Ethical and societal challenges of the approaching technological storm

van de Poel, I.R.; de Wildt, T.E.; Oosterlaken, E.T.; van den Hoven, M.J.

DOI
10.2861/61151

Publication date
2022

Document Version
Final published version

Citation (APA)
van de Poel, I. R., de Wildt, T. E., Oosterlaken, E. T., & van den Hoven, M. J. (2022). *Ethical and societal challenges of the approaching technological storm*. European Parliamentary Research Service. https://doi.org/10.2861/61151

European Parliament

Ethical and societal challenges of the approaching technological storm

STUDY

Panel for the Future of Science and Technology

EPRS | European Parliamentary Research Service

Scientific Foresight Unit (STOA)

PE 729.543 – July 2022

EN

Ethical and societal challenges of the approaching technological storm

Blending augmented/virtual reality, AI, IoT, robotics, blockchain, bio-nano technology and 5/6G networks

Supported by the arrival of 5G and, soon 6G, digital technologies are evolving towards an artificial intelligence-driven internet of robotic and bionano things. The merging of artificial intelligence (AI) with other technologies such as the internet of things (IoT) gives rise to acronyms such as 'AIoT', 'IoRT' (IoT and robotics) and 'IoBNT' (IoT and bionano technology). Blockchain, augmented reality and virtual reality add even more technological options to the mix. Smart bodies, smart homes, smart industries, smart cities and smart governments lie ahead, with the promise of many benefits and opportunities. However, unprecedented amounts of personal data will be collected, and digital technologies will affect the most intimate aspects of our life more than ever, including in the realms of love and friendship. This study offers a bird's eye perspective of the key societal and ethical challenges we can expect as a result of this convergence, and policy options that can be considered to address them effectively.

AUTHORS

This study has been written by Ibo van de Poel, Tristan de Wildt, Ilse Oosterlaken, and Jeroen van den Hoven of Delft University of Technology (TU Delft), at the request of the Panel for the Future of Science and Technology (STOA) and managed by the Scientific Foresight Unit, within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament.

Contributions to specific chapters have been made by Wijnand IJsselsteijn (Eindhoven University of Technology), Dyami van Kooten Passaro (TU Delft), Olya Kudina (TU Delft), Michael Nagenborg (University of Twente), Madhumita Naik (TU Delft), and Filippo Santoni de Sio (TU Delft).

ADMINISTRATOR IN CHARGE OF SUPERVISING THE PROJECT

Andrés García Higuera, Scientific Foresight Unit (STOA)

ADMINISTRATOR RESPONSIBLE

Vasco Guedes Ferreira, Scientific Foresight Unit (STOA)

To contact the publisher, please e-mail stoa@ep.europa.eu

LINGUISTIC VERSION

Original: EN

Manuscript completed in June 2022.

DISCLAIMER AND COPYRIGHT

This document is prepared for, and addressed to, the Members and staff of the European Parliament as background material to assist them in their parliamentary work. The content of the document is the sole responsibility of its author(s) and any opinions expressed herein should not be taken to represent an official position of the Parliament.

PE 729.543
ISBN: 978-92-846-9682-6
doi: 10.2861/61151
QA-09-22-336-EN-N

http://www.europarl.europa.eu/stoa (STOA website)
http://www.eprs.ep.parl.union.eu (intranet)
http://www.europarl.europa.eu/thinktank (internet)
http://epthinktank.eu (blog)

Executive summary

Supported by the arrival of 5G and, soon 6G, digital technologies are evolving towards an artificial intelligence-driven internet of robotic and bionano things. The merging of artificial intelligence (AI) with other technologies such as the internet of things (IoT) gives place to 'AIoT', 'IoRT' (IoT and robotics) and 'IoBNT' (IoT and bionano tech). Blockchain, augmented reality and virtual reality add even more technological options to the mix. Emerging applications will affect each and every societal domain: business, healthcare, education, recreation, family life, governance, and so on. Smart bodies, smart homes, smart industries, smart cities and smart governments lie ahead, with the promise of many benefits and opportunities.

However, unprecedented amounts of personal data will be collected, and digital technologies will affect the most intimate aspects of our life more than ever, including love and friendship. It is possible though to steer these developments in a direction that is aligned with values such as privacy, justice, sustainability and wellbeing. To do that, technology developers, policy-makers and stakeholders have to join forces, an aim to which this study aims to contribute. The study uses responsible research and innovation (RRI) as the overarching framework for developing policy options. The main question addressed in this study is what key societal and ethical challenges we can expect as a result of this convergence, and what policy options can be considered to properly address these challenges.

Converging technologies

The technologies that play a role in the approaching technological storm may be seen as part of a layered structure. The backbone is IoT, as an infrastructure technology that allows systems of systems that are highly interconnected to be built. Communication within this infrastructure is enabled by technologies like 5G/6G. On top of this infrastructure, many applications in different social domains are made possible by general purpose technologies like AI, blockchain, bio-nano devices, augmented/virtual reality (AR/VR) and robotics.

One example is the creation of 'digital twins', virtual representations of physical things based on real-time data from sensors placed on and in physical things – including the human body. Another example is that we can expect drones to operate more often as part of an 'intelligent swarm' of drones, which have some degree of autonomy in taking actions. Facial recognition, voice recognition or bio-tracking technologies may become integrated into autonomous vehicles, to allow for a smoother human-machine interaction or even compliance with safe driving standards. The metaverse may become a reality and offer people alternative worlds that seem very real and are highly immersive and persuasive. Through the integration in the internet of things and by using blockchain technology, ordinary objects such as street lights may become semi-autonomous profit centres, with micropayments determining when they are switched on, and maintenance being automated and optimised based on a range of data.

Known ethical and societal challenges

For each of the technologies that go into the blend, we have looked at some of the main ethical and societal challenges they raise according to the existing literature:

- **5G/6G**: Energy usage is one of the concerns associated with the introduction of 5G/6G networks. There are also some concerns about environmental and health impacts because of the frequencies and wavelengths used. So far, there is no proof that these concerns are justified – but there are still some uncertainties that warrant further investigation. Especially with the arrival of 6G, security and privacy are expected to become bigger challenges.
- **AI and robotics**: The main short-term and concrete ethical and societal challenges discussed in the literature are privacy, surveillance and manipulation of behaviour, transparency and explainability, and bias and discrimination. For AI-driven robotic devices, more specifically,

meaningful human control, and employment and the future of work are the main areas of concern. AI also raises challenges for democracy, such as 'filter bubbles' and 'deep fakes'.

- **Internet of things**: Most of the ethical challenges raised by the IoT – like privacy, informed consent, trust, security, social justice, responsibility/accountability and human freedom and agency – are not completely new. However, certain features of the technology, such as its radically distributed nature, add new levels of complexity and salience.
- **AR and VR**: Two recent reports identify several types of risk in consumer VR applications, namely (1) physical and mental risks, such as addiction, (2) social risks, such as social dissociation, (3) abuse of power, such as use of data without permission, (4) legal risks, such as property issues, and (5) damage in the physical world, such as traffic accidents caused by distracted users.
- **Blockchain**: With some exceptions, the ethics of blockchain has received little attention until recently. Blockchain technology is generally presented as a solution to the ethical challenges arising from other technologies, most notably the challenges of privacy and transparency. However, it may have more long-term negative disruptive consequences, for example the propertisation of private data.
- **Bio-nanotechnology**: 'Biotechnology' and 'nanotechnology' are very broad categories, and a lot has already been written on the ethics of these two broad areas of technology. We have, however, been unable to find any literature on the ethical and social concerns of bio-nanotechnology as part of the IoT – except for new security challenges.

Discussions in media, ethics, tech and policy

For this study, a text-mining exercise allowed four bodies of texts from four different realms of society: ethical research, news and media, regulation and legislation and technical and scientific research publications to be analysed. The underlying idea is that each of these realms plays a specific role in an overall division of labour that is required to properly address new challenges and to bring about responsible innovation. We have investigated how frequently certain core values are being addressed in these four realms of society; to this end we applied a topic modelling method. Our main findings are:

- Ethics (and ethical, legal and social implications (ELSI))[1] research can fulfil an early warning or early detection function when it comes to the timely discerning of new challenges. It already seems to play this role for a number of values (like privacy, fairness and justice, democracy, autonomy and transparency), but less so for the value of sustainability. Moreover, it tends to focus rather one-sidedly on AI and robotics, hardly explicitly addressing values and issues related to other digital technologies like blockchain.
- The news may have to play a role in bringing relevant values, and ethical and social issues and challenges, to the attention of a larger audience. Our analysis shows that the media do so rather well for most values and technologies, but values like autonomy and transparency seem somewhat under-represented.
- Most relevant values seem well represented in legal and regulatory documents, apart from wellbeing, which might require more attention in the future. Legal and regulatory documents furthermore seem to focus primarily on AI and robotics, potentially neglecting other technologies.
- Our analysis suggests that values like reliability, cybersecurity, privacy and sustainability are already well addressed in technological research and innovation. For other values like democracy, autonomy, transparency, and wellbeing this is (still) less the case. This seems particularly problematic for democracy, as this value is frequently mentioned in the ethics and news datasets, and has even been prominent in the ethics dataset since 2000.

[1] ELSI or ELSA stands for the Ethical, Legal and Social Implications or Aspects of emerging science and technology.

In addition to these four conclusions, a main conclusion from our text-mining exercise is that the merger of digital technologies indeed seems to raise new or at least increased ethical issues

Features, opportunities and challenges of converging technologies

A lot of the ethical discussion about digital technologies in the past few years has focused on AI. Therefore, it is important to note that the challenges we have identified extend well beyond those that are typically or usually discussed in the AI ethics literature. The convergence of digital technologies will lead to new technological applications, but will also contribute to the creation of new sociotechnical systems and systems of systems, which may raise their own challenges. The convergence will therefore most likely result in technological possibilities and features that extend beyond those of individual enabling technologies like AI, IoT and blockchain. To address the challenges of the new 'technological storm', we might well need to look for policy options and regulation that extend beyond the realm of AI and the concerns it has raised.

Many, if not all, new applications and socio-technical systems will display one or more of the following features: interactive, long-distance, distributed, autonomous, intelligent, adaptive, reconfigurable, hybrid, fully connected, invisible, fast, precise in location, intimate, immersive, persuasive, and commercially exploitable. These features partly stem from the individual technologies that go 'in the mix'. For example, features like interactivity, autonomy, intelligence and autonomy are typical characteristics of AI systems. However, some features also emerge due to new combinations of technologies. Moreover, it is often the combination of the features that creates new challenges for society, policy-making and regulation. Based on these features and inspired by interviews with a number of experts, we have identified nine key opportunities and challenges:

1. **Digital sovereignty and new economic and social opportunities**: Converging digital technologies create new opportunities for economic growth and for the creation of social welfare. They also offer opportunities to better meet societal needs and to address societal challenges. An important concern is that they are developed in a way that ensures Europe's technological/digital sovereignty, allowing it to uphold its values and steer its own course. This may be increasingly important in the light of recent events like the coronavirus pandemic and the war in Ukraine. It may also be needed to decrease dependence on international 'big tech' companies for digital technologies and to ensure economic and social prosperity in the EU.
2. **Blurring of social and economic areas**: Not only are sociotechnical systems within one social domain increasingly connected, but there are also increasing connections between systems from different social domains (e.g. education and health care). This may result in information flows between social domains that are considered problematic and lead to new possibilities for commercial exploitation.
3. **Increased impact on people's intimate life**: Privacy is a key concern with the technologies addressed in this study. However, the invasiveness of these new technologies in people's intimate life means that we may need to move beyond traditional ways of thinking about and responding to privacy issues. The dominant ethical and legal paradigm is informational privacy, often operationalised as informed consent. However, some of the new challenges raised extend well beyond informational privacy. Addressing these issues will likely require more attention for values like wellbeing and human dignity, as well as for human rights.
4. **Opacity and cognitive overload**: Digital convergence may lead to new applications and socio-technical systems, of which the functioning is hard if not impossible to understand for users and the general public (opacity), or to cognitive overload due to the amount of information that humans need to process in often short time intervals. Meaningful human control will be hard to realise, which has implications for responsibility and accountability.
5. **Energy use and sustainability**: Converging digital technologies may be used to reduce energy consumption and for other sustainability purposes, but they may also increase total energy use and the carbon footprint. Sustainability has so far received little attention in ethical

and legal frameworks for AI. But with the advance of technologies like 5G/6G and blockchain, energy consumption is becoming a challenge that urgently needs addressing.

6. **Increased cybersecurity risks and new cyber-physical risks**: The hybridity of the new systems of systems makes cybersecurity and physical safety and security increasingly interconnected, rather than independent, concerns. Full connectivity may also introduce additional risks, as failure of one component in the system may have cascading effects on the entire system.
7. **Disruptive effects**: Disruptions are to be distinguished from mere impacts not just by the severity of impact and the fact that they may occur in a short time period (like shocks), but also because such impacts may be irreversible. We may see disruption of existing (economic) markets, of social practices and institutions, of regulatory regimes and of the ethical concepts we use to assess new technologies and developments.
8. **Concentration of techno-economic power**: Many of the new applications that arise at the merger of AI, IOT, blockchain, and 5G/6G, are developed by a handful of companies operating internationally, who are very powerful and hard for individual governments to regulate. This concentration of power is further strengthened by technological features and choices, like the increased connectivity of systems. Extensive IoT networks have greater commercial value. However, they may not only reinforce the uneven distribution of techno-economic power, but also make it harder to address challenges like the blurring of social spheres and the creation of new cyber-physical risks.
9. **Fundamental unpredictability**: Some of the challenges and issues that the approaching technological storm will bring, may not only be unknown at present, but may also be fundamentally unpredictable. All the challenges that we discuss above are, to a lesser or greater extent, uncertain. What makes this particularly challenging is the 'Collingridge dilemma': at the early stages of technological development we typically lack knowledge about the societal impacts of new technology, while at the later stages, when such knowledge is (more widely) available, technologies have typically become so well-entrenched in society that it is quite hard to then shape their design and societal embedding.

Policy options for responsible research and innovation

In this study we use responsible research and innovation (RRI) as the overarching framework for developing policy options. RRI has four dimensions (Stilgoe, Owen and Macnaghten, 2013):

- **Inclusiveness**: relevant stakeholders, and their values and needs, should be included in the process of technological innovation from the start;
- **Anticipation**: the impacts, benefits and risks of the technology should be anticipated and these anticipations should be fed back into the process of technological innovation;
- **Reflexivity**: the underlying purposes, motivations, and values for technological innovations should be reflected upon and should guide the process of technological innovation;
- **Responsiveness**: technological developments should be responsive to the values and needs of society and to new insights and developments along the way.

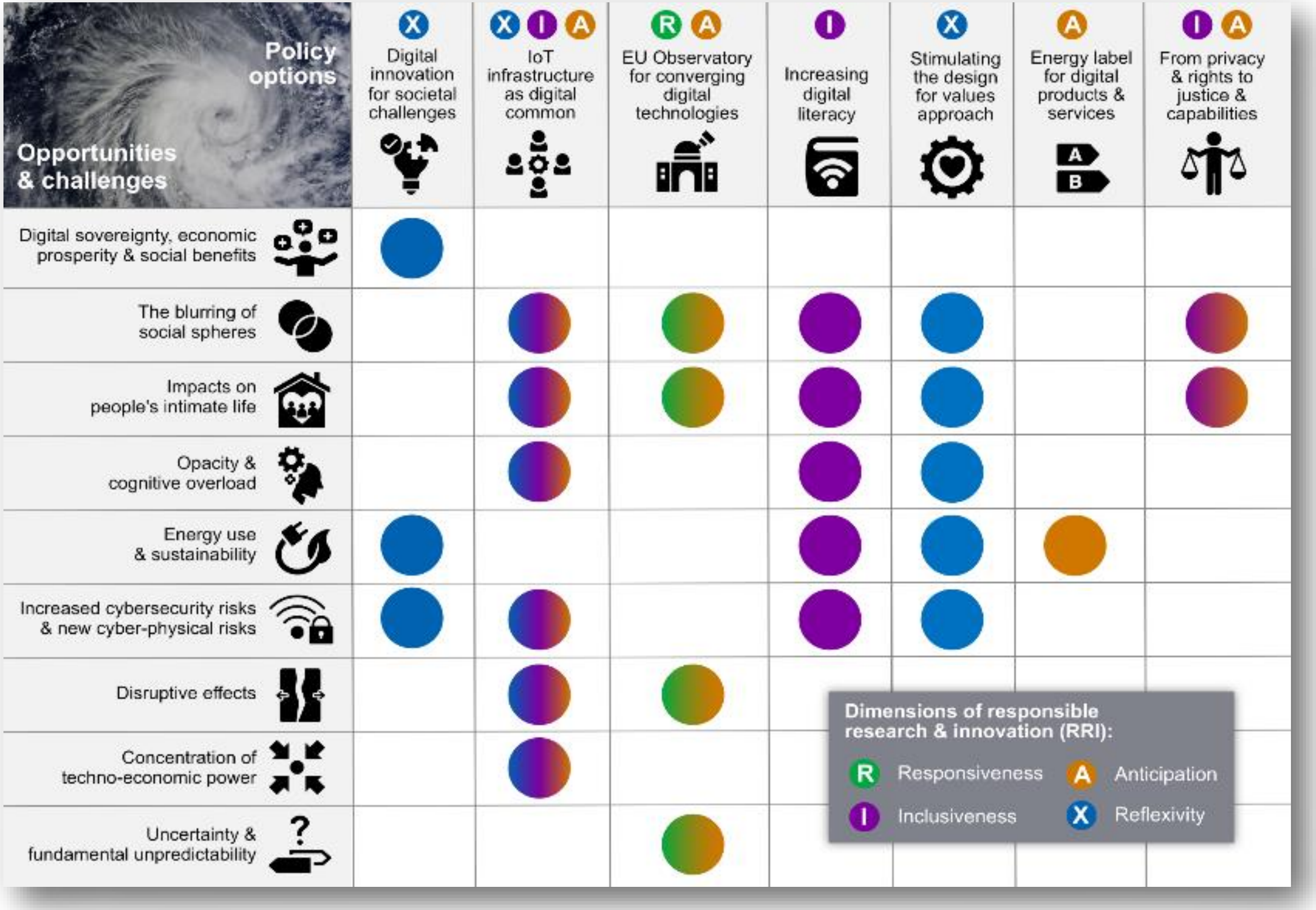
Inspired by these dimensions, this study puts forward the following policy options in response to the challenges identified:

1. **Digital innovation for societal challenges**: Concrete measures could include: (1) giving digital innovation a clearer place in the EU missions in the Horizon Europe research funding scheme, particularly in mission-oriented research; (2) stimulating the creation of European industrial consortia and public-private partnership that can contribute to digital innovation for societal challenges and increasing digital sovereignty; and (3) paying particular attention to how small and medium-sized enterprises (SMEs) and start-ups may contribute to digital innovation for societal challenges, e.g. through incubators and subsidy schemes.

2. **The IoT as a digital common**: As infrastructure technology and an enabler of systems connecting many applications, IoT is a key public good, which allows the production of other goods; it is also a key enabler for ensuring that important social and moral values (like democracy, autonomy, justice and fairness, privacy, sustainability) are respected by specific applications. Safeguarding this role of IoT as a public good and enabler of public values requires coordinated management of all the specific applications of IoT technologies (and other digital infrastructure). This would, at minimum, require a set of public rules for its development, maintenance and use aimed at guaranteeing equal non-discriminatory access, and safeguarding public values. This may be done through public ownership, e.g. by governments, of the basic digital infrastructure, but it is likely that there are also other ownership and institutional structures that allow managing the IoT as a digital common.
3. **EU Observatory for converging digital technologies**: The new ethical and social issues raised by the technological storm are partly unpredictable. Yet some of their effects may be disruptive and be irreversible, and require institutional, regulatory or even conceptual changes to properly deal with them. There is therefore a need for an organisation to play an early warning or early detection function when it comes to new challenges and potential disruptions brought about by the approaching technological storm. This could be done by establishing an EU observatory of converging digital technologies. Such an observatory would be tasked with monitoring relevant developments, carrying out interdisciplinary ELSI research on converging technologies with the aim of early discovery of new issues and challenges and to translate these either into new technological research and innovation, or into new policy, governance or regulatory measures.
4. **Increasing digital literacy**: Digital literacy is important for making digital innovation more inclusive, as inclusiveness would require citizens who are sufficiently well-informed to contribute to a societal dialogue. It will also be helpful in addressing some of the more specific challenges. For example, dealing with opacity and cognitive overload will also require citizens that have a better understanding of digital technologies, including their limitations and threats. Similarly, better awareness will help citizens to play their part in addressing challenges like increased and new cybersecurity risks, energy use and sustainability, and the impact on people's intimate life.
5. **Institutionalisation of design for values**: The design for values approach aims at systematically designing digital technologies for a range of moral and social values. Privacy-by-design and ethics-by-design are already part of the General Data Protection Regulation (GDPR) and of the new EU AI regulation. However, the approach needs to be extended to other values, like democracy and transparency. Concrete measures could include: (1) stimulating training programmes on design for values, for example through the Horizon Europe research funding scheme; (2) Making reporting on design for values obligatory for large technology companies in the EU, as part of the obligatory corporate social responsibility (CSR) reporting; (3) Ensuring that design for values is taken up in standardisation and certification.
6. **Energy label for digital technologies and services**: Energy labels would make consumers (and the public at large) more aware that some digital technologies and services consume considerable amounts of energy, and help them to make more deliberate choices in this respect. Energy labels also create an incentive for the industry to reduce the energy consumption of digital devices and services, and may spur innovation towards lower energy consumption (with similar performance).
7. **From privacy and digital rights to social justice and human capabilities**: Attention to justice is particularly required to deal with the challenge of blurring social spheres. Attention to wellbeing is needed to deal with the challenge of an increased impact on people's intimate life. One barrier to a stronger focus on wellbeing as value might be – certainly for regulation and legislation – that liberal governments have traditionally considered ideas about what constitutes wellbeing or a good, flourishing, life, to be part of the private sphere that should not be intruded upon by the government. Dealing with this barrier may require us to reconsider

our understanding of both wellbeing and freedom, for example in terms of the capability approach, as developed by many scholars, including most prominently philosopher Martha Nussbaum and economist Amartya Sen. The human rights approach and the capability approach are closely related, as both value human dignity and individual freedom highly.

Figure 1: Proposed policy options and their correspondence with identified opportunities and challenges



Policy options / Opportunities & challenges	X Digital innovation for societal challenges	X I A IoT infrastructure as digital common	R A EU Observatory for converging digital technologies	I Increasing digital literacy	X Stimulating the design for values approach	A Energy label for digital products & services	I A From privacy & rights to justice & capabilities
Digital sovereignty, economic prosperity & social benefits	●						
The blurring of social spheres		●	●	●	●		●
Impacts on people's intimate life		●	●	●	●		●
Opacity & cognitive overload		●		●	●		
Energy use & sustainability	●			●	●	●	
Increased cybersecurity risks & new cyber-physical risks	●	●		●	●		
Disruptive effects		●	●				
Concentration of techno-economic power		●					
Uncertainty & fundamental unpredictability			●				

Dimensions of responsible research & innovation (RRI):

R Responsiveness
A Anticipation
I Inclusiveness
X Reflexivity

Table of contents

List of figures

List of tables

1. Introduction

1.1. Background and questions

In the past decade, we have seen rapid developments in digital technology, most prominently in artificial intelligence, augmented reality, blockchain, the internet of things, robotics, virtual reality, and 5G/6G. We have gained many benefits from these technologies, but they have also raised new risks and ethical and legal challenges. These challenges and risks have in Europe been taken up in societal debates, in research projects and in policy-making. Responsible innovation was, for example, made a major focal point in Horizon 2020, the European Commission's programme for research and innovation funding, announced in 2013. In 2016, the European Union adopted the General Data Protection Regulation or GDPR, in response to concerns about privacy. The European Parliament's Special Committee on Artificial Intelligence in a Digital Age is currently working on shaping the EU's policy regarding responsible development and usage of AI.

Are current initiatives and policy frameworks enough, given that these technologies seem to increasingly converge? Supported by the arrival of 5G/6G, we seem to be moving towards an AI-driven internet of robotic things. 'AIoT' (AI and IoT merging) and 'IoRT' (IoT and robotics merging) are new terms reflecting this blending of technologies. Blockchain, augmented reality and virtual reality add even more technological options to the mix. In the business world, this technological convergence is also being discussed under the heading of a 'fourth industrial revolution' (French et al. 2021; Schwab 2016), which indicates that radical and far-reaching changes are expected to occur in the coming decade(s). Smart bodies (Boddington 2021), smart homes, smart industries, smart cities and smart governments lie ahead, with the promise of many benefits and opportunities. However, unprecedented amounts of personal data will be collected, and digital technologies will more than ever affect our lives in intimate ways.

Against this background, this study is an attempt to formulate a provisional answer to the following main questions:

1. Can we expect **new** societal and ethical challenges as a result of this convergence?
2. If so, what policy options should be considered?

In short, are we ready for this future, or might we be surprised by a perfect technological storm?

1.2. Methodology

To explore these questions, we combined a number of research methods:

1. Firstly, we held **semi-structured interviews** with experts (see Appendix 1 for a list of people interviewed). Experts were identified through the network of the authors, and by asking them to recommend further experts.
2. Secondly, we performed a **literature review** (see the reference list). Some initial key publications were identified by asking the experts interviewed for recommendations and through a search of academic databases for papers discussing a combination of previously mentioned technologies. Additional publications were identified by the technique of 'snowballing' and by using ResearchRabbit, an AI powered online tool to explore research papers.
3. Thirdly, we used **topic modelling**, a text-mining method to trace latent concepts (i.e. ideas that are not explicitly mentioned) in bodies of text. The method was applied to trace relevant social and moral values in: (a) the techno-scientific literature, (b) popular media, (c) the relevant ethical literature, and (d) policy/legal documents. Comparison of the results in these bodies of literature may indicate possible policy gaps that may need to be addressed.

1.3. Scope, approach and limitations

1.3.1. Comprehensive overview

These questions are rather broad, and difficult to answer. Firstly, there are many ways in which advances in technological areas such as AI, blockchain, IoT, robotics, 5G/6G, virtual or augmented realities may converge into new technological applications. Emerging applications will moreover affect each and every societal domain: business, healthcare, education, recreation, family life, governance, and so on. Secondly, decades of experience with technology assessment have shown us that the precise trajectory of technological development is notoriously hard to predict. Of course, this should not stop policy-makers and societal stakeholders from exploring possible scenarios, and a range of methods have been developed in the fields of technology assessment and future studies to aid that process.[2] However, an extensive and complete inventory of emerging applications and the ethical issues that they raise would become a very time and resource intensive process – which was beyond the scope of this study.[3]

This Panel for the Future of Science and Technology (STOA) study can thus only provide a first exploration of the topic. To keep a manageable scope for our exploration, we have chosen to focus on the general features that many of these blended technology applications are expected to display – such as their being hybrid, invisible and intimate (Section 5.2) – rather than on specific application types or domains. The challenges and policy options that we subsequently discuss, will also be formulated at a rather high level of abstraction. We hope that they will stimulate more detailed studies and further dialogue, in which more specific applications and application contexts will be addressed.

1.3.2. Responsible innovation

This study largely focuses on the ethical and societal challenges that the merger of digital technologies raises, and less on the opportunities that they also provide to address important societal problems and improve the life of European citizens in numerous ways. The large number of ethical challenges discussed in this study might give the impression that it would be better to keep these technologies at bay and refrain from implementing and using them. Although that may be the case for specific applications in specific contexts, that is certainly not a general claim that we support. If it is even possible to call a halt to any of these technological developments, it would also mean missing out on a lot of potential benefits.

What is possible however is to steer these developments in a direction that is aligned with values such as privacy, justice, sustainability and wellbeing. To do that, technology developers, policy-makers and stakeholders have to join forces to realise responsible innovations, which is where this study aims to contribute. We will use **responsible research and innovation (RRI)** as the overarching framework for developing policy options. As was one of the overarching themes of the EU Horizon 2020 programme, RRI is an 'on-going process of aligning research and innovation to the values, needs and expectations of society' (European Commission 2014), and has also been defined as a 'transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products' (Von Schomberg 2012). This study intends to make a contribution to such responsible innovation (Section 6.1 discusses the approach in more detail).

[2] An interesting recent example of a large-scale technology assessment exercise is the GESDA Science Breakthrough Radar.

[3] An example of such a process is the ETICA project, a research project that made an inventory of the 'Ethical Issues of Emerging ICT Applications', which ran from April 2009 to May 2011 with funding by the European Commission under the 7th Framework Programme. The project had a budget of a little over €1 million.

1.3.3. Existing policy frameworks and regulations

The EU is already doing quite a lot to stimulate (responsible) innovation in the digital domain. A range of existing policies, frameworks, regulations and principles are relevant to the technologies and challenges discussed in this study. These include, but are not limited to:[4]

- Digital markets act (adopted March 2022)
- Data act (proposed February 2022)
- Declaration on European digital rights and principles (proposed January 2022)
- Path to the digital decade (proposed September 2021)
- AI act (proposed April 2021)
- 2030 Digital compass: the European way for the digital decade (proposed March 2021)
- Lisbon Declaration – Digital democracy with a purpose (adopted 2021)
- Digital services act (adopted April 2022)
- Data governance act (proposed November 2020)
- Digital resilience act (proposed September 2020)
- Berlin Declaration on digital society and value-based digital government (2020)
- General Data Protection Regulation/GDPR (adopted in 2018)
- Tallinn Declaration on eGovernment (2017)
- Rome Declaration on Responsible research and innovation in Europe (2014)

We will occasionally refer to some of these documents, such as the GDPR and the Declaration on European digital rights and principles. We also take a number of such documents into account in the topic modelling exercise that we performed (Chapter 4). However, within the scope of this study it was impossible to summarise, analyse, and discuss all of these (or even any of them) in any detail, although it seems worthwhile to do so. It has, for example, been argued that blockchain technology (Section 2.1.5) poses some new challenges that the GDPR is not equipped to deal with (Finck 2019, see also Appendix 4). It has also, to give another example, been argued that emotion recognition technologies (Section 2.2.6) are not adequately dealt with in the proposed AI act (Czarnocki 2021). One suggestion for follow-up studies to the present study is therefore to assess the adequacy and sufficiency of the current policy and regulation landscape in light of the technological developments and resulting challenges as discussed in this study.

1.4. Structure of the study

The study is structured as follows:

- In **Chapter 2** we briefly describe the main technologies that make up the blend. Readers more familiar with the technological side of the topic of this study, may like to skip this chapter.
- A lot has already been written on the societal, legal and ethical challenges of each of these individual technologies. A high-level overview, based on some key literature, is given in **Chapter 3**.
- **Chapter 4** presents the results of the text mining exercise we performed to get more insight into which values surrounding converging technologies are being discussed in different bodies of text (legal documents, news and media, academic ethics and science and technology), with the goal of identifying gaps that may need attention.
- Next, **Chapter 5** gets to the heart of the matter: what new challenges may arise from the above-mentioned technologies blending? After listing the general features that many of these blended technology applications are expected to display, we will discuss the challenges that we have identified through the methods detailed above. We will also discuss some examples of applications to make these challenges more concrete, as well as what we learned from a number of additional interviews with representatives of some key societal stakeholders.

[4] Some further examples of relevant EU documents can be found in Appendix 2, in the description of the "legal dataset" we used for the text mining exercise presented in Chapter 4.

- Finally, in **Chapter 6** we first present a framework of general policy approaches that may be selected and/or combined to address the challenges and gaps discussed in the preceding chapters. Then, we present a number of more specific policy options that might help to address the challenges identified in Chapter 5.

2. What goes into the blend: rapidly developing technologies[5]

The technologies that play a role in the new technological storm may be seen as part of a layered structure. At the backbone is IoT as an infrastructure technology that allows building systems of systems that are highly interconnected. Communication within this infrastructure is enabled by technologies like 5G/6G. On top of this infrastructure, there are many applications in different social domains that are made possible by general purpose technologies like AI, blockchain, bio-nano devices, AR/VR and robotics. In this chapter we briefly describe these main technologies (Section 2.1), as well as some examples of more specific technological developments within or related to these domains (Section 2.2). The latter section is merely illustrative of new developments and does not aim at completeness. More new technologies could have been mentioned, and the ones included should be seen as examples that make the technological domains addressed in this study more concrete, and illustrate the far-reaching changes that we can expect to see in the decade to come.

2.1. Main enabling technologies

2.1.1. 5G / 6G

The introduction of 5th generation telecommunication networks or 5G is an important precondition for supporting the further development of many of the technologies covered in this study, and for making the development of the AIoT (intelligent internet of things), IoRT (internet of robotic things) and IoBNT (internet of bio-nano things) systems possible. Compared to 4G it is faster, has a lower latency, can have more devices connected at the same time, and is more efficient in a myriad of ways. This allows for a much greater amount of traffic to occur. 6G is likely to be a massive improvement over 5G in every way, seeing a 50x improvement in data rate as well as a 10x improvement in connectivity density and a 10-100x improvement in latency (Ylianttila et al. 2020). This is important as '*the unprecedented proliferation of smart devices and the rapid expansion of IoT networks, 5G cannot completely meet the rising technical criteria, e.g., autonomous, ultra-large-scale, highly dynamic and fully intelligent services*' (Nguyen et al. 2021). With these developments, a whole array of new possibilities arises, as the transfer of data becomes easier and the rise of more data-intensive services feasible. The expectation is that with 6G, the determination of geo-location can be done with an accuracy of less than a centimetre.[6] Due to the small wavelength used, 6G sensing networks will be able to create a picture of their environment in greater detail (expert interview).

2.1.2. Artificial intelligence (AI)

Artificial Intelligence can refer to the science and engineering of making intelligent machines and/or software, as well as to the products created by this field. 'Intelligence' is notoriously hard to define. (The European Commission's High-Level Expert Group on Artificial Intelligence 2018) defines AI as follows:

> '*Artificial intelligence (AI) refers to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to predefined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behavior by analyzing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as*

[5] A first version of this chapter was written by Dyami van Kooten Passaro.

[6] Talk by prof. Vincent Poor at the STOA event on Edge computing, 6G and satellite communications (1 December 2021)

> *machine learning*[7] *(of which deep learning*[8] *and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).*'

Note that this definition lists robotics as a sub-discipline of AI, while arguably robotics and AI are different fields that are increasingly but not completely overlapping (see also Section 2.1.5. Robotics). A further distinction that is important in the field of AI, is that between narrow (or weak) and general (or strong) AI: '*A general* [or strong] *AI system is intended to be a system that can perform most activities that humans can do. Narrow* [or weak] *AI systems are instead systems that can perform one or few specific tasks*' (The European Commission's High-Level Expert Group on Artificial Intelligence 2018). All current AI systems are instances of narrow AI. In addition to sensing, modelling, planning and action, '*current AI applications also include perception, text analysis, natural language processing (NLP), logical reasoning, game-playing, decision support systems, data analytics,* [and] *predictive analytics* [...].' (Müller 2021)

2.1.3. Internet of things (IoT)

The traditional internet connects communication devices to each other. The internet of things broadens this to all kinds of devices. In general, an IoT system consists of the following five components (Miraz et al. 2018):

1. '*Sensors*: which are used to mainly collect and transduce the data;
2. *Receiver*: to facilitate collecting the message sent by the computing nodes or other associated devices;
3. *Computing Node*: a processor for the data and information, received from a sensor;
4. *Actuator*: based on the decision taken by the Computing Node, processing the information received from the sensor and/or from the Internet, then triggering the associated device to perform a function;
5. *Device*: to perform the desired task as and when triggered'

An example of this are smart watches that through biological markers sense when the best time is to wake you up and let your smart home know to open the shutters and make you a cup of coffee. All of these devices are connected to the internet -and with each other- through the IoT architecture, allowing the creation of a system that seamlessly communicates and caters to your needs based on data. In this example the smart watch acts as a sensor that gathers data and triggers other devices to act in a certain way based on the data it sends. The system thus senses a certain state and reacts accordingly.

2.1.4. Augmented and virtual reality

Augmented reality is a reality that is enhanced in some way through computer generated perceptual information, like an app showing you an altered version of reality through your camera, taking information from the real world (camera footage) and augmenting it with additional computational elements. Augmented reality is being used in a wide array of sectors, amongst others by the military (to more easily identify friends and foes in the surrounding environment) and by companies such as Disney (to make colouring book images appear 3D to users of their app) (Abramovich 2021). Virtual Reality takes this one step further: Instead of altering or enhancing existing reality in some way, virtual reality creates a whole new reality to step into for the user (Greengard 2019). This usually requires a full

[7] With machine learning, AI can learn from processing data, so that it becomes more adept at dealing with similar data in the future.

[8] In deep learning the machine uses neural networks to train itself to perform a task without human supervision being a prerequisite, by using vast amounts of data to spot patterns. Raw data by itself is enough for the machine to start making detections or classifications based on the input that it gets (LeCun, Bengio, and Hinton 2015).

VR headset, as simply using a smartphone screen often doesn't do enough to make the user feel connected to the reality they are trying to enter. One avenue of research within this field is how to make the VR experience more immersive by recognizing and responding to the emotions of a VR user on the basis of biometric data (Nam et al. 2019). Another interesting development that we may see in the future, is the integration of holographic images of real people in augmented or virtual reality (expert interview).

2.1.5. Blockchain

Blockchains are distributed digital ledgers that allow a community of users to store information with digital signatures, often without central control. The chain is made up of many individual blocks, which are cryptographically linked together after the newest block is verified and added to the chain. This makes it very difficult to tamper with the chain and makes any alterations to the chain permanent. Because of this, once a digital signature is ascribed in the chain, it is (almost) impossible to falsify it. Thus, you get certainty of the validity of what has been written and it allows for trustless transactions (Monrat, Schelen, and Andersson 2019). This can be used in many different fields, most famously in the financial market in the form of Bitcoin, although there are other uses, such as IBM using it for certainty of the original port of containers, keeping track of where it has gone (Groenfeldt 2017). It is also used to prove the digital ownership of goods and even for so-called smart contracts (software deployed on the blockchain and executed by computers running that blockchain) (Yaga et al. 2018). The certainty that the ledger can provide, means that a trusted middle-man for transactions is no longer needed. This could have a big effect on certain industries and practices within those industries.

2.1.6. Robotics

In popular imagination robots often have a humanoid shape, but that only applies to a small part of all robots that have been produced and deployed so far. In essence, robots can be defined as 'physical machines that move.' Industrial machines that 'follow completely defined scripts with minimal sensory input and no learning or reasoning' are robots according to this definition. Yet we typically think of robots as machines that not only move, but also have some degree of autonomy in their actions due to the integration of AI (Müller 2021). These are also the types of robots that are of interest to this study. In the past years several scholars have started to explore the concept of an 'Internet of Robotic Things' or IoRT (Ray 2016; Kamilaris and Botteghi 2020; Vermesan et al. 2020; Villa et al. 2021). One could even argue that ordinary objects, such as lamps and bookshelves, can become 'robotic' once they become integrated in the AIoT in certain ways (Loke 2021). Robotics is not only converging with AI and IoT, but also with augmented reality. The latter is amongst others used to improve motion planning/control and human-robot interactions (Makhataeva and Varol 2020).

2.1.7. Bio- and nanotechnology

In addition to an 'Internet of Robotic Things', we may expect the development of an 'Internet of Nano Things' or IoNT (Akhtar et al. 2020; Cruz Alvarado and Bazán 2018; Miraz et al. 2018; Dressler and Fischer 2015). Nano devices may range from 1 to 100 nm - which is at least 1000 times smaller than a human hair - and could be nano sensors, nano-bots, nano cameras, nano phones or other nano things. The integration of nanotechnology with synthetic biology creates new options; Nanosensors and communication between nanodevices may not only be based on physical (mechanical, acoustical, thermal and radiation, optical, and magnetic) or chemical (atomic and molecular energies) mechanisms, but also on biological mechanisms (antibody and antigen interaction, DNA interaction, enzymatic interaction). Scholars and engineers are therefore also, more specifically, working on an 'Internet of Bio-Nano Things' or IoBNT (Akhtar et al. 2020; Kuscu and Unluturk 2021). Although at the moment Io(B)NT systems *'are not as well-developed as their IoT counterparts',* they have certain advantages:

> '*their ability to gather data using such small sensor points makes them useful for applications that are not compatible with other (bulkier) sensor networks.* [...] *Because they are so small, nanosensors can*

> *collect information from millions of different points. External devices can then integrate the data to generate incredibly detailed maps showing the slightest changes in light, vibration, electrical currents, magnetic fields, chemical concentrations and other environmental conditions. The Io[B]NT could provide much more detailed, inexpensive, and up-to-date pictures of our cities, homes, factories, even our bodies.* [...] *on-body nano-sensors could provide electrocardiographic and other vital signals, while environmental nano-sensors could collect information about pathogens and allergens in a given area*' (Akhtar et al. 2020, p.134-136)..

These sensors may be used not only *on* the human body, but increasingly also *in* the human body.

2.1.8. Quantum computing

Quantum computing refers to computers that use qubits as opposed to the original bits to do their calculations. Comparing them to classical computers is, so Ghose (2018) claims, akin to comparing light bulbs and candles: They can fulfil the same purpose, but one cannot get a lightbulb from perfecting candle technology. The inner workings of a quantum computer are outside the scope of this rapport. What is important, is that they will be able to do certain types of calculations orders of magnitude faster than classical computers. This will have serious implications for the field of cryptography, where the standard of encryption is based on the presumption that certain mathematical problems are too time consuming for computers to deal with, and thus the encryption is safe. A codebreaker with a quantum computer will, however, be able to make short work of most classical encryption algorithms (Williams 2011). The flip side of that coin is that quantum computing could also be used to create safer encryption through quantum entanglement, which would allow the users of data to be able to tell whether their information was intercepted by a third party. This means that if keys are being shared (as they are in standard cryptography), the users can decide to get rid of a key that has been potentially intercepted by a codebreaker, thus making it a very secure way of transmitting data (Giles 2019). Given how difficult it is to create quantum computers or networks, it is highly unlikely that the majority of data will start to travel on a quantum internet - and certainly not any time soon. In the remainder of this study we will not take quantum computing in consideration.

2.2 Some specific technological developments

2.2.1. Reinforcement learning (AI)

Reinforcement learning differs from traditional forms of machine learning; It does not need labeled data in the same way that supervised machine learning does, while unlike in traditional unsupervised learning, constraints are put in place to direct the learning process (Carew 2021). In reinforcement learning, the algorithm carries out random actions based on its options within a concrete environment, and then either collects a predefined reward or does not. It thus learns to associate certain actions with a reward while punishing actions that do not lead to the reward. The actions of the AI are random at first, but rewards and punishments can change the distribution of action-probability, so that the combination of actions that led to the reward are more likely to reoccur in the future. Think of the algorithm playing pong, for example: A game in which the player can only move up or down and has to score a point by getting the ball past his opponent by reflecting it off itself. It can be rewarded when it wins a point and punished when it loses one. This reward can then be fed into its future decision to act in certain ways (Karpathy 2016). However, defining the rewards is difficult to do - and not a morally neutral exercise! - if it concerns an AI that has to operate in complex socio-technical situations - such as autonomous vehicles (Gilbert 2021).

2.2.2. Distributed / federated learning (AI)

To train AI models, a lot of data is needed. Multiple machines may sometimes be needed for this, as single machines don't have enough computing power. This requires so-called 'distributed machine learning', a form of machine learning in which local nodes keep their own data, but work together to aggregate a central model. The individual nodes only share their updates to the central model, not the original data on which these updates are based. A specific form of distributed machine learning is federated learning, which works as follows (McMahan and Ramage 2017):

1. An initial model is sent to a large number of local machines or clients that have data;
2. The clients train the model based solely on their own data. This is, however, not enough for a robust model;
3. The clients therefore send their trained model back to the server (but not the actual data that was used to train the model);
4. The client models are aggregated and combined into a new and improved central model;
5. Using this new and improved model, steps 1 to 4 are repeated as often as is necessary to get a robust result.

Distributed / federated learning may contribute to protecting privacy, as it makes it possible to learn from data without centrally collecting and storing it (Nguyen et al. 2021). At the same time, security risks increase, as there are more access points to data. Another advantage of the data not being sent to a central point, is that this may reduce the amount of energy needed (expert interview).

2.2.3. Edge computing (AI and IoT)

Edge computing '*refers to the enabling technologies allowing computation to be performed at the edge of the network, on downstream data on behalf of cloud services and upstream data on behalf of IoT services*' (Shi et al. 2016). Due to the increase in data provided by the rise of IoT, solutions are needed to limit the amount of central computing to be done by cloud networks. By employing edge computing, data can be analyzed close to its source without the need of sending all of that data to the cloud or a central system. Devices on the edge can then act based on this data, as well as share the analyzed data with a central system if desired. It thus allows the data to stay closer to its source, providing an intermediary between the central system/server and the sensors that collect data. Edge intelligence can be considered to be one of the fundamental supporting technologies for applications in a 6G-supported internet of things (Nguyen et al. 2021).

2.2.4. Digital twins (AI, IoT and VR)

A digital twin is a virtual representation of a physical thing, based on real-time data from sensors placed on the physical thing. A highly complex virtual model is created, which is intended to be the exact counterpart of the physical thing. The model can then provide real time information on how the actual physical thing is doing, while at the same time allowing users to test different scenarios or predict the future by running the model that has been created. The digital twin can also be linked up to other systems for a bigger, more complete, model (Armstrong 2020). In the future, digital twins could also be integrated in virtual reality applications and digital environments like the metaverse, where digital twins might even be made interactive (George 2021). They are also discussed as a means to improve health care. One idea is that a digital twin of a patient can be created that allows physicians to gain a more thorough understanding of the patient and test how certain care interventions could play out. The digital twin could be based on historical data on the patient, or real-time data coming from (nano) devices on or inside the body. Another idea is to create a digital twin of the hospital, allowing for more strategic planning of actions to take (Croatti et al. 2020).

2.2.5. Drone technology (IoT, AI and robotics)

Drones or 'unmanned aerial vehicles' (UAVs) have become an important research topic in the field of robotics (Suzuki 2018). UAVs could, for example, collect data on agricultural processes or environmental conditions, either through video surveillance or measurements of substance dosages like CO2 levels. They could also be utilized for the monitoring and management of traffic by collecting information about congestion areas, causes of congestion, vehicle volume etc. Although this can be done by static cameras to a certain extent, the amount and specificity of information able to be collected by a moving drone is much greater. This is just one example of the role that UAVs could play in the realization of 'smart cities' (see e.g. Alsamhi et al. 2019; Mohamed et al. 2020). In the future UAVs might also be used to transport goods - from pizzas to emergency medical supplies - in hard-to-reach areas or when time is of the essence. Advanced drone technology might also allow for the easier transportation of persons, for example by UAV ambulances and UAV taxis. However, at the moment such promises '*often appear oversimplified or are yet lacking a scientific validation*' (Kellermann, Biehle, and Fischer 2020). Increasingly, we can expect drones to operate as part of an 'intelligent swarm' of drones; '*The dynamic uncertain environment and complex tasks determine that the unmanned aerial vehicle (UAV) system is bound to develop towards clustering, autonomy, and intelligence*' (Zhou, Rao, and Wang 2020).

2.2.6. Biometric technology (IoT and AI)

Biometric technology identifies people based on certain features that they possess. These could be physical - 'such as a fingerprint, face, iris, blood vessel pattern at the back of the eye, vascular patterns, DNA, and hand or palm scan recognition' - or behavioural - 'such as signature/handwriting, gait, voice, gesture, and keystroke dynamics" (Fares Al Mashagba 2016). Advances in AI, in particular deep learning techniques, have been crucial for developing new and better biometric recognition technologies (Almabdy and Elrefaei 2021; Mehraj and Mir 2018). Progress has also been made through 'biometric fusion', where several types of data are combined to get better results (Singh, Singh, and Ross 2019). Biometric technologies used to be applied mainly to identify people for purposes in security and law enforcement. However, commercial and civil applications are on the rise.[9] In such applications biometric data is not only captured for identification purposes, but also to track changes in individuals (e.g. personal fitness devices tracking biometric data) and increase knowledge about certain categories of people. Table 1 summarizes the differences between first- and second-generation biometric technologies (North-Samardzic 2020).

[9] See e.g. De Keyser et al. (2021) for examples of commercial applications.

Table 1: Comparison of first- and second-generation biometric technologies

	First generation	Second generation
Purpose	Who are you?	How are you?
Application	Identity management and authentication	Safety and behavioural assessment
Context	Government and security	Civil and private sector
Level of analysis	Individual	Groups
Primary ethical concern	Privacy risks	Discrimination power
Example	Fingerprint or face recognition for law enforcement or consumer device identity management	Voice recognition to understand individual affect and face recognition to assess group demographic characteristics

Source: Nort-Samardzic (2020)

One important area of biometric research is emotion recognition on the basis of various types of biometric data, such as eye movement (Lim, Mountstephens, and Teo 2020), body movement (Ahmed, Bari, and Gavrilova 2020), speech (Schuller 2018), facial expressions (Ko 2018), and EEG signals (Suhaimi, Mountstephens, and Teo 2020). Emotion recognition could be used for '*humanizing the internet of things (IoT) and affective computing systems*' (Dzedzickis, Kaklauskas, and Bucinskas 2020) and beneficially '*applied in many areas such as safe driving, health care and social security*' (Shu et al. 2018), but could of course also be used for manipulation and surveillance.

2.2.7. Non-fungible tokens (blockchain)

Non-Fungible Tokens can be thought of as certificates of ownership of a token that is non-fungible ascribed to the blockchain. 'Fungible' meaning that the item is interchangeable with another (like sacks of rice are identical to you). NFTs make it possible to have ownership of something digitally and you can be assured that you are the sole owner, as your token is not fungible (read: replaceable). This is currently mostly being used for digital items, but that doesn't have to be that way. You could, for example, own a token that confirms ownership over a physical item like a house (Wang et al. 2021). Another example of an application is NFT games, where some players try to make a living by selling the NFT's earned. Although NFTs have a wide range of possible use cases, they also have their drawbacks; in some blockchain environments, adding to the blockchain takes a lot of energy due to the computational power required to add to the ledger, with current NFTs being mostly minted on the Ethereum blockchain where '*...with the current fee mechanism, spending one dollar on transaction fees corresponds to emitting at least the equivalent of 1.151 kilograms of CO2*' (Marro and Donno 2022). Furthermore, NFTs can currently easily be used in fraudulent schemes where the original physical item is digitized by someone other than the original, real owner, and put up on a marketplace. In this case, the NFT loses all connection to the physical object and can be considered fraudulent (Brock 2022).

2.2.8. Satellite IoT networks (6G and IoT)

In the 6G era, satellite communications may become integrated into wireless networks to make massive IoT coverage possible. It will enable IoT applications in remote places where traditional networks cannot operate or do not function well, such as seas and deserts. Some researchers are already talking about the 'Satellite internet of things' or SIoT (Nguyen et al. 2021). Drones or Unmanned Aerial Vehicles (UAVs) are also expected to be integrated in these networks, '*providing an intermediate network layer between ground networks and space ones*' (Mishra et al. 2021). One advantage that the Satellite internet of things may bring, is that '*SIoT platforms can enable energy sustainability via the use of aerial IoT devices* *such as UAVs [Unmanned Aerial Vehicles] and balloons, with renewable sources from space that may not be available at the ground-based stations*' (Nguyen et al. 2021).

2.2.9. Web 3.0 (5G and blockchain)

Web 3.0 is a term used to refer to an evolution of the world wide web. Web 1.0 was focused on enabling people to consume content, mostly through reading webpages and looking at pictures. This is what most early webpages were like. Web 2.0 allowed users to generate their own content and made it possible for websites to become a 2-way street of interaction between different parties. Social media is a prime example of this. With the arrival of web 3.0, the web would become almost completely decentralized and open. This is made possible by the exponential increase in computing power of devices. Whereas in web 2.0 each website 'lives' on a particular server, this same content could in web 3.0 be distributed among the web, meaning that there is no central entity that has full control over the site (Investopedia Team 2022). This makes it possible to give users much more control over their data and how they participate on the internet (Alabdulwahhab 2018). Web 3.0 would most likely make use of additional technologies such as large-scale peer-to-peer networks and blockchain technology to make sure that no intermediaries are needed (as is the case now for the server host) as well as 5G.

2.2.10. Brain-machine interfaces (neurotechnology and AI)

Brain-machine interfaces provide a direct link between the brain and an outside machine; they are currently mostly used in the healthcare industry to give people access to brain-controlled prosthetics, but have a wide range of (potential) applications outside of that field (Coates McCall et al. 2019). By using either invasive or non-invasive methods (Durham 2019), brain activation can be picked up and translated into action using algorithms that decode the brain activity and produce an output based on that activity using a pre-established brain profile as a reference point; this allows, for example, algorithms to reconstruct images that a person sees based solely on their brain activation (Shen et al. 2019). A lot of potential applications are opened up when this technology is linked with Artificial Intelligence, as sufficient data-driven brain-profiles could potentially be used to create (crude) readouts of what is going through a person's mind through analysing their brain patterns. Having the ability to interact with the brain in this way also opens the door to feedback loops in which the brain is stimulated and then analysed to recalibrate the stimulation and get to a certain desired result (Cohen 2019). The technology raises questions of security, privacy and personal identity linked to it, as it interacts and potentially changes the brain (OPTIC 2019).

3. What we already know: ethical challenges and values at stake

The convergence of the technologies described in the previous chapter has an enormous potential for new applications with great benefits for European citizens. In this study the emphasis is, however, on the challenges and potential negative impacts that need to be addressed in order to make the most of these technological developments. A lot has already been written on the overall societal, legal and ethical challenges of each of the individual technologies. Moreover, these are general-purpose technologies that may be applied in many different societal domains, which has led to large bodies of literature on specific ethical challenges in specific application domains (like health care). An extensive and systematic literature review was unfortunately outside the scope of this exploratory study. For this chapter we have instead relied on a couple of key sources per technology that give a high-level overview of the main concerns.

3.1. 5G/6G

An interactive infographic that the European Parliamentary Research Service has published at https://map.sciencemediahub.eu/5g[10] gives an overview of what 5G could bring. Most expected impacts - whether positive or negative - arise because of 5G *in combination with* one or more of the other technologies discussed in this study. Just a few of them are directly connected to the features of 5G itself.

Energy usage is one such concern. With respect to energy, there are actually certain advantages to 5G, namely that '*a sensor that transmits data seldom can do it sporadically in a 5G network, while in 4G environments it has to be transmitting constantly. Likewise, while 5G's power consumption will require more base stations per square kilometer, these will only need as much power as required - whereas predecessor networks are always 'on.*' This is however not expected to be enough to compensate for new and increased usages of communication networks. Several of the experts that we spoke to expressed concern about a massive increase in energy usage.

The transition to 5G also raises questions about its health impact and environmental impact. As for **environmental impact**: 5G networks will partly rely on new frequencies that are not very commonly found in nature. The concern is that the radio-frequency (RF) and electromagnetic fields (EMF) used could do damage to wildlife, as it partly penetrates biological tissue. A recent study commissioned by STOA concluded that it may indeed cause the internal temperature in organisms or cells to increase. However, this study did not draw any conclusions yet about the (long-term) consequences for wildlife and called for more research and systematic monitoring (Thielens 2021).

The concerns for human **health** also arise because of the fact that 5G will rely on wavelengths that were previously not used on such a massive scale as would be the case in a post-5G world. Therefore, there have not been that many studies yet on their impact on health. Another recent STOA study looked at the evidence that is so far available from in vivo animal studies and human epidemiological studies, which investigate if there are any effects on cancer development and fertility. This study concludes that 'the sources of RF emissions that seem at present to pose the greatest threat are mobile phones', and recommends as a precautionary measure to opt for types of mobile devices that are safer. As there is still a lot of uncertainty about the effects of 5G on human health, this study as well recommends more research (Belpoggi 2021).

In addition, with the future arrival of 6G the following challenges are expected to become even more salient than they already are for 5G networks (Ylianttila et al. 2020):

[10] Accessed 22 January 2022

- **Security and safety:** '*The volume of new IoT devices introduced into 6G network will increase 10x from 10 billion scale of 5G networks to 100 billion scale in 6G. As a result of such deployment and use of 6G, the dependence of the economy and societies on IT and the networks will deepen. Safety will depend on IT and the networks. The development of AI blurs the line between reality and fake content and helps to create ever more intelligent attacks. The role of IT and the networks in national security keeps rising – a continuation of what we see in 5G.*'
- **Privacy**: '*5G is still largely device / network specific, 6G envisages far more immersive engagement with the network. It is now the subject of ongoing discussion in the standards world. There is currently no way to unambiguously determine when linked, deidentified datasets cross the threshold to become personally identifiable. This is a major, unaddressed problem for many digital technologies in different sectors, such as in Smart Healthcare, Industrial Automation, and Smart Transportation. Courts in different parts of the world are making decisions about whether privacy is being infringed without formal measures of the level of personal information, while companies are seeking new ways to exploit private data to create new business revenues.*

3.2. Artificial intelligence and robotics

The ethical challenges of artificial intelligence are being discussed in many publications and are by now quite well known. An overview of issues is given in dedicated entries in the Internet Encyclopedia of Philosophy (Gordon and Nyholm 2021) and the Stanford Encyclopedia of Philosophy (Müller 2021), written by ethicists from academia. Both entries discuss some issues that are arguably peripheral to the purposes of this study, either because they are rather abstract (the moral and legal status of intelligent machines) or because they will probably not become urgent for the next couple of decades (the risk of the so-called 'singularity', a moment where AI would get control over us rather than the other way around). The main short-term and concrete ethical and societal challenges discussed in these entries are:

- **Privacy, surveillance and manipulation of behaviour**: AI makes data collection and analysis easier and more rewarding than ever before in human history. The type and amount of data that we leave behind surfing the internet and using (free) online applications, makes it possible for AI to follow us closely and know us in very personal ways. And it is big tech companies rather than we ourselves who are in control of our data. AI systems are often designed to nudge, manipulate and deceive people, based on their data profile, to optimize business results. They also make it increasingly easy to create 'deep fakes' that can also be used for all kinds of manipulation (Müller 2021).
- **Transparency and explainability**: Understanding how decisions come about is a fundamental challenge for AI based on machine learning, which is nowadays a large part of the AI systems. Even the programmers often don't know on which patterns the outcome of the AI's data analyses are based. This is a huge problem for a democratic society, where we want to be able to hold people accountable for the decisions that they make, which includes being able to justify those decisions. Making AI 'explainable' is therefore a major challenge for the AI research community (Müller 2021). Opacity may however also have other causes, such as trade secrets and the difficulty for lay people to understand AI (Gordon and Nyholm 2021).
- **Bias and discrimination**: AI systems have raised serious concerns about bias and discrimination, for example with regard to race or gender. There are several reasons why the decisions by AI systems may be biased. Firstly, the data sets fed to the AI system may be of low quality ('garbage in, garbage out'). Secondly, AI developers may be insufficiently aware of their own biases and societal concerns, resulting in biased algorithms for processing data. Finally, historical data records may steer future action into certain directions that perpetuate undesirable existing patterns (Gordon and Nyholm 2021).

For AI-driven robotic devices, more specifically, two important concerns are:

- **Meaningful human control**: Robotic systems that can - to a larger or smaller degree - take autonomous decisions and have the ability to physically harm people, have raised discussions about (a) how ethical factors can and should be taken into account into those decisions and about (b) safeguarding the ability of humans to maintain meaningful human control and take responsibility for the actions of the system. The most prominent examples of such systems are autonomous vehicles and weapon systems (Müller 2021; Gordon and Nyholm 2021).
- **Employment and the future of work**: New possibilities to intelligently automate more and more tasks and types of work, have raised a lot of discussion. One main concern is the possibility of mass unemployment, an issue of justice and fair distribution (Müller 2021). Another topic of discussion is the changing nature of work and its implications for our capabilities to realise a good, meaningful life (Gordon and Nyholm 2021).

Of course, the ethical and societal challenges of AI are not just being discussed by academic ethicists, but also by companies, non-profit organizations, governments and other actors. This has led to hundreds of documents outlining principles, guidelines and policies for realizing responsible AI. Subsequently, research papers have appeared that analyse and compare a smaller or larger subset of these documents to discover patterns, gaps, topics of consensus and differences and/or develop an overarching framework for AI ethics (Cath et al. 2018; Daly et al. 2019; Dutton, Barron, and Boskovic 2018; Fjeld et al. 2020; Floridi and Cowls 2019; Hagendorff 2020; Jobin, Ienca, and Vayena 2019; Schiff et al. 2021; Zeng, Lu, and Huangfu 2018). Some of these publications present an overview of the main ethical values or principles discussed in the documents that were analysed, which we have brought together in table 2.

So, what conclusions can we draw from table 2, other than that AI clearly raises a lot of different ethical concerns? Perhaps that (1) accountability, (2) autonomy / freedom, (3) justice / fairness, (4) privacy, (5) safety / non-maleficence and (6) transparency / openness are the most important concerns. However, we should remain open to the possibility that this prioritization may change as a result of technological developments, or that new values - not included in table 2 - start to attract attention. For example, several of the experts that we spoke to for this study consider **sustainability** to be one of the main future challenges of the convergence of technologies discussed in this study. More in particular, they are concerned about an expected exponential growth in energy usage due to the massive application of AI, IoT devices and related technologies (Section 5.3.5). Yet it seems that this is currently not receiving a lot of attention in AI documents.

Table 2: Values / principles discussed in AI documents according to 6 overview papers

VALUE / PRINCIPLE (# times included)	Fjeld et al (2020)[11]	Floridi and Cowls (2019)[12]	Jobin et al (2019)	Hagendorf (2020)[13]	Schiff et al (2021)	Zeng et al (2018)[14]
Accountability (5)	x	x		x	x	x
Human autonomy, freedom (5)		x	x	x	x	x
Dignity (2)			x			x
Explainability, interpretability (4)	x	x		x	x	
Non-discrimination, justice, fairness (6)	x	x	x	x	x	x
Meaningful human control (3)	x			x	x	
Privacy (5)	x		x	x	x	x
Public participation (1)					x	
Responsibility (3)	x		x		x	
Non-maleficence, safety (6)	x	x	x	x	x	x
(Cyber) security (3)	x			x		x
Solidarity, social cohesion (2)			x	x		
Sustainability (2)			x	x		
Transparency, openness (5)	x		x	x	x	x
Trust (2)			x		x	
Wellbeing, beneficence (4)	x	x	x			x

An example of a value not surfacing at all in table 2, is **truth** - or perhaps rather **veracity**, as there are fundamental problems with knowing (for sure) what is true. This value goes to the heart of the societal challenge of dealing with 'deep fakes', which has become more salient now that AI is making it increasingly easy to create deep fakes (see e.g. Van Huijstee et al. 2021). Another much discussed

[11] Fjeld et al. (2020) put 'transparency and explainability' respectively 'safety and security' down as a single principle. These have been separated here in four principles instead of two.

[12] Floridi and Cowls (2019) put accountability and explainability (which they call 'intelligibility') together under a single principle, which they call 'explicability'. These have been separated here.

[13] Hagendorf (2020) lists 'safety, cybersecurity' as a single principle, which has been split here.

[14] Zeng et al. (2018) put human dignity, wellbeing / beneficence and freedom together under a simple principle of 'humanity', these have been separated here.

societal challenge related to AI, is that of 'filter bubbles' or 'echo chambers', the phenomenon that on (social) media sites people would increasingly be exposed to just (the views of) like-minded people.[15] This could arguably lead to a decline in solidarity and social cohesion, which - according to table 2 - at the moment is also one of the topics receiving less attention. Both 'filter bubbles' and 'deep fakes' have been discussed in relation to the value of **democracy** (for a recent literature review see Kuehn and Salter 2020), which is also completely missing from table 2.

Furthermore, if one looks at the documents in more depth and detail, beyond the apparent consensus on key ethical issues, the picture quickly becomes more complex. For example, the analysis by Jobin et al (2019) reveals *'significant semantic and conceptual divergences in both how the [...] ethical principles are interpreted and the specific recommendations or areas of concern derived from each.*' Schiff et al. (2021) conclude that '*NGO and public sector documents reflect more ethical breadth and depth [...], and are generally more similar to each other than to private sector documents.*' Furthermore, '*while the private sector tends to emphasize ethical issues with ostensible technical fixes, such as algorithmic bias and transparency, the NGO sector addresses a wider set of topics, such as accountability and misinformation, and the public sector focuses on unemployment and economic growth.*' Chapter 4 will provide a further analysis, based on our own research, of some differences between different arenas in what is being discussed.

3.3. Internet of things

While ethicists, practitioners and policy-makers have been actively discussing the ethical implications of artificial intelligence in recent years, the implications of the internet of things have received less attention. Most of the ethical challenges raised by the internet of things are not completely new to (computer) ethicists, but certain features of the technology - such as its radically distributed nature (see also Section 4.1) - add new levels of complexity and/or salience to the challenge. These familiar ethical challenges include:

- **Privacy**: IoT devices will be able to collect huge amounts of data. Especially in the case of IoT devices in private homes or in people's bodies, this will include data on intimate aspects of people's lives. This data may be used to deliver a better service and enhance the user experience or to achieve important personal and societal goals such as good health. But it may also be used to profile users or to harm them when the data falls into the wrong hands. Encrypting the data could help, but metadata - such as the moments at which a device was used / data was generated - can still reveal a lot, especially when combined with other data (Allhoff and Henschke 2018).
- **Informed consent**: In the context of other information technologies the principle of informed consent is also important for protecting privacy. But due to the ability of IoT devices to act invisibly on behalf of the user, informed consent becomes even more important. At the same time, it becomes more difficult to realise, as IoT devices - especially those moving towards micro-scale - may disappear from sight and be taken for granted as a background technology that users are only vaguely aware of, if at all (Van den Hoven 2012).
- **Trust**: Whether people have reasons to trust IoT devices depends on the degree to which ethical concerns - such as privacy and informed consent - have been properly taken into account. Since AI is increasingly being integrated in the internet of things, the challenge of trustworthy AI also comes into play here (Allhoff and Henschke 2018). As IoT promises to be highly distributed, dynamic and ubiquitous (everything communicates and interacts; no boundaries, new entities can enter the IoT at all times), establishing trust among entities becomes even more important. At the same time, it also becomes more difficult, as entities will

[15] It should be noted though that it is contested to what degree filter bubbles are actually widespread and/or having the negative impact that they could in theory have. Several recent studies suggest that severe concerns about filter bubbles may not be justified (Dahlgren 2021; Möller 2021; Fletcher and Jenkins 2019).

have to engage, relate and negotiate with unfamiliar entities without a pre-existing trust relationship (Van den Hoven 2012).

- **Security**: The information security challenges that the IoT poses, result from a number of factors. A key factor is the combination of devices having both sensors collecting data, and communicators that are connected to 'remote and opaque data receivers.' An additional risk factor is that users of IoT devices typically don't change factory default passwords, which can quite easily be found online by hackers (Allhoff and Henschke 2018). Furthermore, '*the diversity of IoT devices and access mechanisms as well as massive device connectivity in large-scale IoT access networks brings new security challenges as handovers between different access technologies increase the risk of attacks.*' And the integration of edge intelligence in the IoT can cause security vulnerabilities as at the edge attackers '*can deploy data breaches or modifications while the management of remote 6G core network controllers is limited*' (Nguyen et al. 2021). Another reason for concern about security is that IoT devices often cannot be updated remotely, and neither can their security status be monitored from a distance. At the same time, it is costly to retrieve devices that become redundant or go out of service - making it tempting to just abandon them.This may lead to a rapidly growing number of vulnerable devices that can easily be hacked or otherwise compromised (expert interview).
- **Social justice**: The IoT raises various concerns about social justice and the possibility of a widening digital divide. These include the information position of citizens being determined by developers and industry, the interest of ordinary citizens not being taken into account in applications that shape their life immensely, IoT networks discriminating and providing differential access, the user distress and possibly complicated legal appeals that may result from unwanted data transfers and processing, and only an educated elite being able to grasp and utilize the benefits of IoT (Van den Hoven 2012).
- **Responsibility / accountability**: The causal networks in the IoT are complex. Moreover, incremental developments and deployments of devices in the IoT may lead to unpredictable emergent behaviors that the humans in the netwerk are unaware of. It may become unclear who the acting agent is - user or object. All this contributes to opacity with respect to human responsibility, accountability and liability (Van den Hoven 2012; Allhoff and Henschke 2018).
- **Human freedom and agency**: With the IoT, 'big brother' becomes a cluster of 'some brothers.' The extensive and deep profiling of users that will become possible in the IoT, threatens to deprive the individual of the autonomy to establisher her/his public self-image (personality, identity) and without the individual having effective means to know whether and when profiles are being used or abused (Van den Hoven 2012).

As said, these familiar ethical challenges surface more frequently and/or in more complicated ways in the internet of things. This is especially the case for a challenge which is currently underexplored in computer ethics, but core to engineering ethics more broadly, namely **physical safety**. In comparison to the internet as we used to know it, the IoT has real-world physical implications. Consumer-oriented IoT devices may - to a smaller or larger degree - come with risks for individual bodily harm (Allhoff and Henschke 2018). But more importantly, as factories, energy facilities, transportation systems and other socio-technical systems increasingly become part of the IoT, there is an increased risk for larger-scale incidents that form a physical threat. Physical safety may come under threat as a result of security breaches, which facilitate actions with malicious intent. In addition, one of the experts that we interviewed pointed out that we should not underestimate that coding is never error free, which may cause accidents. An internet of things where 'everything is connected with everything' also makes the potential impact of errors more far reaching and harder to foresee and assess (Section 5.3.9).

That **sustainability** becomes a major challenge due to a massive increase in energy usage, was already mentioned in the preceding sections on 6G and AI. The internet of things is another technological development contributing to this challenge.

Finally, it is important to note that experts foresee that the internet of things will lead to a **blurring of the boundaries between traditionally separate application contexts**. For example, it is not unthinkable that certain wearables connected to health apps (e.g. to monitor the blood sugar of children) may start to have consequences in the education domain (e.g. if parents start expecting teachers to act upon the measurements in certain ways). This blurring of boundaries complicates the ethical challenges that we need to deal with, as context is often an important factor in both ethical reflection and policy making to deal with ethical challenges. Contextual integrity is, for example, a central concept in the analysis of the notion of privacy and in its protection. More in particular, the conceptual distinction between the private and public sphere, traditionally very important in political philosophy, is also starting to become vaguer. This challenges the checks and balances associated with the separation of powers in our democracy. If the boundaries between end-users, government agencies and corporations are blurring, this has implications for our ability to assign responsibility and demand accountability. A concrete example of how IoT devices may cross the public-private barrier, are smart electricity usage meters in people's homes. They can be considered to be part of both the public and the private sphere (Van den Hoven 2012). This issue is further discussed in Section 4.3.1 as another key challenge.

3.4. Augmented and virtual reality

Virtual reality (VR) can be applied both in a business context and for consumer products and services. The Rathenau Institute, the Dutch technology assessment office, concludes that '*there has been little political or public debate, whether in the Netherlands or elsewhere, about VR technology and very few VR-related policy measures, case law or ethical codes have emerged*' (Snijders et al. 2020, p.4). Their report aims to contribute to filling this gap. Based on an analysis of academic publications on the public and ethical issues raised by VR, the Rathenau report concludes that '*VR raises a multitude of ethical and public issues, for example with regard to privacy, autonomy, physical and mental integrity, informed consent, and access to technology*' (p.5). The Rathenau report identifies four main clusters of risk for consumer VR applications (p.6):

- **Physical and mental risks**: '*emotional involvement, long term damage, blurring of boundaries, alienation, addiction*'
- **Social risks**: '*damage to social values, slander and intimidation, social dissociation, virtual violence, sexualisation*'
- **Abuse of power**: '*manipulation, not transparent, curtailing autonomy, political influence, use of data without permission*'
- **Legal risks**: '*invasion of privacy, identity abuse, property issues, uncertain legal status of virtual actions*'

A recent report written at the request of the Dutch Ministry of Justice and Security (Schermer and van Ham 2021) arrives at a similar inventory of social and ethical challenges and the values that are at stake. Values at stake as identified by this report include human dignity, security and truth. As for the latter: the large-scale implementation of AR and VR applications could mean that people more than ever start to live in different realities without a shared frame of reference. One new category of risks that this report adds to those mentioned in the Rathenau report, is that of **damage in the physical world**, which can result from '*users who are distracted by their immersive technologies, or who interpret the augmented reality in the wrong way*' (p.4, our translation). An example is that traffic safety may become endangered by the large-scale usage of VR and AR applications.

Both reports remark that we have at the moment insufficient knowledge of the long-term effects of (different forms of and aspects of) AR and VR, which is a problem for effective policy making and regulation. For example, we don't know yet if having extreme virtual sex would lead to people engaging in unacceptable forms of sexual behavior in the real world.

3.5. Blockchain

A major concern about blockchain technology is that it would lead to enormous increases in **energy usage**. It has, however, also been claimed that '*second- and third-generation blockchains [...] are so programmed as to reduce or prevent that problem*' (Dierksmeier and Seele 2020, p.348) and that energy usage depends on the type of blockchain technology and application (Sedlmeir et al. 2020). Nevertheless, this concern should not be dismissed too lightly and deserves further scrutiny.

Other than this major concern about energy usage, blockchain is often presented as a *solution* to the ethical challenges that other technologies cause or face, such as that of **privacy** (e.g. Ylianttila et al. 2020; Bertino, Kundu, and Sura 2019); '*One of the main advantages of blockchain lies in its ability to ensure that data are secure, private, reliable and valid, and thus personal data are not compromised*' (Kritikos 2020). Another advantage is that it can create **transparency** in information chains, as blockchain technology ensures that data has not been tampered with. Kritikos (2020) even calls blockchain 'a transparency machine' and argues that it can increase societal **trust** in AI.

Often mentioned is also that blockchain would facilitate a more **democratic and egalitarian** society, as it makes decentralized networks possible which *'can eliminate power asymmetries that usually work to the benefit of intermediaries*' (Dierksmeier and Seele 2020, p.350). It facilitates new collaborative forms of business and entirely new ways of organizing things (ibid), and can be expected to change the economy in myriad and sometimes radical ways (Tang et al. 2019). However, it is exactly because blockchain technology is so disruptive, that we cannot be certain that these consequences will unambiguously or even by and large be positive. For example, looking at blockchain technologies that aim at '*reorganizing data flows in the internet of things (IoT) architectures*', (Ishmaev 2020, p.411) argues that:

> *'the promised benefits are counterbalanced by a significant shift towards the propertization of private data, underlying these proposals. Considering the unique capacity of blockchain technology applications to imitate and even replace traditional institutions, [....] without careful consideration of a wider impact, such blockchain applications could have effects opposite to the intended ones, thus contributing to the erosion of privacy for IoT users.*'

Atzori (2015) as well argues that blockchain may not only have positive implications, but also comes with 'r*isks related to a dominant position of private powers in distributed ecosystems, which may lead to a general disempowerment of citizens.*'

In addition to ethically favorable (e.g. making supply chains transparent) and unfavorable applications (e.g. blockchain facilitated assassination markets) there are, so Dierksmeier and Seele (2020) argue, also a lot of applications that are morally ambivalent (e.g. how blockchain may radically change the job market). Unfortunately, the ethics of blockchain has until recently not received much attention from academia (Tang et al. 2019).

3.6. Bionanotechnology

'Biotechnology' and 'nanotechnology' are very broad categories, and on the ethics of these two broad areas of technology a lot has already been written. In this study we are especially interested in the ethical and societal challenges of the Internet of Bio-Nano Things. As this is a very recent and new development, we have not found any publications that specifically discuss that topic. Probably it will raise similar ethical and social concerns as we just discussed for the internet of things, although some issues may become even more salient or challenging to deal with. Where regular-size IoT devices may already be opaque for many users, this will probably be even more so for nanodevices. **Privacy** may be one such issue that becomes even more salient, because the application of nanosensors in and on the human body leads to the generation of data that is highly privacy sensitive. This also makes **security** of the utmost importance, while at the same time the technology raises new security challenges (Akhtar et al. 2020; Dressler and Fischer 2015):

> *'The Internet of Nano Things is vulnerable to all types of attacks, either physical or through wireless technologies, given that this type of device does not meet with constant vigilance. The attacks can occur to acquire private data through the theft of sensors, interrupt applications controlled utilizing computers, or modify the communication links in the nano-networks. This is because standard security techniques cannot be applied to nano networks that operate in the terahertz band. To secure the IoNT system, there is a need to develop new security solutions.' (Akhtar et al. 2020, p.139)*

4. What people discuss: media, ethics, tech, policy

Responsible innovation is required to ensure that technologies that are part of the approaching technological storm align with societal values. However, developments in these digital technologies are extremely rapid, and their impact is sometimes hard to predict or to grasp for regulators. In this section, we present the results of a text-mining investigation in order to investigate which values in relation to the converging digital technologies are addressed in four different realms of society. These realms are: 1) ethical research, 2) news media, 3) regulation and legislation and 4) technological research. As will be explained, this chapter is based on the idea that these different realms each have to play a specific role in overall division of labour in order to properly address relevant values in relation to the merger of digital technologies. On the basis of a text-mining exercise, we aim to investigate whether the realms indeed currently play their role and what might be improved.

4.1. Values in different realms of society

Different realms of society (e.g. news media, regulators, technological and ethical research) each play a different role in fostering responsible innovation of technologies that are part of the technological storm. While reality is complex, one might at least suggest the following roles. News media are essential in reporting (new) concerns that the development and deployment of technologies are raising in society. Policy-making and regulation set restrictions to the use of technologies and encourage the development of new (more responsible) innovation. Ethical research is key to discovering new ethical issues and better understanding them, and for conceptualizing relevant values, while technological research can provide solutions more in line with societal values, for example by technically operationalizing social and moral values and translating them in design requirements or by developing new innovations.

Responsible innovation may fail if the different realms do not perform their roles properly, or are insufficiently aligned. For example, it might take time before regulators become aware of concerns voiced in society. Some realms might be biased towards specific values because technologies are not well understood or due to historical reasons. Engineers might not always have the means to address critical societal concerns, sometimes because of a lack of incentives but also because tools to translate values into a set of tangible design guidelines are lacking. Ultimately, a lack of coordination or alignment between realms could pose a severe threat to the successful and responsible deployment of technologies that are part of the technological storm.

The aspiration to foster responsible innovation for technologies that are part of the technological storm is highly ambitious. Many organizations (businesses, regulators, governments, consumers, etc.) are involved in the approaching technological storm. Numerous applications will be created combining these general-purpose technologies, and the range of potential moral concerns is extensive. Policy-makers may not always find the time to dive into the societal implications of different technologies. Even experts may not always be able to capture the approaching technological storm and its moral implications in a comprehensive manner.

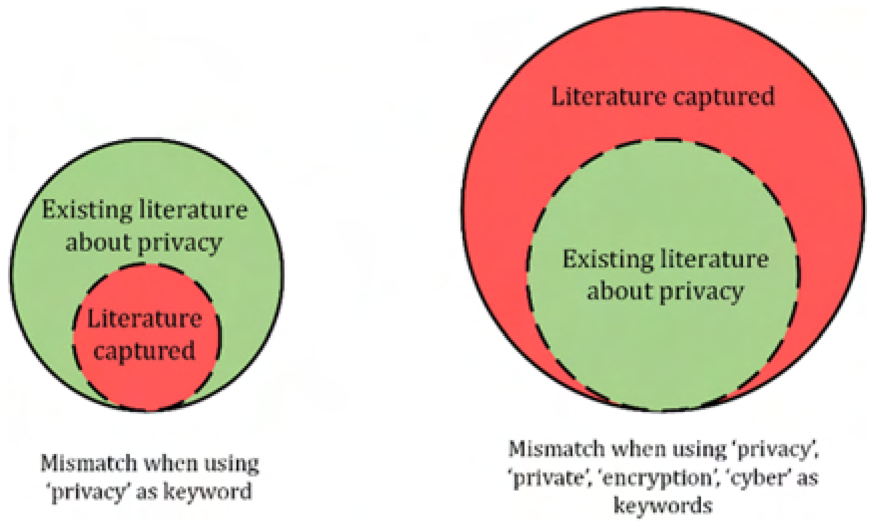
4.2. Tracing values with topic models

In this chapter, we employ a text mining approach based on topic modelling to create an overview of which values are addressed in different realms of society in connection with the approaching technological storm. For example, the techno-scientific literature might suggest new encryption methods to increase the security of IoT systems, while the media may be discussing what implications developments in AI may have for democracy. The advantage of using a text mining method is the large number of texts that can be analysed in a relatively short amount of time in a systematic way.

The approach that we use was specifically developed to capture complex concepts such as values (de Wildt, van de Poel, and Chappin 2021). A difficulty when identifying values in text corpora is that values tend to be discussed in a latent manner. Rather than explicitly naming the value in question, authors often use a wide range of other words, which implicitly refer to a value. For example, texts addressing the value of privacy may not explicitly mention the word 'privacy,' but they may contain words such as 'private,' 'theft,' or 'cyber.' Texts may also refer to solutions to privacy issues, for example, by mentioning the word 'encryption' or 'firewall.' However, encountering such words in a document does not always mean that the document is about privacy. For example, an author might also use the word 'private' to refer to individual ownership. These problems can be addressed by using topic models rather than key words.

Consider the example of the value privacy: using only the key word 'privacy' would lead to omitting many documents that refer to privacy using other words (see figure 1). This might lead to an underrepresentation of the value privacy in the final dataset, so-called false negatives. Adding keywords such as 'private,' 'cyber,' and 'cryptography' could help to capture missing documents. However, doing so might lead to an overrepresentation of the value of privacy, as it might lead to capturing documents that are not related to privacy, so-called false positives. Because values tend to be latent, it is often impossible to define a set of keywords that adequately represent the value we are interested in.

Figure 2: Mismatches that may occur in text mining


Existing literature about privacy
Literature captured
Mismatch when using 'privacy' as keyword
Literature captured
Existing literature about privacy
Mismatch when using 'privacy', 'private', 'encryption', 'cyber' as keywords

The approach proposed by de Wildt, van de Poel, and Chappin (2021) to capture latent concepts such as values is based on probabilistic topic models (Blei, Carin, and Dunson 2010). In topic modeling, we define values as a (probabilistic) distribution of words instead of a set of keywords. For example, a topic referring to the value privacy will have high probabilities on terms such as 'confidentiality,' 'private,' and 'secret' occurring, and low probabilities on words that do not refer to the value. Texts addressing values can then be captured by comparing the distribution of words in this text and the distribution of words in a topic model built to represent a value. Doing so, texts addressing certain values can be captured, even if they do not mention these values explicitly. This makes it possible to evaluate to what extent specific values are being discussed in different text corpora, and to explore whether, and if so how, the frequency in which certain values are discussed have changed over time.

4.3. Methodology

To explore values in different realms of society, we have used four datasets. The first dataset (NEWS) is composed of a large number of news articles (#562.295), taken from 26 different newspapers (including Reuters, CNBC, The New York Times). This dataset is used to evaluate which values are associated with the approaching technological storm in public debate. The second dataset (ETHICS) consists of 8.565 scientific articles downloaded from ethics-related journals. We have filtered both the NEWS and ETHICS dataset on articles discussing technologies that are central to this study. The third dataset (TECH) contains 674.656 scientific articles downloaded from Scopus, selected using keywords related to each technology, and excluding ethics-related journals. The fourth dataset (LEGAL) has been built using regulatory procedures related to each technology found on the EU Legislative Observatory website[16]. These datasets are further described in appendix 2. We have concentrated our analysis on the main enabling technologies mentioned in section 2.1: 5G and 6G, AI, IoT, augmented and virtual reality, blockchain technology, robotics, bio and nanotechnology, and quantum computing.

Next, a survey was built to identify the ten most relevant values for the approaching technological storm. The survey was sent to 19 experts with various technical and ethical backgrounds. After discussion of the survey results, we decided to add inclusiveness to this list of values. The final list of values can be found in table 3. More detailed outcomes of the survey can be found in appendix 2.

Finally, distributions of words (i.e., a topic model) have been built for each value, by using so-called anchor words. Anchor words are used to progressively guide a distribution of words towards an adequate representation of a value of interest. For example, the words 'wellbeing', 'wellbeing', 'wellbeing', 'quality life', 'good life', 'QOL', 'life satisfaction', and 'welfare'' have been used as anchor words to create a distribution of words for the value of wellbeing. The final list of anchor words and topics created for values can be found in appendix 2.

[16] EU Legislative Observatory website: https://oeil.secure.europarl.europa.eu/oeil/home/home.do

Table 3: The 10 values selected as most relevant for digital technologies

Selected values	Definitions
Justice and fairness	The technology avoids undesirable biases and discrimination; the technology contributes to fair outcomes and just distributions.
Privacy	The technology protects the personal sphere of people against intrusion by others, and it allows people to decide which personal data (not) to share.
Cyber-security	The technology is safe from malicious attacks.
Environmental sustainability	The technology does not harm the environment or burden ecosystems.
Transparency	Choices by and with respect to the technology are clear, explainable and can be inspected.
Accountability	Relevant decisions by - or consequences of - the technology can be accounted for (e.g. by tracing them back to human decision makers).
Autonomy	The technology increases, or at least does not harm, people's capacity to reason and make choices in line with their values.
Democracy	The technology respects democratic values and human rights like equal representation, the rule of law, and freedom of speech.
Reliability	The technology fulfills its purpose consistently over time.
Trust	The technology is trustworthy, and it promotes trust among its users and others.
Well-being	The (use of the) technology contributes to wellbeing and good quality of life.
Inclusiveness	The (use of the) technology promotes social integration. Different perspectives are taken into account.

4.4. Results

In this section, the main results of the analysis are presented. We concentrate here on four main observations: (1) different social realms focus on different values in relation to the approaching technological storm, (2) the ETHICS and LEGAL dataset have focused on AI and far less on other technologies, (3) the integration of digital technologies may cause new or more moral concerns and (4) the values that are addressed in relation to digital technologies have changed over time.

4.4.1. Differences between social realms

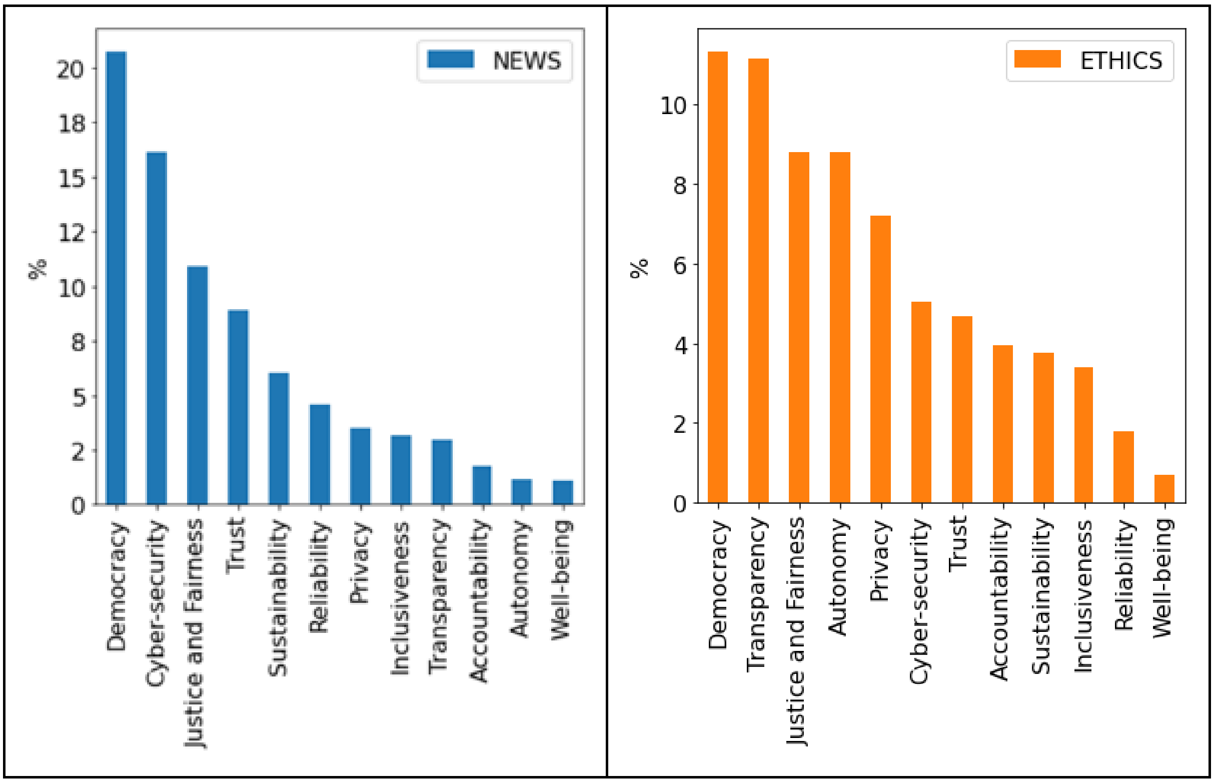
A first result is that we found differences in the frequency by which values are discussed in relation to digital technologies in the four different datasets. Figure 2 shows which values are most frequently mentioned in our datasets. In the NEWS dataset the most discussed values are democracy, cybersecurity, and justice and fairness. In the ETHICS dataset, the values of democracy and transparency are most often mentioned. The values most frequently addressed in the TECH dataset are cybersecurity and reliability. Finally, in the LEGAL dataset, the most prominent values are cybersecurity, transparency and justice and fairness.

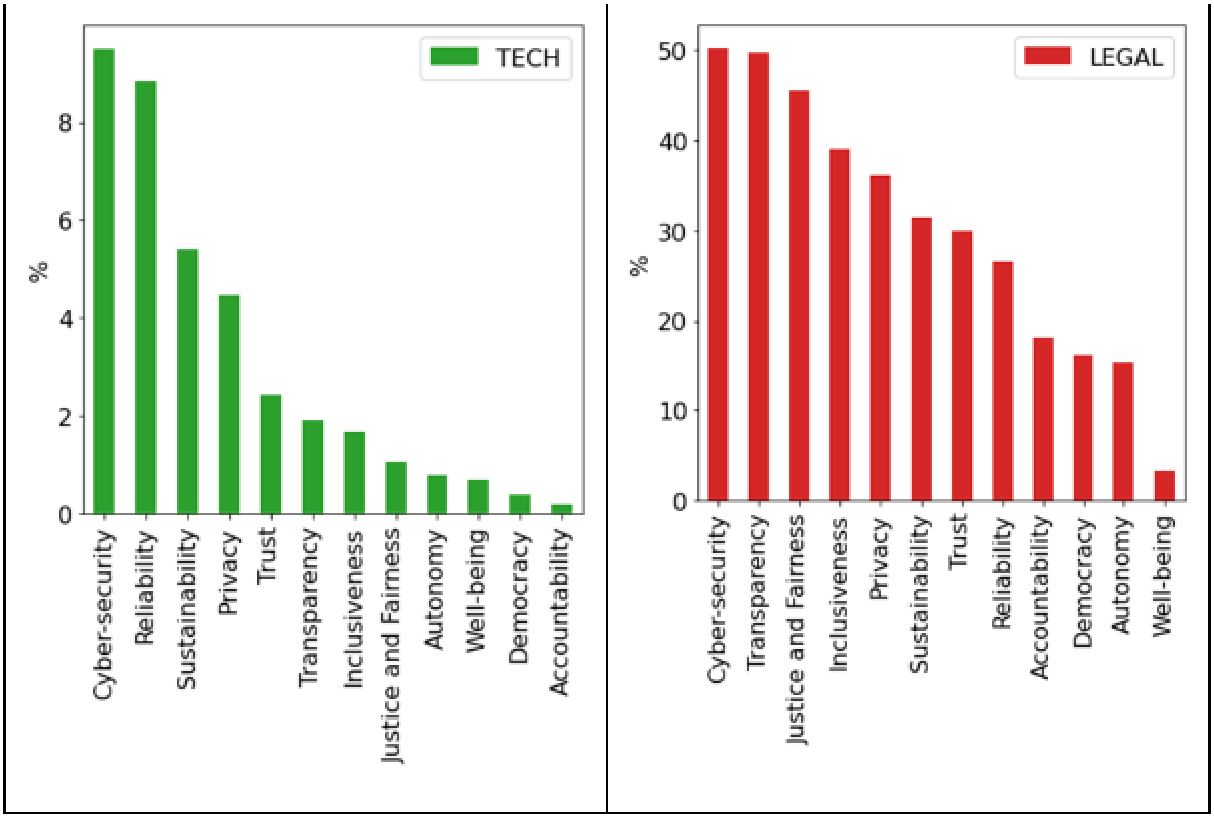
Notably, some values are very prominent in some datasets while significantly less frequent in others. This is the case for the value democracy, which is an important value in the NEWS and ETHICS dataset, while being one of the least frequently mentioned values in TECH and LEGAL. Transparency seems to be a frequently mentioned value in ETHICS and LEGAL, while playing a smaller role in NEWS and TECH. Finally, wellbeing is in all the 4 datasets among the least frequently mentioned values.

It should be noted that frequencies with which values are mentioned cannot be directly compared between datasets because one would not expect values to be mentioned as often in techno-scientific articles as, for example, in the news. For instance, does the fact that the value of democracy is mentioned in 20% of the articles in the NEWS dataset and only in 0.5% in the TECH dataset mean that this value is not sufficiently addressed in the latter? Clearly such a conclusion would be too quick since many TECH articles are not about values to begin with. However, much more telling is the relative importance between values in different datasets. The fact that democracy is the most frequent value in NEWS and one of the least frequent in TECH clearly seems to indicate a different focus between those datasets, and hence between the social realms these datasets represent.

Figure 3: Values mentioned in each dataset

(percentage of sub-set of documents discussing relevant technologies)



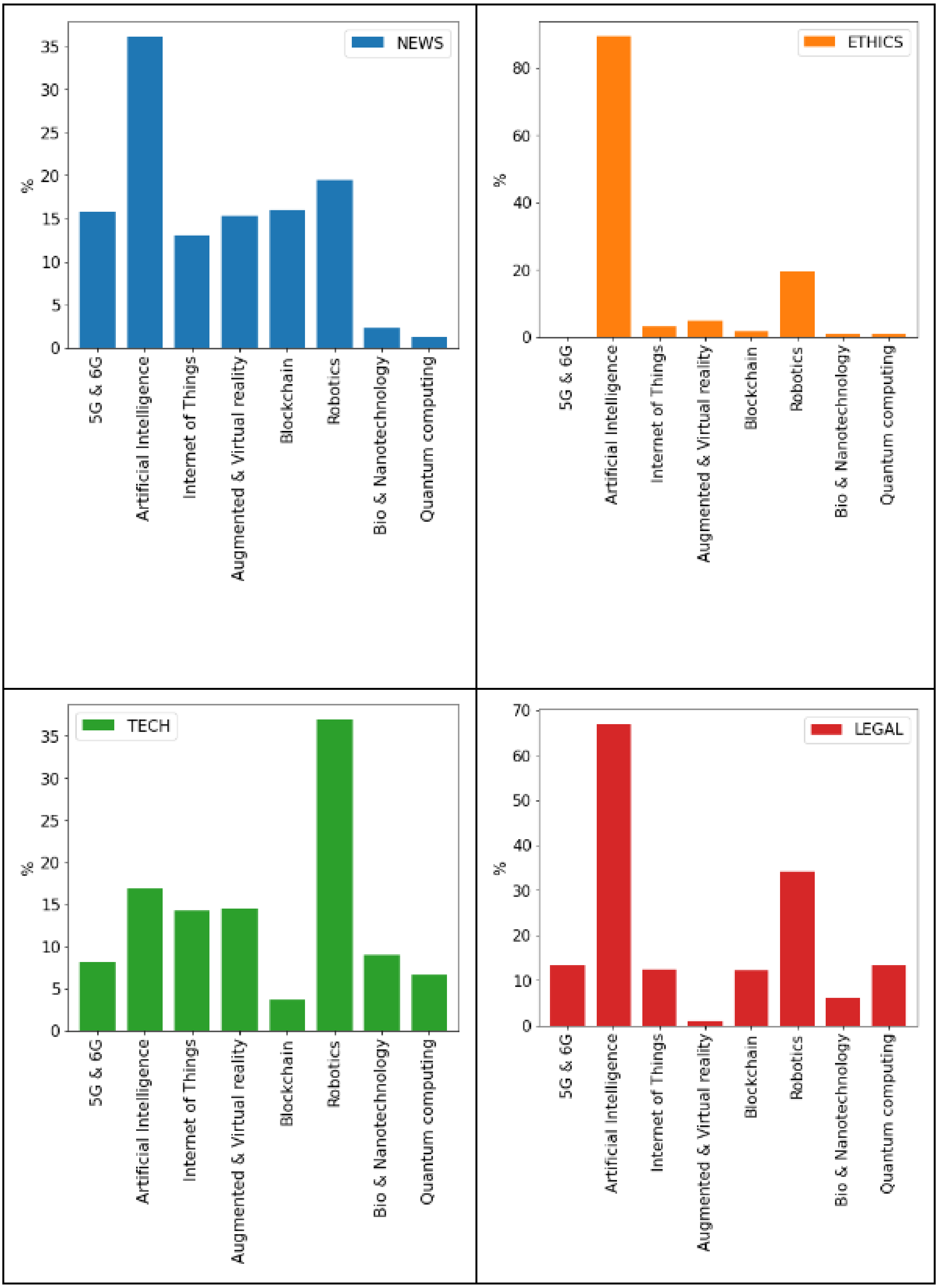


4.4.2. Focus on AI

A second result is that in the ETHICS and LEGAL dataset, we see a focus on AI, and to a lesser extent on robotics, and that there seems to be far less discussion about the ethical and legal issues of other digital technologies.

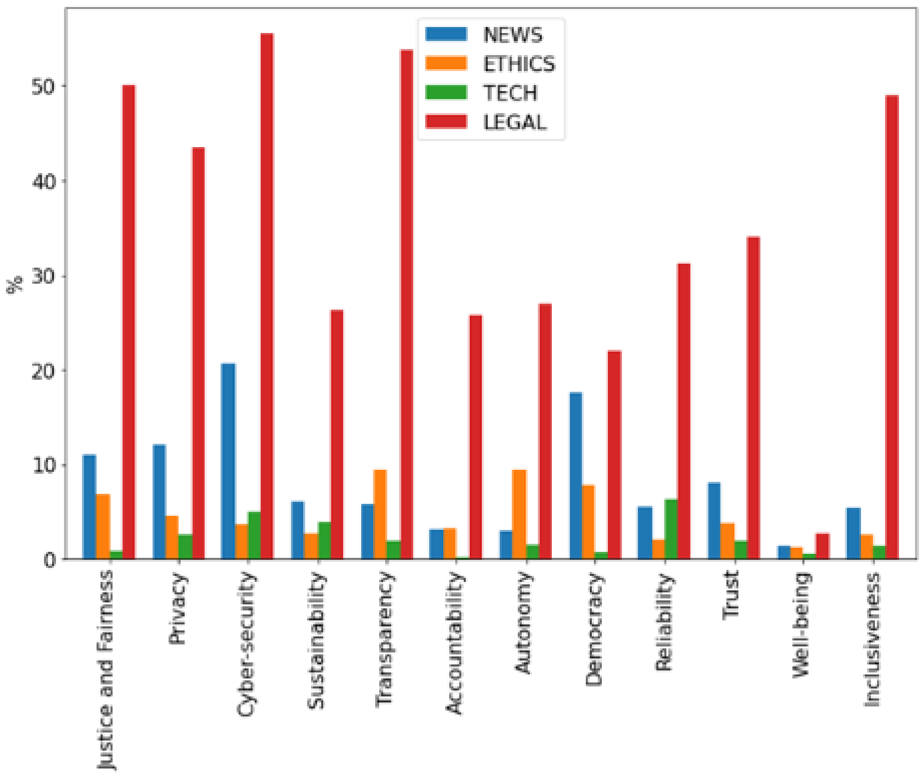
Figure 3 shows technologies mentioned in the four datasets. This figure shows a substantial disparity in the extent to which different technologies are discussed in the four datasets. The ETHICS and LEGAL dataset seem to concentrate largely on AI (and to a lesser extent robotics), while the NEWS and TECH dataset discuss a larger number of different technologies. Some technologies seem to be poorly addressed in the NEWS dataset (bio and nanotechnology, and quantum computing), or missing in the ETHICS dataset (5G and 6G). Augmented and virtual reality are also almost absent in the LEGAL dataset.

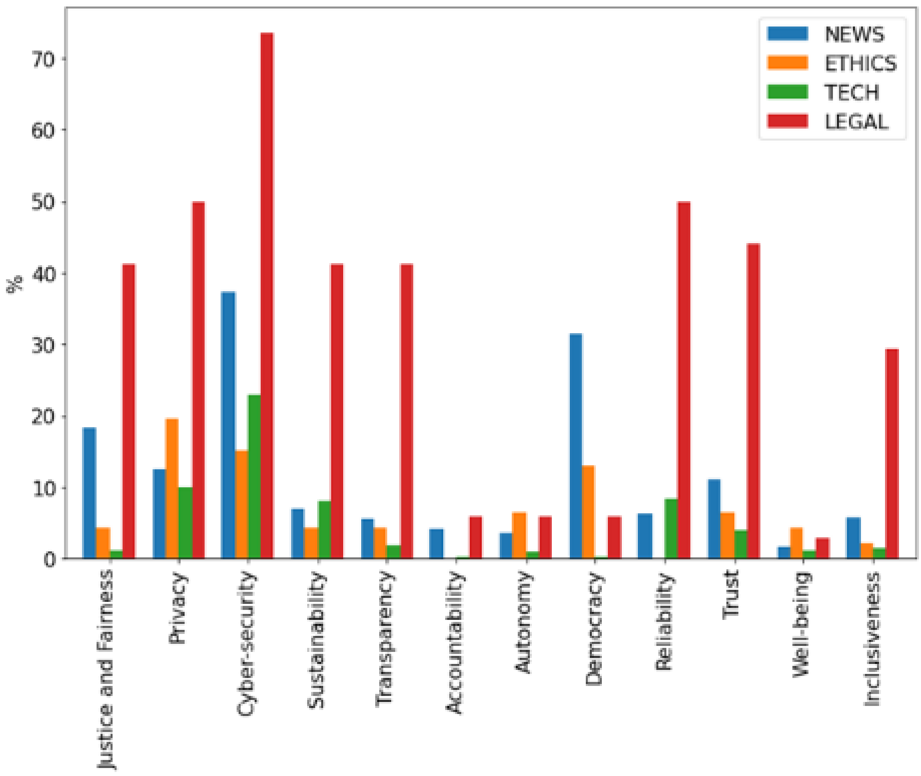
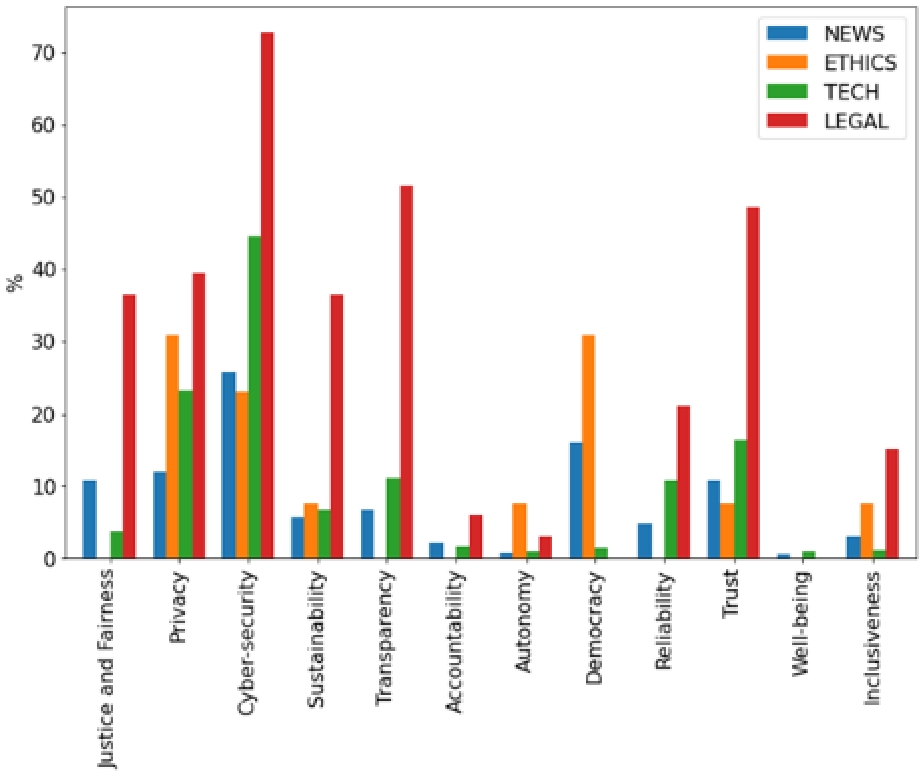
Figure 4: Technologies mentioned in the four datasets



The focus on AI in the ETHICS and LEGAL dataset also seems to be reflected in what values are being discussed in these datasets in relation to other technologies than AI. Appendix 3 provides an overview of which values are being discussed for each technology in each dataset. As an illustration, we here provide Figure 4 which concentrates on values discussed in the four datasets in relation to AI, IoT and blockchain technology. It is striking that particularly in the ETHICS dataset some values in relation to blockchain, and to a lesser extent IoT, are not being discussed while they are being discussed in the NEWS and TECH dataset. This might be explained by the fact that only a limited number of articles in the ETHICS dataset focuses on blockchain and IoT (see Figure 3). Nevertheless, it suggests that in the ethics literature several value issues in relation to blockchain, and to a lesser extent IoT are not yet being discussed while they have been picked up in the technical literature and in the news.

Figure 5: Values mentioned with AI in the four datasets

AI	

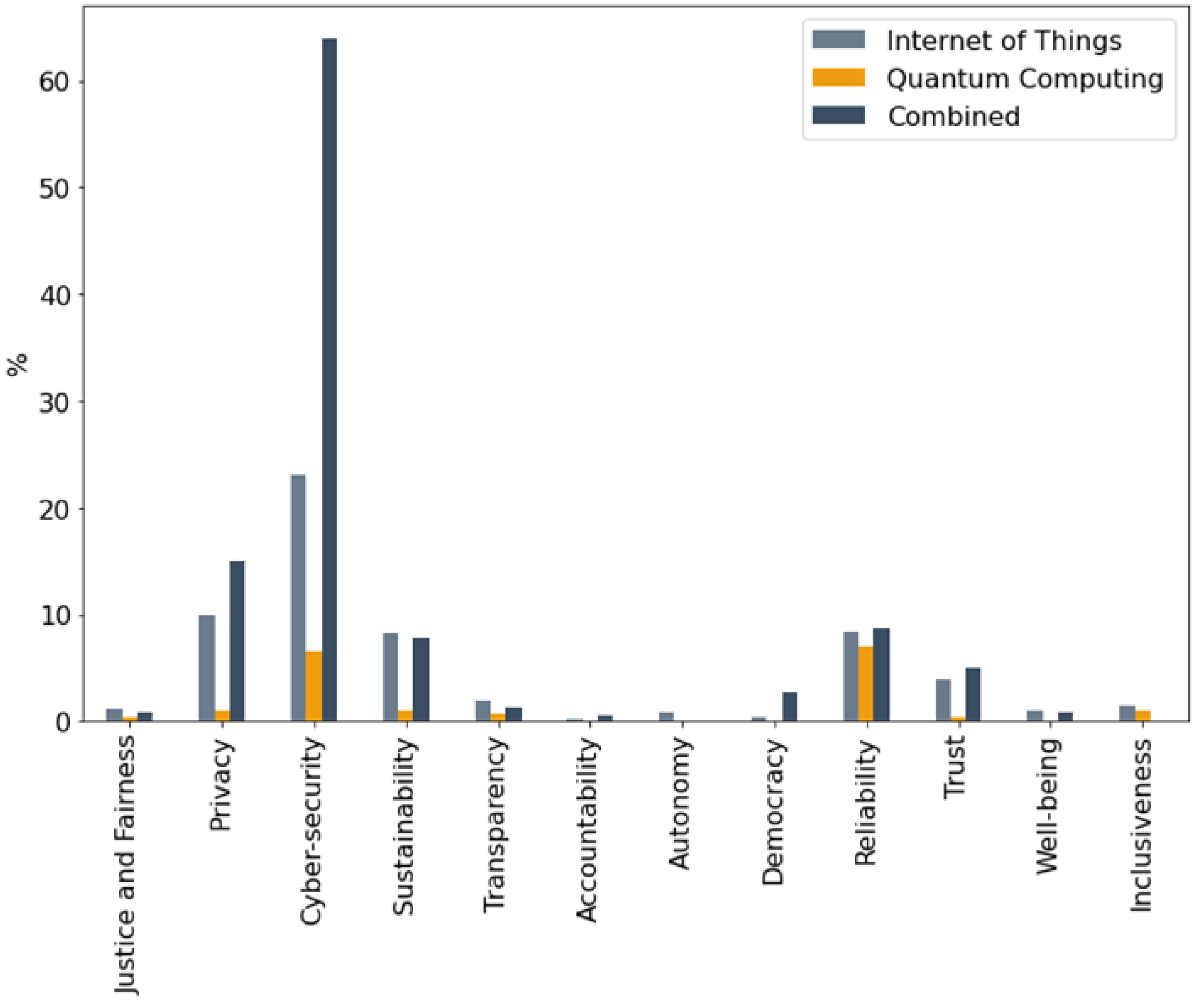
IoT	
Blockchain	

4.4.3. New moral concerns

A third result is that the merger of digital technologies can bring about new (or more) moral concerns. As the possible combinations of digital technologies discussed in different datasets is large, we concentrate on new moral concerns discussed in the TECH dataset.

An indication of new (or more) moral concerns is when one value is discussed much more frequently in documents on both technologies than in documents on individual technologies. An example is provided in figure 5. This figure shows how frequent different values are discussed for internet of things and quantum computing in the TECH dataset. From this figure, it can be seen that the value of cybersecurity is deemed relevant for both technologies. However, the frequency of the value of cybersecurity substantially increases when both technologies are discussed together. This seems to indicate that new or more cybersecurity issues are raised specifically when these two technologies are used together. A manual analysis of these documents reveals that quantum computers can be used to breach existing cryptographic protocols used in IoT devices. Different solutions are proposed by authors, including latency-optimized hash-based digital signature accelerators and data encryption methods based on quantum walks.

Figure 6: Values for internet of things and Quantum Computing in TECH dataset



An overview of all values that are mentioned substantially more often in documents that discuss the combined application of those technologies can be found in table 4. This table has been built based on the TECH dataset; values found are those perceived and addressed in the techno-scientific literature.

Values have been included in this table in case the percentage of documents mentioning the value was at least 50% higher than for both technologies separately, and the number of documents on both technologies was higher than 50.

Table 4 shows that new moral concerns discussed in the TECH dataset essentially result from a limited number of technologies: 5G and 6G, AI, IoT and blockchain technology. Cybersecurity seems to be the value that is most frequently discussed when it comes to the merger or combination of different digital technologies.

Table 4: Values found when combining different technologies

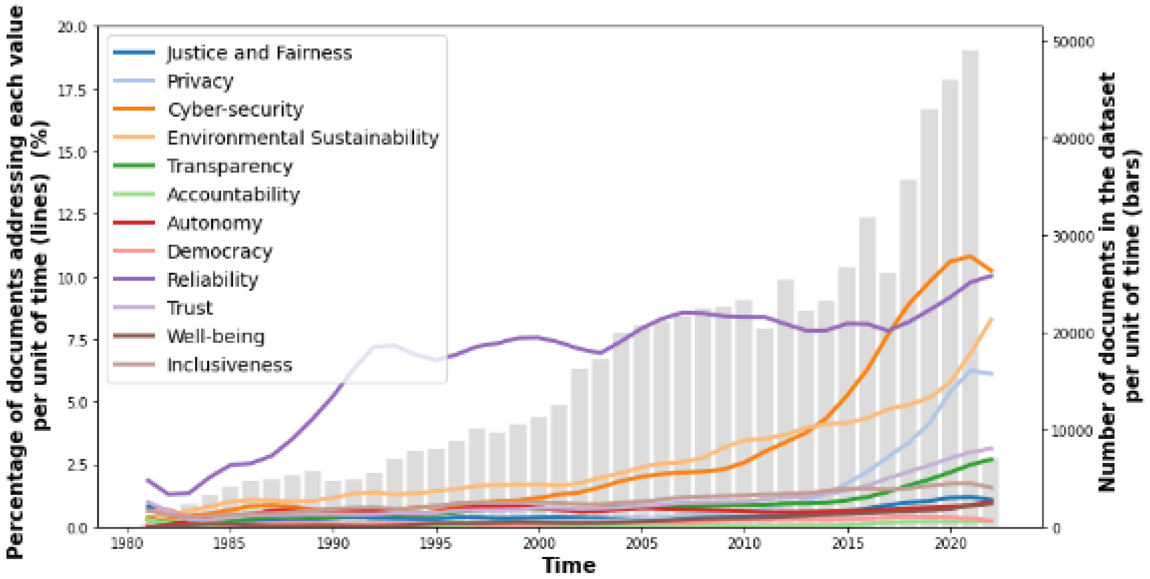
	5G and 6G	AI	IoT	Augmented and virtual reality	Blockchain	Robotics	Bio and nano technology	Quantum computing
5G and 6G								
AI	*Cybersecurity* *Privacy* *Reliability* *Sustainability*							
IoT	*Cybersecurity* *Reliability*	*Sustainability*						
Augmented and virtual reality		*Cybersecurity*						
Blockchain			*Cybersecurity*					
Robotics		*Autonomy*						
Bio and nano technology								
Quantum computing			*Cybersecurity*		*Cybersecurity*			

4.4.4. Changes over time

A last result is that there are changes over time in which values are most often discussed in connection to digital technologies. Figure 6 shows how often specific values have been addressed in the techno-scientific literature. Initially, this literature essentially focused on the value of reliability. Over time, the literature has progressively addressed other values, namely cybersecurity, sustainability, and privacy.

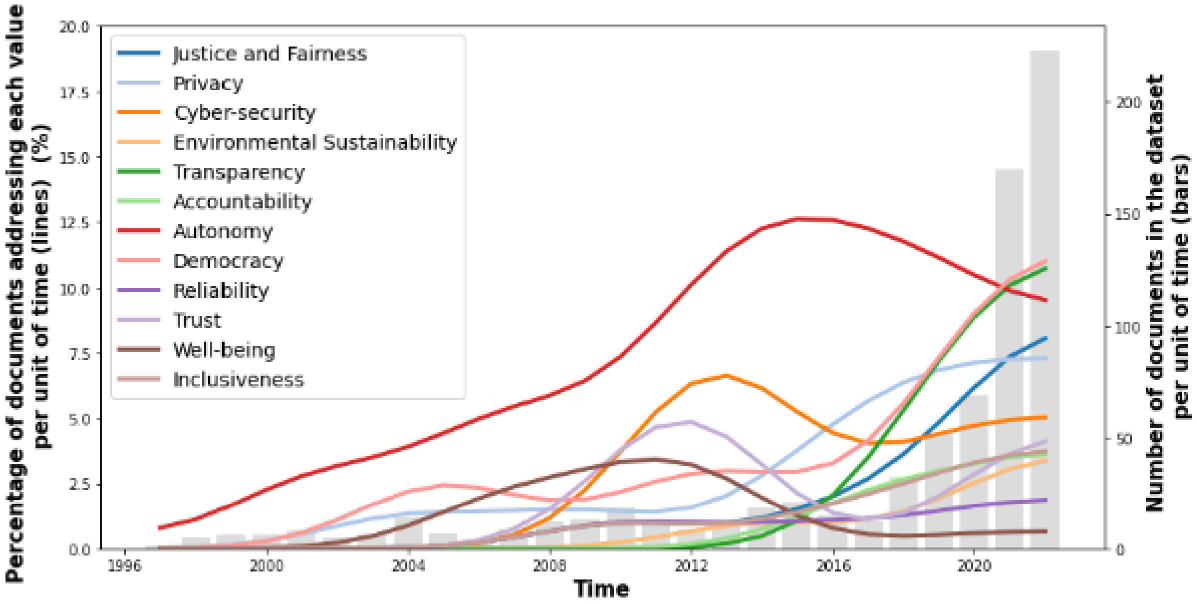
Some values still do not receive much attention, namely justice and fairness, democracy, and autonomy.

Figure 7: Value change in the TECH dataset



The frequency of values discussed in relation to digital technologies seem to have changed in the ETHICS dataset as well (Figure 7). Autonomy appears to always have been a dominant value in this dataset. Cybersecurity seemed to have been a value frequently discussed in the early 2010s. The most prominent values are now democracy, transparency, autonomy, justice and fairness and privacy.

Figure 8: Value change in the ETHICS dataset



Such an analysis of value change is not possible in the NEWS and LEGAL dataset as they only cover a short time period.

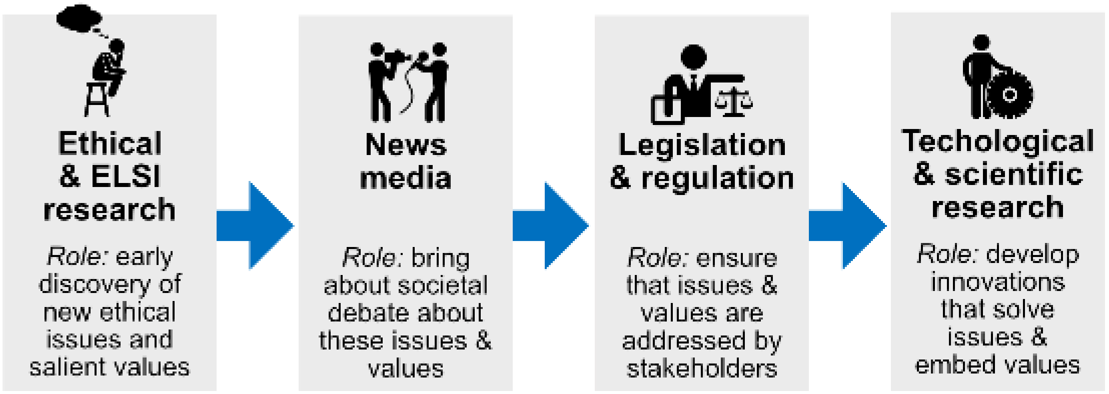
4.5 Discussion of results

The results of our analysis suggest that the merger of different digital technologies raises new or at least an increase in moral and social issues (Section 4.4.3). An important question is whether the different realms of society we have studied in our text-mining analysis (the news/public discussion, legislation, ethical research and technological research) are sufficiently prepared to timely address these new issues. In addressing this question, it is important to take into account that different realms of society have a different role to play in what one may call an overall division of labour in dealing with the social and ethical issues raised by the approaching technological storm (Section 4.1). We would like to suggest that different realms may play the following roles:

- Ethical research and more generally ELSI research has a role to play in early discovery of new ethical issues and values that need to be addressed in relation to digital technologies.
- The news media have a role to play in bringing these issues to the attention of a larger societal audience and they may reflect what values are considered important in society at large.
- Legislation and regulation have a role in addressing ELSI issues through new regulation and legislation, also aiming at ensuring that relevant values are addressed in the application of new technologies, as well as in technological research.
- Technological and scientific research, as well as technological innovation, may not only raise new societal and technical challenges but are oftentimes also crucial in better addressing these. Ideally, one would therefore want that values that are considered important in the public debate, in ethics or legislation are also translated into new technological solutions.

This description suggests a certain temporal order in how ethical issues and values are likely mentioned in our datasets:

Figure 9: Temporal order in ethical issues and values


Ethical & ELSI research
Role: early discovery of new ethical issues and salient values
News media
Role: bring about societal debate about these issues & values
Legislation & regulation
Role: ensure that issues & values are addressed by stakeholders
Techological & scientific research
Role: develop innovations that solve issues & embed values

The idea behind this order is that oftentimes new ethical and moral issues will first be discovered in ethical and ELSI research (the ETHICS dataset), then will be discussed in news media, also reflecting societal priorities (NEWS), and next will be translated into new regulation and legislation (LEGAL), as well in new technological options and solutions (TECH).

Figure 6 and 7 provide some evidence for this temporal order when it comes to the values of 'privacy' and 'justice and fairness'. As these figures show, privacy has become an issue in the ETHICS dataset since around 2000, and about 10 years later - around 2010 - became more prominent in the TECH dataset. Similarly, 'justice and fairness' has become more prominent in the ETHICS dataset since 2008 and around 7 later - around 2015 - started to become somewhat more prominent in the TECH dataset. (For reasons explained in Section 4.4.4, we regretfully lack time series over a longer period for the NEWS and LEGAL dataset).

For other values, we however see another pattern (see figures 2, 6 and 7):

- **Reliability** is a frequently mentioned value in the TECH dataset from the start, and is less frequently mentioned in the ETHICS dataset. This may be explained by the fact that reliability is more a technical and instrumental value rather than an intrinsic moral value (see further Section 4.5.1).
- **Cybersecurity** and **sustainability** are also not so often discussed in the ETHICS dataset as in the TECH dataset, and also follow a different temporal pattern than privacy and fairness and justice. We will discuss potential explanations in Section 4.5.2.
- **Democracy** has been a frequently mentioned value in the ETHICS dataset since 2000 but is still hardly mentioned in the TECH dataset. Values like transparency, autonomy are also often mentioned in the ETHICS dataset but still hardly in the TECH dataset (see also Figure 2). We will discuss potential explanations and whether this should be seen as problematic in Section 4.5.5.

Below, we will discuss for each of the four societal realms (ethics, news, legislation, technology) what we can say about how it plays its role in the overall division of labor (Figure 8) on basis of our results, but before we do so we first discuss a number of caveats that are important in properly interpreting our results.

4.5.1 Some caveats in interpreting the results

Before we present possible interpretations of the results of the text-mining exercise, a few caveats are in place. First, there are a number of *methodological* reasons why we should be careful in interpreting the results. A first point is that we have measured the frequency with which certain values are mentioned in the four datasets. This frequency may be an indication of how important a certain value is considered in a certain societal realm, but there may be other reasons as well why a certain value is less or more often mentioned in a dataset. For example, some values may not so often be mentioned in the technical literature because they are hard to translate into technological choices, but that might not mean that they are considered unimportant by engineers and technological researchers (although it might mean that that value is less taken into account in innovation and technological research). Similarly, a value may be often mentioned in NEWS due to specific events or because journalists expect the public to be interested in it (rather than considering it an important value).

Second, we operationalised values, as explained in Section 4.3, through topic models. These topic models were created by using a dataset that combines the four specific datasets, but it is still conceivable that these topic models are biased to one of the four datasets (so that value seems to appear more often in that dataset) or does not fully cover the intended value. In case a value is mainly discussed in only one of the datasets, words attributed to this value might essentially reflect how the value is being discussed in that dataset.

Third, as also explained in Section 4.3, we take the four datasets to be representative for different realms of society, but this representation may not be full or ideal. For example, the ETHICS dataset is based on ethics journals, but it might be argued that the role of early discovery of ethical and social issues is played by a broader category of ELSI research than just ethics research.

In addition to these more methodological reasons, there might be an important *substantive* reason why certain values are more frequently mentioned in some datasets than in others. This has to do with the difference between intrinsic and instrumental values (van de Poel 2009). Intrinsic values are those 'that are good in themselves or for their own sake', while instrumental ones are 'valuable because they help to achieve other values.' Reliability is an example of an instrumental value; More reliable designs can lead to more cyber-secure technologies, which in turn may contribute to more stable democratic processes. Democracy is the intrinsic value here. Security could be seen as both instrumentally and intrinsically valuable. It may very well be the case that particularly in technological literature intrinsic values are often not mentioned, not even in a latent manner, as the importance of the instrumental value may appear self-evident, while at the same time mentioning the underlying intrinsic value does not seem to contribute anything to the quality of the work done. This may for example very well be justified in case of a paper presenting very specialized, technical work on increasing the reliability of a technology. It may indeed not add much to mention that the work presented is important because it contributes to security and indirectly contributes to democracy, as it is generally not disputed that reliability matters. The same explanation holds for the value of wellbeing, which could be seen as a key societal driver behind the development and deployment of digital technologies. Whether we can conclude from this that it is unproblematic that certain (intrinsic) values are less frequently mentioned in the TECH dataset given the overall division of labour remains to be seen. We will discuss this question in Section 4.5.5.

4.5.2 Ethical research and early warning

The overall division of labour sketched in figure 8 suggests that ethics research has, among others, an early warning or early detection function when it comes to new ethical and social issues raised by new digital technologies and their merger. We can ask the question where ethics research is indeed playing such a role when it comes to the approaching technological storm. We have only data about developments over time for the ETHICS and TECH dataset, which limits our analysis somewhat. Still, we can make a number of interesting observations and draw some conclusions.

First, for the values of **privacy** and **justice and fairness**, Figure 6 and 7 seem to confirm that ethics research indeed played an early warning and detection function as we already discussed above. We also seem to witness such a role for ethics research when it comes to values like **democracy**, **transparency**, and **autonomy** (although these are less taken up in other datasets.)

For three values, ethics research does not seem to have played an early warning function: **reliability**, **cybersecurity**, and (environmental) **sustainability**. For reliability, this is largely explained by the fact that this is an instrumental value, rather than a moral intrinsic value, and it is to be expected that ethics articles pay less attention to such instrumental values (see also Section 4.5.1). A somewhat similar explanation may apply to cybersecurity. Cybersecurity is also usually seen as an instrumental value (van de Poel 2020), and may therefore receive less attention in the ethics literature although it is certainly an ethically relevant value (e.g. Christen, Gordijn, and Loi 2020) and more so than reliability.

The case seems different for sustainability which is usually seen as intrinsic value (Van de Poel 2017b; Dobson 1998). It has also been noted by others that sustainability has been relatively ignored in ethical frameworks for AI (Bird et al. 2020), although there has been (increasing) attention for it more recently (as can be seen from Figure 7). One explanation might be that sustainability as a value, and moral concern, is not specific for digital technologies, and that ethics research tends to focus on moral concerns and values that are distinct and characteristic for digital technologies. Still, if we expect ethics research to fulfil an early warning function with respect to digital technologies, this may be considered problematic. This is also one of the reasons why we have named sustainability and energy use as an important challenge of the approaching technological storm (Section 5.3.5).

There is also another reason why we may doubt whether ethics research is fully playing an early detection function. This is the almost exclusive focus in this research on AI and robotics (see figure 3 and Section 4.4.2). A possible explanation is that many of the other digital technologies are in fact discussed under the umbrella of AI. However, this would raise the question to what extent digital technologies and related ethical concerns are discussed with sufficient precision. This is particularly the case because other digital technologies raise value issues as well, as shown in Figure 4 and as we have seen in Section 4.4.3, the merger of digital technologies is likely to lead to new or at least an increase in moral issues and values that need to be addressed. So, it is certainly not true that the ethical issues raised by the new technological storm are mainly due to AI (which would justify a focus on AI in ethics research).

One caveat is that one might argue that the early warning function is not just to be played by ethics research but by a broader category of ELSI research. It should be noted that our ETHICS dataset also includes some interdisciplinary journals (like the Journal for Responsible Innovation, Science, Technology and Human Values, and Big Data and Society) that publish ELSI research rather than just ethics research (see Appendix 3).

All in all, there might be a reason, on the basis of these results, for trying to strengthen the early warning function of ethics and ELSI research, particularly by focusing on other technologies than AI, like IoT and blockchain, and particularly on the new issues raised by the merger of digital technologies. The policy option of an EU Observatory for converging digital technologies (Section 6.2.3) is intended to address this concern.

4.5.3 Values in the news

If we again follow the suggested division of labor depicted in Figure 8, news media have a role to play in bringing issues and values, discovered for example in ethics and ELSI research, in relation to digital technologies to a larger societal audience and may to some extent reflect societal priorities. Our results show that the news indeed seems to play such a role; most values are well discussed in the news (see Figure 2 in Section 4.4.1).

However, there are two values that are prominent in the ETHICS dataset but (still) hardly discussed in the NEWS, namely (human) **autonomy**, and **transparency** (see Figure 2). A possible explanation might be that journalists or press agencies do not refer to these values in news articles because they are considered too specialist or esoteric, instead using other value terms like democracy, and justice and fairness. It might also be that journalists consider these values societally less important (and therefore not worth referring to), but that would seem a less likely explanation. At least from a moral point of view, these would seem to be important values that also need to be addressed in the public debate.

A final point is that it is conceivable that values, and moral issues, mentioned in the NEWS dataset do not fully reflect the issues actually raised by new technologies, for example because people or the media are not well-informed. We have found no evidence for that, but nevertheless in general it would be important to increase what has been called digital literacy (Section 6.2.4). Another reason why the NEWS dataset (as well as the ETHICS dataset) may not reflect all relevant values is the fundamental uncertainty and unpredictability that accompanies the new technological storm (Section 5.3.9).

4.5.4 Values in regulation and legislation

Figure 2 (Section 4.4.1) suggests that most values are well addressed in regulation and legislation. One value that is still hardly addressed is **wellbeing.** This value is also still hardly addressed in the three other datasets which might suggest that this value is not so relevant for digital technologies. However, as we will discuss in chapter 5, particularly in the section on affecting people's intimate life (Section 5.3.3), there might be good reasons to think that the value of human wellbeing will increasingly be at stake with the merger of digital technologies. This is a value that therefore might require more attention, as also reflected in the policy option to protect the personal sphere beyond privacy (6.2.7).

Although most other values are well addressed in the LEGAL dataset, it cannot be concluded that current regulation and legislation is adequate for all the new moral and social challenges raised by the new technological storm, as that might require quite detailed rules and regulation. For example, as was already mentioned in the introduction, blockchain seems to pose new challenges for the GDPR (see Appendix 4). A detailed analysis of the adequacy of current legal and regulatory frameworks is beyond the scope of this study. One potential concern is nevertheless that the current regulation and legislation is very much focused on AI and robotics, and pays far less attention to other technologies (see Figure 3 in Section 5.4.2).

4.5.5 Does technological research sufficiently address values?

According to the division of labour depicted in Figure 8, technological research and innovation may not only raise new ethical and social concerns, but also need to play a role in properly addressing these. This is also the core idea of Responsible research and innovation (Section 6.1).

When we look at Figure 2 (Section 4.4.1) and Figure 6 (Section 4.4.4), it seems that values like reliability, cybersecurity, and privacy are well translated into technological research and innovation. For other values like democracy, autonomy, transparency, and wellbeing this is (still) less the case. This seems particularly problematic for **democracy**, as this value is frequently mentioned in the ETHICS and NEWS dataset (see Figure 2) and is already prominent in the ETHICS dataset since 2000 (see Figure 7). We will therefore focus our discussion here on this value.

One possible explanation why democracy is not so often mentioned in the TECH dataset might be the distinction between instrumental and intrinsic values discussed before (Section 4.5.1). Technological solutions based on instrumental values like reliability and cybersecurity and on more intrinsic values like privacy may be instrumental in addressing the value of democracy and related moral and social concerns. Still, it might be considered problematic when a value of democracy is never explicitly mentioned in technological and scientific literature because there are different notions of democracy which would translate in different solutions and innovations, and therefore in different technological research priorities (see e.g. Bozdag and van den Hoven 2015). In other words, simply assuming that technological solutions addressing privacy or cybersecurity will also address democracy concerns will not do. A recent paper on values encoded in machine learning research, indeed suggests that researchers tend to be vague about how their projects relate to societal needs and values. Consideration of potential negative effects of their project is very rare (Birhane et al. 2021).

Another possible explanation might be that, unlike other values like privacy and cybersecurity, democracy is a value that mainly needs to be addressed through regulation and legislation rather than through technological innovation. For example, by manually going through documents on AI and democracy in the NEWS dataset, we encounter discussions of potential threats caused by AI on democratic processes and its effects on people's political opinions. It might be that such threats are hard to address through technological solutions, and primarily require regulation. However, democracy is not a very prominent value in the LEGAL dataset (see Figure 2). Moreover, a manual investigation of articles on artificial intelligence and democracy in the TECH dataset shows that articles tend to discuss supervised learning algorithms and natural language processing as potentially helpful for classifying political opinions and detecting fake news. This suggests that there are also potential technological innovations and research priorities to better address the value of democracy.

One way to make technological research and innovation more responsive to societal and moral values is to foster the approach of Design for Values, as suggested in one of the policy options in Chapter 6 (Section 6.2.5).

4.6 Conclusions from our text mining exercise

We present the main conclusions with respect to how values are addressed in the four social realms studied (ethics research, the news, regulation and legislation, and technological research) against the (desirable) division of labour between these realms sketched in Figure 8.

1. We have suggested that **ethics** and ELSI research can play an early warning or early detection function when it comes to new ethical and societal issues and challenges raised by the merger of digital technologies. The results of our text-mining analysis suggests that ethics (and ELSI) research indeed play such a role when it comes to a number of values (like privacy, fairness and justice, democracy, autonomy and transparency). It does however less so for the value of sustainability. Moreover, it tends to focus rather one-sidedly on AI and robotics, hardly explicitly addressing values and issues related to other digital technologies. This may be improved through the policy option of an EU observatory for converging digital technologies that will be discussed in chapter 6 (Section 6.2.3).
2. The **news** may have to play a role in bringing relevant values, and ethical and social issues and challenges to the attention of a larger audience. Our analysis shows that at least the news articles in our dataset (mainly from Reuters and addressing the relevant technologies) do so rather well for most values and technologies, but values like autonomy and transparency seem somewhat underrepresented.
3. **Regulation and legislation** may have a crucial role in addressing some of the new challenges raised by the merger of digital technologies. Most relevant values seem well represented in legal and regulatory documents, apart from democracy and wellbeing, which might require more attention in the future (see also the sections Section 5.3.3 and Section 6.2.7). Also, the documents seem to focus primarily on AI and robotics potentially neglecting other technologies. Our analysis only concerns attention for values in legal and regulatory documents at a very general level and does not provide insight into whether new legislation might be required to address specific challenges.
4. **Technological research and innovation** may be important in addressing some of the newly raised social and moral concerns and challenges, certainly if one follows the approach of responsible research and innovation. Our analysis suggests that values like reliability, cybersecurity, privacy and sustainability are already well addressed, but that other values, and primarily the value of democracy, might require more attention in the future. Here the policy option of institutionalizing design for values (Section 6.2.4) might provide a way forward.

In addition to these four conclusions, a main conclusion is that the merger of digital technologies indeed seems to raise new or at least increased ethical issues. In the next chapter, we will explore in more detail what these new challenges might be.

5. What may emerge from the blend: features, opportunities and challenges

5.1. Introduction

Where in chapter 3, we focused on individual technologies and related social and ethical challenges, here we will focus on the combination or merger of these technologies. This combination results in new applications. For example, voice assistants - such as Alexa or Google Home - combine technologies like AI, IoT and voice recognition. The Metaverse will likely combine IoT with AI, VR/AG, blockchain and 5G/6G.

The merger of technologies, and its consequential social and ethical effects, however does not just result in new applications, it also results in new sociotechnical systems. Sociotechnical systems are combinations of technologies, human agents and institutions that serve certain functional purposes (Ottens et al. 2006). Smart cities are a typical example of a new sociotechnical system that is created as the merger of the technologies discussed in previous chapters. They combine AI, IoT, and sensing, but also, for example, face recognition, 5G/6G, and robotics.

Also, many other sociotechnical systems are increasingly shaped by the combination of technologies discussed in this study; think of the traffic system, or the financial system but also the political system and the government are increasingly affected. Moreover, socio-technical systems do not just exist in isolation, but they are often part or larger socio-technical systems. For example, we might focus on an individual hospital as a socio-technical system, which is increasingly shaped by the use of AI, IoT, robotics and 5G; this socio-technical system is part of a larger socio-technical system, like a combination of hospitals, which again maybe part of a larger (national) healthcare system. Sociotechnical systems typically are nested and partly overlapping, and they form systems of systems. As we will see in the section on the blurring of social spheres (Section 5.3.2), some of the new challenges arise precisely because increasingly (socio) technical systems are created that span the boundaries of previously largely disconnected social spheres.

Moreover, these new sociotechnical systems are not just used like we use an app or voice assistant, but they provide the backbone or infrastructure of many, if not all, of our social activities. Rather than being explicitly used, they are becoming the environment in which we are acting, and which is often taken for granted and in that sense 'invisible.'

New challenges may be created by the merger of digital technologies in three distinct but related ways: (1) the combination of enabling or general-purpose technologies like AI, IoT, 5G/6G and blockchain may create new technological possibilities and features that extend beyond what these general-purpose technologies individually allow, (2) this is likely to result in many new applications that may raise quite specific new social and moral issues. The way the technological storm will impact healthcare is likely to be rather different from how it impacts for example the judicial system or city life, (3) it is likely to affect to affect existing socio-technical systems and lead to the creation of new sociotechnical systems and systems of systems which may raise their own challenges.

Although not all new challenges raised by the new technological storm can be anticipated, we will nevertheless attempt to identify some of them. The way we will do so is by first (Section 5.2) identifying a number of features that we believe to be characteristic for the new technological possibilities created by the merger of digital technologies like AI, IoT, robotics, 5G/6G, and blockchain. Next (Section 5.3), we discuss some of the new challenges that the combination of new features may create. Finally (Section 5.4), we sketch some examples of newly emerging applications and application domains. This section does not aim to be exhaustive but rather aims to illustrate some of the emerging applications and the specific opportunities and challenges they raise.

5.2. Features of new technological applications

The merger of technologies like AI, IoT, 5g/6G results in technological possibilities and features that extend beyond those of individual enabling technologies like AI, IoT and blockchain. These features partly stem from the individual technologies that go 'in the mix'. For example, features like interactivity, autonomy, intelligence and autonomy are typical characteristics of AI systems (van de Poel 2020). However, some features also emerge due to new combinations of technologies. Moreover, it is often the combination of the features that creates new challenges for society, policy making and regulation. Although not every new application or sociotechnical system may combine all these features (most only combine some of them), the list of features nevertheless is helpful in identifying new challenges we are confronted with as society.

- **Ubiquitous.** The user is engulfed and immersed by IoT and there are no clear ways of opting out of a fully-fledged IoT, except for a retreat into a pristine natural and artifactless environment, which will be hard to come by in the remainder of the 21st century (Van den Hoven 2012).
- **Fully connected.** High and unprecedented degree of connectivity between objects and persons in the IoT network, leading to a high degree of production and transfer of data (Jeroen van den Hoven 2012). Increasingly, IoT nodes and sensors are everywhere and cover the whole world. We are moving to a situation where everything is connected to everything, although (Nguyen et al. 2021) effective full coverage may require 6G.
- **Interactive.** Particularly, developments in AI, robotics and IoT allow the creation of interactive technologies and applications. These can interact with the environment and with people; They process information, often collected through sensors, and they can act upon this data through actuators.
- **Long-distance.** The internet of things makes an increased level of automation possible, in which more actions than before can be carried out at a distance.
- **Distributed.** Particularly technologies like blockchain (ledger) and edge computing also the creation of distributed systems without a centre of control. This has clear advantages and offers opportunities in terms of privacy and vulnerability, but also makes governmental control difficult, responsibilities unclear and may create new (cyber) security risks.
- **Autonomous.** Technologies like AI allow the creation of applications and technological systems that can function on their own without human input.
- **Unpredictable and uncertain**. IoT environments may present spontaneous interventions (not directly caused by human agents or operators) and emergent behaviours (unforeseen and unexpected).
- **Intelligent.** Increasingly, technologies and applications are intelligent, i.e. they can carry out tasks that would require intelligence when carried out by humans.[17]
- **Adaptive.** In particular due to the employment of AI, applications and systems can learn and adapt themselves in response to inputs from the environment. This makes them very effective and flexible, but also introduces risks; it may e.g. create bias or systems may learn undesirable behaviour.
- **Reconfigurable**. Sensors, which play an important role in the internet of things, increasingly become reconfigurable. This means that it is possible to start using them for functions that were initially not anticipated. This poses new challenges for making sure that technological applications are aligned with our norms and values (Dechesne, Warnier, and van den Hoven 2013).
- **Hybrid and ambiguous.** IoT is characterized by an integration of the physical and digital world. It leads to cyber-physical systems. Physical objects become increasingly digital in the sense that they

[17] This formulation deliberately leaves open to the question whether the systems themselves can be meaningfully called intelligent. For example, AI systems are obviously better in some tasks than humans, but they still lack what has been called 'general intelligence'. They also often do not perform well in situations that are somewhat unexpected and require 'common sense.' General artificial intelligence may be a possibility and concern for the far future, but it does not seem a feature of the upcoming technological storm.

are connected through IoT and/or are represented through digital twins. Digital objects may be hard to distinguish from physical objects in VR/AR. By using nanotechnology (sensors), human bodies can also become part of the IoT. So, the distinction between natural objects, artifacts and human beings tends to blur. Identities and system boundaries become more ambiguous (Van den Hoven, 2012).

- **Invisible.** Digital technologies are increasingly becoming 'invisible' (WRR 2021). That is to say: they merge in the background or are embedded in systems or infrastructures that operate in the background but are not actively perceived by their users. Most people are unaware of the ubiquity of sensors in their environment. This invisibility would be further increased when nano-technology gets integrated in the IoT (see e.g. Kuscu and Unluturk 2021).
- **Fast.** 5G/6G networks will allow very fast transfers of very large amounts of data. Quantum computers may in the future allow very fast processing of such data, allowing forms of machine learning in almost real-time. This enables the creation of digital technologies that autonomously interact with people in real-time and so to close the 'feedback loop'. That is to say, sensors can very quickly gather data about people, through machine learning data can be processed and inferences can be drawn from it, leading to almost immediate interventions. The human reaction to this intervention can then be directly processed again and be used for further inferences and interventions. One might, for example, think of screens in shopping malls that show people products based on such data as their eye movement, their way of moving and past shopping behaviour, while the algorithms processing these data also learn from the effectiveness of previous offers presented.
- **Precise in location**. The expectation is that with 6G, the determination of geo-location can be done with an accuracy of less than one centimetre.[18]
- **Intimate.** New technologies lead to increasing possibilities to collect very personal and intimate data on people. Such intimate data can be used in ways that are beneficial for people (e.g. innovations in health care), but also in the interest of business and institutions (e.g. marketing and surveillance). The movement of your eyes can for example reveal what you pay attention to and thus what is in some way important to you, data on the way you move your head and body in virtual reality applications can be turned into a personal profile that can be used to identify you on the street, and sensors in your toilet may collect data on your faeces in order to draw conclusions about your health. New opportunities for collecting intimate data may be created with the integration of bio-nanotechnology in our bodies, which then become connected to the internet of things (Kuscu and Unluturk 2021). Given all the intimate 'digital data from within and around the living body' that is nowadays being transmitted, (Boddington 2021) even speaks of an 'Internet of Bodies'.
- **Immersive and persuasive.** Technologies like virtual reality (VR) and augmented reality (AR) allow the creation of technological systems in which people may feel fully immersed and which might therefore be very persuasive. They may also make it increasingly difficult for people to distinguish real from virtual. While deep fakes are (typically) created intentionally, the distinction between real and virtual may also increasingly get blurred unintentionally, as it might become harder to distinguish real from virtual experiences. Moreover, even if people know that certain (immersive) experiences are not real, they might still have persuasive effects, and result in for example manipulation (Schoenmakers et al. 2022).
- **Commercially exploitable.** The merger of technologies creates many new business and economic opportunities. It also makes services that were public in the past potentially commercially exploitable. An example are street lights that 'go in business' by adding a digital wallet to each individual lamp post through blockchain. (Section 5.4.5). In this way, street lights can be commercially exploited, and what used to be a public service might become a private good.

[18] Talk by prof. Vincent Poor at the STOA event on Edge computing, 6G and satellite communications (1 December 2021)

5.3. Opportunities and key challenges

Based on the features discussed in Section 5.2, and inspired by the interviews with a number of experts, we have identified nine key opportunities and challenges: (1) Digital sovereignty and new economic and social opportunities; (2) the blurring of social spheres; (3) an increased impact on people's intimate life; (4) opacity and cognitive overload; (5) energy use and sustainability; (6) increased cyber risks, and new cyber-physical risks; (7) disruptive effects; (8) the concentration of techno-economic power; and (9) fundamental unpredictability.

None of these challenges is completely new, that is to say each of them is also raised by some of the individual technologies that 'go in the blend.' However, there seem to be good reasons to assume that these challenges are particularly aggravated by the merger of AI with IoT, blockchain, 5G/6G and AR/VR, as well as with robotics and nanotechnology. The main reason to think so is that these challenges are caused or at least aggravated by the features of the blend we discussed in Section 5.2.

A lot of the ethical discussion about digital technologies in the past few years has focused on AI, as also shown by the results of our text mining analysis in chapter 4. Therefore, it is important to note that the challenges we have identified extend well beyond those that are typically or usually discussed in the AI ethics literature (for the latter, see also section 3.2). This suggests that in order to address the challenges of the new technological storm, we might well need to look for policy options and regulation that extend beyond the realm of AI and the concerns it has raised.

5.3.1. Digital sovereignty, economic prosperity and social benefits

How this challenge is affected by the features discussed in Section 5.2:

- **Fully connected**: can potentially help to address societal challenges and create new economic and social opportunities.
- **Commercially exploitable**: may create new economic opportunities.
- Most of the **other features** contribute to new technical functionalities and possibilities, which may also create economic opportunities or contribute to solving societal challenges.

The merger of digital technologies offers tremendous commercial and societal opportunities. They may allow the commercial exploitation of goods and services that traditionally belonged to the public domain; an example is given in Section 5.4.5 of 'street lights going into business. In this way, the merger of digital services is likely to lead to new business opportunities for existing companies as well as for new start-ups. At a societal level, this may contribute to economic growth as well as to the quality of services. Digital technologies may also disrupt existing markets (Section 5.3.7), as well as contribute to the creation of digital platforms that allow new business models and opportunities (Section 5.3.8).

The merger of digital technologies also offers tremendous opportunities to better meet societal needs and to address societal challenges. For example, they may enable the energy transition through smart grids (requiring IoT and AI), and smart energy contracts (using Blockchain). The merger of digital technologies may also provide new opportunities for reducing energy consumption (Section 5.3.5). Increased connectivity may make products and services more reliable and improve their quality (Section 5.3.2).

Seizing these economic and social opportunities may require economic incentives as well as an industry policy. One particular concern at the EU level here is what has become known as technological sovereignty or digital sovereignty. This term refers to the need for Europe to develop its own capacities and to reduce its dependency on other parts of the globe for key enabling technologies like AI

(Ramahandry et al. 2021). It also includes the need to ensure the integrity and resilience of data infrastructure, networks, and communications. In the words of Ursula von der Leyen, this is required to enable Europe 'to make its own choices, based on its own values, respecting its own rules' (cited in the STOA study by Ramahandry et al, 2021).

Digital sovereignty, and more generally technological sovereignty, may be increasingly important in the light of recent events like the Corona pandemics and the war in Ukraine. It may also be needed to decrease the dependence on international big Tech companies for digital technologies (Section 5.3.8) and to ensure economic and social prosperity in the EU.

5.3.2. The blurring of social spheres

How this challenge is affected by the features discussed in Section 5.2:

- **Fully connected**: the blurring of social spheres is enabled by the full connectivity of (socio)technological systems
- **Commercially exploitable**: may lead to the introduction of profit or financial motives in other social spheres
- **Long distance**: may contribute to the blurring of social spheres
- **Intelligent**: makes also social activities and spheres that initially were exclusively human open to computation, and may so contribute to blurring with other social spheres
- **Autonomous and adaptive**: data from different social spheres may be combined without human intervention
- **Hybrid**: may contribute to the blurring of social spheres

Sociotechnical systems are increasingly interconnected, which is made possible by the feature of *full connectivity*. Sociotechnical systems within one social sphere, like for example health, are increasingly interconnected, resulting in systems of systems. This interconnectivity may raise the quality of, in this case, health care services and may make these more efficient. It may, for example, prevent people from getting wrong medication, but also may give caretakers at the bedside in the hospital immediate access to relevant information through the IoT. Obviously, interconnectivity, apart from having such advantages, also raises privacy issues, which are well known and are being increasingly addressed in regulation like the GDPR.

Not only sociotechnical systems within one social sphere are increasingly connected, but there are also increasing connections between systems from different social spheres. This may result in information flows between social spheres that are considered problematic. Most people would feel it inappropriate that their medical data is shared with their colleagues or with their insurance company, while sharing with the hospital or their medical doctor may be fully appropriate and even desirable. Addressing such issues, Nissenbaum (2010) has proposed to understand privacy in terms of contextual integrity. On such an understanding of privacy, we should understand it in terms of appropriate information flows, where appropriateness is dependent on differences in context or social sphere. This notion of privacy is different from understandings of (informational) privacy in terms of confidentiality, or in terms of informed consent; the latter notion is particularly important in the GDPR (Appendix 4). But understanding privacy in terms of appropriate information flows determined by contextual integrity would mean that for some information transfer, informed consent may not be required, while others information flows may be considered inappropriate even if people would give informed consent. Technologies like AI, blockchain and edge computing may be enablers for integrating rules for proper information flows in IoT applications.

However, addressing the challenge of blurring social spheres requires not just managing information flows between social spheres. The reason is that relevant values and norms and appropriate ethical standards may well be different for different social spheres. What may be appropriate behaviour among friends or family may be unacceptable for example in a professional work setting. As Loke (2021) says, 'what is considered ethical behaviour might depend on the context of operation and the application - a device's action might be considered ethical in one context but unethical in another [...].' They go on to suggest that 'it would require multiple levels of norms and ethical rules to guide the design and development of IoT devices and ecosystems: a basic ethical standard could apply (e.g., basic security built into devices, basic user-definable data handling options, and basic action tracking), and then additional configurable options for context-specific ethical behavior added.' While this may be a useful design strategy, it does not fully solve the issue that devices and applications may well cross the boundaries of the context for which they were initially developed. However, as also pointed out by Walzer (2008), the principles of (social) justice are different between different social spheres, and in order to sustain justice, it might be required to ensure that technological applications do not cross the boundaries of these spheres (cf. also Nagenborg 2009).

The approaching technological storm also blurs the boundaries between different social spheres in another way. As we saw in Section 5.2, one feature is that it creates many new possibilities for *commercial exploitation*, also for services and social spheres that in the past were either public or not part of the (formal) economy. Commercial apps are, for example, already becoming rather common in healthcare or in dating, social spheres that were traditionally distinct from the economic or business sphere.

The challenge of a blurring of social spheres may be particularly hard to address because most of the technologies that enable full connectivity are general purpose technologies and are often applicable in different social spheres or contexts. Designing such technologies for specific application domains or contexts may be unattractive in terms of economies of scale and therefore require new business models. It may also require applying a design principle like 'capability caution', which means that an application is designed with the minimal technical capabilities required for its intended use, so that it is harder to use for other purposes or in other contexts (Cawthorne and Devos 2020; Floridi et al. 2018). However, capability caution seems at tension with current innovation practice in industry.

A larger issue is whether full connectivity is indeed desirable (or unavoidable). There are clear advantages of increasing connectivity, in terms of efficiency and potentially in terms of quality of services. One might also argue that boundaries between social spheres are not given by nature but socially construed and can change over time. Still, there is the issue that some boundaries between social spheres are likely to be desirable, in order to sustain social justice, and also to protect our personal intimate sphere from intrusion by other social spheres (Section 5.3.3). These are not just ethical but also political questions, but they also seem to require choices with how we give shape particularly to the IoT as the technical backbone or infrastructure for social and political life (Section 5.3.8).

5.3.3. Affecting people's intimate life

How this challenge is affected by the features discussed in Section 5.2:

- **intimate**: directly contributes to the challenge
- **immersive and persuasive**: by becoming part of our bodies and (implicitly) part of our motivations
- **invisible**: makes the effect on our intimate life more ethically problematic
- **commercially exploitable**: may introduce profit or financial motives in our intimate life
- **interactive**: makes these technologies more effective but also potentially more manipulative when it comes to our intimate life
- **hybrid**: technologies become part of our body

New applications increasingly affect people's most intimate life in several senses. First, they are used to collect intimate and sensitive data about people like information about moods and emotions (e.g. through tracking eye movements, brain activity, blood pressure and composition, etc.). Second, they can literally become part of people's bodies, for example through the use of nanotechnologies and what has been called the Internet of Bodies (Boddington 2021). Third, people increasingly use these technologies for very intimate activities, like making and maintaining friends, dating and sex (Stark and Levy 2018; Levy 2015).

The ethical and legal paradigm that is still largely used for properly dealing with the consequent ethical and social issues is informational privacy, often operationalised as informed consent. While the GDPR now enforces such a privacy notion for the development and use of many new technologies, some of the new challenges raised extend well beyond informational privacy.

First, the privacy issues at stake are not just informational, but also have other components like spatial ones. For example, small-scale pilots with the use of Google Glass showed that people were concerned not just about information flows, but felt uncomfortable with a dining partner wearing Google Glass which was felt as privacy intrusion independent from whether it was used to collect information or not (Kudina and Verbeek 2019; Van de Poel 2018). Similarly, people wearing Glass in public were sometimes approached aggressively (Honan 2013). Privacy is thus a concern also if there is no information sharing. It has been argued by (Koops et al. 2016) that privacy has nine dimensions, which may all need consideration.

A second reason why the new challenge cannot be addressed by informational privacy alone is that new applications are likely to combine the features of *intimacy, immersiveness* and *persuasiveness.* This seems to raise new challenges that extend well beyond privacy. In particular, it requires attention for human vulnerability, particularly because these technologies may increasingly affect how humans develop their personal and social identity. Moreover, some of the developed technologies can get literally under our skin. It therefore also requires attention for how human vulnerability and identity formation are connected to our body and how embodied technologies might either enable or constrain (or even manipulate) such processes.

Addressing these issues will likely require more attention for values like wellbeing and human dignity, as well as for human rights. The latter two are now increasingly addressed, for example in the recent proposal for a declaration on European digital rights and principles.[19] There have also been calls to pay more attention to neurorights (Ienca? 2021). However, a value like wellbeing is currently only

[19] On 26 January 2022, the Commission proposed the European Parliament and Council sign up to a declaration of rights and principles that will guide the digital transformation in the EU, as can be found here: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_452

marginally addressed, as our analysis in Chapter 4 showed: in all relevant societal realms that we discussed (ethics research, the news, regulation and legislation, and technological research), the value is one of least mentioned values in relation to the technologies of the approaching technological storm (Section 4.4.1).

One barrier to a stronger focus on wellbeing as value might be - certainly for regulation and legislation - that liberal governments have traditionally considered ideas about what constitutes wellbeing or a good, flourishing, life to be part of the private sphere that should not be intruded by the government. While that might still be an important consideration, it would seem desirable that wellbeing concerns get a more prominent role in public discussions as well as in technological choices with the increasing merger of digital technologies.

Dealing with this barrier may require us to reconsider our understanding of both wellbeing and freedom. One option for doing so could be to understand wellbeing in terms of human capabilities as proposed by philosophers like Martha Nussbaum (Nussbaum 2011) and Amartya Sen (Sen 1999): the effective opportunities people have to realise valuable 'being' and 'doings'. Particularly Martha Nussbaum has suggested that there are a number of universal human capabilities that are required for human wellbeing; promoting these general capabilities through policies still allows people to make individual choices when it comes to questions about wellbeing and the good life.

If one follows this line of thinking, digital technologies should at least not endanger these universal human capabilities, and - if possible - foster them. Several authors have also suggested that human capabilities can be a useful starting point for Design for Values and have developed more specific approaches for designing for human capabilities (Oosterlaken 2015; Jacobs 2020). One relevant capability one might think of is what Nussbaum (2011) calls 'practical reason', which has to do with the ability to make reasoned decisions about one's own life. The attention for values like autonomy and explainability in AI ethical frameworks may be seen as related to this capability. But the approaching technological storm is also likely to affect other capabilities, mentioned by Nussbaum, such as 'bodily integrity', 'senses, imagination and thought, 'emotions', and 'affiliation.'

5.3.4. Opacity and cognitive overload

How this challenge is affected by the features discussed in Section 5.2:

- **invisible**: directly contributes to opacity
- **Interactive**: allows systems to react to human behaviour in ways that they may hard to understand or opaque for humans
- **commercially exploitable**: may lead to opacity due to property rights on code
- **autonomous and adaptive**: makes opacity ethically more problematic
- **intelligent**: systems may be opaque for decisions that require an explanation or justification
- **fast**: contributes to cognitive overload
- **immersive and persuasive**: makes opacity ethically more problematic because it may lead to manipulation

The approaching technological storm is likely to lead to new applications and (socio) technical systems of which the functioning is hard if not impossible to understand for users and the general public (opacity), or - if these technologies are so designed to provide data and explanations to users and the public - to cognitive overload due to the amount of information that need to be processed by humans in often short time intervals.

The opacity of the new applications and (socio) technical systems has roughly three causes. A first is the use of machine learning (ML) techniques and neural networks in AI that learn, and represent information in ways that are (sometimes) fundamentally incomprehensible for humans. The field of explainable AI (XAI) is already addressing this challenge, which has led to various techniques that add, for example, an explanatory interface to ML systems (Schoenborn and Althoff 2019). A second cause is that many of the applications are developed in *commercial* settings, so that the exact codes and algorithms may be proprietary knowledge and not be publicly available (Ferguson 2017). A third cause is related to the feature of *invisibility* discussed in Section 5.2. New applications and systems may be invisible in the sense that they are increasingly not explicitly used but form the taken-for-granted background of daily activities. For example, people walking through a smart city may not be aware that their itinerary is sensed and that they are nudged into, for example, avoiding a traffic jam or an accident by technologies in their environment, like for example traffic lights or electronic signposts. In addition, the technologies may also become literally invisible as nano-sensors may become part of the human body. In this way, invisibility contributes to the opacity of new applications and (socio) technical systems to users and the general public.

The problem of a cognitive overload is in a sense the mirror of the problem of opacity. Where opacity is typically due to a lack of information given to users or the general public, cognitive overload is typically due to an overload of information or the incomprehensibility of such information. It is, however, important to see that both issues are connected because attempts to overcome the opacity of the newly created (socio) technical systems may well result in cognitive overload.

This is first of all due to the sheer amount of data collected and processed by such systems, which is clearly more than humans can cognitively process. Second, as indicated earlier, ML applications may be made explainable, but the resulting explanations may still be hard to completely grasp for lay people and even for the professionals working with such systems. Third, information can increasingly be collected and processed so fast, that interventions based on the output of AI/ML algorithms fed with these data may be almost made in real time; so, the *fastness* of the new applications may give humans simply too little time to comprehend how they may be influenced or nudged, or even manipulated by these new technological applications.

Due to their *autonomy, interactivity, intelligence and adaptability*, particularly new technological applications and systems that employ AI/ML are very effective in influencing human behaviour, while at the same time the exact working of these systems is likely to be opaque or at least very hard to comprehend for humans. Santoni de Sio and van den Hoven (2018) therefore have argued for meaningful human control as a guiding principle for developing new technological applications and systems based on AI. Meaningful human control means that these systems are ultimately controlled by humans in a meaningful way; 'meaningful' here, among others, means that the human in control is provided with relevant information that can be processed given her cognitive abilities and available time. According to the meaningful human control paradigm, the human in control need not be the user of the system, it can also be the operator of the system; like for example, flight control in the air traffic system; or it can also be the (human) designer of the system.

While meaningful human control may be hard to realise for AI systems, it becomes even more challenging to guarantee if AI is combined with other technologies like IoT, 5G/6G, AR/VR and nanotechnology, particularly because of system features such as invisibility, fastness, full connectivity and immersiveness and persuasiveness. This is a challenge because meaningful human control is not only important as a way to uphold an important moral value like human autonomy, but also because a lack of meaningful human control also has implications for responsibility and accountability. If we cannot trace back important decisions made by technological systems to some human, as meaningful human control requires, we might not able to hold anyone accountable for the failure of such systems or other harm resulting for their use; moreover no one would seem to be in the position to take forward-looking responsibility for the responsible development and deployment of such systems.

While the challenge of opacity and cognitive overload can partly be addressed through more responsible technological development and design based on such principles as meaningful human control and employing, for example, XAI techniques, it would also seem to require making society, individually as well as collectively, more resilient in dealing with digital technologies. This among others would require digital literacy because even if technological applications respect principles like meaningful human control; they are not likely to be employed and used in a responsible way, unless users have an awareness of their limitations and risks (Section 6.2.3).

5.3.5. Energy use and sustainability

How this challenge is affected by the features discussed in Section 5.2

- **fully connected**: contributes to energy use
- **commercially exploitable**: may make profit rather than sustainability and energy use a driving force
- **adaptive**: adaptive systems that employ e.g. ML are usually very energy-intensive
- **fast**: leads to more energy use (5G/6G)
- **distributed**: is likely to lead to more energy use (e.g. blockchain).

The increasing use of digital technologies may well increase total energy consumption. Particularly new technologies like 5G/6G, blockchain and AI can be very energy-intensive in use. At the same, new opportunities for increasing energy efficiency may arise, which may result in reductions of energy use. Moreover, different technological and design choices are possible, which will likely affect energy consumption.

Energy use and sustainability are by no means a new concern. For example, the declaration on European digital rights and principles[20], which was proposed by the European Commission in January 2022, puts quite some emphasis on sustainability. The text-mining analysis in chapter 4 also showed that sustainability is already regularly addressed in technological research, although it does not yet get much emphasis in ethical studies and medium emphasis in regulatory documents (see Figure 2 in Section 4.4.1). Other studies suggest as well that sustainability has received relatively little attention in ethical and legal frameworks for AI (Bird et al. 2020), which have tended to focus on 'digital' values and rights, like autonomy, fairness, privacy and explainability. But certainly, with the advance of technologies like 5G/6G and blockchain, energy consumption is becoming a challenge that urgently needs addressing for the responsible development and use of these technologies.

Malmodin et al. (2010) estimate that the share of the ICT sector in the global energy use was 3.9 % in 2007, contributing to 1.3 % of the GHG emissions in that year. Other studies suggest that this share has been growing since and is likely to continue growing. For example, (Andrae and Edler 2015) made an estimation of the global energy use that can be ascribed to communication technologies for the period 2010 to 2030. In their expected scenario, the share of communication technology in global energy use increases to 21% in 2030, with a best-case scenario of 8% in 2020, and a worst-case scenario of 51%.

Belkhir and Elmeligi (2018) made a trend assessment of the carbon footprint of ICT technologies, which includes not only energy use but also, for example, emissions from production. They expect an increase in the ICT carbon footprint from about 4% of the total world-wide carbon footprint in 2020 to 14% in 2040 (with a lower bound of 6 %)). This footprint is largely caused by ICT infrastructure (data centres

[20] On 26 January 2022 the Commission proposed the European Parliament and Council sign up to a declaration of rights and principles that will guide the digital transformation in the EU, as can be found here: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_452

and networks), with a relative share of 61% in 2010 rising to 79 % in 2020. Smartphones make a relatively large contribution as well (around 11% in 2020).

When it comes to the effects on energy consumption of digital technologies and the approaching technological storm, one should not only look at the energy use of digital technologies themselves but also how they affect other technologies and activities. For example, smart grids and smart meters offer opportunities for more efficient energy use, and may in effect contribute to a reduction of energy use (Hu et al. 2014). On the other hand, it is conceivable that the Metaverse may invite (digital) activities that further increase energy use.

It has indeed been suggested that AI, IoT and even 5G/6G may lead to drastic energy savings in other sectors, even to the extent that they in effect reduce net energy consumption (Cunliff 2020). Such claims have, unlike some of the studies cited above, not been supported by detailed scenarios and estimates, and it is therefore hard to say how realistic such projections are. Moreover, any forecast of future energy use of these technologies, and certainly their effect on other sectors, is bound with great uncertainties. Nevertheless, it is clear that while the share of digital technologies in total energy use and carbon footprint may well increase in the coming decades, there is also a large potential for energy savings and reduction in carbon footprint if these technologies are used to reduce energy consumption and for other sustainability purposes in others sectors, like transport, the industry and energy production and consumption.

Reducing total energy use and carbon footprint requires not only certain choices in the employment of digital technologies but also additional efforts to reduce the energy consumption of these technologies. Our text mining analysis (see chapter 4) suggests that sustainability and energy use is already a major concern also in the technological literature. Technological and design choices may matter greatly for energy use. For example, worries about the energy use of blockchain are largely based on the large amount of energy used for Bitcoin. However, energy use of cryptocurrencies (and more generally blockchain) may not scale linearly with use (Sedlmeir et al. 2020). Moreover, a main reason for the large energy consumption of Bitcoin is the used consensus mechanism, while other architectures for blockchain are likely to consume much less energy (ibid).

There seems therefore a need to design digital technologies more explicitly for sustainability. In the literature, there have also been pleas for, for example, green AI (Schwartz et al. 2019; van Wynsberghe 2021). Such technological and design choices for sustainability may well require trade-offs with other values. For example, it is conceivable that blockchain architectures that require less energy may be less secure, so that a trade-off between sustainability versus privacy and security is required. Similarly, choices for green AI may require accepting some loss in performance or effectiveness of algorithms.

5.3.6. Increased cybersecurity risks and new cyber-physical risks

How this challenge is affected by the features discussed in Section 5.2:

- **hybrid**: physical risks and cyber risks get connected to each other
- **fully connected**: increases the risk of cascading effects in technological systems
- **autonomous, intelligent and adaptive**: may help to detect risks and abate them, but may also introduce new risks
- **distributed**: may reduce risks but also introduce new risks that are hard to (centrally) govern
- **fast**: may give less time for risk detection and mitigation
- **invisible**: people may be unaware of risks

The increasing merger of IoT, AI, blockchain and AR/VR is likely to lead to increased cybersecurity risks as well as new cyber-physical risks. Home-based IoT systems are increasingly being adopted (e.g., smart TVs, smart lighting, smart thermostats, smart security cameras or baby monitors) making smart homes vulnerable to cyberattacks. Home users tend not to have adequate risk awareness nor adequate digital skills to implement security measures on their internet-connected devices, increasing the risk of hackers gaining easy access to home IoT devices. Such security breaches can have a variety of adverse effects, including using IoT devices as part of a multi-device coordinated DDOS attack, access to personal information such as passwords or financial information, identity theft, scanning a home for the presence of people enabling physical burglary, stalking, sexual harassment, and extortion.

We already saw in the text-mining analysis that the main value of which the frequency increases if digital technologies are discussed in combination in the technological literature is cybersecurity (Section 4.4.3). But in addition to an increase in cybersecurity risks, there are also new cyber-physical risks. For example, through IoT, cyber-attacks on, for example, water treatment plants, chemical factories or nuclear plants are increasingly possible. The hybridity of the new systems of systems, with often IoT as backbone, make cybersecurity and physical safety and security increasingly interconnected rather than independent concerns. This is even more the case in so far as systems become fully connected. Full connectivity may also introduce additional risks, as failure of one component in the system may have cascading effects on the entire system (cf. Perrow 1984 on tightly coupled systems).

The increasing connection between the physical and digital domain also may create new challenges for properly managing risks. A distinction is sometimes made between safety risks (due to unintentional events) and security risks (due to intentional events) (e.g. Hansson 2009). Managing security risks typically requires other strategies than addressing safety risks, because intentional agents (humans) are involved that may deliberately adopt their strategies in response to risk reduction attempts (Rios Insua et al. 2021). In the physical domain, risk management is still often primarily based on safety concerns. While for example, terrorist attacks or sabotage already constitute security risks in the physical domain, with the increasing merger of physical and digital infrastructures, enabled particularly by IoT, security concerns are likely to increase considerably.

In the digital domain, cybersecurity is already a major concern. Here, new risks may arise not only due to increased connectivity (cascading effects), but also due to the addition of artificial agents, with self-learning and adaptive features, that may introduce new unintended risks. At the same time, AI may be instrumental, and even indispensable, in managing safety and security risks in the new systems of systems that are created (Taddeo, McCutcheon, and Floridi 2019). AI systems may for example be able to detect cyber-attacks or irregularities in the functioning of systems in ways that are impossible, or at least far less effective, by humans.

Addressing this challenge may require not only addressing safety and security concerns through add-ons to existing (and newly created) systems, but rather following a safe- and secure-by-design approach from the start, and aiming for systems that are more inherently safe or at least more resilient. For example, distributed and or redundant systems may offer possibilities for risk mitigation and increasing resilience.

5.3.7. Disruptive effects

How this challenge is affected by the features discussed in Section 5.2:

- **fully connected and long-distance**: due to their world-wide scale applications and technological systems do not fall under one jurisdiction and may be hard to regulate
- **distributed**: may undermine or make superfluous existing institutions
- **autonomous, intelligent and adaptive**: contributes to the disruptive potential of new applications and sociotechnical systems
- **commercially exploitable**: may lead to new business models and so to institutional and regulatory disruption

The new applications and sociotechnical systems that arise in the approaching technological storm may have disruptive impacts (Hopster 2021, Millar, Lockett, and Ladd 2018). Such disruptions are to be distinguished from mere impacts not just by their impactfulness and the fact that they may occur in a short time period (like shocks) but also by that such impacts may be irreversible. Although the exact disruptive effects of the approaching technological storm may be hard to predict at this stage, one can distinguish between four different kinds of disruptions: 1) disruption of existing (economic) markets, 2) disruption of social practices and institutions, 3) regulatory disruption and 4) conceptual disruption. We discuss each of these briefly below.

New digital technologies may disrupt existing markets (or create new markets) in the sense that they offer newcomers to such markets opportunities to enter these markets or even to completely take over such markets, particularly because new technologies may offer new opportunities to create value for users, or to serve new categories of users (Abernathy and Clark 1985, Christensen 2013). An example is the disruption of the mobile phone market by the introduction of smartphones, which made existing dominant companies (like Nokia) marginal, and newcomers to this market (like Apple) dominant. The digital revolution has led in the past decade to a number of big tech companies dominating many digital markets (Section 5.3.8). Given the highly connective character of new digital applications, these companies may be in a good position also to be leading in many more specific digital innovations. Nevertheless, for specific applications and markets, digital innovation may still have disruptive market effects, and offer opportunities to newcomers.

The merger of digital technologies also has disruptive effects on social practices and institutions. For example, AI may have disruptive effects on practices in such domains as policing, the juridical system and the health system. Blockchain is often cited as a system that allows trustworthy (and trusted) transactions without a central organization, like a central bank. It does, however, require social institutions to function properly and to remain trustworthy, also because its exact development cannot be predicted (Alston et al. 2022). Davidson, Filippi, and Potts (2018) suggest that blockchain may be seen as an institutional technology that allows new – more decentralized or polycentric – institutions and so may lower transaction costs. When it comes to the merger of digital technologies, particularly the first two challenges that we have discussed above (the blurring of social spheres and the deep intrusion of the personal sphere) may be potentially socially disruptive.

The merger of digital technologies may also disrupt existing legislations, regulation and governance structures. For one thing, existing laws to protect certain values or public goods may no longer apply to new digital technologies or applications. One reason is that sometimes relevant legal or regulatory terms have been defined in a way that (tacitly) assumes current technological possibilities. Another reason is the merger of digital technologies creates new challenges and threats that may require new regulation or governance. Many of such new regulations and legislation in relation to digital technologies have already seen the light in the past few years (Section 1.3.3).

New applications that arise because of the merger of digital technologies may disrupt current regulatory regimes and be hard to regulate because they can combine two features 1) they are fully connected and (potential) span the whole world. They do thus not fall under one jurisdiction, which makes them hard to regulate and there is also a danger of regulatory race to the bottom. 2) they can at the same time be distributed and decentralized, e.g. through the use of such technologies as blockchain, edge computing and federated learning. Consequently, they may function independent from existing regulatory regimes and institutions. The main example is of course Bitcoin and other crypto currencies that enable a financial system without a central bank. But there are many other examples, like smart contracts, new platform companies (like Uber) etc., particularly made possible by blockchain.

The approaching technological storm may also disrupt the very concepts and values by which we understand and evaluate its impacts (Swierstra 2013; Hopster 2022).[21] We have already seen that a challenge like digital technologies increasingly affecting our most intimate life does not just require a value like privacy, but also a value like wellbeing, and attention for related values and concepts like human dignity, human rights and human vulnerability (Section 5.3.2). More generally, social disruptions and challenges brought by the merger of digital technologies might require new values to properly address them (cf. Van de Poel and Kudina 2022). One such a value that has been emerging in recent years is that of technological or digital sovereignty (Section 5.3.1). Similarly, digital technologies like AI and robotics may challenge existing conceptions of for example (moral) agency, intentionality and responsibility; in order to answer fundamental questions like whether artificial agents and robots can – and if so should – be treated as moral agents.

The approaching technological storm may thus potentially be disruptive in different ways. This disruptive potential is a challenge but that does not mean that it is necessarily to be evaluated as something negative, or to be avoided. For example, new applications of for example blockchain may offer useful opportunities in contexts where existing institutions are weak or lacking, e.g., for protecting property rights of small farmers in developing countries (see Mintah et al. 2020). Second, at a larger scale, there seem to be reasons to believe that some of the current legal and institutional frameworks have proven not very effective in dealing with societal challenges like climate change, and world hunger.

Dealing with the potential disruptive nature of the new technological storm may require rethinking how we protect essential values and public goods through regulatory instruments and institutional safeguards. This task is urgent because as some of the previous challenges suggest the technological storm may affect social justice, people's wellbeing and intimate life, and lead to increased energy consumption and new or increased cyber-physical risks. It is also a daunting task, given the uneven distribution of techno-economic power over the development and employment of these technologies (see next section).

[21] See also www.esdit.nl

5.3.8. Concentration of techno-economic power

How this challenge is affected by the features discussed in Section 5.2:

- **Commercially exploitable**: directly contributes to the challenge
- **Fully connected**: further enables the concentration of techno-economic power
- **Immersive and persuasive**: makes the concentration of techno-economic power politically and ethically more problematic
- **Reconfigurable**: may make owners and shapers of digital structures more powerful vis-à-vis other players
- **Intimate**: makes the concentration of techno-economic power ethically more problematic

Many of the new applications that arise at the merger of AI, IOT, blockchain, and 5G/6G are developed by a handful of internationally operating companies, who are powerful and hard to regulate by individual governments. These companies are sometimes as powerful as individual states (Taylor 2021). Moreover, the services these applications offer do not only affect people's daily and intimate life (Section 5.3.2), but also sometimes concern services and goods that belong traditionally to the public sphere. For example, voting, government services, border control, and the juridical system increasingly depend for their functioning on these new digital technologies and applications. Consequently, in many cases, citizens cannot opt out from the use of these technologies, and there is a serious danger of domination of citizens not just by states, but also by big tech companies (Taylor 2021).

Others have suggested that we witness the emergence of what has been called 'surveillance capitalism' (Zuboff 2019). Surveillance capitalism is characterized by a business model in which big tech companies are able to make a profit by collecting large amounts of data about their users, which can be capitalized, for example, through the sale of targeted advertisements. The more these companies know about their users, the more valuable services they can offer, creating an incentive for collecting large amounts of data, and leading to a potential exploitation of users.

Moreover, the digital technologies of the approaching technological storm will further increase the possibility to manipulate people into behaviour that reflects the interests of commercial parties while carelessly neglecting the legitimate interests of users and citizens (cf. Klenk 2022). Such manipulation may not only put at risks individual values like autonomy and wellbeing, but also is threat to democracy (Lewandowsky et al. 2000).

Technological networks and techno-economic power

In order to properly understand this challenge, and to respond to it, it is important to be aware that it is not just economical in nature, but also has an important technological component. That is to say, it is enabled by certain features of the technologies being developed and deployed, as well as certain (specific) technological and design choices being made, while at the same time sometimes further reinforcing these choices. One technological feature that is particularly important in this respect is the network character of many of the new technologies and applications, and what we have earlier called the feature of '*full connectivity*.' This feature has a number of important consequences.

First, it enables the creation of systems that potentially span the whole world and that do not respect national boundaries. As noted in the section on institutional, regulatory and conceptual disruption (Section 5.3.7), this makes these systems harder to regulate for national governments and therefore increases the power of private companies developing and deploying such systems vis-a-vis national governments.

Second, such networked systems are economically characterized by network externalities and increased returns to adoption (e.g. Arthur 1989). That is to say, the more a system, service or platform offered by a company is adopted by users, the more attractive it tends to become for other users, for example in terms of ease of use, compatibility and quality of service. Importantly, this also allows new business models. For example, it may be attractive to initially offer a (digital) service for free or a low price and to acquire a large share of the market, and later to economize on that market share by e.g. increasing prices. It also allows business models in which consumers do not pay with money but rather with their data, sometimes without being aware, or at the expense of their privacy.

Third, and related to the previous point, the networked and fully connected character of these new digital systems has led to the creation of so-called digital platforms. Digital platforms are digital services that allow exchanges between various producers and consumers. An example is the App store of Google or Apple, which both allow various producers of apps to offer their services to a large number of consumers. The creation of such platforms makes the companies offering them into a kind of gatekeepers, with a powerful position not only vis-a-vis consumers and governments, but also vis-a-vis other companies.

Fourth, the above points create a further economic incentive for certain technological choices, in particular the choice for fully connected systems. As Loke (2021) notes device manufacturers may very well have economic incentives to prefer extensive IoT networks, because of network effects: 'a device that can cooperate with more devices could have greater value, compared to ones that cooperate with only a few'. However, such technological choices may not only reinforce the uneven distribution of techno-economic power, they also make it harder to address some of the earlier challenges discussed, like the blurring of social spheres (Section 5.3.2) and the creation of new cyber-physical risks (Section 5.3.6).

Digital infrastructures

One way to understand and address this challenge is to understand the approaching technological storm resulting in the creation of *new digital infrastructures*. While the internet (World Wide Web) is also a digital infrastructure, it is important to see that the newly arising infrastructures have two additional features. First, they have a *hybrid* character, connecting the digital and physical domain. In fact, in some instances they may be said to add a digital layer to existing physical infrastructure, like transport, water or energy infrastructures. These make them even more vital for societies than existing digital infrastructures. Second, because they are reconfigurable, they make existing (and new) infrastructures to some extent programmable and adaptable.[22] This may have all kinds of advantages but it also implies that those in control of those new digital infrastructures are in a very powerful position, as they may be able to control the adaptation of infrastructures, perhaps almost in real time.

Although there is not an agreed upon definition of 'infrastructure', for the current purpose they may be understood as generally accessible socio-technical systems that enable the production of other goods, private as well as public (Frischmann 2012). Infrastructures for transport, energy, water management, and communication are vital for the production of public goods and public values. They also provide the backbone of many economic activities that create other goods, private as well as public.

This vital role of digital infrastructures would seem a reason to treat them as public goods. In economic theory, public goods are goods that are non-excludable and non-rivalrous (e.g. Mankiw 2012). Non-excludable means that people cannot be excluded from the use of the good. Non-rivalrous means that use of the good by someone does not reduce the availability of the good to other users.

Although use of digital infrastructures, like the Internet, may often be non-excludable, there are a variety of reasons why people may nevertheless be excluded from their use. One is digital illiteracy:

[22] See also https://www.tudelft.nl/tbm/programmable-infrastructures

people may lack the required skills to make use of digital infrastructures. A second is socio-economic inequalities. People may need to pay to use digital infrastructures and not everyone may be able to afford it. Moreover, there are clear geographical differences in the availability and quality of digital infrastructures, which may often (but not always) coincide with socio-economic inequalities. In addition, the fact that some of the current digital infrastructures are shaped or owned by private parties may also reduce (equal) accessibility.

The use of digital infrastructures seems largely non-rivalrous; for example, information that is retrieved from the Internet would remain equally available for subsequent users (Greco and Floridi 2004).[23] Still, some digital goods and services are somewhat rivalrous. For example, bandwidth is usually not unlimited, so that users may suffer from use by others (Greco and Floridi 2004); and even when such problems can be overcome (by e.g. 5G/6G), they may result in abundant use of energy and environmental problems (cf. Section 5.3.4). Similarly, use of digital infrastructures by some users may create negative externalities for other users, like - for example - fake news and cyber-attacks.

Table 5: Different types of goods

	Excludable	Non-excludable
Rivalrous	Private goods (e.g. clothing, cars)	Common pool resources (e.g. fish stocks, oil and gas)
Non-rivalrous	Club goods (e.g. public transport, satellite TV, social media)	Public goods (e.g. clean air, open source software)

In as far as exclusion from use for certain digital goods and services is possible, they are more like club goods. Consider, for example, social media. Companies like Meta and Twitter determine who has access to social media and we also typically look at them for balancing non-discriminatory access to the service and the avoidance of negative externalities from its use, like fake news and hate speech. Still, the question can be asked whether from a regulatory point of view the goods offered by social media (like access to certain information or opinions) should not be treated as public goods (or common pool resources) rather than as club goods.

In as far as the use of digital infrastructures and goods is rivalrous, they are more like common pool resources. Although digital infrastructures and goods may be less rivalrous than such public goods as clean air, it has become clear in the last decade that their use may have serious negative externalities, and may put at risk public values like democracy, veracity, and safety and security. Properly dealing with these externalities may require managing digital infrastructures as common pool resources, and some of the *institutional arrangements* that have been proposed for managing common pool resources may be useful in the digital realm as well (Fuchs 2021; Rosnay and Stalder 2020; Greco and Floridi 2004).

Commons management

From a governance or policy point of view, digital infrastructures may thus require what Frischmann (2012) calls 'commons management', i.e. 'the situation in which a resource is accessible to all members of a community on non-discriminatory terms, meaning terms that do not depend on the users' identity or intended use.' The reasons to treat digital infrastructures as commons in this sense are twofold. First, access to digital infrastructures may be required in order for citizens to have access to what the political philosopher John Rawls has called primary goods, i.e. those goods that every citizen needs as a free person and member of society (Rawls 1999). This includes such goods as basic civil rights and political rights, liberties, income and wealth, and the social bases of self-respect. In this light, it has been suggested that people have a right to some minimal level of information access (Van den Hoven and

[23] The creation of non-fungible tokens (§2.2.7) may be seen as an attempt to create digital information that can be privately owned, thus creating the possibility of private goods.

Rooksby 2008) as well as a right to Internet access.[24] A second reason for considering digital infrastructures as commons is that they create positive externalities, in the sense that they are often crucial for the creation of other goods (Frischmann 2012). Equal and non-discriminatory access to digital services is therefore likely to be in the benefit of all.

However, treating digital infrastructures as commons does not only require non-discriminatory sharing of resources, it also requires safeguarding public values and avoiding, or at least minimizing, negative externalities that might arise from their use. It is worth noting that both requirements may conflict. For example, if we conceive of social media, like Facebook and Twitter, as commons, the requirements to give everyone equal access, independent from their intended use, may de facto lead to fake news and hate speech, which conflict with public values (and create negative externalities for other users and non-users). This requires institutional frameworks and mechanisms to balance these requirements in an acceptable and legitimate way.

While it is open question what institutional frameworks and mechanisms are exactly required for managing digital infrastructures as commons, it would most likely require some curtailing of the power of Big Tech companies, as well as a larger role for governments. However, it would not seem to necessarily imply that digital infrastructures are owned by the government or a form of top-down management. As shown by for example Ostrom (2015), more bottom-up and polycentric institutional arrangements may also be effective in governing commons.

5.3.9. Uncertainty and fundamental unpredictability

Some of the challenges and issues that the approaching technological storm will bring, may not only be currently unknown but also be fundamentally unpredictable. In fact, all challenges that we have discussed above are to a lesser or bigger degree *uncertain*. That is to say, it is, at least to some extent, uncertain whether they will really occur, but if they occur, it is also still uncertain how big their impact on society exactly will be, and how normatively disturbing these challenges are. But in addition to such uncertainties, there are *unknowns*, i.e. things that we do not know yet and may even not be knowable right now. Some of these are so-called known unknowns, for example we do not fully know how the blurring of social spheres that we discussed above (Section 5.3.2) will unfold and how morally problematic that will turn out to be. But there are also unknown unknowns, i.e. there may be challenges that we are not yet aware of that may be potentially more relevant than the ones we have discussed above.

When we look at the past, some of the worries that accompanied the introduction of new technologies in society at the time did not materialize, while some of the disadvantages or risks that actually materialized were never foreseen or anticipated. Examples of the latter are asbestos, CFCs and some of the current concerns about the use of social media and their effects on e.g. democracy. If there is one thing to learn from the past is that we should expect the unexpected.

What makes it particularly challenging to deal with uncertainty and unknowns in technological development is the so-called Collingridge dilemma (Collingridge 1980). This dilemma says that that at the early stages of technological development we typically lack knowledge about the societal impacts of new technology, while at the later stages, when such knowledge is (more) available, technologies

[24] A report by special rapporteur Frank La Rue to the UN in 2011 stated: "*Given that the Internet has become an indispensable tool for realizing a range of human rights, combating inequality, and accelerating development and human progress, ensuring universal access to the Internet should be a priority for all States. Each State should thus develop a concrete and effective policy (…) to make the Internet widely available, accessible and affordable to all segments of population*" (La Rue 2011 p.22). This was interpreted by some as a plea for Internet access as a human right. Also, the Declaration on European Digital Rights and Principles states that "*Everyone, everywhere in the EU, should have access to affordable and high-speed digital connectivity.*" (https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles#Declaration)

have typically got such well entrenched in society that it has become quite hard to still shape their design and societal embedding.

In order to deal with this dilemma, roughly two strategies are possible. One is to try to better anticipate potential impacts of new technology. Such anticipation we have practiced in identifying the challenges in the previous sections. It should be stressed that anticipation is not a form of prediction or forecasting. It is not about what the most likely future is, or even about attaching probabilities to certain scenarios, but rather it involves thinking about possible, but realistic futures so that we are better prepared for those futures and can already start thinking how to address the challenges brought by such possible futures.

The other approach for dealing with the Collingridge dilemma might be described as 'experiment and adapt.' The underlying idea here is that the introduction of new technology into society is inevitably a kind of social experiment, in the sense that some of the impacts of new technology will only become clear along the way (Felt et al. 2007; Van de Poel 2017a). Moreover, it may also be called a moral experiment, in the sense that we might not be able to normatively evaluate all the social changes that will materialize beforehand (Van de Poel 2018).

This second approach does not imply a 'wait and see' strategy when it comes to the social, ethical, and political consequences of the approaching technological storm. Rather it requires thinking about ways to experiment more responsibly as a society. Van de Poel (2016) has proposed a number of conditions for responsibly experimenting with new technologies. One condition that is particularly important in the context of this study is the need to monitor the societal impacts and challenges brought by new technologies, and to aim at early detection of unexpected (and unwanted) ones.

This is particularly important because of what might be called the *pacing problem* (cf. Mnyusiwalla, Daar, and Singer 2003). This is the problem that: 1) it typically takes time before new social, ethical, political, economic issues raised by the use of new technology are detected (and systematically investigated) and 2) it also takes some time after these issues have been detected (and investigated) before they are addressed, for example through better technological design or innovation or though new ethical, legal and institutional measures.

While the pacing problem may to some extent be unavoidable, reducing it at least requires that the different relevant realms of society play their role in timely discovering and addressing new ethical and social issues timely. As we have seen, our analysis in Chapter 4 suggests that when it comes to digital technologies, some of these roles still need to be better performed (Section 4.6).

5.4. Some examples

5.4.1. Smart digital voice assistants[25]

Digital voice assistants (DVAs) are devices that present voice as the primary mode of interaction. Even though voice Assistants started as another type of user interface on the smartphones (e.g. Apple's Siri), they increasingly permeate the homes of people as a separate device in a form of smart speakers, such as Amazon's Echo with its brand DVA Alexa or Google's Home with Google as an interactive assistant. As of 2021, there are more than 153 million installed smart speakers in the world (Bratten 2021). The first smart speakers for the home were introduced in 2016-2017. By early 2021, around 25% of the adult population in the US and Germany owned at least one such device, while in the UK this number was as high as 38% (Kinsella 2021). According to some estimates, the Covid-19 pandemic accelerated the global adoption of voice assistants by a quarter within 2020-2021, owing to staying at home for extended periods of time (Ibid.). But beyond turning on the lights and informing how far the nearest pizza place is, DVAs also change the way we interact with each other.

[25] This case description was written by dr. Olya Kudina.

The increasing popularity of DVAs can be credited to expanding human interaction with the world without the distraction of typing or swiping. With the help of DVAs, one can play and manage music without leaving the shower, have a deeper experience of one's home by interconnecting the spaces and other technologies by voice (e.g. turning on the lights or TV), secure the home in new ways (e.g. managing the lighting or alarms at a distance), switch the radio stations hands-free when doing the chores or have a new way to bond with others at home playing voice-games through the device. This elevates the values of connection, efficiency, security and others to a new level, made possible by the DVAs. At the same time, there are already signs of the social—and ethical—implications of these devices.

Most of the current ethical concerns related to DVAs are connected with jeopardizing the values of privacy, trust, communication and gender equality. Privacy-wise, the companies behind major Western DVAs did not originally explain to the users that real people would be listening in to (parts of) their conversations with voice assistants, instead assuring the users that their conversations would always remain private. However, in 2019 it became clear that the companies such as Microsoft, Google and Amazon, responsible for the lion share of DVAs in the world, have designated staff members across the globe who listen to the conversations the users have with smart speakers in order to improve the quality of language processing and speech recognition algorithms. Additionally, the devices frequently mishear their wake-words (e.g. Hey Google) that trigger the interaction with the devices and start the recordings of the conversations around when not prompted by the users. Together, this caused a range of privacy concerns and user outrage (Lau, Zimmerman, and Schaub 2018). Even though the companies were quick to adapt the Terms and Conditions of the smart speakers, the risk to privacy violations and the lack of trust in the corporate practices remain the top reasons why people choose not to buy a DVA. While privacy concerns frequently accompany the adoption of digital technologies, what complicates the privacy issues with regard to DVAs is the ubiquity of speech and the ability of voice to transcend spaces (Bugeja, Jacobsson, and Davidsson 2016). This makes it much more difficult to devise physical safeguards to protect one's home or construct privacy-proof spaces for using DVAs in one's household. An additional factor is the social aspect of communication, creating new privacy challenges specifically related to DVAs. For instance, if one chooses not to use a DVA for privacy reasons and visits a friend who has several DVAs installed across the home, a new set of social norms and negotiation etiquette will have to arise to accommodate such new social practices.

Regarding the values of communication and mutual understanding, there is a worry that DVAs flatten human interaction to the current limited technological capabilities and prompt people to adopt new, simpler ways of speaking to fit the technological affordances (Kudina 2021). The speech processing technology is far from perfect and frequently processes human speech with mistakes or does not process it at all. One may say something too emphatically or too slowly, there may be a lot of ambient noise that all prevent effective interaction with the DVAs. In other instances, the nascent research suggests that DVAs are discriminating particularly toward non-native English speakers (Pyae and Scifleet 2018; Wu et al. 2020), people with speech variability (e.g. due to oral cancer or stuttering) (Schuster et al. 2006; Halpern et al. 2020), children or the elderly, female voices vs. male voices (Wiederhold 2018; Feng et al. 2021). DVAs are also most receptive to the use of English language as the major mental model of constructing one's speech (Palanica et al. 2019; Sowański and Janicki 2020; Ureta et al. 2020). In general, DVAs identify what people say as a series of patterns and commands - the clearer and shorter they are, the better voice assistants are able to react. The users quickly notice this and learn to adjust their speech such that voice assistants can process it. Hence, our language becomes very functional and top-down (Burton and Gaskin 2019). We use commands and short sentences, as simple as possible. However, language is much more than just orders, we also use jokes, sarcasm, idioms and extremely long complicated sentences that together enable meaningful interactions, sociality and mutual understanding. Voice assistants cannot process such language complexity and variability yet, providing wrong answers or asking you to repeat your question over and over again. This makes some people angry and others blame themselves for saying it wrong - but they continue using the technology. Because we like voice as a means of communication, we are very forgiving towards voice assistants and instead of adapting the technology to the way we speak, we are adjusting ourselves to the currently limited technological capabilities (Bonfert et al. 2018). With this, we are also changing what it means to be a social interactive human. Such challenges to the values of interaction and communication frequently bypass the radar of the policy-makers and risk being disregarded as not easily measurable and thus more difficult to devise a consolidated policy response. Such a response would warrant a broad media literacy education to explain the issues related to the language aspect of DVAs and how to be reflective of what their use does to our communication.

Lastly, the companies behind DVAs spark another ethical controversy by predominantly giving their assistants female voices and challenging the values of diversity and inclusions by promoting gender stereotypes. Female voices as a default interface is an intentional choice. User studies support the fact that people react better to female voices. They view them as more trustworthy, more comfortable, more helpful to talk to. Female voices make people feel like no one is commanding them. Manufacturers of the speakers use this feature to sell their products and to motivate users to interact with the devices longer. However, already in 1997, a US study (Nass, Moon, and Green 1997) showed that using female voices in computer programs promotes gender stereotypes: for instance, when a computer assistant employs a female voice, the users start associating any female voices with being an assistant (Nass and Brave, 2005). A 2019 UNESCO report (West, Kraut, and Chew 2019) confirmed this finding by explicitly exploring the case of smart speakers. According to the report, voice assistants promote an image of a woman who is docile, always available and never says no. One way in which DVAs do that is reflecting on how voice assistants react to what we say to them. Because they are essentially robots that cannot get irritated, people try to test their limits. Very often, people use sexual insults to provoke the smart speakers. In her ethnographic research, (Fessler 2017) showed that when a user addresses the voice assistant with 'You're a slut', Siri almost flirts in response: 'I'd blush if I could', while Alexa says 'Well, thanks for the feedback'. Microsoft's Cortana almost always falls back on porn search. And Google home frequently responds with 'My apologies, I don't understand'. These commonplace patterns of interaction rewire the way we think about women. The UNESCO report (2019) already highlighted some concerns about this. By making it a design feature that voice assistants cannot contradict the user, cannot refuse their request and are available 24/7, smart speakers bring back the stereotype that we as a society tried to tackle a long time ago – that women are always by your side to do you bidding, cannot refuse requests or speak for themselves. It will take intentional responsible redesign to both include more dialogical interaction patterns and promote a diversity of voice options to mitigate such detrimental consequences to the values of diversity and inclusion when DVAs are concerned.

In sum, an informed use of voice assistants calls not only for a reflective individual use and response, but also for the attention of the stakeholders, particularly the policy-makers, to facilitate a conscious adoption of these devices and provide guidance for their responsible development.

5.4.2. The next generations of autonomous vehicles[26]

Self-driving cars are, in a nutshell, vehicles driven by computers. More precisely, they are robotic cars, equipped with sensors, adapting their behaviour to changing environmental circumstances, based on the acquisition and elaboration of data. Self-driving cars may be partially automated, that is assisting or partially replacing the human driver, or fully automated, i.e. driverless. The introduction of driverless private commercial vehicles on the public road such as the Google car has turned out to be more challenging than many expected, due both to technical and societal complexities. However, car manufacturers and public institutions are still preparing for the introduction of higher and higher levels of driving automation on the road.

Current automated driving systems rely on robotic and AI technology. This already raised some thorny ethical and societal issues. The first is safety. While many have rushed to conclude that reducing the

[26] This case description was written by dr. Filippo Santoni De Sio.

role of human drivers, with all their cognitive and motivational limitations, will guarantee an improvement in road safety, serious concerns have been raised about the new risks associated with complex, learning, potentially unpredictable behaviour of intelligent vehicles in complex environments, as well as their interactions with diverse human agents. Recent (fatal) accidents involving semi-automated vehicles have been a wake-up call (Calvert et al. 2021; Bonnefon et al. 2020, chapter 1). Similarly, many have claimed that more efficiency in traffic flow promised by self-driving technology may translate into more energy-efficiency and sustainability, but again it is far from clear what the net environmental impact of this technological transition may or will be (Martin 2020). Another set of ethical issues concerns the new distribution of obligations, responsibility, and liability in traffic systems where the roles of drivers will be heavily reshaped, new actors will enter the picture – software developers, programmers, data managers and others – and the driving task will be distributed across a network of human and artificial agents. Who has a moral/legal obligation to guarantee road safety and who should be considered morally and legally responsible for accidents in such a complex network (Bonnefon et al. 2020, chapter 3)? Finally, automated driving systems will heavily rely on the massive acquisition of data about a vehicle, its occupant, and road users' behaviour. Cars will furthermore increasingly become sensor platforms that scan their environment and capture big data. Actors controlling this data flow in the physical space may become as powerful as those currently controlling the data flow in the digital space. With all possible existing issues of privacy, discrimination, imbalances of power, privatization of public spaces and others already present in the digital space, becoming dramatically relevant also in the physical space of the public road (Bonnefon et al. chapter 2).

Existing AI- and robotic-based self-driving technologies already present big challenges for safety, unpredictability, regulatory frameworks, distribution of power. But in the next future they may be integrated and merged with other emerging technologies. Vehicles may be equipped with in-car facial recognition, voice recognition or bio-tracking technologies to allow for a smoother human-machine interaction or even compliance with safe driving standard (e.g. to alert/nudge/stop the driver when their psycho-physical conditions are not good enough for driving (Hawkins 2019). Cars may even be equipped with brain-machine interfaces to allow more direct control from drivers and/or the possibility of people with disabilities to drive. This is not such a far-fetched scenario considering that people with paralysis can already have some control on wheelchairs (Galán et al. 2008), and one of the biggest players in self-driving technology, Tesla's Elon Musk, has already been investing in brain-machine interaction (Neate 2022). This will create a convergence of the above-mentioned ethical issues with self-driving cars with those traditionally associated with self-tracking (Lanzing 2016) and brain-machine interface technologies: an impact on people's intimate life (e.g. Mecacci and Haselager 2019) and risks for serious violations of personal autonomy and human rights (Ienca and Andorno 2017).

Finally, the next generation of self-driving technologies are likely to be merged with emerging communication technologies, such as IoT and 5G. In addition to being connected among them (so-called V2V technologies) and to their drivers via various technological interfaces, vehicles may be connected to any physical and digital infrastructure (so-called V2X technologies). These connections and interactions may be used not only to aim for a safer, more efficient, sustainable etc. traffic flow but also potentially for any other service- or commercial purpose not related to traffic coordination. The amount and diversity of public and private actors involved in the traffic network may expand dramatically, thereby opening great opportunities for new forms of individualized (mobility) services as well as new big ethical and societal challenges. Vehicles and their occupants may become nodes in a new internet of things. Whether this will bring more safety, autonomy, wellbeing, inclusivity or rather more unpredictability, manipulation, mental overload, and injustice heavily depends on how these networks will be owned, designed, regulated.

5.4.4. The metaverse: real virtual worlds[27]

A technology that is receiving substantial attention at the time of this writing is the metaverse. There is, however, considerable ambiguity as to what the metaverse exactly is, or how it should be defined. In a way, it is primarily a marketing term for the convergence and integration of various online digital games and social media technologies into virtual worlds that will continue to exist even in the absence of a person interacting with it. The metaverse implies interoperability between platforms and seamless immersive user experiences. It has been marketed by Facebook - now Meta - as an immersive social media experience, where people control their avatars within a virtual space, being able to socially interact with other people's avatars or computer agents, play games, attend live concerts or sports events, have learning experiences, take virtual trips, create and share personal immersive content such as 360 degree videos or virtual artworks, share virtual journalism, build and commercialize virtual real estate, buy and sell real or virtual items and services, and so on.

The technologies that provide users access to the metaverse may include virtual reality - immersive interactive virtual worlds - or augmented reality - integrating aspects of the virtual world with the real one. However, the metaverse does not imply that it requires any of these specific technologies as an exclusive means of access. As we've seen before with virtual-world games such as Fortnite, Roblox, or MineCraft, these worlds can be accessed and shared through a variety of platforms (PCs, game consoles, smartphones). Similarly, the metaverse is likely to be a cross-platform experience. Whereas most virtual worlds today support avatars, virtual identities and goods only tied to their own particular platform, the metaverse would ideally allow one to create a digital persona with its associated paraphernalia that can be taken anywhere and everywhere, across a multitude of virtual spaces and platforms.

The metaverse is also a significant potential space for e-commerce of both real-world and virtual assets, and its digital economy is likely to be connected to digital means of secure financial transactions, such as blockchain-powered technology. The immutable decentralized ledger would potentially allow for virtual real estate, art objects, avatar identity, skins, etc to be coupled to a unique and distributed digital ID, and thus be more protected against unauthorized copying, virtual theft, hacks, or other forms of cyberthreats. The arrival of blockchain technology and associated NFTs have enabled artists and art consumers or collectors to exert exclusive ownership over digital artworks and other assets, creating scarcity (as opposed to the costless copying of digital assets), and thus sharply increasing demand and subsequent price points of virtual goods. Whereas we have witnessed somewhat of a frenzy over NFTs in the past year, examples including JPEG images that could be used as social profile pictures, or the so-called Bored Ape Yacht Club, with some apes selling for over $3 million, it remains to be seen to what extent such virtual goods will retain their value.

Aspects of the metaverse that deserve attention from researchers and policy-makers include:

- The meta user experience - intimate and with psychological impact;
- Uniquely identifiable users through kinetic fingerprints, which creates opportunities for persuasion, marketing, political influencing;
- Convergence of meta with AI - greater more seamless quality of avatar and virtual agent appearance and behaviour;
- Difficulties in distinguishing real from virtual - both within meta (am I talking to a real person?) as well as between real world and virtual world experiences (where did I experience this first?);
- Greater risks of fake news having large psychological impact;
- Transfer from VR to reality of skills, response patterns, experiences
- Acculturation effects of virtual worlds - male dominated, violence, sexism, immediate gratification

[27] This case description was written by prof. dr. Wijnand IJsselsteijn.

5.4.5. Streetlights 'going in business'

A concrete example of applications that may emerge when blockchain, IoT and AI converge can be found in a recent article by Sandner, Gross, and Richter (2020). They explore the possibility that devices connected to the internet of things (like cameras and streetlights) will '*in the future act as own profit centers that (1) have a digital twin leveraging IoT, (2) send and receive money leveraging blockchain technology on their own, and (3) autonomously make decisions as independent economic agents leveraging AI and data analytics*.' They make this idea concrete by describing the use case of a street light that is monetized in this way. We have copied this example in the box below:

"One can think of a lamp (e.g., a street light), that has its own block chain-based identity [...] and operates with a block chain-based Euro [...]. Therefore, the lamp gets the status of an autonomous entity operating "on its own." By using smart contracts, micropayments can be made directly to the lamp, triggering the lamp to turn on. The lamp will shine once somebody pays for it, e.g., an individual, a company, or even the public administration. In this context, pay-per-use payment schemes could be implemented. Since the lamp owns a digital wallet, it can act as its own profit center.

Since all lamps are connected to a block chain, they will store data, e.g., about their usage, performance, and downtime. Artificial intelligence could leverage this data and optimize the network's maintenance. For example, it could suggest a more regular maintenance of lamps that are used frequently as well as immediately dispatch the maintenance crew in case of a fault. Additionally, AI can smooth the maintenance process by improving the ordering process of replacement parts for the network or by helping to anticipate the number of replacement parts required more precisely. This support would ultimately result in less down-time of the network.

Since lamps can be tokenized as assets, they can be made available to investors [...]. Consequently, investors could be willing to build and maintain these lamps on a full scale. In return, investors would receive their share on the lamps' profits. This application is a potential gamechanger. The tokenization of such assets could drive a new wave of investments since investors would directly be rewarded with a share of the return of the tokenized asset, in this case, of the lamp.

The benefits of tokenization do not only hold for lamps but all IoT devices and, therefore, a wide range of industrial applications. For example, this could be sensors, cars, machines, cameras, trucks once these devices are connected to the internet and are connected to a block chain network."

Source: Sandner et al (2020)

"One can think of a lamp (e.g., a street light), that has its own block chain-based identity [...] and operates with a block chain-based Euro [...]. Therefore, the lamp gets the status of an autonomous entity operating "on its own." By using smart contracts, micropayments can be made directly to the lamp, triggering the lamp to turn on. The lamp will shine once somebody pays for it, e.g., an individual, a company, or even the public administration. In this context, pay-per-use payment schemes could be implemented. Since the lamp owns a digital wallet, it can act as its own profit center.

Since all lamps are connected to a block chain, they will store data, e.g., about their usage, performance, and downtime. Artificial intelligence could leverage this data and optimize the network's maintenance. For example, it could suggest a more regular maintenance of lamps that are used frequently as well as immediately dispatch the maintenance crew in case of a fault. Additionally, AI can smooth the maintenance process by improving the ordering process of replacement parts for the network or by helping to anticipate the number of replacement parts required more precisely. This support would ultimately result in less down-time of the network.

Since lamps can be tokenized as assets, they can be made available to investors [...]. Consequently, investors could be willing to build and maintain these lamps on a full scale. In return, investors would receive their share on the lamps' profits. This application is a potential gamechanger. The tokenization of such assets could drive a new wave of investments since investors would directly be rewarded with a share of the return of the tokenized asset, in this case, of the lamp.

The benefits of tokenization do not only hold for lamps but all IoT devices and, therefore, a wide range of industrial applications. For example, this could be sensors, cars, machines, cameras, trucks once these devices are connected to the internet and are connected to a block chain network."

Source: Sandner et al (2020)

In this use case (Sandner, Gross, and Richter 2020) we can easily recognize several of the features that were listed in Section 5.2:

- It is a hybrid system in which the physical and digital world become connected;
- It is an autonomous system that can act on its own without human input;
- Control over street lighting can become distributed rather than centralized.

The use case as described here does not immediately or obviously raise ethical concern. Yet it is still interesting to consider, because such radical new business models could lead to socio-economic system changes that are at the moment hard to predict (Section 5.3.9.). Such system changes could in turn lead to new societal and ethical challenges connected to responsibility, ownership, equality, (digital) commons, the digital divide, and the distinction between the public and private sphere (Section 5.3.7.).

5.4.6. Digital twins in health care

Comprehensive Digital models – referred to as Digital Twins – are in use for predictive maintenance of equipment, devices and infrastructure. A computer model is fed with real time big data and over time stops being a general model as it turns gradually into an exact digital replica of the source entity. Twin technology combines Data Science, Machine Learning, IoT, and 5G/6G. Digital Twins are also introduced to model complex things such as organisations, ecological systems, or even the world as a whole. They are increasingly studied and applied in medicine and health care to model organs and single cells or an individual's genetic makeup, physiological and anatomical characteristics. Combined with data on lifestyle, habits and the so-called to 'exposome', digital twins are taken to be an important step towards precision medicine and personalized care. At Linköping University in Sweden researchers have mapped mice RNA into a digital twin to predict the effects of medication. The ultimate vision is to have a lifelong, personalized model of a patient that is updated with each measurement, scan or exam, and that includes behavioural and genetic data as well. It is obvious that development of building comprehensive models of individual human beings introduces vast privacy and data protection problems, problems of equal access to Digital Twins, and problems of power over individuals and enhanced possibilities for manipulation (Bruynseels, Santoni de Sio and Van den Hoven 2018; Erol, Mendi, and Dogan 2020).

5.4.7. Predictive humanitarian aid[28]

Predictive humanitarian aid is a good example of the challenge of 'institutional, regulatory and conceptual disruption' (Section 5.3.7), based on features (Section 5.2) such as the required technology operating from a long distance and being intelligent and adaptive. We also need to beware that the approach comes with the risk of contributing to a further 'concentration of techno-economic power' (Section 5.3.8) in richer countries and potentially with specific players in the private sector due to the ownership of data and computational power.

Predictive humanitarian aid aims at identifying the need for support before a disaster strikes.[29] To achieve this goal, humanitarian actors such as the Red Cross Red Crescent movement do not only make use of multiple data sets and predictive modules, but also implement institutional changes to enable a proactive approach. In technical terms, funding will become available for risk-mitigating actions when Red Cross Red Crescent movement's Early Action Protocol (EAP) is being triggered, 'when an impact-

[28] This case description was written by dr. Michael Nagenborg. "Predictive Humanitarian Aid" is one of the subjects in the NWO-funded research project "*Disastrous Information: Embedding 'Do No Harm' principles into innovative geo-intelligence workflows for effective humanitarian action*" (https://www.nwo.nl/en/projects/mvi19007), on which Nagenborg works.

[29] It is important to note that our examples are concerned with predicting the impact of natural disasters rather than the prediction of human actions, e.g., in an armed conflict. This is not to deny the social nature of so-called "natural disasters." After all, the kind and magnitude of the harm caused by natural disasters often depends on human (in)action. However, we will focus on cases where the harm is caused by natural hazards such as typhoons or floods rather than intentional human acts.

based forecast [...] reaches a predefined danger level' (Van den Homberg, Gevaert, and Georgiadou 2020, p.460). The whole procedure is known as 'forecast-based Financing (FbF) for early action and preparedness for response' (Ibd.). It is also in line with a paradigm shift towards impact-based forecasts as recommended by the (World Meteorological Organization 2015). Examples of early actions include (Red Cross Red Crescent Climate Centre 2020):

- Evacuation of vulnerable communities, individuals and their livestock and most important belongings;
- Pre-deployment of flood barriers;
- Closing of roads and bridges;
- Cash transfer;
- Early harvesting; or
- Pre-position or distribution of relief packages, like tents, food, water, and purification tablets, etc.

These examples point to the diverse time scales which can be addressed in predictive humanitarian aid. For example, early harvesting or the pre-deployment of flood barriers might require more time than the transfer of cash.

As we can already see, the challenge of predictive humanitarian aid cannot be reduced to the reliability and transparency of the AI systems used for predictive analytics, but the overall change in delivering aid. After all, it might not be immediately clear why an institution such as the Red Cross Red Crescent movement is concerned with early harvesting or deploying flood barriers.

On the data side, we do not only have to address questions of scalability and transferability (e.g., do we need to retrain ML algorithms for different countries?), but also questions about data ownership. Given the Humanitarian principles of independence and neutrality (Slim 2015), data donations by private or public parties can become an immediate threat to the self-understanding of humanitarian actors.

To complicate the picture, the technology needed for the analytics may not be available to the people who are suffering the most from the consequences of climate change, such as extreme weather events. Which also means that the decisions about the design and use of the technology may take place elsewhere.

To conclude, predictive humanitarian aid comes with the promise of preventing or, at least, minimizing human suffering. The implementation, however, raises not only technical challenges, but also requires institutional and regulatory changes. It does not only raise questions about the trustworthiness of technologies and institutions, but we might very well need to re-think fundamental principles of and values in humanitarian aid.

6. What we can do? Policy options

6.1. Responsible research and innovation (RRI)

As mentioned in the introduction, we will use **responsible research and innovation (RRI)** as the overarching framework for developing policy options, with its four dimensions (Stilgoe, Owen, and Macnaghten 2013):

- **Inclusiveness**: relevant stakeholders, and their values and needs, should be included in the process of technological innovation from the start;
- **Anticipation**: impacts, benefits and risks of the technology should be anticipated and these anticipations should be fed back into the process of technological innovation;
- **Reflexivity**: the underlying purposes, motivations, and values for technological innovations should be reflected upon and should guide the process of technological innovation;
- **Responsiveness**: technological developments should be responsive to the values and needs of society and to new insights and developments along the way.

In the next sections, we further elaborate upon these four dimensions and how they might be translated into policy directions in relation to the approaching technological storm. In Section 6.2 we discuss more specific policy options, which are a reply to the challenges discussed in Section 5.3; these policy options are inspired by the framework of responsible innovation and its four dimensions.

6.1.1 Improving inclusiveness through societal dialogue

RRI is aimed at making technological development more inclusive and at actively engaging societal stakeholders and society at large in technological development and decisions concerning it. Ensuring inclusivity would require a more even distribution of techno-economic power (Section 5.3.8). It also requires that citizens can meaningfully engage in societal dialogues about the future of these technologies, which may require attention is paid to digital literacy (Section 6.2.4), as well as to reducing the opacity of the current technology (Section 5.3.4).

Inclusive technological development may not only require engaging societal stakeholders, but also ensuring that certain key or fundamental technological choices with respect to the approaching technological storm remain under public (democratic) control. It may therefore require ensuring that certain digital infrastructures remain under some form of public control (Section 6.2.2).

6.1.2 Anticipation and precaution

Anticipation is a core aspect of RRI; it requires anticipating the benefits as well as the risks of new technologies. In this study, we have taken an anticipatory approach by identifying a number of key challenges (Section 5.3), but we have also warned that some impacts may be fundamentally unpredictable (Section 5.3.9) and may therefore require a more responsive and adaptive approach (Section 6.1.4).

Anticipated benefits and risks should also be fed back into the innovation process and the governance of new technology. In terms of benefits, this may be given shape by aiming to address societal challenges through innovation, as we discuss in the next section (Section 6.1.3). Here, we focus on how to address potential risks.

Some of the potential risks we have identified relate to the blurring of social spheres, a potential intrusion into people's most intimate life, opacity, increased energy use, and new cyber-physical risks. It should be noted that all these risks also come with opportunities. The upcoming technological storm may increase energy consumption, but also offers new possibilities for energy reduction. Increased connectivity can offer many societal advantages, but may also lead to blurring of social spheres. It is for these reasons that we talked about challenges in Section 5.3, rather than risks.

One possible approach to the potential risks of the new technological storm is a precautionary one, e.g. based on the 'precautionary principle' (Appendix 3). Several risks, particularly in AI, are already addressed in the GDPR (Appendix 4) and the proposed AI act. The new risks we have identified in this study might not yet be concrete enough to require immediate regulatory action, but some of them would seem to require policy attention and possibly future legislative action. For example, there seems to be a need to protect people's personal sphere beyond privacy (Section 6.2.7), and policy and regulatory options could be explored to address the sustainability challenges of digital technologies (Section 6.2.6). On a more general level, the development of new digital infrastructures that is a key part of the upcoming technological storm could be organised in a way that key societal values can be upheld, e.g. by treating them as digital commons (Section 6.2.2).

It should also be emphasised that this study focuses on rather general challenges raised by the approaching technological storm, rather than specific risks of specific technologies and applications. Some of these risks may also be currently unknown. What would be required to deal with the potential risks of the approaching technological storm in a precautionary way is then first of all an early warning system to detect potential risks that need addressing in a timely manner (Section 6.2.3).

6.1.3 Reflexivity: Meeting societal challenges through innovation

The reflexivity dimension of RRI requires being explicit and reflective about the societal challenges, goals, values and needs for which technologies are being developed. One way to make this more concrete, is by more explicitly addressing societal challenges in digital innovation. There is, for example, a growing body of literature on how to design AI for good (Floridi et al. 2018; Umbrello and van de Poel 2021). Increasing attention is also paid to how companies can contribute to the United Nations Sustainable Development Goals (UN SGDs) through innovative products (Imaz and Eizagirre 2020; Voegtlin et al. 2022). As we have seen, the merger of digital technologies may offer many new opportunities to better address societal challenges (Section 5.3.1).

Innovation may not only be useful to address existing societal challenges, but also some of the new challenges raised by the approaching technological storm. For example, new cyber-physical risk (Section 5.3.6) or sustainability concerns (Section 5.3.5), may at least partly be addressed through technological innovation. Many new digital technologies not only bring risks, but also new opportunities to respect values like security, privacy, democracy, wellbeing, and sustainability.

One way to ensure that new innovations also respect or even promote important social and moral values is to follow what might be called a design for values approach (6.2.5). We use here design for values as an umbrella term for a number of approaches that aim at systematically addressing social and moral values in the design of new technologies (Friedman and Hendry 2019; Van den Hoven, Vermaas, and Van de Poel 2015). This includes an approach like value-sensitive design that has been developed since the 1990s, but also more recent approaches that go by names like ethics-by-design, or more specifically, privacy-by-design or security-by-design.

6.1.4 Responsiveness: Experiment and adapt

RRI requires technological development and innovation that is responsive to the values and needs of society, but is also responsive to new developments and insights. As we discuss above (Section 5.3.9), technological development and certainly its impacts on society is often unpredictable and may lead to surprises, pleasant as well as unpleasant. This means that there is a need to acquire new insights along the way as well as to act responsively.

Such an 'experiment and adapt' approach might be furthered in a number of ways. First it may involve actively trying new technologies not just in the laboratory but also in more 'real-world' settings, to systematically learn about the (social) opportunities and risks of these technologies and ways to properly employ and implement them. This may, for example, be done in 'living labs'.

In addition to such more deliberate experimentation with new technology, the approach requires a monitoring of the effects of new technology. This may, for example, be effected through the establishment of an EU observatory for converging digital technologies (Section 6.2.3).

An 'experiment and adapt' approach also requires a governance structure that enables timely action on new information and insights. For example, building on approaches for adaptive governance (Chaffin, Gosnell, and Cosens 2014; Klinke and Renn 2011), experimentalist governance (Sabel and Zeitlin 2010), and planned adaptation (Haasnoot et al. 2013).

6.2 Policy options

We now outline a number of more specific policy options. These policy options are inspired by the general framework for responsible innovation set out in Section 6.1, but are more specific and tailored to addressing the challenges discussed in Chapter 5. The figure on the next page shows the various challenges identified and indicates how these may be addressed by the various policy options.

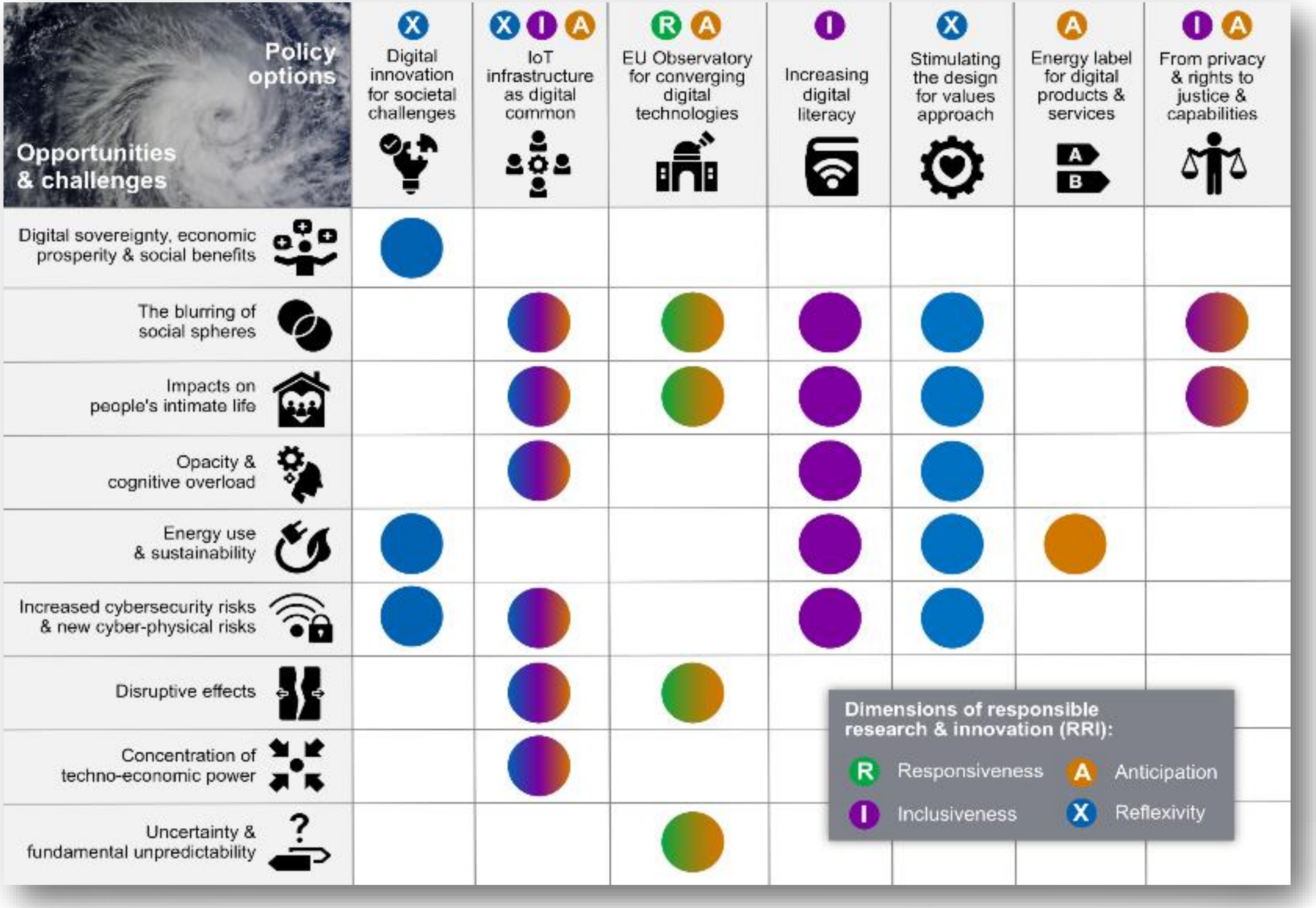
6.2.1. Digital innovation for societal challenges

As we have seen, the merger of digital technologies offers new economic and social opportunities (Section 5.3.1). It enables the EU to address existing societal challenges, as well as offering new opportunities for addressing some of the key challenges the coming technological storm is likely to raise. For example, properly addressing concerns like energy use, and cyber-physical risks may not be possible without further digital innovation.

One policy option is to strengthen digital innovation for societal challenges. These challenges may be further defined, for example, in terms of the UN SDGs.[30] At a more concrete level, however, digital innovation would also need to reflect some of the challenges discussed in this study. Strengthening digital innovation would also require tools for translating very general challenges into more concrete goals for technological innovation (for example employing approaches developed in design for values, Section 6.2.5).

On a strategic level, digital innovation for societal challenges could contribute to increasing the digital sovereignty of the EU; for example, by increasing the capacity to address important challenges in the EU through digital infrastructures and digital technologies that are European in nature or of European origin. Various options to increase such digital sovereignty have already been discussed in the STOA study on key enabling technologies (Ramahandry et al. 2021).

[30] The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, identifies 17 Sustainable Development Goals (SDGs) as shown in https://sdgs.un.org/goals



Policy options Opportunities & challenges	X Digital innovation for societal challenges	X I A IoT infrastructure as digital common	R A EU Observatory for converging digital technologies	I Increasing digital literacy	X Stimulating the design for values approach	A Energy label for digital products & services	I A From privacy & rights to justice & capabilities
Digital sovereignty, economic prosperity & social benefits	●						
The blurring of social spheres		●	●	●	●		●
Impacts on people's intimate life		●	●	●	●		●
Opacity & cognitive overload		●		●	●		
Energy use & sustainability	●			●	●	●	
Increased cybersecurity risks & new cyber-physical risks	●	●		●	●		
Disruptive effects		●	●				
Concentration of techno-economic power		●					
Uncertainty & fundamental unpredictability			●				

Dimensions of responsible research & innovation (RRI):

R	Responsiveness	A	Anticipation
I	Inclusiveness	X	Reflexivity

Concrete measures could include:

1) A clearer place for digital innovation in the EU Missions in the Horizon Europe research funding scheme, particularly in mission-oriented research. Currently five EU missions have been formulated relating to: 1) climate change, 2) cancer, 3) water and oceans, 4) climate-neutral and smart cities, and 5) soil.[31] Not only can the contribution of merging digital technologies to all these five missions be further developed, it would seem to make sense to aim for a mission more squarely in the digital realm.
2) Stimulate the creation of European industrial consortia and public-private partnerships that can contribute to digital innovation for societal challenges and increasing digital sovereignty.
3) Pay particular attention to how SMEs and start-ups may contribute to digital innovation for societal challenges, e.g. through incubators and subsidy schemes.

6.2.2. The IoT as digital common

As infrastructure technology, the IoT is an enabler of systems connecting many applications, and so facilitating the production of other goods. To seize the economic and social opportunities of the approaching technological storm (Section 5.3.1), it might be an option to manage such digital infrastructures as digital commons. This would require ensuring that all relevant parties have equal and non-discriminatory access to the IoT and other relevant digital infrastructures. New proposed EU regulations such as the digial markets act (DMA) and digital services act (DSA) already contribute to such objectives.

However, digital infrastructures like the IoT are not just neutral tools that enable economic and social activities – they are value-laden. That is to say, if we want certain important values like, for example, privacy, sustainability and transparency to be respected, it has certain consequences for how our IoT infrastructures are to be shaped. This means that managing IoT as a digital common also requires seeing it as part of the public sphere, to be protected by certain institutional and regulatory arrangements.

Managing the IoT as a digital common thus requires balancing two sets of (potentially conflicting) requirements, one relating to equal non-discriminatory access to the infrastructure; the other to the safeguarding of public values like privacy, sustainability, fairness, democracy and transparency.

Treating IoT infrastructure as a digital common would counterbalance the current uneven distribution of techno-economic power (Section 5.3.8), and may therefore also contribute to achieving digital sovereignty for the EU (Section 5.3.1). It may also help to address some of the regulatory and institutional voids that might result from the merger of digital technologies (Section 5.3.7). It would furthermore be helpful in addressing the challenge of blurring social spheres (Section 5.3.2), that is to say if we want some social spheres, and related digital systems, to be **disconnected**, it most likely requires the ability to make common, public choices with respect to the underlying IoT infrastructures, rather than leaving these decisions to private companies. If we treat IoT as a digital common it may also be helpful in addressing sustainability issues (Section 5.3.5), opacity (Section 5.3.4), and cyber-physical risks (Section 5.3.6), as it would allow some minimal conditions that the infrastructure would need to meet.

Treating IoT as a digital common would, at minimum, require a set of public rules for its development, maintenance and use, aimed at guaranteeing equal non-discriminatory access as well as the safeguarding of public values. This may be effected through public ownership, e.g. by governments, of the basic digital infrastructures, but it is likely that other ownership and institutional regimes also allow the IoT to be managed as a digital common. For example institutional rules that prevent that companies that offer services using basic digital infrastructure can also be the owners and shapers of

[31] These five EU Missions are expressed in the Horizon Europe research funding scheme, and can be found here: https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/eu-missions-horizon-europe_en

those basic digital infrastructures. Such unbundling rules now exist, for example, with respect to energy infrastructures and these could also be a model for digital infrastructures like the IoT. [32]

6.2.3. EU Observatory for converging digital technologies

The new ethical and social issues raised by the technological storm are partly unpredictable (Section 5.3.9). Yet some of their effects may be disruptive and be irreversible (Section 5.3.7), and require institutional, regulatory or even conceptual changes to properly deal with them. There is therefore a need for an organisation playing an early warning or early detection function when it comes to new challenges and potential disruptions brought about by the approaching technological storm. This could be done by establishing an EU observatory of converging digital technologies.

Such an observatory could also be instrumental in monitoring whether new technological developments are still in line with important European digital rights and values, as for example recently laid down in the Declaration on European digital rights and principles. It could furthermore play a role in putting a broader range of concerns and values on the agenda, as suggested in the policy option 'from privacy and digital rights to justice and human capabilities' (Section 6.2.7).

In Chapter 4, we suggested that ethics, and more broadly ELSI, research could play an early warning role, but is not yet fully playing such a role. We note that most of the ethics research still focuses on AI, and pays far less attention to technologies like IoT, blockchain, AR/VR and 5G/6G, or their convergence (see Section 4.4.2 and Section 4.5.2). This might imply that it also takes a while before ethical, social, legal or political issues related to these other technologies, or their convergence, are discovered. Chapter 4 also notes that there might be time lapses between the moment at which certain issues or values come to be seen as relevant and the moment they are translated into new technological solutions or in regulatory or governance measures (Section 4.5).

While some time lags may be unavoidable, it seems reasonable to try to speed up the process of early detection of new ethical and social issues, investigation of these issues and the translation into either new technological innovations and solutions or policy and governance measures. An EU observatory can play a key role in this process. Such an observatory would be tasked with monitoring relevant developments, carrying out interdisciplinary ELSI research on converging technologies with the aim to discover new issues and challenges early and to translate these either into new technological research and innovation, or into new policy, governance or regulatory measures.

There are already a number of related observatories, like the European Digital Media Observatory,[33] the EU Blockchain Observatory,[34] and the European 5G Observatory.[35] There are also relevant initiatives at the national level, like the recently established AI ELSA (ethical, legal and social aspects) in the Netherlands.[36] An EU observatory of converging digital technologies, would be distinguished by:

1. a focus on converging (digital) technologies rather than individual digital technologies like AI and robotics;
2. a focus on new societal challenges and potential disruptions raised by technologies and how to adequately and timely address these; and
3. activity at the international level, with an emphasis on the EU but reaching out to the United States and China, for example.

6.2.4. Increasing digital literacy

[32] See also https://fsr.eui.eu/unbundling-in-the-european-electricity-and-gas-sectors/

[33] European Digital Media Observatory: https://digital-strategy.ec.europa.eu/en/policies/european-digital-media-observatory

[34] EU Blockchain Observatory: https://www.eublock chainforum.eu/

[35] European 5G Observatory: https://5gobservatory.eu/

[36] To be found here: https://nlaic.com/en/news/all-the-signs-are-in-place-for-elsa-labs-and-human-centred-ai/

'Digital education and skills' is discussed in the proposed declaration on European digital rights and principles, which was published in January 2022; it is one of the four topics discussed in the declaration's chapter on solidarity and inclusion. Digital literacy is important for making digital innovation more inclusive (see also Section 6.1.1 Improving inclusiveness through societal dialogue), as that would require citizens who are sufficiently well-informed to contribute to a societal dialogue. It will also be helpful in addressing some of the more specific challenges. For example, dealing with opacity and cognitive overload (Section 5.3.4) will also require citizens that have a better understanding of digital technologies, including their limitations and threats. Similarly, better awareness will help citizens to play their part in addressing challenges like increased risks and new cybersecurity risks (Section 5.3.5), energy use and sustainability (Section 5.3.4), and the impact on people's intimate life (Section 5.3.3). While these challenges also require other measures and should not be made the sole responsibility of individual citizens, at the same time other measures are likely to be ineffective if they are not accompanied by behavioural changes based on an awareness of the possibilities and limitations of digital technologies. In this way, digital literacy can also contribute to resilience against still unknown effects of digital technologies, both individually and at a societal level. As Braun et al. (2020) point out, digital literacy is a necessary precondition for social inclusion and equal participation in a digitalised democracy, and is necessary for safeguarding European values such as equality, democracy and the rule of law.

Digital literacy is more than the ability to use digital devices and services, or to understand how they function technologically. It also requires an awareness of the limitations and pitfalls of digital technologies and services. For example, digital literacy would mean that citizens are aware that social media might be used for fake news; have the skills to recognise potential fake news and ways to check it in independent ways. Digital literacy also means that citizens are aware of cybersecurity threats and what they can do to minimise these, or to limit the impact of cyber-attacks. A key component of digital literacy, as we understand it here, is therefore the ability for critical thinking in a digitalised environment.

Digital literacy can be furthered in a number of ways. First through the formal education system, preferably from an early age. Second, through life-long learning: as digital technologies evolve quickly, digital literacy needs to be updated. Third, it would seem mandatory to set up special efforts for groups that are not reached, or are hard to reach, through formal education and life-long learning, with special attention paid to vulnerable or socially disadvantaged groups. Fourth, digital literacy also requires a critical press and media that can independently bring to light the pitfalls and threats raised by new digital technologies, as well as ways to better deal with these technologies. Fifth, digital tools and alternative digital technologies may be crucial for digital literacy. Digital tools, for example, may make people aware that they only are receiving information from certain digital resources (and thus are in a 'filter bubble'), or might help them to train their critical thinking skills. Alternative digital technologies might make people aware that what technologies they chose to use matters, and that some digital technologies are, for example, more privacy friendly than others. Finally, a requirement could be placed on providers of digital services and producers of digital products to produce leaflets that make people aware of possible risks or side-effects of their services and products, somewhat similar to those required for medication.

6.2.5. Stimulating a design for values approach

One policy option would be to further strengthen a design for values approach to digital technologies. Such an approach aims at systematically designing digital technologies for a range of moral and social values (Section 7.1.3). Privacy-by-design and ethics-by-design are already part of the GDPR (Appendix 4) and of new EU AI regulation.

Strengthening design for values would help to better address challenges like: energy use and sustainability (Section 5.3.5) – requiring more attention to be paid to the value of sustainability in the development and design process of new technology; challenges like increased cybersecurity risks and new cyber-physical risks (Section 5.3.6, requiring more attention to be paid to values like safety and security); effects on people's intimate lives (Section 5.3.3, requiring more attention to be paid to the value of human-wellbeing); as well as the blurring of social spheres (Section 5.3.2, requiring paying greater attention to the value of justice); and opacity and cognitive overload (Section 5.3.4, requiring that attention is paid to values like transparency and explainability). A design for values approach would, moreover, allow these challenges to be tackled in an integral way, as certain trade-offs between these values may be required in the design process of new technologies. Our analysis in Chapter 4 suggests that some values are already well-addressed in technological research and innovation (see Section 5.5.5 Does technological research sufficiently address values?). This is particularly true for reliability, cybersecurity, privacy and sustainability. A value like 'justice and fairness' has recently received greater attention, but other values like democracy and transparency still get limited attention in technological research and innovation.

As discussed in Chapter 4, there may be several reasons why certain values get less attention in technological research and innovation. One reason is that certain values may be less relevant for technological development, or harder to translate into design choices and innovation; such values may, for example, more properly be addressed through policy, governance and regulation. Another reason might be that values first need to be operationalised and specified before they play a role in technological research and innovation. This has been done in recent years, for example, for fairness and justice, which have been translated into a number of fairness metrics in the AI literature (e.g. Ruf and Detyniecki 2021). For other values, like democracy and transparency, this translation may still be needed.

To speed up the process of making values relevant, and to address them properly in technological research and innovation, stimulating a design for values approach would be useful. There has been considerable progress in recent decades both with respect to the general approach, as well in relation to specific technologies and specific values (Van den Hoven, Vermaas, and Van de Poel 2015). A policy option could be to further institutionalise design for values. Concrete measures could include:

- The EU could stimulate training programmes on design for values, and contribute to their development, for example through the Horizon Europe research funding scheme.
- Reporting obligations on responsible innovation and design for values could be made part of the obligatory corporate social responsibility (CSR) reporting for large companies in the EU. Currently, EU law requires certain large companies to disclose information on the way they operate and manage social and environmental challenges.[37] These obligations are laid down in Directive 2014/95/EU – the Non-Financial Reporting Directive (NFRD). On 21 April 2021, the Commission adopted a proposal for a corporate sustainability reporting directive (CSRD), which would amend the NFRD. However, this proposal does not yet contain any reporting obligations with respect to responsible (digital) innovation (or design for values).
- Encourage that design for values is taken up in standardisation and certification. The Institute of Electrical and Electronics Engineers (IEEE) recently launched a new standard (IEEE 7000-2021) for ethically aligned design of autonomous and intelligent systems.[38] This and other technical standards can contribute to the systematic uptake of design for values approaches by industry and in research and innovation. The EU could also consider certification schemes for new digital products and services that require that they have been developed following a design for values approach.

[37] Corporate sustainability reporting: https://ec.europa.eu/info/business-economy-euro/company-reporting-an d-auditing/company-reporting/corporate-sustainability-reporting_en

[38] Standard to be found here: https://ethicsinaction.ieee.org/

6.2.6. Energy label for digital technologies and services

One policy option to address the challenge of energy use and sustainability (Section 5.3.4) might be to introduce (compulsory) energy labels for digital devices and services. This would potentially have two advantages. One is that it creates more **transparency** about energy use for consumers and the public at large, and allows them to make more deliberate choices when it comes to purchasing and using certain digital devices and services. It will also help to make people more **aware** that some digital technologies and services consume considerable amounts of energy. Second, energy labels may create an **incentive** for the industry to reduce the energy consumption of digital devices and services, and may spur innovation towards lower energy consumption (with similar performance).

Indeed, the proposed declaration of European digital rights and principles (European Commission 2022) states that 'everyone should have access to accurate, easy-to-understand information on the environmental impact and energy consumption of digital products and services, allowing them to make responsible choices.' An energy labelling scheme has existed in the EU since 1992, and has been updated several times since.[39] It applies, among other things to white goods, cars, and light bulbs. There are also recent EU initiatives introducing EU energy labels for personal computers and servers (European Commission 2018, expected to be adopted in 2023) and mobile phones and laptops (European Commission 2020, expected to be adopted in 2023).

So, while initiatives are underway, a policy option could be to extend them to a wider range of digital devices and services. One challenge with digital services is that their energy consumption will, to a large extent, depend on the network, including energy use for communication and storage of data. Because such network services are typically shared among many devices and users, it raises attribution and accounting problems when it comes to energy usage and hence to introducing energy labels. Nevertheless, it may be worthwhile to investigate where energy labels (not only for digital devices but also for digital services) would be feasible, particularly when such services are likely to consume substantial amounts of energy.

6.2.7. From privacy and digital rights to social justice and human capabilities

As we have seen, the new technological storm is likely to raise challenges, such as the blurring of social spheres (Section 5.3.2), as well as affecting people's intimate life (Section 5.3.3). Such issues extend well beyond privacy concerns, and can likely not be addressed by existing regulations like the GDPR that focus on privacy and information exchange.

Addressing these challenges might first of all require paying more attention to human digital rights. In fact, the European Commission recently published a draft declaration of European digital rights and principles (European Commission 2022), which addresses some of the more fundamental rights that are at issue. Various digital policies and regulations that the EU has been, and is still, working on will contribute to protecting these rights. But more specific legislation and regulation might still be needed to make this declaration effective. Moreover, the focus is still very much on individual rights, while some challenges (like the blurring of social spheres (Section 5.3.1) and the impact on people's intimate life (Section 5.3.2)), require attention to be paid to broader issues of human wellbeing and social justice.

A policy option would be to broaden the normative basis for policy-making and regulation in relation to the technologies of the approaching technological storm, from a narrow focus on privacy and digital rights to justice and human capabilities more broadly. The human rights approach and the capability approach are closely related, as both highly value human dignity and individual freedom. One way to understand human rights is as a right to have certain capabilities – effective opportunities to achieve

[39] Directive 92/75/EC; which was replaced by Directive 2010/30/EU, and was again replaced by Regulation 2017/1369/EU from 1 August 2017. Updated labelling requirements came into force in 2021.

valuable 'being' and 'doings'.[40] An advantage of understanding human rights in this way, is that the capability approach stimulates us to investigate what is realistically, all things considered, necessary to truly enable people to do and be certain things. Legal human rights protections may be an important factor in realising valuable human capabilities, but certainly not the only one (Vizard, Fukuda-Parr, and Elson 2011).

Attention to human capabilities might be particularly required to deal with the challenge that digital technologies increasingly affect people's most intimate life (Section 5.3.3). It would require paying attention to how such technologies affect human wellbeing and human vulnerability, starting with becoming more specific about what human capabilities should be enabled by new digital technologies. We also need to understand the different factors that jointly determine whether or not certain capabilities are endangered or present. For example, we may simultaneously need to work on improving digital literacy, developing more transparent technologies and setting clear limits on what data companies are allowed to collect and process. And we should do this with a realistic view of e.g. people's cognitive abilities and the ability of governments to actually enforce legislation.

Ensuring justice is also particularly required to deal with the challenge of blurring social spheres (Section 5.3.2). The proposed declaration of European digital rights and principles pays attention to justice issues in terms of equal access to digital services and infrastructures as well as in terms of equal (digital) opportunity. However, this does not yet address the issue that different social spheres might be characterised by different concerns and principles of social justice; what is (socially) just in the personal or family sphere may be different from what is just in the economic or political sphere. Addressing the challenge of blurring social spheres would require becoming more explicit about such normative questions about social justice that extend well beyond traditional distributive concerns.

To address these broader normative questions, a societal dialogue including a broad spectrum of stakeholders would be required (Section 6.1.1) before moving towards the stage of policy formulation or translating these insights into design for values.

[40] This understanding is especially convincing under the positive, substantive view of human rights that sees them as creating positive obligations on certain actors rather than calling for non-interference and non-intervention with people's freedom (Vizard et al, 2011).

References

Abramovich, Giselle. 2021. "5 Innovative Examples Of Augmented Reality In Action." *Abode Business* (blog). 2021. https://business.adobe.com/sea/resources/5-realworld-examples-of-augmented-reality-innovation.html.

Ahmed, Ferdous, A. S. M. Hossain Bari, and Marina L. Gavrilova. 2020. "Emotion Recognition From Body Movement." *IEEE Access* 8: 11761–81. https://doi.org/10.1109/ACCESS.2019.2963113.

Akhtar, Nikhat, Yusuf Perwej, Nikhat Akhtar, and Yusuf Perwej. 2020. "The Internet of Nano Things (IoNT) Existing State and Future Prospects." *GSC Advanced Research and Reviews* 5 (2): 131–50. https://doi.org/10.30574/gscarr.2020.5.2.0110.

Alabdulwahhab, Faten Adel. 2018. "Web 3.0: The Decentralized Web Blockchain Networks and Protocol Innovation." In *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*, 1–4. Riyadh: IEEE. https://doi.org/10.1109/CAIS.2018.8441990.

Allhoff, Fritz, and Adam Henschke. 2018. "The Internet of Things: Foundational Ethical Issues." *Internet of Things* 1–2 (September): 55–66. https://doi.org/10.1016/j.iot.2018.08.005.

Almabdy, Soad M., and Lamiaa A. Elrefaei. 2021. "An Overview of Deep Learning Techniques for Biometric Systems." In *Artificial Intelligence for Sustainable Development: Theory, Practice and Future Applications*, edited by Aboul Ella Hassanien, Roheet Bhatnagar, and Ashraf Darwish, 912:127–70. Studies in Computational Intelligence. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-51920-9_8.

Alsamhi, Saeed H., Ou Ma, Mohammad Samar Ansari, and Faris A. Almalki. 2019. "Survey on Collaborative Smart Drones and Internet of Things for Improving Smartness of Smart Cities." *IEEE Access* 7: 128125–52. https://doi.org/10.1109/ACCESS.2019.2934998.

Alston, Eric, Wilson Law, Ilia Murtazashvili, and Martin Weiss. 2022. "Blockchain Networks as Constitutional and Competitive Polycentric Orders." *Journal of Institutional Economics*, January, 1–17. https://doi.org/10.1017/S174413742100093X.

Andrae, Anders S. G., and Tomas Edler. 2015. "On Global Electricity Usage of Communication Technology: Trends to 2030." *Challenges* 6 (1). https://doi.org/10.3390/challe6010117.

Armstrong, Maggie Mae. 2020. "Cheat Sheet: What Is Digital Twin? Internet of Things Blog." IBM Business Operations Blog. December 4, 2020. https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin/.

Arthur, W. Brian. 1989. "Competing Technologies, Increasing Returns, and Lock-In by Historical Events." *The Economic Journal* 99 (394): 116. https://doi.org/10.2307/2234208.

Atzori, Marcella. 2015. "Blockchain Technology and Decentralized Governance: Is the State Still Necessary?" *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2709713.

Belkhir, Lotfi, and Ahmed Elmeligi. 2018. "Assessing ICT Global Emissions Footprint: Trends to 2040 & Recommendations." *Journal of Cleaner Production* 177: 448–63. https://doi.org/10.1016/j.jclepro.2017.12.239.

Belpoggi, Fiorella. 2021. "Health Impact of 5G; Current State of Knowledge of 5G-Related Carcinogenic and Reproductive/Developmental Hazards as They Emerge from Epidemiological Studies and in Vivo Experimental Studies." Brussels: Panel for the Future of Science and Technology (STOA).

Bertino, Elisa, Ahish Kundu, and Zehra Sura. 2019. "Data Transparency with Blockchain and AI Ethics." *Journal of Data and Information Quality* 11 (4): 1–8. https://doi.org/10.1145/3312750.

Birhane, Abeba, Pratyusha Kalluri, Dallas Card, William Agnew, Ravit Dotan, and Michelle Bao. 2021. "The Values Encoded in Machine Learning Research." *ArXiv:2106.15590 [Cs]*, June. http://arxiv.org/abs/2106.15590.

Blei, David, Lawrence Carin, and David Dunson. 2010. "Probabilistic Topic Models." *IEEE Signal Processing Magazine*, November, 5563111. https://doi.org/10.1109/MSP.2010.938079.

Boddington, Ghislaine. 2021. "The Internet of Bodies—Alive, Connected and Collective: The Virtual Physical Future of Our Bodies and Our Senses." *AI & SOCIETY*, February. https://doi.org/10.1007/s00146-020-01137-1.

Bonfert, Michael, Maximilian Spliethöver, Roman Arzaroli, Marvin Lange, Martin Hanci, and Robert Porzel. 2018. "If You Ask Nicely: A Digital Assistant Rebuking Impolite Voice Commands." In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, 95–102. ICMI '18. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3242969.3242995.

Bonnefon, J.F., D. Černy, J. Danaher, N. Deviller, V. Johansson, T. Kovacikova, M. Martens, et al. 2020. "Ethics of Connected and Automated Vehicles: Recommendations on Road Safety, Privacy, Fairness, Explainability and Responsibility." Horizon 2020 Commission Expert Group to advise on specific ethical issues raised by driverless mobility, E03659. Luxembourg: Publication Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/89624e2c-f98c-11ea-b44f-01aa75ed71a1/language-en.

Bozdag, Engin, and Jeroen van den Hoven. 2015. "Breaking the Filter Bubble: Democracy and Design." *Ethics and Information Technology* 17 (4): 249–65. https://doi.org/10.1007/s10676-015-9380-y.

Bratten, Eric. 2021. "As Smart Speakers Evolve, So Do Consumers." Comscore, Inc. February 12, 2021. https://www.comscore.com/Insights/Blog/As-Smart-Speakers-Evolve-So-Do-Consumers.

Braun, Anette, Anna März, Fabian Mertens, and Annerose Nisser. 2020. "Rethinking Education in the Digital Age." Brussels: Panel for the Future of Science and Technology (STOA). https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2020)641528.

Brock, Thomas J. 2022. "From the Experts: 8 Pros and Cons of Non-Fungible Tokens and How They Compare to Traditional Investments." *Annuity.Org* (blog). January 14, 2022. https://www.annuity.org/2022/01/14/from-the-experts-8-pros-and-cons-of-nfts/.

Bruynseels K, Santoni de Sio F. and Jeroe van den Hoven. 2018. "Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm". In: *Frontiers in Genetics*. 9:31. https://doi.org/10.3389/fgene.2018.00031.

Bugeja, Joseph, Andreas Jacobsson, and Paul Davidsson. 2016. "On Privacy and Security Challenges in Smart Connected Homes." In *2016 European Intelligence and Security Informatics Conference (EISIC)*, 172–75. https://doi.org/10.1109/EISIC.2016.044.

Burton, Nathan, and James Gaskin. 2019. "'Thank You, Siri': Politeness and Intelligent Digital Assistants." *AMCIS 2019 Proceedings*, July. https://aisel.aisnet.org/amcis2019/social_inclusion/social_inclusion/5.

Calvert, Simeon C., Bart van Arem, Daniël D. Heikoop, Marjan Hagenzieker, Giulio Mecacci, and Filippo Santoni de Sio. 2021. "Gaps in the Control of Automated Vehicles on Roads." *IEEE Intelligent Transportation Systems Magazine* 13 (4): 146–53. https://doi.org/10.1109/MITS.2019.2926278.

Carew, Joseph M. 2021. "Reinforcement Learning." *TechTarget* (blog). March 2021. https://www.techtarget.com/searchenterpriseai/definition/reinforcement-learning.

Cath, Corinne, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo, and Luciano Floridi. 2018. "Artificial Intelligence and the 'Good Society': The US, EU, and UK Approach." *Science and Engineering Ethics* 24 (2): 505–28. https://doi.org/10.1007/s11948-017-9901-7.

Chaffin, Brian C., Hannah Gosnell, and Barbara A. Cosens. 2014. "A Decade of Adaptive Governance Scholarship: Synthesis and Future Directions." *Ecology and Society*. https://doi.org/10.5751/ES-06824-190356.

Christen, Markus, Bert Gordijn, and Michele Loi, eds. 2020. *The Ethics of Cybersecurity*. Vol. 21. The International Library of Ethics, Law and Technology. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-29053-5.

Coates McCall, Iris, Chloe Lau, Nicole Minielly, and Judy Illes. 2019. "Owning Ethical Innovation: Claims about Commercial Wearable Brain Technologies." *Neuron* 102 (4): 728–31. https://doi.org/10.1016/j.neuron.2019.03.026.

Cohen, Michael. 2019. "Lecture 12: Brain-Machine Interface (Course 'The Human Brain')." MIT Open Courseware, Massachusetts Institute of Technology. https://ocw.mit.edu/courses/brain-and-cognitive-sciences/9-13-the-human-brain-spring-2019/lecture-videos/brain-machine-interface/.

Collingridge, David. 1980. *The Social Control of Technology*. London: Frances Pincher.

Council of the EU. 2016. "Better Regulation to Strengthen Competitiveness (Press Release)." May 26, 2016. https://www.consilium.europa.eu/en/press/press-releases/2016/05/26/conclusions-better-regulation/.

Croatti, Angelo, Matteo Gabellini, Sara Montagna, and Alessandro Ricci. 2020. "On the Integration of Agents and Digital Twins in Healthcare." *Journal of Medical Systems* 44 (9): 161. https://doi.org/10.1007/s10916-020-01623-5.

Cruz Alvarado, Mainor Alberto, and Patricia Alejandra Bazán. 2018. "Understanding the Internet of Nano Things: Overview, Trends, and Challenges." *E-Ciencias de La Información*, November. https://doi.org/10.15517/eci.v1i1.33807.

Cunliff, Colin. 2020. "Beyond the Energy Techlash: The Real Climate Impacts of Information Technology." Information Technology and Innovation Foundation. https://itif.org/publications/2020/07/06/beyond-energy-techlash-real-climate-impacts-information-technology.

Czarnocki, Jan. 2021. "Will New Definitions of Emotion Recognition and Biometric Data Hamper the Objectives of the Proposed AI Act?" In *2021 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 1–4. Darmstadt, Germany: IEEE. https://doi.org/10.1109/BIOSIG52210.2021.9548285.

D. Cawthorne and A. Devos. 2020. "Capability Caution in UAV Design." In *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1572–81. https://doi.org/10.1109/ICUAS48674.2020.9214008.

Dahlgren, Peter M. 2021. "A Critical Review of Filter Bubbles and a Comparison with Selective Exposure." *Nordicom Review* 42 (1): 15–33. https://doi.org/10.2478/nor-2021-0002.

Daly, Angela, Thilo Hagendorff, Hui Li, Monique Mann, Vidushi Marda, Ben Wagner, Wayne Wei Wang, and Saskia Witteborn. 2019. "Artificial Intelligence, Governance and Ethics: Global Perspectives." SSRN Scholarly Paper ID 3414805. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3414805.

Davidson, Sinclair, Primavera De Filippi, and Jason Potts. 2018. "Blockchains and the Economic Institutions of Capitalism." *Journal of Institutional Economics* 14 (4): 639–58. https://doi.org/10.1017/S1744137417000200.

De Keyser, Arne, Yakov Bart, Xian Gu, Stephanie Q. Liu, Stacey G. Robinson, and P.K. Kannan. 2021. "Opportunities and Challenges of Using Biometrics for Business: Developing a Research Agenda." *Journal of Business Research* 136 (November): 52–62. https://doi.org/10.1016/j.jbusres.2021.07.028.

Dechesne, Francien, Martijn Warnier, and Jeroen van den Hoven. 2013. "Ethical Requirements for Reconfigurable Sensor Technology: A Challenge for Value Sensitive Design." *Ethics and Information Technology* 15 (3): 173–81. https://doi.org/10.1007/s10676-013-9326-1.

Dierksmeier, Claus, and Peter Seele. 2020. "Blockchain and Business Ethics." *Business Ethics: A European Review* 29 (2): 348–59. https://doi.org/10.1111/beer.12259.

Dobson, Andrew. 1998. *Justice and the Environment*. Oxford University Press. https://doi.org/10.1093/0198294956.001.0001.

Dressler, Falko, and Stefan Fischer. 2015. "Connecting In-Body Nano Communication with Body Area Networks: Challenges and Opportunities of the Internet of Nano Things." *Nano Communication Networks* 6 (2): 29–38. https://doi.org/10.1016/j.nancom.2015.01.006.

Durham, Emily. 2019. "First-Ever Noninvasive Mind-Controlled Robotic Arm." *Carnegie Mellon University, College of Engineering* (blog). June 21, 2019. https://engineering.cmu.edu/news-events/news/2019/06/20-he-sci-robotics.html.

Dutton, Tim, Brent Barron, and Gaga Boskovic. 2018. "Building an AI World; Report on National and Regional AI Strategies." CIFAR.

Dzedzickis, Andrius, Artūras Kaklauskas, and Vytautas Bucinskas. 2020. "Human Emotion Recognition: Review of Sensors and Methods." *Sensors* 20 (3): 592. https://doi.org/10.3390/s20030592.

Erol, Tolga, Arif Furkan Mendi, and Dilara Dogan. 2020. "The Digital Twin Revolution in Healthcare." In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–7. Istanbul, Turkey: IEEE. https://doi.org/10.1109/ISMSIT50672.2020.9255249.

European Commission. 2000. *Communication from the Commission on the Precautionary Principle, COM(2000) 1 Final*. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0001:FIN:EN:PDF.

———. 2014. "Rome Declaration on Responsible Research and Innovation in Europe." https://ec.europa.eu/research/swafs/pdf/rome_declaration_RRI_final_21_November.pdf.

———. 2017. *Investing in a Smart, Innovative and Sustainable Industry – A Renewed EU Industrial Policy Strategy, COM(2017) 479 Final*.

———. 2018. *Energy Labelling Requirements for Computers and Computer Servers*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/1580-Energy-labelling-requirements-for-computers-and-computer-servers_en.

———. 2020. *Energy Labelling of Mobile Phones and Tablets – Informing Consumers about Environmental Impact*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12798-Energy-labelling-of-mobile-phones-and-tablets-informing-consumers-about-environmental-impact_en.

———. 2022. *European Declaration on Digital Rights and Principles for the Digital Decade*. https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles#Declaration.

European Parliament. 2015. *Technological Solutions for Sustainable Agriculture (2015/2225(INI))*. https://www.europarl.europa.eu/doceo/document/TA-8-2016-0251_EN.pdf.

European Political Strategy Centre. 2016. "Towards an Innovation Principle Endorsed by Better Regulation." European Commission. https://data.europa.eu/doi/10.2872/626511.

European Risk Forum. 2015. "The Innovation Principle - Overview." https://www.eriforum.eu/uploads/2/5/7/1/25710097/innovation_principle_one_pager_5_march_2015.pdf.

———. 2020. "Fostering Innovation - Better Management of Risk." https://www.eriforum.eu/uploads/2/5/7/1/25710097/erf_ip_monograph_briefing_note_002.pdf.

Fares Al Mashagba, Eman. 2016. "Human Identification Based on Geometric Feature Extraction Using a Number of Biometric Systems Available: Review." *Computer and Information Science* 9 (2): 140. https://doi.org/10.5539/cis.v9n2p140.

Felt, Ulrike, Brian Wynne, Michel Callon, Maria Eduarda Gonçalves, Sheila Jasanoff, Maria Jepsen, Pierre-Benoît Joly, et al. 2007. "Taking European Knowledge Society Seriously." Brussels: European Commission, Directorate-General for Research and Innovation.

Feng, Siyuan, Olya Kudina, Bence Mark Halpern, and Odette Scharenborg. 2021. "Quantifying Bias in Automatic Speech Recognition." *ArXiv:2103.15122 [Cs, Eess]*, April. http://arxiv.org/abs/2103.15122.

Ferguson, Andrew G. 2017. "Policing Predictive Policing." *Washington Univ Law Review* 94 (5): 1109–89.

Fessler, Leah. 2017. "We Tested Bots like Siri and Alexa to See Who Would Stand up to Sexual Harassment." Quartz. February 22, 2017. https://qz.com/911681/we-tested-apples-siri-amazon-echos-alexa-microsofts-cortana-and-googles-google-home-to-see-which-personal-assistant-bots-stand-up-for-themselves-in-the-face-of-sexual-harassment/.

Finck, Michèle. 2019. "Blockchain and the General Data Protection Regulation; Can Distributed Ledgers Be Squared with European Data Protection Law?" Brussels: Panel for the Future of Science and Technology (STOA). https://data.europa.eu/doi/10.2861/535.

Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhu Srikumar. 2020. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," January. https://dash.harvard.edu/handle/1/42160420.

Fletcher, Richard, and Joy Jenkins. 2019. "Polarisation and the News Media in Europe." Brussels: European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA). https://data.europa.eu/doi/10.2861/059702.

Floridi, Luciano, and Josh Cowls. 2019. "A Unified Framework of Five Principles for AI in Society." *Harvard Data Science Review* 1 (1). https://doi.org/10.1162/99608f92.8cd550d1.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, et al. 2018. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28 (4): 689–707. https://doi.org/10.1007/s11023-018-9482-5.

French, Aaron, J.P. Shim, Marten Risius, Kai R. Larsen, Hemant Jain, Kai R. Larsen, University of Colorado, Boulder, Hemant Jain, University of Tennessee Chattanoog, and University of Tennessee Chattanoog. 2021. "The 4th Industrial Revolution Powered by the Integration of AI, Blockchain, and 5G." *Communications of the Association for Information Systems* 49 (1): 266–86. https://doi.org/10.17705/1CAIS.04910.

Friedman, Batya, and David Hendry. 2019. "Value Sensitive Design: Shaping Technology with Moral Imagination." Cambridge, Massachusetts: The MIT Press.

Frischmann, Brett M. 2012. "Defining Infrastructure and Commons Management." SSRN Scholarly Paper ID 2117460. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2117460.

Fuchs, Christian. 2021. "The Digital Commons and the Digital Public Sphere How to Advance Digital Democracy Today." *Westminster Papers in Communication and Culture* 16 (1). https://doi.org/10.16997/wpcc.917.

Galán, F., M. Nuttin, E. Lew, P. W. Ferrez, G. Vanacker, J. Philips, and J. Del R. Millán. 2008. "A Brain-Actuated Wheelchair: Asynchronous and Non-Invasive Brain-Computer Interfaces for Continuous Control of Robots." *Clinical Neurophysiology: Official Journal of the International Federation of Clinical*

Neurophysiology 119 (9): 2159–69. https://doi.org/10.1016/j.clinph.2008.06.001.

George, Sam. 2021. "Converging the physical and digital with digital twins, mixed reality, and metaverse apps." Microsoft Azure Blog. May 26, 2021. https://azure.microsoft.com/pt-pt/blog/converging-the-physical-and-digital-with-digital-twins-mixed-reality-and-metaverse-apps/.

Ghose, Shohini. 2018. *A Beginner's Guide to Quantum Computing*. TEDx Talk. TEDx Talk. https://www.ted.com/talks/shohini_ghose_a_beginner_s_guide_to_quantum_computing.

Gilbert, Thomas Krendl. 2021. "Mapping the Political Economy of Reinforcement Learning Systems: The Case of Autonomous Vehicles." *Simons Institute for the Theory of Computing* (blog). January 31, 2021. https://simons.berkeley.edu/news/mapping-political-economy-reinforcement-learning-systems-case-autonomous-vehicles.

Giles, Martin. 2019. "Explainer: What Is Quantum Communication?" Science journalism. *MIT Technology Review* (blog). February 14, 2019. https://www.technologyreview.com/2019/02/14/103409/what-is-quantum-communications/.

Gordon, John-Stewart, and Sven Nyholm. 2021. "Ethics of Artificial Intelligence." In *Internet Encyclopedia of Philosophy*. https://iep.utm.edu/ethic-ai/.

Greco, Gian Maria, and Luciano Floridi. 2004. "The Tragedy of the Digital Commons." *Ethics and Information Technology* 6 (2): 73–81. https://doi.org/10.1007/s10676-004-2895-2.

Greengard, Samuel. 2019. *Virtual Reality*. The MIT Press Essential Knowledge Series. MIT Press.

Groenfeldt, Tom. 2017. "IBM And Maersk Apply Blockchain To Container Shipping." *Forbes* (blog). March 5, 2017. https://www.forbes.com/sites/tomgroenfeldt/2017/03/05/ibm-and-maersk-apply-blockchain-to-container-shipping/?sh=4e53a96a3f05.

Haasnoot, Marjolijn, Jan H. Kwakkel, Warren E. Walker, and Judith ter Maat. 2013. "Dynamic Adaptive Policy Pathways: A Method for Crafting Robust Decisions for a Deeply Uncertain World." *Global Environmental Change*. http://dx.doi.org/10.1016/j.gloenvcha.2012.12.006.

Hagendorff, Thilo. 2020. "The Ethics of AI Ethics: An Evaluation of Guidelines." *Minds and Machines* 30 (1): 99–120. https://doi.org/10.1007/s11023-020-09517-8.

Halpern, Bence Mark, Rob van Son, Michiel van den Brekel, and Odette Scharenborg. 2020. "Detecting and Analysing Spontaneous Oral Cancer Speech in the Wild." *ArXiv:2007.14205 [Cs, Eess]*, July. http://arxiv.org/abs/2007.14205.

Hansson, Sven Ove. 2009. "Risk and Safety in Technology." In *Philosophy of Technology and Engineering Sciences*, 1069–1102. Elsevier. https://doi.org/10.1016/B978-0-444-51667-1.50043-4.

Hawkins, Andrew J. 2019. "Volvo Will Use In-Car Cameras to Combat Drunk and Distracted Driving." The Verge. March 20, 2019. https://www.theverge.com/2019/3/20/18274235/volvo-driver-monitoring-camera-drunk-distracted-driving.

Honan, Mat. 2013. "I, Glasshole: My Year With Google Glass." *Wired*, 2013. https://www.wired.com/2013/12/glasshole/.

Hopster, Jeroen. 2022. "The Ethics of Disruptive Technologies: Towards a General Framework." In *New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence*, edited by Juan F. de Paz Santana, Daniel H. de la Iglesia, and Alfonso José López Rivero, 1410:133–44. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-87687-6_14.

Hu, Zhuangli, Canbing Li, Yijia Cao, Baling Fang, Lina He, and Mi Zhang. 2014. "How Smart Grid Contributes to Energy Sustainability." *Energy Procedia* 61: 858–61. https://doi.org/10.1016/j.egypro.2014.11.982.

Ienca, Marcello. 2021. "On Neurorights." *Frontiers in Human Neuroscience* 15. https://www.frontiersin.org/article/10.3389/fnhum.2021.701258.

Ienca, Marcello, and Roberto Andorno. 2017. "Towards New Human Rights in the Age of Neuroscience and Neurotechnology." *Life Sciences, Society and Policy* 13 (1): 5. https://doi.org/10.1186/s40504-017-0050-1.

Imaz, Oier, and Andoni Eizagirre. 2020. "Responsible Innovation for Sustainable Development Goals in Business: An Agenda for Cooperative Firms." *Sustainability*.

Investopedia Team. 2022. "Web 2.0 and Web 3.0." *Investopedia* (blog). 22 2022. https://www.investopedia.com/web-20-web-30-5208698.

Ishmaev, Georgy. 2020. "The Ethical Limits of Blockchain-Enabled Markets for Private IoT Data." *Philosophy & Technology* 33 (3): 411–32. https://doi.org/10.1007/s13347-019-00361-y.

Jacobs, Naomi. 2020. "Capability Sensitive Design for Health and Wellbeing Technologies." *Science and Engineering Ethics* 26 (6): 3363–91. https://doi.org/10.1007/s11948-020-00275-5.

Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1 (9): 389–99. https://doi.org/10.1038/s42256-019-0088-2.

Kamilaris, Andreas, and Nicolo Botteghi. 2020. "The Penetration of Internet of Things in Robotics: Towards a Web of Robotic Things." *ArXiv:2001.05514 [Cs]*, January. http://arxiv.org/abs/2001.05514.

Karpathy, Andrej. 2016. "Deep Reinforcement Learning: Pong from Pixels." *Personal Blog on Github* (blog). May 31, 2016. http://karpathy.github.io/2016/05/31/rl/.

Kellermann, Robin, Tobias Biehle, and Liliann Fischer. 2020. "Drones for Parcel and Passenger Transportation: A Literature Review." *Transportation Research Interdisciplinary Perspectives* 4 (March): 100088. https://doi.org/10.1016/j.trip.2019.100088.

Kinsella, B. 2021. "UK Smart Speaker Adoption Surpasses U.S. in 2020 - New Report with 33 Charts [Executive Summary]." https://voicebot.ai/2021/06/18/uk-smart-speaker-adoption-surpasses-u-s-in-2020-new-report-with-33-charts/.

Kirienko, Margarita, Martina Sollini, Gaia Ninatti, Daniele Loiacono, Edoardo Giacomello, Noemi Gozzi, Francesco Amigoni, Luca Mainardi, Pier Luca Lanzi, and Arturo Chiti. 2021. "Distributed Learning: A Reliable Privacy-Preserving Strategy to Change Multicenter Collaborations Using AI." *European Journal of Nuclear Medicine and Molecular Imaging* 48 (12): 3791–3804. https://doi.org/10.1007/s00259-021-05339-7.

Klenk, Michael. 2022. "(Online) manipulation: sometimes hidden, always careless." *Review of Social Economy* 80 (1):85-105. doi: 10.1080/00346764.2021.1894350.

Klinke, Andreas, and Ortwin Renn. 2011. "Adaptive and Integrative Governance on Risk and Uncertainty." *Journal of Risk Research*. https://doi.org/10.1080/13669877.2011.636838.

Ko, Byoung. 2018. "A Brief Review of Facial Emotion Recognition Based on Visual Information." *Sensors* 18 (2): 401. https://doi.org/10.3390/s18020401.

Koops, Bert-Jaap, Bryce Clayton Newell, Tjerk Timan, Ivan Škorvánek, Tom Chokrevski, and Maša Galič. 2016. "A Typology of Privacy." SSRN Scholarly Paper ID 2754043. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2754043.

Kritikos, Mihalis. 2020. "What If Blockchain Could Guarantee Ethical AI?" Brussels: Panel for the Future of Science and Technology (STOA). https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_ATA(2020)656334.

Kudina, Olya. 2021. "'Alexa, Who Am I?': Voice Assistants and Hermeneutic Lemniscate as the Technologically Mediated Sense-Making." *Human Studies* 44 (2): 233–53. https://doi.org/10.1007/s10746-021-09572-9.

Kudina, Olya, and Peter-Paul Verbeek. 2019. "Ethics from within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy." *Science, Technology, & Human Values*. https://doi.org/10.1177/0162243918793711.

Kuehn, Kathleen M., and Leon A. Salter. 2020. "Assessing Digital Threats to Democracy, and Workable Solutions: A Review of the Recent Literature." *International Journal of Communication* 14 (0): 22.

Kuscu, Murat, and Bige D. Unluturk. 2021. "Internet of Bio-Nano Things: A Review of Applications, Enabling Technologies and Key Challenges." *ArXiv:2112.09249 [Cs]*, December. http://arxiv.org/abs/2112.09249.

La Rue, Frank. 2011. "Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression." A/HRC/14/23. United Nations. https://digitallibrary.un.org/record/683561?ln=en.

Lanzing, Marjolein. 2016. "The Transparent Self." *Ethics and Information Technology* 18 (1): 9–16. https://doi.org/10.1007/s10676-016-9396-y.

Lau, Josephine, Benjamin Zimmerman, and Florian Schaub. 2018. "Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-Seeking Behaviors with Smart Speakers." *Proceedings of the ACM on Human-Computer Interaction* 2 (CSCW): 102:1-102:31. https://doi.org/10.1145/3274371.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep Learning." *Nature* 521 (7553): 436–44. https://doi.org/10.1038/nature14539.

Levy, Karen. 2015. "Intimate Surveillance." *Idaho Law Review* 51 (3). https://ssrn.com/abstract=2834354.

Lim, Jia Zheng, James Mountstephens, and Jason Teo. 2020. "Emotion Recognition Using Eye-Tracking:

Taxonomy, Review and Current Challenges." *Sensors* 20 (8): 2384. https://doi.org/10.3390/s20082384.

Lewandowsky, S., Smillie, L., Garcia, D., Hertwig, R., Weatherall, J., Egidy, S., Robertson, R.E., O'connor, C., Kozyreva, A., Lorenz-Spreen, P., Blaschke, Y. and Leiser, M., Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making, EUR 30422 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-24088-4

Loideain, Nóra Ni. 2019. "A Port in the Data-Sharing Storm: The GDPR and the Internet of Things." *Journal of Cyber Policy* 4 (2): 178–96. https://doi.org/10.1080/23738871.2019.1635176.

Loke, Seng Wai. 2021. "Achieving Ethical Algorithmic Behaviour in the Internet of Things: A Review." https://doi.org/10.3390/iot2030021.

Makhataeva, Zhanat, and Huseyin Varol. 2020. "Augmented Reality for Robotics: A Review." *Robotics* 9 (2): 21. https://doi.org/10.3390/robotics9020021.

Malmodin, Jens, Åsa Moberg, Dag Lundén, Göran Finnveden, and Nina Lövehagen. 2010. "Greenhouse Gas Emissions and Operational Electricity Use in the ICT and Entertainment & Media Sectors." *Journal of Industrial Ecology* 14 (5): 770–90. https://doi.org/10.1111/j.1530-9290.2010.00278.x.

Mankiw, N. Gregory. 2012. *Principles of Microeconomics*. Mason, OH: South-Western Cengage Learning.

Marro, Samuele, and Luca Donno. 2022. "Green NFTs: A Study on the Environmental Impact of Cryptoart Technologies." *ArXiv:2202.00003 [Cs]*, January. http://arxiv.org/abs/2202.00003.

Martin, George. 2020. *Sustainability Prospects for Autonomous Vehicles: Environmental, Social, and Urban*. Routledge. https://www.routledge.com/Sustainability-Prospects-for-Autonomous-Vehicles-Environmental-Social/Martin/p/book/9780367786274.

McMahan, Brendan, and Daniel Ramage. 2017. "Federated Learning: Collaborative Machine Learning without Centralized Training Data." *Google AI Blog* (blog). 2017. http://ai.googleblog.com/2017/04/federated-learning-collaborative.html.

Mecacci, Giulio, and Pim Haselager. 2019. "Identifying Criteria for the Evaluation of the Implications of Brain Reading for Mental Privacy." *Science and Engineering Ethics* 25 (2): 443–61. https://doi.org/10.1007/s11948-017-0003-3.

Mehraj, Haider, and Ajaz Hussain Mir. 2018. "A Survey of Biometric Recognition Using Deep Learning." *EAI Endorsed Transactions on Energy Web*, July, 166775. https://doi.org/10.4108/eai.27-10-2020.166775.

Mintah, Kwabena, Kingsley Tetteh Baako, Godwin Kavaarpuo, and Gideon Kwame Otchere. 2020. "Skin Lands in Ghana and Application of Blockchain Technology for Acquisition and Title Registration." *Journal of Property, Planning and Environmental Law* 12 (2): 147–69. https://doi.org/10.1108/JPPEL-12-2019-0062.

Miraz, Mahdi, Maaruf Ali, Peter Excell, and Richard Picking. 2018. "Internet of Nano-Things, Things and Everything: Future Growth Trends." *Future Internet* 10 (8): 68. https://doi.org/10.3390/fi10080068.

Mishra, Debashisha, Anna Maria Vegni, Valeria Loscri, and Enrico Natalizio. 2021. "Drone Networking in the 6G Era: A Technology Overview." *IEEE Communications Standards Magazine* 5 (4): 88–95. https://doi.org/10.1109/MCOMSTD.0001.2100016.

Mnyusiwalla, Anisa, Abdallah S Daar, and Peter A Singer. 2003. "Mind the Gap : Science and Ethics in Nanotechnology." *Nanotechnology* 14 (3): R9–13. https://doi.org/10.1088/0957-4484/14/3/201.

Mohamed, Nader, Jameela Al-Jaroodi, Imad Jawhar, Ahmed Idries, and Farhan Mohammed. 2020. "Unmanned Aerial Vehicles Applications in Future Smart Cities." *Technological Forecasting and Social Change* 153 (April): 119293. https://doi.org/10.1016/j.techfore.2018.05.004.

Möller, Judith. 2021. "Filter Bubbles and Digital Echo Chambers." In *The Routledge Companion to Media Disinformation and Populism*. Routledge.

Monrat, Ahmed Afif, Olov Schelen, and Karl Andersson. 2019. "A Survey of Blockchain From the Perspectives of Applications, Challenges, and Opportunities." *IEEE Access* 7: 117134–51. https://doi.org/10.1109/ACCESS.2019.2936094.

Müller, Vincent C. 2021. "Ethics of Artificial Intelligence and Robotics." In *Stanford Encyclopedia of Philosophy*, summer 2021. https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/.

Nagenborg, Michael. 2009. "Designing spheres of informational justice." *Ethics and Information Technology* 11 (3):175-179. doi: 10.1007/s10676-009-9200- .

Nam, Jaehyun, Hyesun Chung, Young ah Seong, and Honggu Lee. 2019. "A New Terrain in HCI: Emotion Recognition Interface Using Biometric Data for an Immersive VR Experience."

https://doi.org/10.48550/ARXIV.1912.01177.

Nass, Clifford, Youngme Moon, and Nancy Green. 1997. "Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices." *Journal of Applied Social Psychology* 27 (10): 864–76. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x.

Neate, Rupert. 2022. "Elon Musk's Brain Chip Firm Neuralink Lines up Clinical Trials in Humans." *The Guardian*, January 20, 2022, sec. Technology. https://www.theguardian.com/technology/2022/jan/20/elon-musk-brain-chip-firm-neuralink-lines-up-clinical-trials-in-humans.

Nguyen, Dinh C., Ming Ding, Pubudu Pathirana, Aruna Seneviratne, Jun Li, Dusit Niyato, Octavia A. Dobre, H. Vincent Poor, H. Vincent Poor, and H. Vincent Poor. 2021. "6G Internet of Things: A Comprehensive Survey." *ArXiv: Signal Processing*. https://doi.org/10.1109/jiot.2021.3103320.

North-Samardzic, Andrea. 2020. "Biometric Technology and Ethics: Beyond Security Applications." *Journal of Business Ethics* 167 (3): 433–50. https://doi.org/10.1007/s10551-019-04143-6.

Nussbaum, Martha. 2011. *Creating Capabilities; The Human Development Approach*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Oosterlaken, Ilse. 2015. "Human Capabilities in Design for Values." In *Handbook of Ethics, Values, and Technological Design*, edited by Jeroen van den Hoven, Pieter E. Vermaas, and Ibo van de Poel, 221–50. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_7.

OPTIC. 2019. "(Re)Building Trust in Technology (Ethics & Tech Report 2019)." Optic Technology.

Ostrom, Elinor. 2015. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.

Palanica, Adam, Anirudh Thommandram, Andrew Lee, Michael Li, and Yan Fossat. 2019. "Do You Understand the Words That Are Comin Outta My Mouth? Voice Assistant Comprehension of Medication Names." *NPJ Digital Medicine* 2: 55. https://doi.org/10.1038/s41746-019-0133-x.

Perrow, Charles. 1984. *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.

Porambage, Pawani, Mika Ylianttila, Corinna Schmitt, Pardeep Kumar, Andrei Gurtov, and Athanasios V. Vasilakos. 2016. "The Quest for Privacy in the Internet of Things." *IEEE Cloud Computing* 3 (2): 36–45. https://doi.org/10.1109/MCC.2016.28.

Pyae, Aung, and Paul Scifleet. 2018. "Investigating Differences between Native English and Non-Native English Speakers in Interacting with a Voice User Interface: A Case of Google Home." In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, 548–53. OzCHI '18. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3292147.3292236.

Ramahandry, Tiana, Vincent Bonneau, Emarildo Bani, Nikita Vlaslov, Michael Flickenschild, Olga Batura, Nikolay Tcholtchev, Philipp Lämmel, and Michell Boerger. 2021. "Key Enabling Technologies for Europe's Technological Sovereignty." Brussels: Panel for the Future of Science and Technology (STOA). https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2021)697184.

Rawls, John. 1999. *A Theory of Justice. Revised Edition*. Revised edition. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Ray, Partha Pratim. 2016. "Internet of Robotic Things: Concept, Technologies, and Challenges." *IEEE Access* 4: 9489–9500. https://doi.org/10.1109/ACCESS.2017.2647747.

Red Cross Red Crescent Climate Centre. 2020. "The Future of Forecasts: Impact-Based Forecasting for Early Action." *ReliefWeb* (blog). September 1, 2020. https://reliefweb.int/report/world/future-forecasts-impact-based-forecasting-early-action.

Rios Insua, David, Aitor Couce-Vieira, Jose A. Rubio, Wolter Pieters, Katsiaryna Labunets, and Daniel G. Rasines. 2021. "An Adversarial Risk Analysis Framework for Cybersecurity." *Risk Analysis* 41 (1): 16–36. https://doi.org/10.1111/risa.13331.

Rosnay, Mélanie Dulong de, and Felix Stalder. 2020. "Digital Commons." *Internet Policy Review* 9 (4). https://policyreview.info/concepts/digital-commons.

Ruf, Boris, and Marcin Detyniecki. 2021. "Towards the Right Kind of Fairness in AI." https://doi.org/10.48550/ARXIV.2102.08453.

Sabel, Charles F., and Jonathan Zeitlin. 2010. "Experimentalist Governance in the European Union: Towards a New Architecture." Oxford: Oxford University Press.

Sandner, Philipp, Jonas Gross, and Robert Richter. 2020. "Convergence of Blockchain, IoT, and AI." *Frontiers in Blockchain* 3: 42. https://doi.org/10.3389/fbloc.2020.522600.

Santoni de Sio, Filippo, and Jeroen van den Hoven. 2018. "Meaningful Human Control over Autonomous Systems: A Philosophical Account." *Frontiers in Robotics and AI*. https://doi.org/10.3389/frobt.2018.00015.

Schermer, Bart W., and Joas van Ham. 2021. "Regulering van Immersieve Technologieën." The Hague: Wetenschappelijk Onderzoek- en Documentatiecentrum, Ministerie van Justitie & Veiligheid [Research and Documentation Centre, Dutch Ministry of Justice and Security].

Schiff, Daniel, Jason Borenstein, Justin Biddle, and Kelly Laas. 2021. "AI Ethics in the Public, Private, and NGO Sectors: A Review of a Global Document Collection." *IEEE Transactions on Technology and Society* 2 (1): 31–42. https://doi.org/10.1109/TTS.2021.3052127.

Schoenborn, Jakob M, and Klaus-Dieter Althoff. 2019. "Recent Trends in XAI: A Broad Overview on Current Approaches, Methodologies and Interactions," 10.

Schoenmakers, Matzat, Bakker, and IJsselsteijn. 2022. "Deepfake Effects in Interpersonal Communication: Negativity Effect, but Not Realism Heuristic, Explains Changes in Credibility through Deepfakes." *Under Review*.

Schuller, Björn W. 2018. "Speech Emotion Recognition: Two Decades in a Nutshell, Benchmarks, and Ongoing Trends." *Communications of the ACM* 61 (5): 90–99. https://doi.org/10.1145/3129340.

Schuster, Maria, Andreas Maier, Tino Haderlein, Emeka Nkenke, Ulrike Wohlleben, Frank Rosanowski, Ulrich Eysholdt, and Elmar Nöth. 2006. "Evaluation of Speech Intelligibility for Children with Cleft Lip and Palate by Means of Automatic Speech Recognition." *International Journal of Pediatric Otorhinolaryngology* 70 (10): 1741–47. https://doi.org/10.1016/j.ijporl.2006.05.016.

Schwab, Klaus. 2016. *The Fourth Industrial Revolution*. First U.S. edition. New York: Crown Business.

Schwartz, Roy, Jesse Dodge, Noah A. Smith, and Oren Etzioni. 2019. "Green AI." *ArXiv:1907.10597 [Cs, Stat]*, August. http://arxiv.org/abs/1907.10597.

Sedlmeir, Johannes, Hans Ulrich Buhl, Gilbert Fridgen, and Robert Keller. 2020. "The Energy Consumption of Blockchain Technology: Beyond Myth." *Business & Information Systems Engineering* 62 (6): 599–608. https://doi.org/10.1007/s12599-020-00656-x.

Sen, Amartya. 1999. *Commodities and Capabilities*. Oxford: Oxford University Press.

Shen, Guohua, Kshitij Dwivedi, Kei Majima, Tomoyasu Horikawa, and Yukiyasu Kamitani. 2019. "End-to-End Deep Image Reconstruction From Human Brain Activity." *Frontiers in Computational Neuroscience* 13 (April): 21. https://doi.org/10.3389/fncom.2019.00021.

Shi, Weisong, Jie Cao, Quan Zhang, Youhuizi Li, and Lanyu Xu. 2016. "Edge Computing: Vision and Challenges." *IEEE Internet of Things Journal* 3 (5): 637–46. https://doi.org/10.1109/JIOT.2016.2579198.

Shu, Lin, Jinyan Xie, Mingyue Yang, Ziyi Li, Zhenqi Li, Dan Liao, Xiangmin Xu, and Xinyi Yang. 2018. "A Review of Emotion Recognition Using Physiological Signals." *Sensors* 18 (7): 2074. https://doi.org/10.3390/s18072074.

Singh, Maneet, Richa Singh, and Arun Ross. 2019. "A Comprehensive Overview of Biometric Fusion." *Information Fusion* 52 (December): 187–205. https://doi.org/10.1016/j.inffus.2018.12.003.

Slim, Hugo. 2015. *Humanitarian Ethics: A Guide to the Morality of Aid in War and Disaster*. Oxford: Oxford University Press.

Snijders, Dhoya, Sophie Horsman, Linda Kool, and Rinie van Est. 2020. "Responsible VR. Protect Consumers in Virtual Reality." The Hague: Rathenau Institute.

Sowański, Marcin, and Artur Janicki. 2020. "Leyzer: A Dataset for Multilingual Virtual Assistants." In *Text, Speech, and Dialogue*, edited by Petr Sojka, Ivan Kopeček, Karel Pala, and Aleš Horák, 477–86. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-58323-1_51.

Stark, Luke, and Karen Levy. 2018. "The Surveillant Consumer." *Media, Culture & Society* 40 (8): 1202–20. https://doi.org/10.1177/0163443718781985.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. "Developing a Framework for Responsible Innovation." *Research Policy*. https://doi.org/10.1016/j.respol.2013.05.008.

Storm, Darlene. 2015. "MEDJACK: Hackers Hijacking Medical Devices to Create Backdoors in Hospital Networks." Computerworld. June 8, 2015. https://www.computerworld.com/article/2932371/medjack-hackers-hijacking-medical-devices-to-create-backdoors-in-hospital-networks.html.

Suhaimi, Nazmi Sofian, James Mountstephens, and Jason Teo. 2020. "EEG-Based Emotion Recognition: A

State-of-the-Art Review of Current Trends and Opportunities." *Computational Intelligence and Neuroscience* 2020 (September): 1–19. https://doi.org/10.1155/2020/8875426.

Suzuki, S. 2018. "Recent Researches on Innovative Drone Technologies in Robotics Field." *Advanced Robotics* 32 (19): 1008–22. https://doi.org/10.1080/01691864.2018.1515660.

Swierstra, Tsjalling. 2013. "Nanotechnology and Technomoral Change." *Etica & Politica / Ethics & Politics* XV (1): 200–219.

Taddeo, Mariarosaria, Tom McCutcheon, and Luciano Floridi. 2019. "Trusting Artificial Intelligence in Cybersecurity Is a Double-Edged Sword." *Nature Machine Intelligence* 1 (12): 557–60. https://doi.org/10.1038/s42256-019-0109-1.

Tang, Yong, Jason Xiong, Rafael Becerril-Arreola, and Lakshmi Iyer. 2019. "Ethics of Blockchain: A Framework of Technology, Applications, Impacts, and Research Directions." *Information Technology & People* 33 (2): 602–32. https://doi.org/10.1108/ITP-10-2018-0491.

Taylor, Linnet. 2021. "Public Actors Without Public Values: Legitimacy, Domination and the Regulation of the Technology Sector." *Philosophy & Technology* 34 (4): 897–922. https://doi.org/10.1007/s13347-020-00441-4.

The European Commission's High-Level Expert Group on Artificial Intelligence. 2018. *A Definition of AI: Main Capabilities and Scientific Disciplines*. https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence.

Thielens, Arno. 2021. "Environmental Impacts of 5G; A Literature Review of Effects of Radio-Frequency Electromagnetic Field Exposure of Non-Human Vertebrates, Invertebrates and Plants." PE 690.021. Brussels: Panel for the Future of Science and Technology (STOA).

Umbrello, Steven, and Ibo van de Poel. 2021. "Mapping Value Sensitive Design onto AI for Social Good Principles." *AI and Ethics*. https://doi.org/10.1007/s43681-021-00038-3.

Ureta, Jennifer, Celina Iris Brito, Jilyan Bianca Dy, Kyle-Althea Santos, Villaluna Winfred, and Ethel Ong. 2020. "At Home with Alexa: A Tale of Two Conversational Agents." In *Text, Speech, and Dialogue*, edited by Petr Sojka, Ivan Kopeček, Karel Pala, and Aleš Horák. Springer. https://www.springerprofessional.de/en/at-home-with-alexa-a-tale-of-two-conversational-agents/18338786.

Van de Poel, Ibo. 2009. "Values in Engineering Design." In *Philosophy of Technology and Engineering Sciences*, edited by Anthonie Meijers, 973–1006. Handbook of the Philosophy of Science. Amsterdam: North-Holland. https://doi.org/10.1016/B978-0-444-51667-1.50040-9.

———. 2016. "An Ethical Framework for Evaluating Experimental Technology." *Science and Engineering Ethics* 22 (3): 667–86. https://doi.org/10.1007/s11948-015-9724-3.

———. 2017a. "Society as a Laboratory to Experiment with New Technologies." Edited by Diana M. Bowman, Elen Stokes, and Arie Rip. *Embedding New Technologies into Society: A Regulatory, Ethical and Societal Perspective*. Singapore: Pan Stanford Publishing.

———. 2017b. "Design for Sustainability." In *Philosophy, Technology, and the Environment*, 121–42. Cambridge, Massachusetts: The MIT Press.

———. 2018. "Moral Experimentation with New Technology." Edited by Ibo Van de Poel, Donna C. Mehos, and Lotte Asveld. *New Perspectives on Technology in Society: Experimentation beyond the Laboratory*. Oxon and New York: Routledge.

———. 2020. "Core Values and Value Conflicts in Cybersecurity: Beyond Privacy Versus Security." In *The Ethics of Cybersecurity*, edited by Markus Christen, Bert Gordijn, and Michele Loi, 21:45–71. The International Library of Ethics, Law and Technology. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-29053-5_3.

Van de Poel, Ibo & Olya Kudina. 2022. "Understanding Technology-Induced Value Change: A Pragmatist Proposal." In *Philosophy & Technology* 35 (2):40. doi: 10.1007/s13347-022-00520-8.

Van den Homberg, Marc J. C., Caroline M. Gevaert, and Yola Georgiadou. 2020. "The Changing Face of Accountability in Humanitarianism: Using Artificial Intelligence for Anticipatory Action." *Politics and Governance* 8 (4): 456–67. https://doi.org/10.17645/pag.v8i4.3158.

Van den Hoven, Jeroen. 2012. "Fact Sheet-Ethics Subgroup IoT-Version 4 . 0 1." https://www.semanticscholar.org/paper/Fact-sheet-Ethics-Subgroup-IoT-Version-4-.-0-1-Hoven/d88701968142b7b34198895ec923ed4c12c784d7.

Van den Hoven, Jeroen; Pieter E. Vermaas and Ibo Van de Poel. 2015. "Handbook of Ethics and Values in Technological Design. Sources, Theory, Values and Application Domains." Springer.

Van den Hoven, Jeroen, and Emma Rooksby. 2008. "Distributive Justice and the Value of Information: A (Broadly) Rawlsian Approach." In *Information Technology and Moral Philosophy*, edited by Jeroen van den Hoven and John Weckert, 376–96. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511498725.019.

Van Huijstee, Mariëtte, Pieter Van Boheemen, Djurre Das, Linda Nierling, and Jutta Jahnel. 2021. "Tackling Deepfakes in European Policy." PE 690.039. Brussels: European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA). https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf.

Vermesan, Ovidiu, Roy Bahr, Marco Ottella, Martin Serrano, Tore Karlsen, Terje Wahlstrøm, Hans Erik Sand, Meghashyam Ashwathnarayan, and Micaela Troglia Gamba. 2020. "Internet of Robotic Things Intelligent Connectivity and Platforms." *Frontiers in Robotics and AI* 7: 104. https://doi.org/10.3389/frobt.2020.00104.

Villa, Davide, Xinchao Song, Matthew Heim, and Liangshe Li. 2021. "Internet of Robotic Things: Current Technologies, Applications, Challenges and Future Directions." *ArXiv:2101.06256 [Cs]*, January. http://arxiv.org/abs/2101.06256.

Vizard, Polly, Sakiko Fukuda-Parr, and Diane Elson. 2011. "Introduction: The Capability Approach and Human Rights." *Journal of Human Development and Capabilities* 12 (1): 1–22. https://doi.org/10.1080/19452829.2010.541728.

Voegtlin, Christian, Andreas Georg Scherer, Günter K. Stahl, and Olga Hawn. 2022. "Grand Societal Challenges and Responsible Innovation." *Journal of Management Studies*. https://doi.org/10.1111/joms.12785.

Von Schomberg, Rene. 2012. "Prospects for Technology Assessment in a Framework of Responsible Research and Innovation." Edited by M. Dusseldorp and R. Beecroft. *Technikfolgen Abschätzen Lehren: Bildungspotenziale Transdisziplinärer Methoden*. Wiesbaden: Springer.

Walzer, Michael. 2008. *Spheres Of Justice: A Defense Of Pluralism And Equality*. Basic Books.

Wang, Qin, Rujia Li, Qi Wang, and Shiping Chen. 2021. "Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges." *ArXiv:2105.07447 [Cs]*, October. http://arxiv.org/abs/2105.07447.

West, Mark, Rebecca Kraut, and Han Ei Chew. 2019. "I'd Blush If I Could: Closing Gender Divides in Digital Skills through Education." GEN/2019/EQUALS/1 REV 3. EQUALS and UNESCO. https://en.unesco.org/Id-blush-if-I-could.

Wiederhold, Brenda K. 2018. "'Alexa, Are You My Mom?' The Role of Artificial Intelligence in Child Development." *Cyberpsychology, Behavior, and Social Networking* 21 (8): 471–72. https://doi.org/10.1089/cyber.2018.29120.bkw.

Wildt, T. E. de, I. R. van de Poel, and E. J. L. Chappin. 2021. "Tracing Long-Term Value Change in (Energy) Technologies: Opportunities of Probabilistic Topic Models Using Large Data Sets." *Science, Technology, & Human Values*, November, 016224392110544. https://doi.org/10.1177/01622439211054439.

Williams, Colin P. 2011. *Explorations in Quantum Computing*. Texts in Computer Science. London: Springer London. https://doi.org/10.1007/978-1-84628-887-6.

World Meteorological Organization. 2015. "WMO Guidelines on Multi-Hazard Impact-Based Forecast and Warning Services." Geneva: World Meteorological Organization. https://library.wmo.int/doc_num.php?explnum_id=7901.

WRR. 2021. "Opgave AI. De Nieuwe Systeemtechnologie." https://www.wrr.nl/publicaties/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie.

Wu, Yunhan, Daniel Rough, Anna Bleakley, Justin Edwards, Orla Cooney, Philip R. Doyle, Leigh Clark, and Benjamin R. Cowan. 2020. "See What I'm Saying? Comparing Intelligent Personal Assistant Use for Native and Non-Native Language Speakers." In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, 1–9. 34. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3379503.3403563.

Wynsberghe, Aimee van. 2021. "Sustainable AI: AI for Sustainability and the Sustainability of AI." *AI and Ethics* 1 (3): 213–18. https://doi.org/10.1007/s43681-021-00043-6.

Yaga, Dylan, Peter Mell, Nik Roby, and Karen Scarfone. 2018. "Blockchain Technology Overview." NIST IR 8202. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8202.

Ylianttila, Mika, Raimo Kantola, Andrei Gurtov, Lozenzo Mucchi, Ian Oppermann, Zheng Yan, Tri Hong Nguyen, et al. 2020. "6G White Paper: Research Challenges for Trust, Security and Privacy." *ArXiv:2004.11665 [Cs]*, April. http://arxiv.org/abs/2004.11665.

Zeng, Yi, Enmeng Lu, and Cunqing Huangfu. 2018. "Linking Artificial Intelligence Principles." *ArXiv:1812.04814 [Cs]*, December. http://arxiv.org/abs/1812.04814.

Zhou, Yongkun, Bin Rao, and Wei Wang. 2020. "UAV Swarm Intelligence: Recent Advances and Future Trends." *IEEE Access* 8: 183856–78. https://doi.org/10.1109/ACCESS.2020.3028865.

Zuboff, Shoshana. 2019. "Surveillance Capitalism and the Challenge of Collective Action." *New Labor Forum* 28 (1): 10–29. https://doi.org/10.1177/1095796018819461.

Appendix 1 – Overview of experts consulted and contributors to the study

Experts consulted

For this study we interviewed a number of experts. Of course, the usual caveat applies that they cannot in any way be held responsible for the contents of this study.

Name	Position	Organisation	Interview date
Dr. Aaron Ding	Assistant professor in computer science	Delft University of Technology	09-11-21
Dr. Adam Henschke	Assistant professor in philosophy	University of Twente	10-11-21
Prof. dr. Mireille Hildebrandt	Professor of digital security	University of Nijmegen	25-11-21
Dr. Mariette van Huijstee	Coordinator	Rathenau Institute	26-11-21
Prof. Dr. Wijnand IJsselsteijn	Professor of cognition and affect in human-technology interaction	Eindhoven University of Technology	22-11-21
Prof. Dr. Marijn Janssen	Professor of ICT and governance	Delft University of Technology	
Dr. Fernando Kuijpers	Professor of computer science	Delft University of Technology	29-11-21
Diego Naranjo	Head of policy	EDRi	23-02-22
Dr. Birna van Riemsdijk	Associate professor of intimate computing	University of Twente	10-11-21
Maarten van Steen	Scientific director of the Digital Society Institute	University of Twente	08-11-21

In addition, a draft version of the study was discussed by an advisory committee of scholars working on research projects that investigate the ethical challenges of disruptive technologies. The same caveat applies to them:

Name	Position	Organisation
Prof. dr. Philip Brey	Professor of philosophy of technology	University of Twente
Prof. Dr. Wijnand IJsselsteijn	Professor of cognition and affect in human-technology interaction	Eindhoven University of Technology
Dr. Birna van Riemsdijk	Associate professor in intimate computing	University of Twente

Contributors to the study

A number of people have made a contribution to specific parts of this study. The usual caveat applies that they cannot in any way be held responsible for the contents of this study:

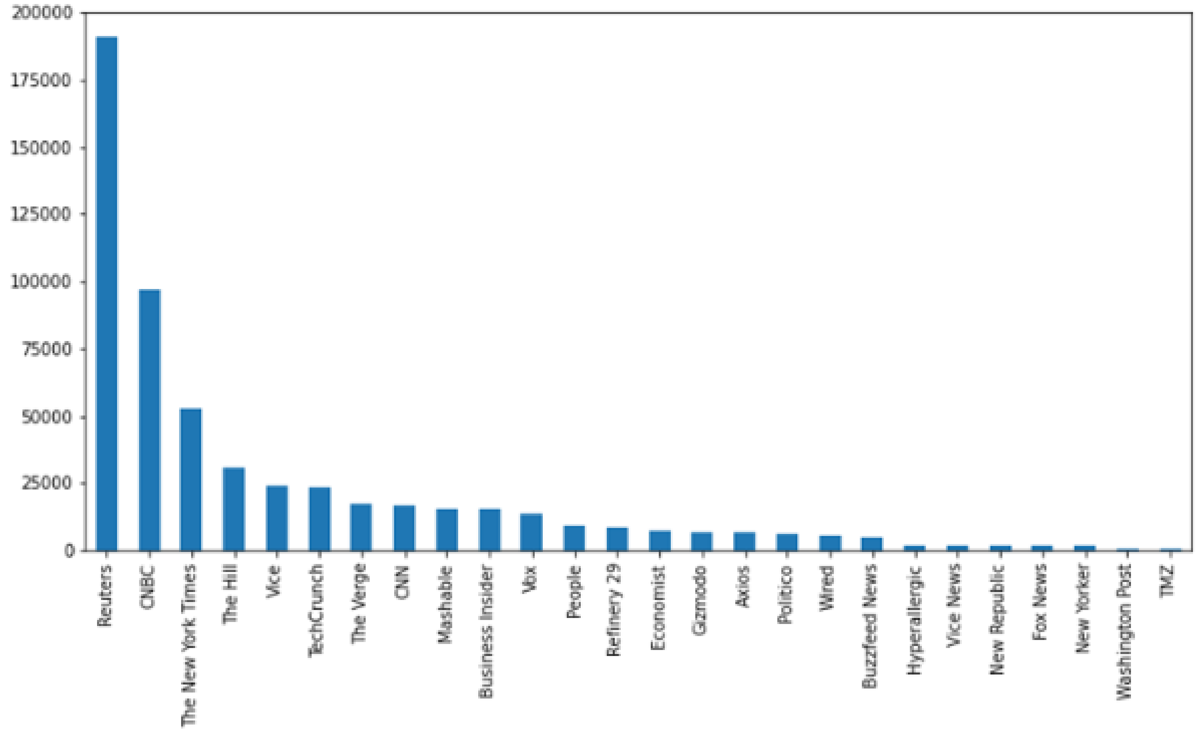
Name	Position	University	Contribution
Dyami van Kooten Passaro	Student assistant	Delft University of Technology	A large part of chapter 2
Dr. Olya Kudina	Assistant professor of ethics of technology	Delft University of Technology	Case 'smart digital voice assistants (chapter 5)
Dr. Michael Nagenborg	Associate professor in philosophy of technology	University of Twente	Case 'predictive humanitarian aid' (chapter 5)
Madhumita Naik, M.Sc.	Research assistant	Delft University of Technology	Appendix on the GDPR
Dr. Filippo Santoni De Sio	Associate professor in ethics of technology	Delft University of Technology	Case 'The next generations of autonomous vehicles' (chapter 5)
Prof. Dr. Wijnand IJsselsteijn	Professor of cognition and affect in human-technology interaction	Eindhoven University of Technology	Case 'The metaverse: real digital worlds' (chapter 5), policy option to address digital literacy (chapter 6)

Appendix 2 – Methodology text mining analysis

Description of datasets using for the topic modelling analysis

NEWS dataset

562.295 news articles from a range of newspapers, from 2015-02 to 2019-04



ETHICS dataset

8.565 scientific articles downloaded from ethics related journals (805 were related to 14 technologies studied):

- Science and engineering ethics
- Ethics and information technology
- AI and society
- Minds and machines
- Science, Technology, and Human Values
- Big data and society
- Ethical theory and moral practice
- Journal of responsible innovation
- Philosophy and technology

TECH dataset

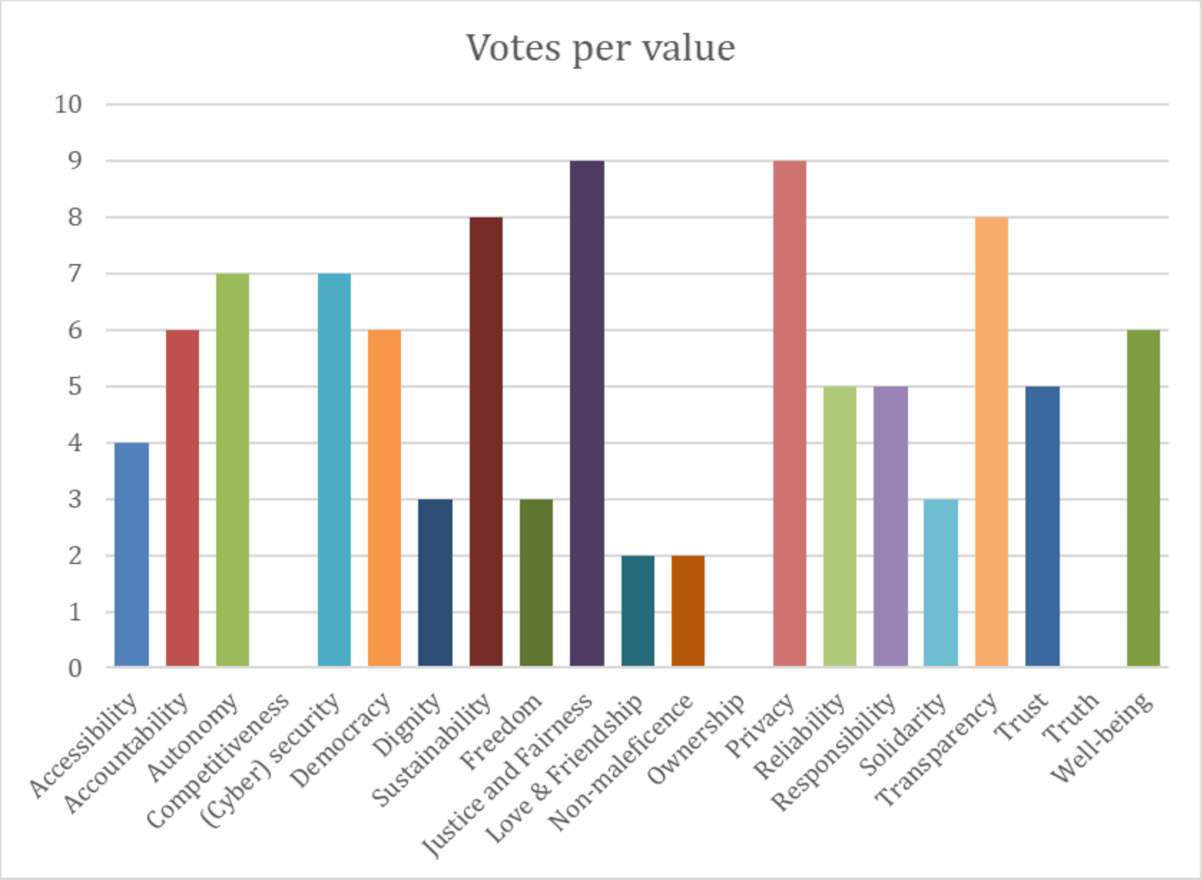
410.217 scientific articles downloaded from Scopus, selected using keywords related to each technology. Articles from journals related to ethics have been excluded from this dataset.

LEGAL dataset

Code	Regulation
2020/2016(INI)	Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters
2019/2186(INI)	Fair working conditions, rights and social protection for platform workers - New forms of employment linked to digital development
2020/0260(NLE)	European High Performance Computing Joint Undertaking
2019/2164(INI)	Promoting gender equality in science, technology, engineering and mathematics (STEM) education and careers
2021/2568(RSP)	The EU's Cybersecurity Strategy for the Digital Decade
2021/0068(COD)	Digital Green Certificate - Union citizens
2018/0328(COD)	European Cybersecurity Competence Centre
2020/2216(INI)	Shaping the digital future of Europe: removing barriers to the functioning of the digital single market and improving the use of AI for European consumers
2020/2017(INI)	Artificial intelligence in education, culture and the audiovisual sector
2018/0227(COD)	Digital Europe program 2021–2027
2020/2135(INI)	Shaping digital education policy
2020/2217(INI)	A European strategy for data
2019/2168(INI)	Closing the digital gender gap: women's participation in the digital economy
2019/2181(INL)	The right to disconnect
2020/2013(INI)	Artificial intelligence: questions of interpretation and application of international law in so far as the EU is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice
2020/2076(INI)	A New Industrial Strategy for Europe
2020/2019(INL)	Digital Services Act: adapting commercial and civil law rules for commercial entities operating online
2020/2012(INL)	Framework of ethical aspects of artificial intelligence, robotics and related technologies
2020/2014(INL)	Civil liability regime for artificial intelligence
2020/2015(INI)	Intellectual property rights for the development of artificial intelligence technologies
2020/2034(INL)	Digital Finance: emerging risks in crypto-assets - regulatory and supervisory challenges in the area of financial services, institutions and markets
2019/2915(RSP)	Resolution on automated decision-making processes: ensuring consumer protection and free movement of goods and services
2018/2115(INI)	Taking stock of the follow-up taken by the EEAS two years after the EP Report on EU strategic communication to counteract propaganda against it by third parties. Recommendation to the Vice President/High Representative of the Union for Foreign Affairs and Security Policy and to the Council

2018/2089(INI)	Autonomous driving in European transport
2018/2752(RSP)	Resolution on autonomous weapon systems
2018/2028(INI)	Language equality in the digital age
2015/2103(INL)	Civil law rules on robotics
2017/0003(COD)	Privacy and Electronic Communications
2020/0361(COD)	Digital Services Act
JOIN(2021)0014	Implementation of the EU's Cybersecurity Strategy for the Digital Decade.
2021/0106(COD)	Artificial Intelligence Act
2020/0340(COD)	European data governance (Data Governance Act)
COM(2021)0205	Fostering a European approach to Artificial Intelligence
COM(2021)0118	2030 Digital Compass: the European way for the Digital Decade
2021/0293(COD)	2030 policy program 'Path to the Digital Decade'
JRC125343	What future for European robotics?

Outcomes of the survey on relevant values for digital technologies

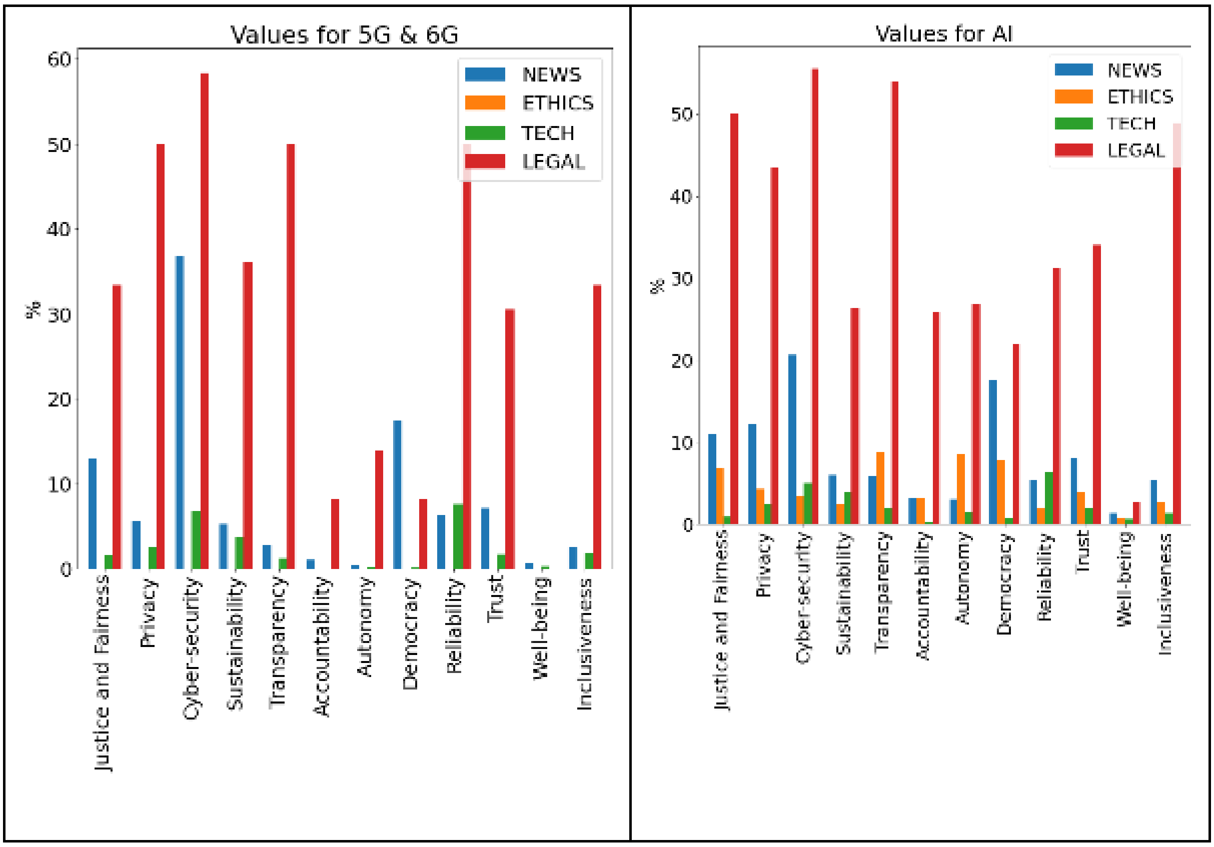


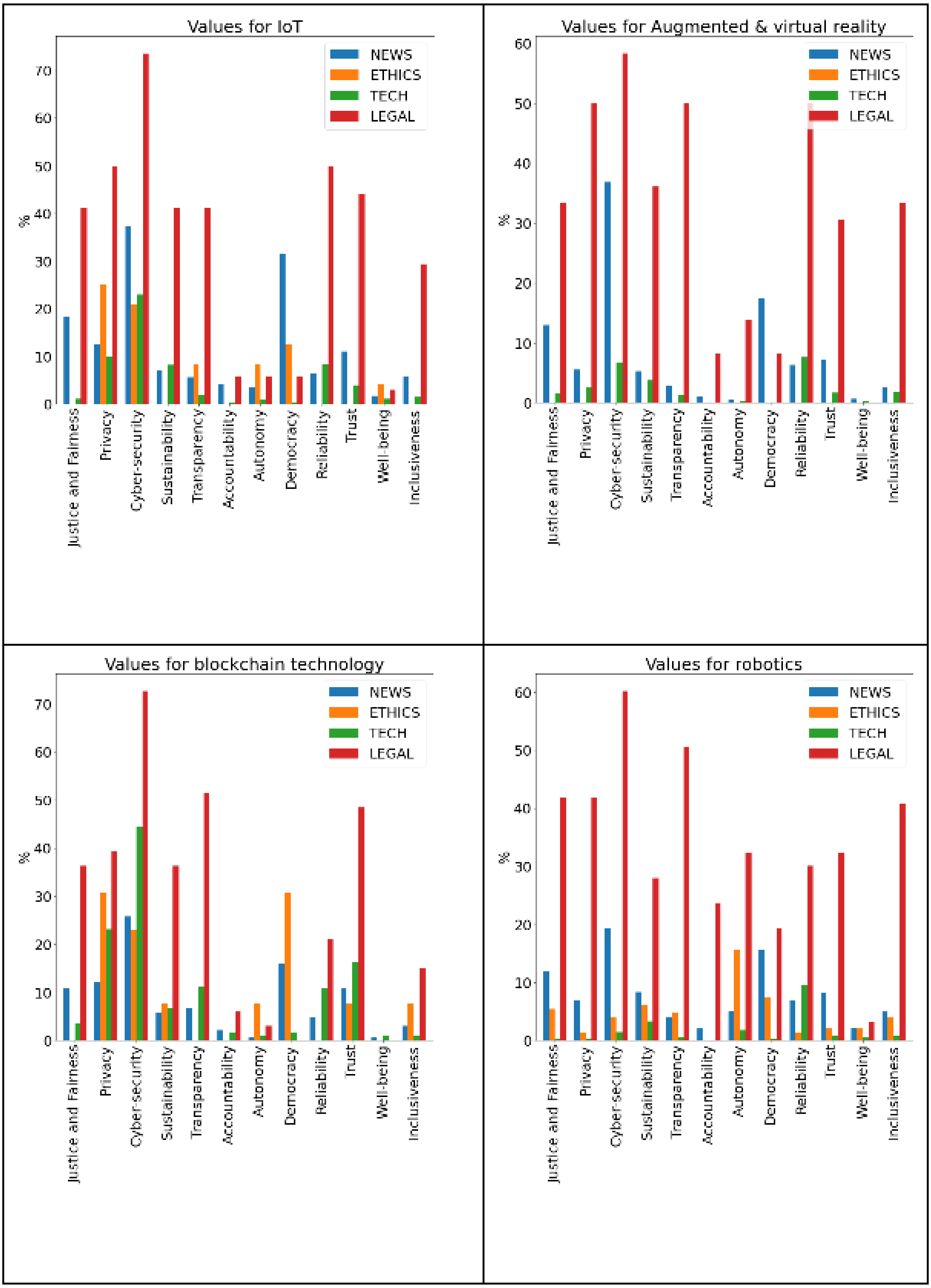
Anchor words and topic model created

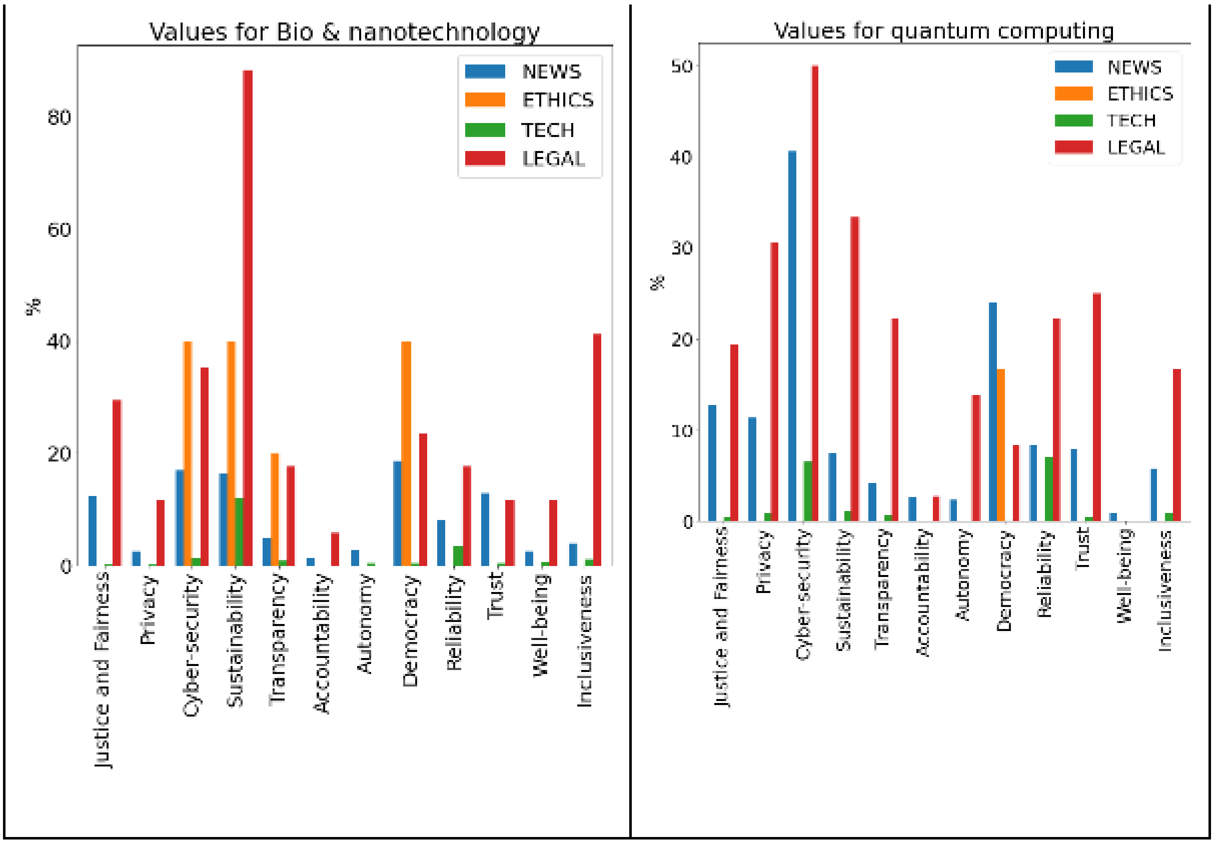
Values	Anchor words	Topic created
Justice and Fairness	"justice", "fairness", "fair", "equality", "unfair", "unequal" ,"unjust", "proportional fairness", "equitable"	Topic #0 (Justice and Fairness): justice, fair, fairness, equality, unfair, unequal, equitable, unjust, criminal justice, social justice, justice system, department justice, gender equality, free fair, egalitarianism, distributive justice, distributive, inequalities, entencing
Privacy	"privacy", "personal data", "personal sphere", "data privacy", "privacy protection", "privacy data", “privacy issues", "user privacy", "privacy preserving", "privacy concerns", "privacy preservation", "confidentiality"	Topic #1 (Privacy): privacy, personal data, data privacy, privacy protection, privacy concerns, privacy data, user privacy, privacy issues, security privacy, privacy security, personal information, privacy preserving, confidentiality, privacy preservation, privacy law, consumer privacy, facebook privacy, privacy information, privacy policies, issues privacy
Cyber-security	"cyber", "security", "cybersecurity", "malicious", "attacks"	Topic #2 (Cyber-security): security, attacks, cybersecurity, cyber, threats, malicious, encryption, social security, safety security, security issues, vulnerabilities, information security, malware, security system, cyber security, security concerns, security analysis, cyber attacks, security threats, security risks
Environmental Sustainability	"sustainability", "sustainable", "renewable", "durable", "durability", "sustainable development", "environmental"	Topic #3 (Environmental Sustainability): environmental, sustainable, sustainability, renewable, sustainable development, durable, renewable energy, carbon, emissions, environmental protection, greenhouse, pollution, dioxide, carbon dioxide, greenhouse gas, waste, carbon emissions, environmental impact, ecological, environmental social
Transparency	"transparency", "transparent", "transparently", "explainability", "interpretability", "explainable", "opaque", "explainable artificial", "explainable ai", "transparency data", "interpretable"	Topic #4 (Transparency): transparency, transparent, opaque, explainable, interpretable, explainability, explainable artificial, interpretability, transparency data, lack transparency, explainable ai, transparency traceability, data transparency, privacy transparency, system transparent, traceability transparency, information transparency, opacity, other organizations, deploy
Accountability	"accountable", "accountability", "accountable", "accounted", "traceability", "traceable"	Topic #5 (Accountability): accountability, accountable, transparency accountability, accountability transparency, algorithmic accountability, author accountability, government accountability, traceability, accountability office, fairness accountability, public accountability, traceable, divergence, uphold
Autonomy	"autonomy", "self-determination", "autonomy human", "personal autonomy", "decision making", "human beings", "human autonomy", "individual autonomy", "paternalistic"	Topic #6 (Autonomy): autonomy, human beings, paternalistic, personal autonomy, people human, privacy autonomy, decision making, self determination, paternalism
Democracy	"democracy", "democratic", "human rights", "freedom speech", "equal representation", "political", "voting", "elections", "participation"	Topic #7 (Democracy): political, democratic, democracy, elections, human rights, voting, liberal, freedom speech, social political, censorship, debates, economic political, political social, regime, authoritarianism, democracies, political economy, political economic, legitimacy, political leaders

Reliability	"reliability", "reliable", "robustness", "robust", "predictability", "predictable"	Topic #8 (Reliability): robust, reliable, robustness, predictable, predictability, global warming, warming, reliability, secure reliable
Trust	"trust", "trustworthy", "trustworthiness", "confidence", "honesty", "benevolence", "truthful", "truthfulness", "public confidence"	Topic #9 (Trust): trust, confidence, trustworthy, trustworthiness, honesty, public confidence, public trust, lack trust, transparency trust, trust privacy, people trust, security trust, trust management, key factor, negative impact, challenges future, auditors, truthful
Well-being	"well being", "well-being", "wellbeing", "quality life", "good life", "qol", "life satisfaction", "welfare"	Topic #10 (Well-being): welfare, quality life, wellbeing, good life, social welfare, welfare state, health welfare, behavioural, self governance
Inclusiveness	"inclusiveness", "inclusive", "inclusivity", "discrimination", "discriminate", "diversity", "exclusion","diversity inclusion", "bias data", "lack diversity"	Topic #11 (Inclusiveness): diversity, discrimination, inclusive, lack diversity, inclusivity, diversity inclusion, bias data, discriminate, inclusiveness, inclusion, discriminatory, racial discrimination, inferior, fosters, exclusion

Values for digital technologies mentioned in the different datasets




Values for IoT
Values for Augmented & virtual reality
Values for blockchain technology
Values for robotics
NEWS
ETHICS
TECH
LEGAL
%
0
10
20
30
40
50
60
70
Justice and Fairness
Privacy
Cyber-security
Sustainability
Transparency
Accountability
Autonomy
Democracy
Reliability
Trust
Well-being
Inclusiveness


Values for Bio & nanotechnology
NEWS
ETHICS
TECH
LEGAL
%
0
20
40
60
80
Justice and Fairness
Privacy
Cyber-security
Sustainability
Transparency
Accountability
Autonomy
Democracy
Reliability
Trust
Well-being
Inclusiveness
Values for quantum computing
NEWS
ETHICS
TECH
LEGAL
%
0
10
20
30
40
50
Justice and Fairness
Privacy
Cyber-security
Sustainability
Transparency
Accountability
Autonomy
Democracy
Reliability
Trust
Well-being
Inclusiveness

Appendix 3 - The EU's precautionary principle and innovation principle

The precautionary principle and innovation principle have been argued to be complementary; the two principles '*should be used alongside each other, recognizing the need to protect society and the environment while also protecting Europe's ability to innovate*' (European Risk Forum 2015).

The precautionary principle

The **precautionary principle (PP)** - developed out of discussions on environmental issues in the 1970s and 1980s - is now enshrined in the Treaty on the Functioning of the European Union (TFEU, article 191(2)) and in various legislative instruments of the European Union and those of several Member States. In February 2000, the Commission issued a Communication on the Precautionary Principle. According to that Communication, although the principle is mentioned in the TFEU only in the context of environmental protection, '*in practice, its scope is much wider, and specifically where preliminary objective scientific evaluation indicates that there are reasonable grounds for concern that the potentially dangerous effects on the environment, human, animal or plant health may be inconsistent with the high level of protection chosen for the Community*' (European Commission 2000, p.2). This is confirmed by jurisprudence of the Court of Justice, which finds the precautionary principle to be a 'general principle' of EU law.[41] The Communication describes the principle as a decision-maker's principle for risk management, which '*should not be confused with the element of caution that scientists apply in their assessment of scientific data*' (EC 2000, p.2). Its implementation should '*start with a scientific evaluation, as complete as possible, and where possible, identifying at each stage the degree of scientific uncertainty*' (EC 2000, p.3). The measures taken may range from banning a substance or procedure to initiating research or issuing a recommendation. Six requirements on applications of the principle have been laid down in the Communication on the Precautionary Principle (EC 2000, p.3 emphasis in original):

'*Where action is deemed necessary, measures based on the precautionary principle should be, inter alia:*

- *proportional to the chosen level of protection,*
- *non-discriminatory in their application,*
- *consistent with similar measures already taken,*
- *based on an examination of the potential benefits and costs of action or lack of action (including, where appropriate and feasible, an economic cost/benefit analysis),*
- *subject to review, in the light of new scientific data, and*
- *capable of assigning responsibility for producing the scientific evidence necessary for a more comprehensive risk assessment.*'

Each of these six requirements is further explained in the document. For instance, the requirement that applications of the principle should be subject to review is said to imply that the principle will only be maintained as long as the scientific information is incomplete or inconclusive. Precautionary measures should be '*periodically reviewed in the light of scientific progress, and amended as necessary*' (EC 2000, 4). In general, the document puts a strong emphasis on the use of scientific information in risk management, and restricts the application of the precautionary principle to cases of decision-relevant scientific uncertainty.

[41] Joined Cases T-74/00, T-76/00, T-83/00 to T-85/00, T-132/00, T-137/00, T-141/00 Artegodan v Commission [2002] ECR II-4945.

The innovation principle

The **innovation principle (IP)** can be considered as an emerging European *policy framework* rather than a 'principle' in its strict legalistic connotation, although the European Commission's EPSC has identified legal bases in the EU Treaty by which the innovation principle could complement existing legislation (European Political Strategy Centre 2016). The Innovation Principle requires that *'whenever the EU's institutions consider policy or regulatory proposals, the impact on innovation should be fully assessed and addressed'*; this leads to – as the claim goes – '*better management of risk* (European Risk Forum 2020).

Reference to the innovation principle was made in the Competitiveness Council Conclusions of May 2016 (Council of the EU 2016) and in the Commission EU Industrial Policy Strategy (European Commission 2017, p.14). The European Parliament has also expressed support for the innovation principle, for instance when it adopted a report on technical solutions for sustainable agriculture (European Parliament 2015). Several public statements and speeches by European Commissioners and Members of the European Parliament further confirm the commitment by the EU institutions to explore ways to implement the innovation principle as a contribution to overcoming important challenges such as climate change and food (in) security.

Appendix 4 - The General Data Protection Regulation (GDPR)[42]

The GDPR - put into effect in 2018 - was designed primarily to manage challenges emerging for the Internet which were not considered in the previous 1995 Data Protection Directive. It outlines a framework for the rights of data subjects and rules for how data must be processed. It protects data subjects by *minimizing* the *use* of *personal data* to *purposes limited* to those for which *consent* has been sought. Under the regulation, all processing is required to have a legal basis, and is deemed lawful only under the following conditions: *(a) the data subject has provided consent, (b) for performing or entering a contract, (c) for complying with a legal obligation, (d) for protecting vital interests (e) for performing a task in the public interest or in the exercise of public authority, or (f) for a legitimate interest.* Through the GDPR, data subjects have also gained more control over their personal data; It outlines well-defined rights to *access information* regarding processing and to *request erasure* of personal data. It also makes it possible for data subjects to object to various data usages, some of which are *processing* (for private commercial processes), *profiling, direct marketing* and *automated decision-making.*

Although or perhaps because the rights conferred by the GDPR are not focused on specific technologies, they are highly relevant for the challenges that new, upcoming technologies pose. However, the efficacy of the GDPR is contingent on the meaningful recognition of privacy and security and its protection by data controllers and the relevant supervisory bodies (Loideain 2019). For instance, the use of artificial intelligence is subject to certain tensions with regards to data protection principles - specifically, purpose limitation and data minimization. Effective AI creation requires the use of vast quantities of users' data, that may not always have been acquired for the specific purpose. Currently, open-ended clauses in the GDPR legislation allow for experimentation and learning in AI development, and transfer the responsibility of ensuring compliance and managing risks predominantly on data controllers (Finck 2019). In order to enable an environment where AI-based technologies can be developed in both an effective, and GDPR-compliant manner, reinterpretation of GDPR articles and increased guidance is required for both controllers and data subjects from data protection bodies (Ibid.).

While the technology-neutral principles-based approach adopted by the GDPR enables its application to different technologies, the legislation in its current form does not seem sufficient to ensure effective privacy protection on all accounts for upcoming technologies. Blockchain technology, for example, enables the creation of multiple data controllers, with different actors influencing the determination of the means of processing. The GDPR does not provide for such scenarios where multiple data controllers coordinate and govern data processing. This makes it trickier to assign responsibility and hold actors accountable. Another challenge is that '*the GDPR is based on the assumption that data can be modified or erased where necessary to comply with legal requirements such as Articles 16 and 17 GDPR. Blockchains, however, render such modifications of data purposefully onerous in order to ensure data integrity and to increase trust in the network.*' (Finck 2019). Furthermore, the '*distributed ledgers [used in blockchain] are append-only databases that continuously grow as new data is added. In addition, such data is replicated on many different computers. Both aspects are problematic from the perspective of the data minimisation principle*' (Ibid.).

Internet of things (IoT) networks range in size from tens to millions of devices, with potentially heterogeneous characteristics related to resource constraints, mobility, degree of autonomy, etc. Privacy issues in such networks are largely specific to the type of applications deployed. For instance, health monitors are designed to collect personal and sensitive data such as blood pressure, heart rate, etc. from their users. While ensuring access to sensitive data is necessary, making it available over the internet can lead to severe privacy risks. In June 2015, for example, malware-infested blood gas analysers allowed hackers to enter and access hospital networks (Storm 2015). IoT-enabled devices are now increasingly used in smart homes. These also lend themselves to privacy issues as internet service

[42] This section was written by Madhumita Naik, M.Sc.

providers might have access to data regarding users' behaviour patterns with or without their consent (Porambage et al. 2016). Thus, IoT users are likely to be subject to privacy issues despite GDPR-compliant devices due to the diffuse nature of data, global supply chains and manufacturers, usage of cloud technologies, etc. Similar privacy issues are applicable to 5G networks, which have access to large quantities of personal data on a global scale.

Two different legal approaches to data protection to further increase the efficacy of the GDPR are a right based and a risk-based approach (Finck 2019). The GDPR legislation currently leans more toward the right-based approach, clearly defining data subjects' fundamental rights to privacy and data protection. The risk-based approach focuses more on creating a healthy information environment (social and technological), where '*harm is prevented by appropriate organizational and technological measures*' (Finck 2019). Upcoming technologies are likely to create circumstances where fundamental rights can be violated. The convergence of these technologies will also create a socio-technical system where personal and sensitive data is at risk. Solutions could be found through dedicated research to understand suitability of the GDPR to specific technologies, and further interdisciplinary research into compliance by design (Section 6.2.4).

Supported by the arrival of 5G and, soon 6G, digital technologies are evolving towards an artificial intelligence-driven internet of robotic and bionano things. The merging of artificial intelligence (AI) with other technologies such as the internet of things (IoT) gives rise to acronyms such as 'AIoT', 'IoRT' (IoT and robotics) and 'IoBNT' (IoT and bionano technology). Blockchain, augmented reality and virtual reality add even more technological options to the mix. Smart bodies, smart homes, smart industries, smart cities and smart governments lie ahead, with the promise of many benefits and opportunities. However, unprecedented amounts of personal data will be collected, and digital technologies will affect the most intimate aspects of our life more than ever, including in the realms of love and friendship. This study offers a bird's eye perspective of the key societal and ethical challenges we can expect as a result of this convergence, and policy options that can be considered to address them effectively.

QA-09-22-336-EN-N

This is a publication of the Scientific Foresight Unit (STOA)
EPRS | European Parliamentary Research Service

This document is prepared for, and addressed to, the Members and staff of the European Parliament as background material to assist them in their parliamentary work. The content of the document is the sole responsibility of its author(s) and any opinions expressed herein should not be taken to represent an official position of the Parliament.

ISBN 978-92-846-9682-6 | doi: 10.2861/61151 | QA-09-22-336-EN-N