7 Governing the Median Estate

Hyper-truth and post-truth in the regulation of digital innovations

Kjetil Rommetveit and Niels van Dijk

Introduction

Post-truth discourse seems to thrive on the assumption that before there was truth in public, whereas now there is not. As testified by the contributions to this book, this assumption is simplistic. Yet, as the book's various contributions also argue, *something* has changed: post-truth discourse prevails and translates into new shapes, territories, and problem domains. A more promising approach is to argue with Foucault and Pellizzoni (2017, this issue) that post-truth denotes intensifications of certain of modernity's core dynamics, especially those concerning uses of science in public. In this chapter, we pursue this intuition into major areas of the production of knowledge, namely (European) legal and regulatory efforts to steer digital innovation and render it more accountable (see also van Dijk, this volume). Here, there are direct connections with post-truth (ibid.), and more indirect ones, by which we refer to the developments of digital innovation and regulation more generally. As a starting point, consider how, according to Evgeny Morozov: "One unappreciated paradox of today's 'digital condition' is that it celebrates post-truth and hyper-truth simultaneously" (Morozov 2019).

Through our descriptions of two cases, *privacy engineering* and *personhood for machines*, we shall make two interlinked points that connect post-truth to the theme of governance and regulation of the digital: (1) alongside post-truth there is also hyper-truth, i.e. innovation policies imagined as so self-evidently true that they cannot be questioned, as captured in Morozov's quote, and these may be more closely related to post-truth than previously recognised; (2) modern western societies rely on different modes of truth-telling, such as those of law, science, markets, technology, and politics. Post-truth entails intensified *blurring and remaking of fundamental boundaries* between these modes (conceptual and institutional), and these are reflected in broad-scale changes to collective imagination through the knowledge and information economy.

DOI: 10.4324/9780429053061-10

Post-truth discourse and digital hyper-truths

As to the first, we point to an omission in the discussions around post-truth, indicated in the introductory quote from Morozov, and topic of several of the contributions to this volume. This is the occurrence of certain digital hyper-truths, or digital imaginations, underpinning agendas such as Internet of Things, Smart developments, and Fourth Industrial Revolution, and granted "automatic authority in public issues" (Wynne 2014). The introductory quote from Morozov posits this as a conundrum: whereas under post-truth conditions seemingly any truth and its wider framing can be questioned, this unfolds alongside digital innovation regimes whose basic assumptions and premises seem so self-evidently *true* that they are almost impossible to question. Whereas it is *possible* to publicly question the reality of human-induced climate change, it actually seems harder to challenge the necessity and desirability of the smart phone, the next generation of cloud-based processing, storage, and networked services, and the digitalisation of evermore aspects of physical and social reality. It is seemingly only when major institutions such as parliaments, courts, and mainstream media become exposed to existential danger by digital technologies, that broad public questioning becomes possible (van Dijk this volume, Rommetveit, this volume). If this is the case, then a likely explanation is that there is a dynamic relation at work, also implied by Morozov, where hyper-truth produces post-truth conditions, as we shall describe in our two cases.

What kind of "truth" is "hyper-truth" in our case? This self-evident type of "truth"[1] can pertain to different things. First, the digital technologies that have been instrumental in undermining the epistemic authority of institutions such as science, politics, and the media through the spread of disinformation, themselves rely upon conceptions of fact and truth that have become black-boxed and taken for granted. These are based in computer science and historically derive from the epistemology of statistics (van Dijk, this volume). Second, Morozov points rather at the attribution of objectivity to information on digital platforms and algorithmic ledgers, such as Wikipedia and Blockchain.[2] These are digital means of producing knowledge and evidence in non-expert related ways. In this chapter, we expand on this diagnosis, to also include main digital imaginaries and innovation agendas for the future of our societies, presented as inevitable collective developments and self-evident public truths (Wynne 2014). Digital technologies framed as smart and enabling, and as contributing to a new industrial paradigm (Industry 4.0), come enshrined in a strongly universalistic rhetoric where *digital* applies to *any thing*, *anywhere* and at *any scale* (i.e. from nano-molecules to smart cities to IBMs *Smart Planet*), any *process* (of work, traffic systems, manufacture value chains, or news feeds), to any *person*, *organisation*, or *collective* (i.e. Facebook "Global Community").

"Truth" in this sense does not refer to the classical (early 20th-century) image of a correspondence between factual representation and reality;

rather, we point to meanings, imaginations, and future pathways enabled through digital technologies. Heidegger (1978) referred to the technological making of new worlds and meanings, as a "revealing that orders".[3] This ordering is embedded in the many "interlocking parts" that make up technological systems and infrastructures, and the activities and practices enabled through them: "unlocking, transforming, storing, distributing and switching about are ways of revealing" (ibid., 322). *What* is revealed and ordered is not merely technology, but societal and human meanings and relations, and the capacity for creating collective orders (Bijker et al. 1987). To Heidegger, this was a specifically futures-oriented mode of *Being-in-the-World*. This future-orientation is part of what has intensified, since it is increasingly also revealed through highly mediated visions and promises of technological, societal, and industrial convergence of today's most powerful technologies,[4] whose function is also strongly political, since the promise is to address societal problems and to remake the global economy. As testified by sociologists (Brown and Michael 2003, Fortun 2008), innovation as futures promise and expectation has increasingly been pushed to the forefront of collective meaning- and market-making. Concomitant with this, *imagination* has long since been recognised as a performative and constitutive force (Appadurai 1996), and the imagined-possible a source of epistemic, industrial, and political authority within intensified networked innovation (Rommetveit and Wynne 2017, Rommetveit, this volume, Wynne, this volume).

The imaginary of Fourth Industrial Revolution is for instance predicated on "a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres" (Schwab 2016). It performatively draws together, *at the level of the imagined-possible*, most of today's powerful technologies, opening up new domains of nature and society to market-making, exploitation, technological and economic development. In ways similar to Heidegger's concept of the technological framework (*Gestell*), it is the creation of future pathways that become inscribed into collective consciousness as destiny[5]: "In its scale, scope and complexity (…) the Fourth Industrial Revolution is unlike anything humankind has experienced before (…) from the perspective of human history, there has never been a time of greater promise or potential peril" (Schwab 2016). Following the disclosure of this future potential, the ensuing task for policy makers and regulators is to create the terrain on which the mission can be carried out. It is a *will* increasingly targeted towards, and predicated upon, the overcoming of barriers in the bio-physical world, namely those that stand in the way, *qua* obstacles, of the expansion of technoscientific potential and realisation. Thus, the basic orientation is ontological (or: ontic) and not epistemic, and the underlying imagination of nature is as investment-resource: it is projected as plastic and amenable to be shaped through technoscience (Pellizzoni 2015). This boundary-blurring and constructivist orientation feeds directly into efforts to regulate, as we now describe.

Blurring boundaries between modes of veridiction

Post-truth characteristics understood as blurring and remaking of boundaries replicate within efforts to regulate digital innovations. Law and regulation are different modes or practices of truth-telling or "veridiction", and are different from scientific truth (Latour 2013). The specific theme on which we focus here is the capacity of digital technologies to *blur major boundary distinctions* taken as constitutive of western societal orders, such as *fact* versus *value*, *human* versus *machine, science* versus *politics*. Ensuing incapacities to work out the different modes of truth-telling is at the core of STS and philosophical discussions of post-truth (Collins et al. 2017, Jasanoff and Simmet 2017, Sismondo 2017, Fuller 2018, Nordmann 2020). Here, we stick with this theme, and point to its intensification into novel domains and practices, with (we claim) unprecedented implications: Whereas such blurring may not lead directly to post-truth conditions, it certainly feeds into them, since there is a lack of easily available categories (epistemic and institutional) for sorting out novel phenomena, and for making collective sense. Furthermore, insofar as post-truth is taken to imply the dominance of value, opinion, and imagination over scientifically proven fact, the strong investments into the shaping of collective imaginations and futures, indicate also more direct relations of mainstream knowledge production to post-truth.

Blurring of boundaries and hybridisation is a major theme of STS, including in works on science and politics (Latour 1993, Weingart 1999, Jasanoff 2004, Bijker et al. 2009). The STS thesis of co-production (Shapin and Schaffer 1985, Jasanoff 2004) states that there are strong intrinsic relations between the ways in which scientific facts, technologies, and societal and cultural values evolve. The thesis demonstrates how these different sources of legitimacy were balanced and kept separate through buffers (Wynne this volume) upholding a "modern framework" (Toulmin 1990). Considerable practical work and resources went into keeping Nature and Culture separate, termed work of purification (Latour 1993), boundary work (Jasanoff 2011), or (balancing hybridisation with purification) coordination work (Bijker et al. 2009). The practical *work* to keep domains separate, and in mutual balance, can be observed in classical studies from STS about the creative relations and tensions between science, technology, and law: technosciences such as biotechnology or climate science introduce novel entities such as the global climate system, or genetically engineered embryos. They upset existing meanings and procedures: are they human or non-human? Nature or Culture? Do they belong within the realm of the scientists and engineers, or those of lawyers and politicians? They trigger efforts to remake order and to accommodate the new entities (hybrids) within existing institutions and meanings. As described by Jasanoff (2011), it has become the task of professional actors such as lawyers and ethicists to reconstitute ontological and institutional boundaries. By sorting things out (Bowker and Star 1999) and giving each thing its proper ontological status, such creative and adaptive boundary work situates the new entities within cultural and institutional meanings and categories, and enables society to go on with its business.

Yet, as already argued: crucial mechanisms of digital technologies and their related forms of socialisation upset these capacities, and sometimes also actively overturn them. They thereby catalyse the blurring of boundaries between central "modern" conceptual distinctions, which has been one of the central themes of the post-truth discourse (see Rommetveit, this volume). As we note in the conclusions, this poses a tricky question: whereas STS has, almost by routine, used hybridisation and the front-staging of non-humans as a critical corrective to overtly idealistic and purified notions of truth, when hybridisation itself is part of the regulatory response by main institutions, this critical repository is no longer available *qua critique*, and may even turn into a reactionary principle.

One paradigmatic case of such intensified blurring and remaking of boundaries are attempts at the intersection of computer science and ethics at building morality and moral agency into robots, since this cuts through both distinctions between facts and values and between humans and machines (Wallach and Allen 2009, Vanderelst and Winfield 2018), as illustrated in this figure:

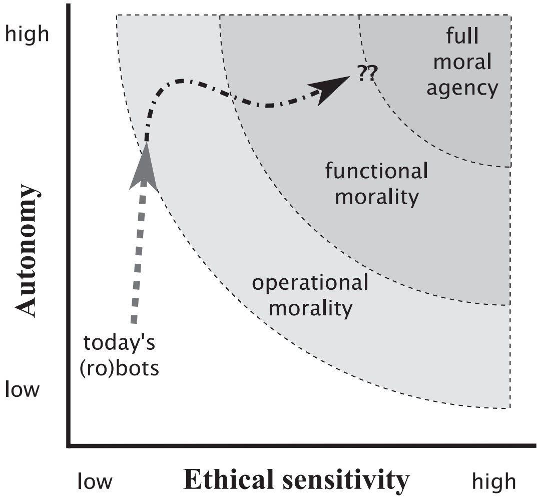

Figure 7.1 Intensification of modernity's forces: Building morality into machines. From: Wallach and Allen (2009), Copyright © 2009 by Oxford University Press, Inc. Reproduced with permission of the Licensor through PLSclear.

We include the case of engineering of morality as emblematic of the problems discussed in this chapter, but we do not analyse it here. Our two cases are however closely related, dealing with privacy engineering and legal personhood for machines. Common to these cases are how digital technologies have become sufficiently powerful, their dynamics so intense, that they not merely infringe on core normative and legal domains, but crucially also *renders them objects of design and engineering interventions*. The effect is, as we said, a blurring of basic distinctions, categories, and institutional arrangements, basic to western orders, with resulting incapacities for sorting things out, for making and upholding the existing social metaphysics.

Governance of the Median Estate

This institutional remake was captured by Lessig's (1999/2006) emblematic and provocative statement: that digital code *is* law. Yet, if this statement is accepted, it means that regulatory practices generally are not up to the task of regulating (since most regulatory practices do not comply with the ideal of becoming code): Law and regulation must be redesigned on a grand scale. This remaking of regulatory practices and law must be seen against the background of quite profound shifts in the political economy of knowledge. It is the distributed nature, complexity, and *speed* of developments that demands new governance mechanisms. Returning to the agenda of Fourth Industrial Revolution,[6] or Industry 4.0, this challenge has been captured through a concept of *agile governance*:

> *Agility* implies an action or method of nimbleness, fluidity, flexibility or adaptiveness. In the software sector, the concept of agile or "agility" has been around since the 1990s. The difference between plan-based methods of policy-making and the concept of agile governance relates to the shift in the value placed on time sensitivity.
>
> (WEF 2018, 4)

This logic of agile, networked governance replicates the "connectivist" logic of ICTs, and mobilises strongly universalist rhetoric taken from cybernetics, systems theory, converging technologies, informatics, and data science (Bowker 1993, Nordmann 2004, Kline 2015). The rhetoric is at once inclusive, since it aims to mobilise the actors and networks (the "multistakeholders") necessary for enabling digital innovation, *and* excluding, insofar as individuals and publics identified with obstacles to innovation are deliberately left out or circumvented. When perceived as standing in the way of the digitally driven networks, the publics can simply be ignored or deleted, resulting in an obstacle model of publics (Rommetveit and Wynne 2017, cf. Welsh and Wynne 2013), and of other modes of public veridiction such as science, law, or morality (Rommetveit et al. 2020).

It is here, in the midst of the technological economy and the tying of technological markets into weakly defined structures of governance, that we locate what we call *the Median Estate*. It denotes a normative change in the nature of the median space between the institutional stratifications of modernity, coinciding with its uptake in governance and innovation frameworks. There is an increasing policy related push for dissolving ontological, disciplinary, sectoral, and societal boundaries within technoscientific innovation networks targeted at addressing societal challenges. Old "trading zones" hereby move from their peripheral or residual positioning to the intermediary region "between" institutional or disciplinary "silos" between human and machine, facts and values, nature and culture, effectively acting as an innovation imperative. Here they become new centres of socio-technical ordering (apparent in imperatives to break silos, be interdisciplinary, cross domains and sectors). It is catalysed by the expansion of digital networked and networking technologies across and into evermore domains of society.[7] Innovation for a long time (i.e. since the early 1970s) belonged mainly within industrial domains. The concept of the Median Estate captures the expansion of the logics and discourses of informational machines into core institutional (even constitutional) domains: morality and legality, and also democracy.

Case studies: law becoming technology, technology becoming law?

We now turn to our two cases, of privacy by design and electronic personhood, where this problematic is described and analysed. Our accounts are based in prior investigations, including document studies, issues mapping, interviews, and focus group consultations with main concerned parties and salient forms of expertise (see van Dijk et al. 2016, 2018, Rommetveit et al. 2018, 2020). In both cases, we observe how fundamental sources of legitimacy as described in this section, are not so much separated, singled out, and relegated to different institutional, ontological, and expert domains, but rather fusioned and brought together, within the same space of representation and intervention.

Data protection by design: inserting a human in the smart loop?

Our first case pertains to the institutionalisation of the fact/value distinction, and its gradual change through networked co-regulation and techno-regulatory approaches, aimed at the creation of new markets and protection of legal rights. We observe how, within these novel forms of governance, new articulations of data protection and privacy emerge. We refer to these articulations and relations as *privacy-by-network*.

The 1990s brought the development of the Internet, through rapid expansions of personal computing power and networking capacities. Whereas initially accompanied by celebratory and highly optimistic visions for democratisation of information, the 2000s saw rapid incursions of corporate and commercial interest into cyberspace. This shaped the emergence of a ubiquitous "surveillance capitalism" (Cohen 2017, Zuboff 2018) predicated on profiling and individualised targeting of consumers behaviours through online platforms. All along, these developments have been accompanied by privacy concerns, as surveillance capitalism is, *as if designed* that way, predicated on the breach of privacy and data protection principles, such as purpose specificity, data minimisation and informed consent. Public regulators were met with demands from critical publics that the privacy concerns be dealt with, but also from corporations in need of predictability and safeguarding of reputations in rapidly expanding digital markets. Upscaled regulatory measures were seen as necessary in the face of a second wave of technological explosion: smart technologies and the Internet of Things, predicated on new data sources across the digital–physical interface (through sensors, social media, handheld devices, etc.), increasing algorithmic capacities to merge data from various sources, and machines that think, learn, connect, and (sometimes) act. Within emerging digital imaginaries *any source* of data may be connected to any other source of data that link it to people in unforeseen ways, and therefore constitute potentially personal sensitive data, triggering rights to protection (Purtova 2018).

The perception has emerged that law was "lagging behind" technological developments (Reidenberg 1998) and needed to catch up by adapting technological characteristics: more pro-active, incorporated into the design of systems, user-centric, and oriented towards futures developments. As stated in a foundational document on privacy by design, privacy by design "anticipates the risks and prevents privacy invasive events *before* they happen" (Cavoukian 2009). This preventive regime increasingly becomes continuous with logics of pre-emption (cf. Pellizzoni, this volume) in highly competitive technological markets, the main point being to demonstrate how privacy concerns have *already* been taken care of through proper procedures.

A new professional field of *privacy engineering* has emerged to take care of and implement these developments in more holistic and cross-cutting manners (Dennedy et al. 2014, Gürses and Del Álamo 2016). Privacy engineering denotes the *integration of privacy concerns into engineering practices for systems and software engineering life cycle processes* (ISO). Privacy engineers work to bridge across law and engineering, seeking to design and engineer legal principles into technological artefacts and infrastructures (Rommetveit et al. 2018). Professionals within this emerging field are envisioned to work across boundaries and scales: the individual technological application,[8] within organisations, and at standardisation and infrastructural levels. Several of these developments come enshrined in the recently (2018) adopted General Data Protection Regulation (GDPR) of the European Union, where both data protection by design (Art. 25) and data protection impact assessments (Art. 35) are mandatory for those who process personal data, and are backed up by enhanced accountability mechanisms. A prime example here is the European Citizen-Centric approach to Data,[9] aimed to design privacy and data protection into emergent infrastructures, technologies, and services at the level of European (smart) cities (Ballo and Vaage, this volume), in ways that are conducive to the needs of citizens and protective of their rights. Here, privacy by design is mobilised for the pro-active making of new technology-driven markets, infrastructures, and living environments.

The first initial steps towards the present practices had already been taken in the mid-1990s under the heading of Privacy Enhancing Technologies (PETs) and were important to Lessig's (2006) formulation that "law is code". These were mainly targeted at self-protective measures by users engaging in "informational self-determination", through techniques such as encryption, anonymisation, and data minimisation (Hes and Borking 2000). Yet, due to technical complexity and widespread proliferations of data, informational self-determination is beyond the capacities of most users. The European General Data Protection Regulation (GDPR) and Privacy by Design before it, therefore, introduced obligations of data controllers to shift responsibilities onto *organisational*, not individual, levels. GDPR Art. 25 prescribes Data Protection by Design to be undertaken by data controllers (not the user), and this is accompanied by requirements to carry out risk assessments, also at the organisational level, so-called Data Protection Impact Assessments (GDPR Art. 33, cf. van Dijk et al. 2016).

Yet, most information flows, and especially those of smart interconnected technologies, cities, and societies, span more than one organisation only. They have to rely on pre-made technologies (made by other actors), network connections and platforms (such as Google), collaborations with other businesses or organisations, not to forget the "downstream" users (professionals, consumers, patients, etc.). There is little sense in Data Protection by Design becoming implemented at the level of one organisation, if these other actors do not engage in similar and coordinated data protective measures. As explained to us by one privacy engineer working for an energy utility: "the discussion should have been taken from the chain point of view. In this way the transparency of the smart meter would have been discussed in an early stage with all the stakeholders that are related in the chain". Weaknesses in one link may cause rapid escalations of risk throughout the entire chain or ecosystem, and this in turn becomes an argument for scaling up.

Efforts and strategies are made therefore to capture and incorporate individual and collective perceptions of privacy threats into the ongoing infrastructural construction work.[10] Privacy has been called a "subjective" value (Solove 2008) with strong contextual variations (Nissenbaum 2004). One of the main "gaps" to be filled is that between the privacy perceptions of users and citizens, and the technical characteristics of emerging systems. This task, however, poses a Catch 22-like situation: the engineer cannot explain the problem to "the user" (who remains unknown) before the technicalities are in place. Yet, the technicalities are considerable, and cannot be encoded before the privacy concerns are known. Thus, as explained by one privacy engineer "Many efforts currently go into putting technical complexity at work…99% focus of technical people is about solving that" (privacy engineer). If technical challenges are not somehow overcome (at early stages of implementation), it is difficult to see how rights can be built into the information infrastructures, in ways that are explainable and acceptable to users and citizens, but also to the courts (cf. van Dijk et al. 2018).

Because of this complexity, main strategies and approaches come to hinge on the creation and use of proxy user profiles (cf. Silvast et al. 2018), and customer management approaches. The concept of "user" as a holder of rights in the emergent Internet of Things, therefore, is predicated on technical and managerial requirements revealed by engineers and consumer studies, rather than those of law. Yet, as also happens inside organisations (van Dijk et al. 2016), privacy becomes a managerial and organisational principle whose main purpose is to provide temporary stabilisation of expectations: setting forth a promise that rights shall be implemented and protected, thereby enabling mutual alignments between involved actors. Within this reconfigured space, privacy takes on decisive characteristics of what STS scholars term *boundary objects* (Bowker and Star 1999), representing different realities within different public, professional, and private domains (cf. Ballo and Vaage, this volume). Thus, a privacy engineer explained to us how privacy becomes a kind of "transversal concern" alongside other values and requirements:

> When we want to take into account privacy and other concerns, we have to take them into account as transversal concerns […] security, privacy, safety, energy consumption or taking into account ethical aspects and things like that. […] we need to be able to engineer transversal concerns and 'capabilities' in things (privacy designer).

In such ways, the overall purpose and scope of data protection expands considerably, not merely as a consequence of technological developments and "data explosion", but as the result of political and industrial motivations to create the internal (European) digital market. *Privacy-by-network*, thus, emerges simultaneously as: a fundamental right, as a regulatory measure (aimed to provide predictability and stability), as a market-making device (aiming to enhance competitiveness), as a matter for engineering, a principle for implementation in large organisations, and as innovation imperative (to create the digital market). This implies that the protection and regulation of the fundamental rights to privacy and data protection, move out of legal-regulatory institutions, and into more privatised, and also technology-centred, environments, captured by terms such as co-regulation, standardisation, and public–private partnerships.

Personhood for machines: new members of society, or threats to human rights?

Our next case concerns more basic (ontic) perceptions of the fundamental distinction between human and machine, and the normative and institutional implications thereby entailed. In this sense, it enters at a more basic level of collective imagination than the previous case. In Europe, robotics come embedded in increasingly institutionalised initiatives through governance and innovation schemes such as Joint Technology Initiatives, technology platforms, and public–private partnerships (PPPs). Through increasing entanglements of innovation agendas with public institutions, robotics is promised to contribute to a number of societal challenges, or "missions", such as caring for the ageing society, remaking European industries in highly competitive global markets (following the 2008 economic downturn), and rendering the world of work more efficient through increased human–machine interactions (cobots) and automation. This can be seen clearly in the field of robotics, and its role within a "Fourth Industrial Revolution" (cf. Schwab 2016, Fuchs 2018). Whereas most digital innovation agendas are predicated on the convergence of a whole host of different technologies, robotics is the literal embodiment of such initiatives: seen as a kind of "discipline of disciplines", it assembles technologies that sense, think, act, and feel into concerted plans for innovation and development. Projects are now under way to implement robotic technologies into self-driving cars, drones operating by themselves, as "cobots" at work, and as robot companions for care and companionship at home. These initiatives go well beyond the confines of research laboratories or factories and are implemented as large-scale industrial and technological remakes in and on society itself (Rommetveit et al. 2020).

Yet, the technological and market structures in which robotics applications would thrive only exist partially, and mainly as promise: they are in need of being built and made. This poses huge challenges on a number of levels, from deep-seated public fear and stigma of machines (going all the way back to the Luddites), to hugely complex legal and technical matters relating to control, and to loss of control. Specifically, since the machines are intended to operate "autonomously", and to take on capacities of learning and adaptation in unstructured environments, their behaviours will be unpredictable. This has issued in a diagnosis of a "responsibility gap" (Matthias 2004) and an "accountability gap", referring to the impossibility of predicting and controlling the behaviours of "intelligent" machines that act (seemingly) autonomously, but also the organisational challenges of implementing responsibility across value- and production chains that cut across national, regional, and institutional boundaries. Yet, it is also a "market gap" and an "innovation gap", since there is a need to provide certainty (about possible adverse consequences) in the face of the insecurities introduced by autonomous machines, among groups as diverse as policy makers, investors, innovators, entrepreneurs, researchers, users, and publics. It was within this broad context that propositions were made (in Europe) for the creation of "electronic personhood" for machines, as one way of addressing the regulatory gaps.

The idea of an artificial personhood had circulated for years in academic and legal scholarship (Solum 1992). Yet, it was introduced to a European legislative agenda by industrial robotics networks, setting themselves up as main actors at a European institutional level: first through the making of a technology platform, then by entering into a public–private partnership with the European Union. The robotics industry has gradually become more assertive and has established itself as a main mover and a key enabler in the making of a European market for robotics (in manufacture, care, and companionship, at work and in public spaces). Following the Covid-19 pandemic, the role and promise of robotic automation of tasks can be expected to increase even more, across fields of application (care, services, infrastructure repair, farming, etc.).

In 2012, the industry network *euRobotics* issued a Green Paper on *Ethical Legal and Societal issues in robotics* (Leroux and Labruto 2012). The main purpose was "to act and find ways to favour the development of European robotics" and this included taking care of "worries about the consequences of introducing robots into society" (ibid., 5). Framing ethical, legal, and societal issues as "obstacles and barriers" to be overcome, preferably *before* they arise and settle in society, a main task of the paper was to mobilise legal and ethical expertise in order to deal with problems of responsibility and liability. If new markets and value chains are to be created around learning machines that act autonomously in people's living and working environments, legal certainty about possible unintended consequences has to be established *first*.

It was to this end that the Green Paper, in a speculative vein, proposed the introduction of *Electronic Personhood* as a new kind of hybrid agency, granting to machines a limited legal status. The robot as a legal entity would be inscribed in a public registry and connected to a fund, paid into by various actors along the production and value chain, such as designers, manufacturers, vendors, professionals (e.g., care workers), owners and end-users (patients, consumers). If the machines are capable of learning, all these actors (and more) will take part in enabling and shaping their behaviours in different ways, and so be co-responsible for their actions and their consequences. This position was arrived at through consideration of different types of legal agency: from machines as physical instalments (i.e. a robot used for manufacture, locked up in the safe environment of the factory), to animals or children capable of moving around freely. In the case of animals and children, the responsible parties would be the guardians, i.e.

the *owners* or the *parents*. Note however that transferring such a scheme by legal analogy to the case of autonomous robots, would mark a transition and distribution of responsibility to users and others, at a stage where the robotic society is still primarily a project belonging to industry and roboticists.[11] In terms of distribution of risks and societal equity, the proposal was shaped by industry interest, and predicated on a machine-centric vision.

Whereas the construct was intended in a limited sense, if implemented, it would break down previous boundaries between machines as physical objects, and humans as (legal) subjectivities. This distinction, between humans and machines, was explicitly commented on and targeted by the Green Paper, as the main legal and constitutional *obstacle* or *barrier* for the entry of robots into society. The paper noted how a "strict differentiation between man and machine ('man-machine – dualism') is no longer acceptable", and that also in an ethico-legal sense "man and machine should be considered simultaneously and their actions should be seen as cooperation" (p. 58). This directly consolidates the cybernetic or hybrid ontology as part of the knowledgebase for subsequent regulatory initiatives.

The proposition, however, breaks with the human-centrism of European and western constitutions, and triggered critical responses. First, a "White Paper" (Bertolini and Palmerini 2014) centred around academic lawyers, emerged as a response to the "Green Paper", and it took a specifically human-centric and fundamental rights-based position. Rather than seeing human-centric constitutions as an obstacle, the lawyers took the position that fundamental rights would have to serve as the baseline for assessing the impacts of robots on society. Eventually, when a proposal was put forward by the European Parliament, it took a mixed approach: fundamental rights were underlined as basic. Yet, the idea of personhood for machines was retained: the EP proposed an insurance scheme, not dissimilar to the Green Paper, and proposed to the Commission the creation of

> a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently.
>
> (EP 2017, 18)

This proposal also met with fierce opposition: A group of more than 150 researchers and lawyers who referred to themselves as "Artificial Intelligence and Robotics experts" signed and submitted an Open Letter to Parliament where they strongly condemned the proposal for legal personhood (Nevejans et al. 2018). The Open Letter claimed that the European Parliaments proposal was based in speculation and science fiction, and furthermore that it would introduce machines to the universe of human rights (see also van

Dijk 2020). It was thus directly opposed to basic human rights, such as dignity, integrity, and citizenship. The proposal was opposed by the European Economic and Social Committee (EESC 2017), which also took a human-centric approach. When the European Commission finally issued an overall strategy for Artificial Intelligence (AI) and robotics (EC 2018), the idea of electronic personhood was nowhere to be seen.[12] In either case, the question denoted a major incursion of machine-centrism into legislative discourse and debate. And, due to the underlying machine-centrism of the technologies, we are quite certain that the problem has not been done away with. It may, for instance, re-emerge at national levels, if national governments would under-cut other countries' governments,[13] aiming to attract innovation, investment and market-makers.

Legitimation: from boundary work to boundary fusion

We started out with the claim that digital technologies and innovations feed into post-truth conditions by blurring *and* reconstituting basic categories (fact–value, human–machine, etc.). This propensity of the digital also feeds into regulatory efforts to stem and steer the technologies and their impacts, possibly triggering an overflowing of the boundaries of western constitutions, and resulting in decreased capacities for working things out. We singled out a specific discourse, one that is predicated on intensified networking across institutional domains through digital means. We argued that this discourse is part of what we have called the Median Estate, as it is targeted towards the making of a world predicated on mediations between previously autonomous institutions: technology, law, regulation, politics, markets, and publics, and with major implications for living, working, and professional environments. This thesis was then explored in two cases: privacy engineering and personhood for machines. We described in rudimentary ways the kinds of hyper-truth constituted by digital technologies, and to which the regulatory efforts have to bend to have an impact. Both were concerned with bridging the digital and the physical, in the case of privacy this referred to the Internet of Things, and in the case of electronic personhood to robotic applications. Compared with classical distinctions in western societies, and to classical STS analyses thereof, we detected significant shifts or intensifications. What are these shifts, what do they signify in terms of broader social ordering and legitimation, and what new starting points seem warranted?

First of all, the above does not mean that suddenly law has become engineering, and human has become machine; this would be too crude an interpretation of "blurring and reconstitution of boundaries". What is entailed is a reconfiguration of the general role of law and science in the creation and upholding of a certain social metaphysics, traditionally described through concepts such as work of purification and boundary work. These concepts would refer to separate domains of reality (nature/culture, science/politics, humans/machines), traditionally enacted within discursively and institutionally distinct domains. Boundary work and coordination work (Bijker et al. 2009) would be concerned with working out the mutual relations and alignments of these domains.

Focusing on the modes of legitimation specific to such new regulatory regimes, then, one may also ask whether a concept of boundary work should be expanded into one of *boundary fusion*. As mentioned in the introduction, the main source of reference for this mode of legitimation emanates from cybernetics and its off-springs (bioinformatics, data science, machine learning, robotics, etc.). Strongly present in the public legitimation of such disciplines is the rhetoric strategy articulated by Bowker (1993) as *legitimacy exchange*: legitimacy and authority are built by merging together powerful sources of authority, such as law and engineering, as opposed to their mutual separation and discursive purification (cf. Rommetveit and Wynne 2017). When expanded to regulatory institutions, boundary fusion refers to a generalised space of possibility, in which citizens law and engineering are situated at *the same plane* of representation and intervention (cf. Pellizzoni 2015).

Concomitant with this, we also see that the *sites for articulation of rights have shifted*: from law and classical regulatory agencies (bureaucracies) and into new places, such as technological artefacts and infrastructures, innovation networks, standardisation bodies and organisations. This was implied by concepts such as “co-regulation”, (distinct from a concept of self-regulation), stakeholder capitalism and “agile governance” (WEF 2018). And, whereas some of these sites may reside in national regulatory agencies, some of which may also become strengthened (i.e. data protection authorities with the GDPR), the developments are increasingly also global and cross-European.[14] This dynamic can be illustrated by shifts of emphasis in salient research programmes in the EU: it goes from embedding science *in* society to embedding society *in* science, where relevant disciplines (ethics, law, social, and humanistic sciences) are brought *into* innovation networks rather than serving the function of embedding innovation networks in broader society.[15]

Within this reconfigured space of possibility, the *meaning of a right* also changes, since it explicitly takes on a more hybrid character: Rights themselves become more material, insofar as they become built and hardcoded into emerging infrastructures. They *also* become more virtual, since these infrastructures are strongly inscribed into the imagined-possible and the speculative drive of Information and Communication Technologies (ICTs), innovation and market-making. Hence, rights also take on much more performative and promissory characteristics. The promise is of more ethics and better rights protection, as co-emergent with, and enabling of a digital market. This kind of promise, embedded in institutions and protected by law is a classical task for legal regulation of stabilising expectations under conditions of uncertainty (cf. Luhmann 1983), and in this sense not novel.

Yet, Luhmann's description presupposed a certain kind of stable environment and embedding within the institution of law. In the Median Estate, however, stabilisation is no longer dependent on law alone, but on its interactions with a number of other actors and institutions, and strongly shaped by technologies that keep changing and upgrading.

Reflecting back on our two cases, then, privacy by design and electronic personhood emerge not so much as efforts to identify and separate a vulnerable subject worthy of protection, as a kind of connecting principle: a boundary-fusion-object for the making of infrastructures for the digital economy (captured by our term privacy-by-network). This was clearly demonstrated in the title of a recent report on ethics for AI set up by the European Commission in the extension of developments described in this article. The name of the report is *Trustworthy AI*, and it is explicitly inscribed in a universe of providing trust and predictability, argued to boost the competitiveness of the European Digital Market *and* protecting fundamental rights (AI HLEG 2019).[16] The shifting modes of legitimation, from boundary work to boundary fusion, thus indicates what, in this volume is referred to as a need for new starting points.

This means that ethics, law, and regulation are not merely embedded within a technological universe, but also within an ecology of fierce (global) economic competition, with the future as investment object. Focusing on the dimension of time, a fundamental motive is that of rendering "the future" an object of intervention through engineering. As described by Nordmann (2010, 5) the future is imagined as "*an object of technical design, the realisation of technical possibility*". Through intensification and proliferation into new areas, this logic now includes ethics and legal regulation within the fold of futures to be designed, engineered, and promised. A main logic here is that of pre-emption, since rendering futures objects of design also simultaneously forecloses other futures, and possible contestations of innovation pathways (Pellizzoni, this volume). We have previously described this as an obstacle model of public issues (Rommetveit and Wynne 2017), pointing to the relational and social dynamics involved.[17]

Conclusions: new starting points?

We first claimed that the post-truth discourse and certain (academic, media, political) responses to it have been too focused on binaries that do not do justice to the underlying dynamics ("before there was truth, now there is not"). We also pointed to omissions of hyper-truth, pertaining to how certain technoscientific imaginaries and agendas are posited as so evidently true that they cannot be questioned. We then claimed that post-truth discourse, since it is about the public uses of science, can also be read as expression of deeper shifts in our societies, and in the political economy of knowledge. These shifts have been identified, in the STS debate and elsewhere, as being concerned with the blurring of distinctions constitutive of modernity, i.e. those between fact and value, science and politics, and humans and machines. But they have not been adequately investigated and analysed, since the debate was not really taken seriously, but rather as something to be avoided (an exception to this is the provocative argument of Steve Fuller). Through our cases of privacy by design and electronic personhood we demonstrated how science and politics, humans and machines, facts and values, become muddled up and actively reconstituted in actual practice, through processes and modes of legitimation referred to as boundary fusion. Such blurring and reconstitution of boundaries has been a main theme in (critical) STS scholarship, whose main innovation was to position itself "in between" science and society, nature and culture, demonstrating their mutual dependence (co-production). This was relied upon as a critical corrective to idealised and purified institutionalisation and practice. Yet, the developments that we describe also demonstrate that main powerful actors are positioning themselves in similar ways. This means that the old recipes for critique are not as strong as they used to be. We argued that this state of affairs underlies and informs the post-truth debate and may well explain the reluctance of some main participants in the debate towards spelling out its full implications (this is why Fuller's critique is misdirected, but also on to something). This indicates to us how the post-truth debate, and some of the real-world phenomena with which it is associated, demonstrate a need for new starting points, taking into account such shifting dynamics of legitimation and ordering.

Notes

[1] As philosophers, the authors are intrinsically sceptical of 'truth', not to say it's adverse, 'post-truth'. We use 'truth' interchangeably with 'publically validated knowledge', underlining that such validation takes place in different ways in different knowledge practices and institutions.

[2] "As narratives get fragmented, allowing competing truths to proliferate, there's also a concurrent effort to deploy bots, ledgers, and algorithms to produce a singular, objective, and eternal truth" (Morozov 2019).

[3] Heidegger had a different notion of truth (Aletheia) as revealing or bringing forth of what is concealed. Whereas modern technology is also a mode of revealing by enframing the world around us in a certain way (for it to work well or efficiently), it conceals other ways in which the world can be revealed, but it also conceals this act of revealing (or truth) itself, in projecting the frame as the real towards which we become predisposed (Heidegger 1978).

[4] Among the main technological application domains Schwab (2016) mentions: implantable technologies, our digital presence, vision as the new interface, wearable internet, ubiquitous computing, a supercomputer in your pocket, storage for all, the internet of and for things, the connected home, smart cities, big data for decisions, driverless cars, artificial intelligence, and decision-making, AI and white collar jobs, robotics and services, bitcoin and the blockchain, the sharing economy, governments and the blockchain, 3D printing and manufacturing, 3D printing and human health, 3D printing and consumer products, designer beings, neurotechnologies.

[5] "The essence of modern technology starts man upon the way of that revealing through which the actual everywhere becomes standing-reserve. "To start upon a way" means "to send" in our everyday language. We shall call the sending that gathers (versammelnde Schicken), that first starts man upon a way of revealing, destining (Geschick). It is from this destining that the essence of all history (Geschichte) is determined" (Heidegger 1978, 329).

[6] According to its authors, this term can be seen as the extension of the Third Industrial Revolution, which was brought by digital networks, the Internet and social media, into physical reality itself. It was pre-figured by RFID chips in the 1990s, and continued in Smart technologies, the Internet of Things, and now, Industry 4.0 (Schwab 2016).

[7] Castells networked society argument in fact joins new networked modes of organisation (in economic and sociological theories) with networked information technologies that intensify this development (Castells 2010).

[8] Through Privacy Enhancing Technologies, PETs.

[9] See https://eu-smartcities.eu/initiatives/2/description

[10] Recall the ISO definition of privacy engineering as *integration of privacy concerns into engineering practices for systems and software engineering life cycle processes*.

[11] A French law professor, Nathalie Lavejans, argued that "By adopting legal personhood, we are going to erase the responsibility of manufacturers" (Delcker 2018).

[12] The reason for its disappearance in the EC proposal is unknown to the authors: it may have come as result of the human-centric criticism levelled at it. It may also have come, as argued by Burri (2018) from the realisation that the capacity to create legal personhood actually does not reside with the European Parliament, but at national and member state level.

[13] Thus, Saudi Arabia (not famous for its record on human rights) bestowed citizenship on the humanoid Sophia.

[14] This happens at the same time as significant societal forces seek to 'take back control', and to build autonomy and sovereignty at national, local, or regional institutional levels. Innovation and techno-regulation, therefore, enter into increasingly conflictive political spaces, and can be read as a pre-emptive strategy for dealing with conflict and antagonism.

[15] Other examples of this dynamic are given by the inclusion of RRI and "Integrated Social and Humanistic Science" as cross-cutting in EU research programs, in ethics-by-design and legal co-regulation.

[16] This was illustrated in a media report, where the following quotes occurred: "Ethics and competitiveness are intertwined, they're dovetailed," (Pekka Ala-Pietilä chairs of the high-level expert group on AI). Similarly, digital commissioner Mariya Gabriel was quoted as saying: "I am personally convinced that ethical guidelines will be enablers of innovation for artificial intelligence" (Politico newspaper 17.03.2019).

[17] To exemplify, one representative of the robotics industry told us how "The obstacles for robots have to be investigated… ELS (Ethical Legal and Social) issues need to be investigated that hinder solutions. European robotics industry has to be made world leader" (Rommetveit et al. 2020). And, a digital rights activist invoked the same logic, from the point of view of publics trying to engage with privacy infringements, but seeing the path as forestalled by the use of privacy risk assessments (PIAs, which frequently feed into design procedures): "'We do a PIA so it is okay'. It is used as a palliative to make it impossible for people opposing, to raise issues that certain developments infringe fundamental rights" (van Dijk et al. 2018, 18).

References

AI HLEG (2019). *Ethics Guidelines for Trustworthy AI. High-Level Expert Group on AI*. Brussels: European Commission.

Appadurai, A. (1996). *Modernity at Large: Cultural Dimensions of Globalization*. Minneapolis: University of Minnesota Press.

Bertolini, A., and Palmerini, E. (2014). Regulating Robotics: A Challenge for Europe. In: Legal Affairs Committee (Ed.) *Upcoming Issues of EU Law*. Brussels: European Parliament.

Bijker, W., Bal, R., and Hendriks, R. (2009). *The Paradox of Scientific Authority. The Role of Scientific Advice in Democracies*. Cambridge: MIT University Press.

Bijker, W., Hughes, T.P., and Pinch, T. (1987). *The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology*. Cambridge: MIT University Press.

Bowker, G. and Star, S. L. (1999). *Sorting Things Out. Classification and Its Consequences*. Cambridge: The MIT Press.

Bowker, G.C. (1993). How to be universal: Some cybernetic strategies, 1943–1970. *Social Studies of Science*, 23, pp. 107–127.

Brown, N. and Michael, M. (2003). A sociology of expectations: Retrospecting prospects and prospecting retrospects. *Technology Analysis & Strategic Management*, 15(1), pp. 3–18.

Burri, T. (2018). The EU is right to refuse legal personality for Artificial Intelligence. Euractiv, 31.05.2018.

Castells, M. (2010). *The Rise of The Network Society*. Vol. 1 of The Information Age, Economy, Society and Culture (2nd edition with a new preface). Oxford: Wiley-Blackwell.

Cavoukian, A. (2009). Privacy by design: The 7 foundational principles. Information and Privacy Commissioner of Ontario.

Cohen, J. (2017). The biopolitical public domain: The legal construction of the surveillance economy. *Philosophy & Technology*, 31(2), pp. 213–233.

Collins, H., Evans, R., and Weinel, M. (2017). STS as science or politics? *Social Studies of Science*, 47(4), pp. 580–586.

Delcker, J. (2018). Europe divided over “robot personhood”. Politico 04.11.2018.

Dennedy, M.F., Fox, J., and Finneran, T.R. (2014). A privacy engineering lifecycle methodology: The privacy engineer’s manifesto – Getting from policy to code to QA to value. ApressOpen.

EESC (2017). Opinion of the European economic and social committee on “Artificial Intelligence—The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society” (2017/C 288/01).

[EP] European Parliament (2017). *Report with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))*. Brussels: European Parliament. http://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html. Accessed 25 May 2019.

European Commission (2018). Communication from the commission to the European parliament, the European council, the council, the European economic and social committee and the committee of the regions: Artificial Intelligence for Europe. SWD (2018) 137 final. Brussels.

Fortun M. (2008). *Promising Genomics: Iceland and deCODE Genetics in a World of Speculation*. Berkeley, Los Angeles and London: University of California Press.

Fuchs, C. (2018). Industry 4.0: The digital German ideology. *TripleC*, 16(1), pp. 280–289.

Fuller, S. (2018). *Post-truth. Knowledge as a Power Game*. London and New York: Anthem Press.

Gürses, S. and Del Álamo. J.M. (2016). Privacy engineering: Shaping an emerging field of research and practice. *IEEE Security & Privacy*, 14(2), pp. 40–46.

Heidegger, M. (1978). *The Question Concerning Technology*. Basic Writings. Abingdon and New York: Routledge.

Hes, R. and Borking, J. (2000). *Privacy-enhancing Technologies: The Path to Anonymity* (Rev. ed.). The Hague: Registratiekamer.

Jasanoff, S. (2004). Ordering knowledge, ordering society. In: Jasanoff, S. (Ed.) *States of Knowledge: The Co-production of Science and Social Order*. New York: Routledge, pp. 25–98.

Jasanoff, S. (2011). *Reframing Rights. Bioconstitutionalism in the Genetic Age*. Cambridge: MIT Press.

Jasanoff, S. and Simmet, H. (2017). No funeral bells: Public reason in a "post-truth" age. *Social Studies of Science*, 47(5), pp. 751–770.

Kline, R.R. (2015). *The Cybernetics Moment. Or Why We Call Our Age the Information Age*. Baltimore, MD: John Hopkins Press.

Latour, B. (1993). *We Have Never Been Modern*. New York: Harvester Wheatsheaf Publisher.

Latour, B. (2013). *An Inquiry into the Modes of Existence*: *An Anthropology of the Moderns*. Cambridge, MA: Harvard University Press.

Leroux, C. and Labruto, R. (2012). Ethical legal and societal issues in robotics D3.2.1. euRobotics.

Lessig, L. (1999/2006). *Code and Other Laws of Cyberspace*. New York: Basic Books.

Luhmann, N. (1983). *Legitimation durch Verfahren*. Frankfurt am Main: Suhrkamp Verlag.

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), pp. 175–183.

Morozov, E. (2019). Can the US government stem the tide of "fake news" in a post-modern world? *The Guardian*, 31 Oct 2019.

Nevejans, N., et al. (2018). Open Letter to the European Commission. Artificial Intelligence and Robotics. Accessed 16.09.2019 at: http://www.robotics-openletter.eu

Nissenbaum, H. (2004). Symposium, privacy as contextual integrity. *Washington Law Review*, 79(1), pp. 119–158.

Nordmann, A. (2004). Converging technologies – Shaping the future of European societies. Report EUR 21357, European Commission Research.

Nordmann, A. (2010). A forensics of wishing: Technology assessment in the age of technoscience. *Poiesis & Praxis: International Journal of Ethics of Science and Technology Assessment*, 7, pp. 5–15.

Nordmann, A. (2020). The advancement of ignorance. In: Sascha Dickel, S., Schneider, C., Maasen, S., et al. (Eds.), *Sociology of the Sciences Yearbook*. New York: Springer, pp. 21–33.

Pellizzoni, L. (2015). *Ontological Politics in a Disposable World. The New Mastery of Nature*. New York: Routledge.

Pellizzoni, L. (2017). Intensifying embroilments: Technosciences, imaginaries and publics'. *Public Understanding of Science*, 26(2), pp. 212–219.

Purtova, N. (2018). The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, 10(1), pp. 40–81.

Reidenberg, J.R. (1998). Lex informatica: The formulation of information policy rules through technology. *Texas Law Review*, 76(3), pp. 553–584.

Rommetveit, K. (2011). Tackling epistemological naivety: Large-scale information systems and the complexities of the common good. *Cambridge Quarterly of Healthcare Ethics*, 20, pp. 1–12.

Rommetveit, K., Tanas, A., and van Dijk, N. (2018). Data protection by design: Promises and perils in crossing the Rubicon between law and engineering. *Springer Series IFIP Advances in Information and Communication Technology*, pp. 25–37.

Rommetveit, K., van Dijk, N., and Gunnarsdottír, K. (2020). Make way for the robots! Human- and machine-centricity in constituting a European public-private partnership. *Minerva: A Review of Science, Learning and Policy*, 58(1), pp. 47–69.

Rommetveit, K. and Wynne, B. (2017). Technoscience, imagined publics and public imaginations. *Public Understanding of Science*, 26(2), pp. 133–147.

Schwab, K. (2016). *The Fourth Industrial Revolution*. Geneva: World Economic Forum.

Shapin, S. and Schaffer, S. (1985). *Leviathan and the Air Pump*. Princeton, NJ: Princeton University Press.

Silvast, A., Williams, R.A., Hyysalo, S., Rommetveit, K., and Raab, C. (2018). Who 'uses' smart grids? The evolving nature of user representations in layered infrastructures. *Sustainability*, 10(10), p. 3738.

Sismondo, S. (2017). Casting a wider net: A reply to Collins, Evans and Weinel. *Social Studies of Science*, 47(4), pp. 587–592.

Solove, D.J. (2008). *Understanding Privacy*. Cambridge, MA: Harvard University Press.

Solum, L.B. (1992). Legal personhood for Artificial Intelligences. *North Carolina Law Review*, 70, pp. 1231–1287.

Toulmin, S. (1990). *Cosmopolis. The Hidden Agenda of Modernity*. Chicago, IL: University of Chicago Press.

van Dijk, N. (2020). In the hall of masks. Contrasting modes of personification. In Hildebrandt, M. and O'Hara, K. (Eds.), *Life and the Law in the Era of Data-Driven Agency*. Cheltenham: Edward Elgar, pp. 230–251.

van Dijk, N., Gellert, R., and Rommetveit, K. (2016). A risk to a right? Beyond data protection risk assessments. *Computer Law & Security Review*, 32(2), pp. 286–306.

van Dijk, N., Tanas, A., Rommetveit, K., and Raab, C. (2018). Right engineering? The redesign of privacy and personal data protection. *International Review of Law, Computers & Technology*, 32(2–3), pp. 230–256.

Vanderelst, D. and Winfield, A. (2018). An architecture for ethical robots inspired by the simulation theory of cognition. *Cognitive Systems Research* (48), pp. 56–66.

Wallach, W. and Allen, C. (2009). *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press.

WEF (2018). Agile governance. Reimagining policy-making in the fourth industrial revolution. World Economic Forum, White Paper. Accessed from: http://www3.weforum.org/docs/WEF_Agile_Governance_Reimagining_Policy-making_4IR_report.pdf

Weingart, P. (1999). Scientific expertise and political accountability: Paradoxes of science in politics. *Science and Public Policy*, 26, pp. 151–161.

Welsh, I. and Wynne, B. (2013). Science, scientism and imaginaries of publics in the UK: Passive objects, incipient threats. *Science as Culture*, 22(4), pp. 539–565.

Wynne, B. (2014). Further disorientation in the hall of mirrors. *Public Understanding of Science*, 23(1), pp. 60–70.

Zuboff, S. (2018). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York: PublicAffairs.