisco, founded by Dawn Song, who is a Professor of the EECS Department at UC Berkeley (Smith 2019). To quote: “(Oasis Labs) is building a platform that can give people the ability to control their data and audit how it is used. She believes that once data is viewed as property, it can propel the global economy in ways unseen before. New business models can be built on this, she said”.

Making such a model happen will require time, perseverance and a change in mindset. However, the public’s increasing attention to topics such as privacy may just act as a catalyst.

4.2.3. Openness and transparency

Black box systems, a nomenclature that applies to many of the digital systems deployed, are by default inexplicable, non-transparent and vulnerable. Errors caused by incomplete or undefined interfaces may give rise to exploitation. and the resulting actions are hard to explain or justify.

This explains the allure of open-source projects, products or initiatives. From opensource.com: “*(They) embrace and celebrate principles of open exchange, collaborative participation, rapid prototyping, transparency, meritocracy, and community-oriented development*”, typically leading to more robust and scalable

systems over time. Bart Preneel, a professor at KU Leuven, says: “*An important part of the solution is a radical choice for open solutions: open hardware and open software are the only way to create independence and to prevent backdoor access,...* (Preneel 2019)”. Open-source software has long been accepted and has formed the cornerstone of many operation systems (UNIX, LINUX) and, more recently, networking software (Open-Network Foundation). Hardware modules (often called IP for Intellectual Property) such as processors have almost always come in the form of a black box. This model is under severe strain with the introduction of the public-domain RISC-V instruction set (and the supporting RISC-V foundation). A similar trend is happening in the data world, where many agencies and publishers now require the data sets underlying a study, research project or publication to be made public. “*Open data* is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control.” (Wikipedia). While this has indeed enabled the creation of enormous data sets that have spurred developments in healthcare, wellness, disaster management, etc., it has also created unexpected side-effects and concerns with respect to privacy, reliability, etc.

4.2.4. Traceability and responsibility

These properties are closely related to the authentication, trust and openness principles elaborated on earlier. The many layers of obfuscation present in today’s complex digital systems often make it extremely hard to answer the simple but important questions that were phrased in the subject of this study: What? How? Who? It seems logical that any responsible digital system should be searchable, and capable of providing insights into at least some aspects of these questions. Two simple examples illustrate the associated challenges:

- AI systems should be able to explain why decisions were made and the reasoning behind them (“explainability”). This is certainly not the case with the current deep-learning systems. The importance of these questions, however, increases dramatically when AI is used in autonomous systems and bears implicit responsibility for the results of its actions. The same holds for any situation where liability comes into play.
- If an image or video is modified and changed and posted online, information on *when, where and by whom* should be embedded in a way that cannot be deleted or modified. This, for instance, could prevent the propagation of “deepfake” videos.

These topics are now burning hot in the Artificial Intelligence community. While progress is being made, further innovation is definitely needed. What is missing are global policies, and the framework to enforce their adoption.

4.2.5. Embedded policies and ethics

The enforcement of policies and rules requires that they are articulated or even exist. Very often they are implicit and baked into the creation of the system, and hence hard to change or adapt. For instance, the operational rules and decision points of an autonomous vehicle are most often fixed and hard coded, and are therefore hard to adjust or adapt in light of changing conditions. The same holds for the operation of the smart grid, wireless networks and traffic management systems. Given the various conditions and states in which a system can operate, operational rules should be adjustable. Consider, for instance, the case of a natural disaster. It would only be logical to give priority access to the wireless network for first-aid providers, changing the normal rules of engagement. The same holds for the traffic management system under these conditions.



Fig. 6: Policies and regulations

Making policies explicit as a set of rules governed by a “policy manager” helps to address this concern. This enables changes in policies, rules of engagements and regulations to happen dynamically and adaptively based on an evaluation of operational conditions and/or through human intervention. Newer generations of some of the above-mentioned distributed systems are already embracing this idea.

Ethics, generally defined as *a branch of philosophy that involved systemising, defending, and recommending concepts of right or wrong conduct* (Wikipedia), form one particular class of such rules, and can hence be addressed in a similar manner. The challenge however is to define, evaluate and measure what is right or wrong. The reality is often complicated and not that black and white. This is a domain that requires interaction between all stakeholders, and should combine philosophical, moral, societal and legal arguments.

A classic example is the 3 Laws of Robots, formulated by Isaac Asimov in 1942 (when actual practical robots were still a distant dream) (Asimov 1942-1950):

- A robot may not injure a human being or, through inaction, allow a human being to come to harm
- A robot must obey the orders given it by human beings except where such orders would conflict with the First Law
- A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws

At first glance, these laws seem to make perfect sense and provide a true sense of ordered decision principles. However, even Asimov published a number of stories shortly afterwards, in which he exploited various loopholes in the formulation of the laws themselves or explored situations that created inherent contradictions.

Another classic example is often quoted to demonstrate how hard it is to define ethical rules that are foolproof: “If a self-driving car is poised for a catastrophic

collision that would kill its passenger, should it execute one of the following actions: (a) swerve left and kill a playing child; (b) swerve right and kill an old lady; or (c) just let the collision happen. I leave it to the reader to answer this.

This set of questions has definitely caught the attention of a number of researchers in all fields involved. It has become a very rich field with many aspects such as biases in AI systems, weaponising of AI, unintended consequences, etc. A good compendium is defined in (White 2019). These discussions may yield truly insightful recommendations. However, the challenge is still to create a digital system that adheres to these rules and cannot be subverted.

4.2.6. Education and opportunities for all

There is no question that all the issues raised above are complex and may seem overwhelming to the general population. The only way to address this is to embark on a mission to teach literacy in digital innovation to all. Fortunately, this is slowly taking off at the college level, where many schools have started to provide courses and programmes in “data science”. These programmes, which are intended for the entire campus population, not only provide the students with basic insights into data statistics and artificial intelligence (and the supporting tools), but also apply them to examples of specific fields (health, arts, law, etc) and look at them in human and ethical contexts. A good example is the Berkeley Data Science Major, a Bachelor of Arts degree (Berkeley 2019).

This is just a start though. Digital innovation should become a part of the curriculum at any level of the educational chain, starting in secondary school. Given its rapid evolution and potential to invoke dramatic changes in the way we work, socialise and play, digital education should never stop, should be lifelong, and should aspire to reach all segments of society.

4.3. Humanity at a crossroads

In summary, it is clear that humanity is at a crossroads again. Throughout its evolution, such crossroads have occurred a number of times as stated in the introduction. This one may be different though, because it may be the first time that humanity itself is determining where it is going, and may change the way we perceive ourselves and our fellow human beings. Hence, we can have a major say in what that future looks like.

The digitalisation of society, as started a number of decades ago, has proven its inherent potential for good, and will undoubtedly continue to do so. We are living longer and healthier, and the percentage of the world population coming out of poverty is increasing despite the continuous growth of that population. At the same time, the potential for major harm and upheaval is definitely there

if left unharnessed. Doing nothing is not a viable option. Neither is the very human desire to turn back time and avoid the often very disturbing side-effects of upheaval. As we have elaborated on in this essay, once identified and understood, solutions can be articulated, or envisioned – as long as they are built on a buy-in from all affected parties. They will certainly require fundamental changes in how individuals, society, government, and companies engage and interact. One other thing is clear: it will need us all, and it will take some bold steps. In a number of ways, we are already too late ... So, let's not waste any more time.

4.4. References

J. Anderson and L. Rainie (July 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 and 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 (Dec. 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 (Nov. 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 (Sept 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 (Nov. 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. Conclusions and recommendations of this Thinkers' Programme

JAN RABAEY, RINIE VAN EST, PETER-PAUL VERBEEK AND JOOS VANDEWALLE

5.1. What?

The theme of this Thinkers' Programme, namely the profound changes in society wrought by digitalisation, was selected at the end of 2018 by the KVAB reflection group Responsible Research & Innovation and Ethics of Science. Since then, this theme has only become more prominent in the media, in politics, in business and generally in the societal discussion in Flanders, the EU and around the world. Evidence of this can be seen in the mass attendance at the end of January 2020 in Antwerp and Kortrijk for the lectures by Yuval Harari⁹ on the AI revolution. The Thinkers have also been able to see this first-hand in the various discussions they have had during the course of 2019¹⁰. The scope of this transformation and its impact in society certainly can't be underestimated, and the full expression and the impact are certainly nowhere in sight. Drawing on their experience and insight, the Thinkers have not projected doom scenarios or scare stories. For them, as for Harari (lesson 18, the future is not what you see in the movies), there is a distancing from the horrific images of science fiction. On the other hand, the Thinkers call for a new relationship between humankind and technology that includes, among other things, a deeper awareness, a cautious attitude, and a responsible approach. The Covid-19 challenge is also confronting us directly with this relationship in the context of working from home, distance learning, online shopping and social distancing. More than ever, an unforeseen problem like this requires 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 high-value recommendations. That should result in a general framework for the ethical guidance of digital technology as a basis for technological citizenship. The recent launch of the Flemish Knowledge Centre Data & Society is a very worthwhile element in this process.

5.2. Who?

All of society, young and old, governments in cities, regions, nations and EU and around the world, education and research, the business world, and social

⁹ Yuval Harari (2018) "21 Lessons for the 21st century," Penguin Random House.

¹⁰ Academieberichten KVAB 15 oktober 2019: Interview with Peter-Paul Verbeek, Jan Rabaey and Rinie van Est.

organisations, should feel they are doing their bit to understand this transition and put it on the right track on the basis of his/her role in society. It has been clearly stated that digital skills are relevant for all professions, for all organisations and businesses and for all citizens. But digital vocational skills are certainly not the only issue here. Citizens need to understand the opportunities and risks of technology so as to shape their own lives. They need to be resilient in the face of risks and take part in the public debate and the political democratic decision-making on the future of the digital society. Digital divides – between pupils, between teachers, between citizens, between digital natives and digital immigrants, between businesses themselves, between businesses and consumers, etc – must be avoided.

5.3. How?

The Three Thinkers are unanimous about the huge potential for digitalisation and AI. But this could just as easily lead to serious harm and unrest in society, if these technologies are used without restraint. Society is at a crossroads. Doing nothing or turning back is never a wise or workable option. There is a path to be followed, where the problems are analysed and possible scenarios for solutions are identified and understood, with input from all parties concerned. We also have to be open to fundamentally different creative ways in which individuals, society, government and business deal and communicate with each other. We all need it and it will require a number of bold steps. But time is running out, so let's not waste another minute.

The government is stimulating techno-economic innovation. Drawing on their own experience, the Three Thinkers deliver the message that value-driven innovation revolves around a constructive interaction between technical and social innovation. In other words, both techno-economic and social innovation must be stimulated alongside each other. Five processes play a role in value-driven innovation: evaluating, experimenting, seizing opportunities, mitigating risks and working and learning together (cf. Rathenau). 'Evaluating' is about assessing the public values and goals that people strive for. 'Experimenting' means making space for trying out new things. 'Seizing opportunities' means being open to the possibilities that digital technologies offer. The government should aim to ensure that our society can benefit as much as possible from the blessings of digitalisation. 'Mitigating risks' involves protecting citizens as far as possible from the risks of digitalisation. Steering digitalisation in the right direction is a joint challenge. It requires coordination and cooperation between various levels of government and between diverse authorities and knowledge institutions, companies, not forgetting citizens. 'Working and learning together' is therefore the fifth principle. There is, for example, an argument for 'co-creation' whereby experiments with new services or systems are set up in a collaboration between citizens and government or business. This can prevent a situation in which the government wields too much

power and there is too much room for surveillance. Citizens are actively involved in the choices that have to be made on the political level, so that the design of the 'smart city' becomes a joint responsibility.

The three Thinkers emphasise the important role of education, from primary through to higher, and equally of continuous education and re-training during the professional period. After all, a supported and shared open vision of the future can stimulate young people throughout their learning process to own their social responsibilities and can also ensure that older people have the flexibility they need to develop and evolve their career. On top of that, digitalisation and AI are creating a great many radical changes in the education process itself. Here we refer to the digital teaching environments that are drastically changing the interaction between students and teachers and between teachers themselves. The online learning experiences that were chalked up during this corona crisis in the infant, primary, secondary and higher education sectors will provide evidence-based insight into the value, possibilities and limitations. One of these attention points is surely the exacerbation of inequalities in society. During the learning activities a host of data can also be collected on student learning behaviour and their interaction with each other and with their teachers, and it is possible to seek correlations that may result in improvements in learning performances, so-called 'learning analytics'. Here we refer to a recent KVAB Position Paper¹¹. The digitalisation of society also places new demands on educational content, such as critical thinking and a transdisciplinary approach.

For the designers of new services and products, there is a need for design principles that will produce a more just, more transparent, safer, in short, a better digital world. The ethics of digitalisation concerns not only abstract principles and grand narratives, but also and above all concrete analyses and action points. The power of the European perspective as a connector between East and West lies precisely in the area of responsible innovation and the ethics of technology. The introduction of corona apps during the exit from lockdown caused by the pandemic is an example of the cautious introduction of digital technology with trade-offs between precision, security, respect for privacy, risk of resurgence of the pandemic, and support and trust among the general public.¹²

¹¹ T. De Laet e.a. (2018) ""Learning Analytics" in het Vlaams hoger onderwijs" Standpunt KVAB.

¹² Communication from the Commission, "Guidance on Apps supporting the fight against COVID 19 pandemic in relation to data protection", Brussels, 16 April 2020.

Appendix 1 – CVs of the Thinkers

Jan Rabaey holds the Donald O. Pederson Distinguished Professorship at the University of California at Berkeley. He obtained a degree in civil engineering and a PhD from the KU Leuven and conducted research at imec. He is the founder of the Berkeley Wireless Research Center (BWRC) and the Berkeley Ubiquitous SwarmLab, and was twice chairman of the department of Electrical Engineering at Berkeley.

Prof. Rabaey has made important contributions in various domains including advanced wireless systems, low power integrated circuits, sensor networks, and ubiquitous computing. His current focus is on the following generation of integrated wireless systems with a wide range of applications, as well as on research into the interaction between the virtual and the biological world.

He has been awarded numerous prizes, including the IEEE Mac Van Valkenburg Award, the European Design Automation Association (EDAA) Lifetime Achievement award, and the Semiconductor Industry Association (SIA) University Researcher Award. He is IEEE Fellow, international member of the Koninklijke Vlaamse Academie van België (KVAB), and has received honorary doctorates from Lund (Sweden), Antwerp and Tampere (Finland). He is also involved in numerous start-ups.

Peter-Paul Verbeek is a university professor in Philosophy of Humans and Technology and scientific co-director of the DesignLab at Universiteit Twente. He is also an honorary professor at Aalborg University, Denmark. His research focuses on the relationships between humans, technology and society as well as philosophical analysis, ethical reflection, and practices in design and innovation. His Dutch publications include *Op de Vleugels van Icarus* (2013), *De Grens van de Mens* (2011) and *De Daadkracht der Dingen* (2000); his English publications include *Moralizing Technology: Understanding and Designing the Morality of Things* (2011) and *What Things Do: Philosophical Reflections on Technology, Agency, and Design* (2005).

Verbeek is chairman of the UNESCO World Commission on the Ethics of Science and Technology (COMEST), vice-chairman of the UNESCO Ad Hoc Expert Group for the Ethics of Artificial Intelligence and member of the NWO Programmaraad Maatschappelijk Verantwoord Innoveren, the board of the Rathenau Institute, the Board of Humanities (KNAW), the Commission for Freedom of Scientific Pursuit (KNAW) and TNO's Regulatory Council.

Verbeek is a member of the Royal Netherlands Academy of Arts and Sciences (KNAW), the Academy of Technology and Innovation (ActI) and the Royal Holland Society of Sciences and Humanities (KHMW). In 2012 he received the Prof. Borghraef prize for biomedical ethics (KU Leuven), in 2016 the World Technology Award in Ethics (World Technology Network). For more information: www.ppverbeek.nl.

Rinie van Est is theme coordinator for the Rathenau Institute and part-time professor in Technology Assessment and Governance at the Technische Universiteit Eindhoven. In the Netherlands the Rathenau Institute stimulates public and political debate on social and ethical issues in relation to science and technology. He coordinates the Smart Society theme. As a physicist and political scientist he is a specialist in the politics of value-driven innovation. He is a worldwide expert in the area of technology assessment, governance and public participation. At the Rathenau Institute he has been occupied for more than twenty years with developments concerning energy transition and digital transition, and in particular the role of upcoming technologies in this area, namely robotics and AI. As a consequence, sustainability of the natural and human environment – respectively ecological and human sustainability – are core facets of his work. In addition, since 2000, Rinie has been working one day a week in the faculty of Industrial Engineering & Innovation Sciences of the Technische Universiteit Eindhoven, currently as a lecturer in Technology Assessment and Governance. He has jointly worked on the following publications: *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 – Members of the Steering Group

Luc Bonte (KTW)
Geert Bouckaert (KMW)
Hugo De Man (KTW)
Inez Dua (KVAB staff)
Freddy Dumortier (KVAB permanent secretary)
Vincent Ginis (Young Academy)
Johan Hanssens (EWI)
Liesbet Lagae (KTW)
Lode Lauwaert (KU Leuven)
Egbert Lox (KTW)
Jan Rabaey (KTW, UC Berkeley, thinker)
Rinie van Est Rathenau, (TU Eindhoven, thinker)
Ine Van Hoyweghen (Young Academy)
Marc Van Hulle (KAGB)
Joos Vandewalle (KTW, coordinator of the KVAB reflectiongroup RRI & science ethics)
Peter-Paul Verbeek (U Twente, thinker)
Pascal Verdonck (KTW)
Simon Verschaeren (EWI)
Paul Verstraeten (KTW)
Bieke Zaman (KU Leuven)

KMW = Class of Humanities

KNW = Class of Natural Sciences

KAGB = Belgian Royal Academy of Medicine

Appendix 3 - Relevant literature

Rathenau Instituut and 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 264 pages, 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>

International academies

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, Mai-2017

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

Ethical codes/legislation

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.

Flanders

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

VokaPaper-Verkiezingsmemorandum_Vlaams_HR_pages.pdf

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

Europe

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, March 2019. <http://www2.itif.org/2019-europe-digital-age-a4.pdf>

Coordinated Plan on Artificial Intelligence (EC Communication - 7 dec 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', 13-12-2018.

"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 pages.

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

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 (Sept 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, Feb 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, Dr Georgios Petropoulos, and Dr 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 - January 2019.

OECD Publishing, 'Digital dividend: policies to harness the productivity potential of digital technologies' OECD Policy paper February 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' March 2018, 228 pages.

Max Tegmark, 'Life 3.0Life 3.0: Being Human in the Age of Artificial Intelligence', August 23, 2017, 384 pages, Random House, ISBN 9781101946596

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



Society is at a crossroads: the digital revolution is proceeding at full speed and nobody knows what the total impact will be. In many sectors, digitalisation is the driving force behind innovation and change. Consumers are using numerous digital services and leaving behind multiple digital traces in the process. The current Covid-19 crisis has shown us, as never before, the advantages of digitalisation. At the same time, however, we should not be blind to the disadvantages and the dangers. There is an urgent need for a well-considered and widely embraced approach that shows respect for our social values.

The topic has a multitude of facets and angles, including many areas of expertise. This Position Paper presents three complementary essays written by international experts. They bring a message for society and policymakers that is focussed on the future. On the basis of profound insights, inspirational practical examples and concrete recommendations, they offer Flanders and by extension all countries a robust foundation for an ethical approach to digitalisation and for the further modelling of what citizenship means in a digital society.

This Position Paper does not waste time on doom scenarios. Rather, it is a heartfelt plea for a new relationship between humankind and technology.

The Academy's Standpunten series (Position papers) contributes to the scientific debate on current social and artistic topics. The authors, members and working groups of the Academy, write in their own names, independently and in full intellectual freedom. The quality of the published studies is guaranteed by the approval of one or several of the Academy's classes.