

Robots and Rights: Reviewing Recent Positions in Legal Philosophy and Ethics

Wolfgang M. Schröder

Contents

Introduction	192
Definitions and Brief Exposition of the Topic	193
The Classical Ontological Stance and the Recent “Relational Turn” in Animal, Robot and Machine Ethics: Mark Coeckelbergh’s Analysis	193
The Juridical Perspective: The “Accountability Gap” Implying a “Responsibility Gap”	194
Recent Juridical Tendencies Towards Advocating Legal Personality for Robots	195
David J. Gunkel’s “Robot Rights” (2018)	196
The EPSRC Paper on “Principles of Robotics”	197
Robots in the Japanese <i>koseki</i> System: Colin P.A. Jones’s Family Law Approach to Robotic Identity and Soft-Law Based Robot Regulation	197
Joanna J. Bryson’s Outline of Roboethics	199
Conclusion	200
References	201

Abstract

Controversies about the moral and legal status of robots and of humanoid robots in particular are among the top debates in recent practical philosophy and legal theory. As robots become increasingly sophisticated, and engineers make them combine properties of tools with seemingly psychological capacities that were thought to be reserved for humans, such considerations become pressing. While some are inclined to view humanoid robots as more than just tools, discussions are dominated by a clear divide: What some find appealing, others deem appalling, i.e.

“robot rights” and “legal personhood” for AI systems. Obviously, we need to organize human–robot interactions according to ethical and juridical principles that optimize benefit and minimize mutual harm. Avoiding disrespectful treatment of robots can help to preserve a normative basic ethical continuum in the behaviour of humans. This insight can contribute to inspire an “overlapping consensus” as conceptualized by John Rawls in further discussions on responsibly coordinating human/robot interactions.

Keywords

Philosophy · Personhood · Rights · Robots · Ethics

W. M. Schröder (✉)
 Professur für Philosophie am Institut für Systematische Theologie,
 Katholisch-Theologische Fakultät, University of Würzburg, Würzburg,
 Germany
 e-mail: wolfgang.schroeder@uni-wuerzburg.de

Introduction

Robots¹—as it seems—are here to stay.² But with which status, and under what normative conditions? Controversies about the moral and legal status of robots in general, and of humanoid (anthropomorphic) robots in particular, are among the top debates in recent practical philosophy and legal theory (Danaher 2017a; Gunkel 2018; Bryson 2019; Dignum 2019; Basl 2019; Nyholm 2020; Wong and Simon 2020; Andreotta 2020). Quite obviously, the state of the art in robotics and the rapid further development of Artificial Intelligence (AI) raise moral and legal issues that significantly exceed the horizon of classic normative theory building (Behdadi and Munthe 2020). Yet what exactly is the problem?

As robots become increasingly sophisticated, and engineers try harder to make them quasi “sentient” and “conscious” (Ekbja 2008; Torrance 2012), we are faced with AI-embedding systems that are ambivalent by design. They combine properties of tools with seemingly “psychological capacities that we had previously thought were reserved for complex biological organisms such as humans” (Prescott 2017: 142).³ Hence there is a growing incentive to consider the ontological status of the robots as “liminal”: robots seem to be “neither living nor simply mechanical” (Sandini and Sciutti 2018: 7:1). Therefore, it is not surprising that humans show inclinations to treat humanoid robots “as more than just tools”, regardless of the “extent to which their machine nature is transparent” (Sandini and Sciutti 2018; see also Rosenthal-von der Pütten et al. 2018). After all, human brains have evolved mainly to understand (and interact with) humans, so

they are likely to be easily “tricked” into interpreting human-like robot behaviour “as if it were generated by a human” (Sandini and Sciutti 2018: 7:1). As a matter of consequence, it is time to come to terms with the question of how “intelligent machines” like robots (especially the humanoid ones) should be categorized and treated in our societies (Walch 2019).

Discussions on that issue have so far been dominated by a clear divide: What some find appealing, others deem appalling: “robot rights” and “legal personhood” (“e-personality”) for AI systems.⁴ While Luciano Floridi, Director of the Oxford Internet Institute, declassifies thinking and talking about the “counterintuitive attribution of rights” to robots as “distracting and irresponsible” (Floridi 2017: 4; see also Coyne 1999; Winfield 2007), *Forbes Magazine* ranks this topic as one of the most important AI ethics concerns in the early 2020s (Walch 2019). However contradicting opinions may be,⁵ the task at stake is undisputed: We need to organize Human–Robot Interaction (HRI) according to ethical and juridical principles that optimize benefit and

¹In absence of a generally accepted definition of what a “robot” is (standards like ISO 8373:2012 only relate to industrial robots), we use the following working definition: “Robots” are computer-controlled machines resembling living creatures by moving independently and performing complex actions (cf. <https://www.merriam-webster.com/dictionary/robot>; Dignum 2019: 31). As a short introduction into robotics, see Winfield (2012).

²According to the latest World Robotic Report (see <https://ifr.org/news/summary-outlook-on-world-robotics-report-2019-by-ifr>), the number of robot installations has never increased so strongly than from 2010 till present. We are in the midst of a “robot invasion”—with no end in sight (cf. Gunkel 2019). See also Bennett & Aly (2020).

³Robotics has progressed from single arm manipulators with motion schemes of limited degrees of freedom to more complex anthropomorphic forms with human motion patterns. Whereas, for security reasons, industrial robots are normally contained in barricaded work cells and are automatically deactivated if approached by a human, humanoid robots seem to blur the boundary between humans and machines. On the one hand, they can be engineered and used as the perfect universal tool, as extensions of industrial robots able to perform menial tasks in the workplace or hazardous work and exploration. On the other, their “human form” suggests that they can have “personhood” and are able to interact with humans. They can utilise devices originally constructed for humans and “inherently suited” to human environments. Thus, humanoid robots seem to be *so much more than just machines*. For future scenarios in robotics, see also Nourbakhsh (2013).

⁴At first glance, it seems that (at least in the European and Western tradition) the sphere of law and rights presents a specifically and exclusively anthropic one (cf. Siewert et al. 2006). In Greek antiquity, Zeus gave humans—and only humans, as opposed to animals—the *nómos* not to eat each other and to use law (*díkē*) instead of force (*bíē*) (Hesiod, Op. 275–279). Coming from the father of the gods, this *nómos* was morally binding for the mortals and corresponds to the “*nomoi* for all of the Hellenes”, frequently mentioned in the fifth cent. BC, or to the “unwritten” *nómoi* that regulated moral conduct, e.g. to honor the gods, parents and strangers, to bury the dead and to protect those who suffer injustice [4]. As early as in Homer (who does not use the word *nómos*) the gods controlled whether there was *eunomia* (good order) or *hybris* (disrespect, arrogance) among mortals (Homer, Od. 17, 487). *Nómos* refers to the proper conduct not only towards fellow human beings but also towards the gods as well, e.g. the obligation to sacrifice (Hesiod, Theog. 417; Hesiod, Fr. 322 M.-W.); thus *nómos* refers to the norms for moral and religious conduct in Greek society. But as Gunther Teubner has pointed out (Teubner 2006), the world of law in medieval and Renaissance Europe and also in other cultures was populated with non-human beings, with ancestors’ spirits, gods, trees, holy shrines, intestines, birds’ flight, to all those visible and non-visible phenomena to which communication could be presupposed and which included the potential to deceive, to lie, to trickster, and to express something by silence. Today, under the influence of rationalizing science, the number of actors in the legal world has been drastically diminished. After the scientific revolution, after philosophical enlightenment, after methodological individualism dominating the social sciences, after psychological and sociological analysis of purposive action, the only remaining plausible actor is the human individual. The rest is superstition. To be sure, the law still applies the construct of the juridical person to organizations and states. But increasingly, especially under the influence of legal economics, this practice has been devalued as merely an “analogy”, a “linguistic abbreviation” of a complex legal relationship between individuals, as a “trap” of corporatist ideologies, at best as a “legal fiction”, a superfluous myth, that should be replaced by the nexus model which conceives the organization as a multitude of contracts between individuals.

⁵See as overview Wallach (2007), Wallach (2010), Wallach and Allen (2010), Robertson et al. (2019) and Dubber et al. (2020).

minimize mutual harm (cf. Floridi 2013; Lin et al. 2014; Nemitz 2018; Scharre 2018; Bremner et al. 2019; Brockman 2019; Loh 2019; Schröder et al. 2021).

My paper takes up this topic from a legal ethics perspective and proceeds in three main steps. It begins with definitions of central terms and an exposition of central aspects under which robots (as AI-embedding machines and AI-controlled agents) can become a topic of moral and juridical discourse. Then follows a brief review of some of the most prominent theses from recent literature on the moral and juridical status of robots. In conclusion, a balanced intermediate result and a modest new proposal are presented and substantiated in recognition of the previous discussion. We will start with definitions.

Definitions and Brief Exposition of the Topic

Rights talk is a prominent issue in both moral and legal theory and praxis. Thus, we need to briefly clarify the different basic meanings that the term “rights” has in the respective contexts.

Ethical discourse on moral norms and rights concerns an enormously broad field of problems and investigation. Yet basically it is an effort in theory building on descriptive or normative abbreviations of commonly acceptable social conduct (Lapinski and Rimal 2005).⁶ Even from the perspective of some theories of natural law, *moral* norms and rights are *cultural products* (like values, customs, and traditions); they represent culturally shaped ideas and principles of what a reasonable grammar of practical freedom should look like (Schröder 2012; Bryson et al. 2017; Spiekermann 2019).

Legal rights and norms differ from purely moral ones by their specifically institutional character. Legal rights and norms “exist under the rules of *legal* systems or by virtue of decisions of suitably authoritative bodies within them” (Campbell 2001). Following the standard Hohfeldian account (Hohfeld 1913), rights as applied in juridical reasoning can be broken down into a set of four categories (“the Hohfeldian incidents” in Wenar 2015)⁷:

– Privileges

⁶First, this discourse makes the distinction between perceived and collective norms and between descriptive and injunctive norms. Second, the article addresses the role of important moderators in the relationship between descriptive norms and behaviours, including outcome expectations, group identity, and ego involvement. Third, it discusses the role of both interpersonal and mass communication in normative influences. Lastly, it outlines behavioral attributes that determine susceptibility to normative influences, including behavioral ambiguity and the public or private nature of the behavior. See also Bicchieri (2006) and Soh & Connolly (2020).

⁷Named after Wesley Hohfeld (1879–1918), the American legal theorist who discovered them. Each of these Hohfeldian incidents has a distinctive logical form, and the incidents fit together in characteristic ways to create complex “molecular” rights.

- Claims
- Powers
- Immunities

In the case of “privileges”, you have a liberty or privilege to do as you please within a certain zone of privacy. As to “claims”, they mean that others have a duty not to encroach upon you in that zone of privacy. “Powers”, in this context, mean that you have the ability to waive your claim-right not to be interfered with in that zone of privacy. And, finally, “immunities” provide for your being legally protected against others trying to waive your claim-right on your behalf.

Obviously, the four above-mentioned categories logically relate to one another: “Saying that you have a privilege to do X typically entails that you have a claim-right against others to stop them from interfering with that privilege” (Danaher 2017b: 1).

The Classical Ontological Stance and the Recent “Relational Turn” in Animal, Robot and Machine Ethics: Mark Coeckelbergh’s Analysis

From the point of view of robot and machine ethics, the question of robot rights refers to the kind of relations we have or will develop with robots as AI-embedding machines.⁸ According to Mark Coeckelbergh,⁹ we need to adapt our practice of moral status ascription to the fact that the number of candidates for moral patients and agents is growing (Coeckelbergh 2011, 2012a; see also Danaher 2019; Birch 1993; Haraway 2008; Hart et al. 2012; Latour 2015). Coeckelbergh criticizes that classificatory thinking in animal and machine ethics is usually one-sidedly property-based: entities are considered in isolation from other entities, thereby reducing ethics to a kind of mechanical thinking.¹⁰ For Coeckelbergh, this raises three major problems: (1) How can we know which property is sufficient, or at least decisive, for ascribing moral status? (2) How could we indubitably establish that an entity

⁸See as an overview Dubber et al. (2020) and the respective online supplement <https://c4ejournal.net/the-oxford-handbook-of-ethics-of-ai-online-companion/>. And Anderson and Anderson (2018).

⁹Mark Coeckelbergh is Professor of Philosophy of Media and Technology at the Department of Philosophy, University of Vienna, and member of the High Level Expert Group on Artificial Intelligence for the European Commissions. He is best known for his work in AI Ethics and *Ethics* of Robotics, yet he has also broadly published in the areas of Moral and Environmental Philosophy.

¹⁰“From a social-philosophical point of view, this approach is individualist, since moral status is ascribed to entities considered in isolation from other entities—including the observer. [...] The modern scientist, who forces nature to reveal herself, is now accompanied by the moral scientist, who forces the entity to reveal its true moral status” (Coeckelbergh 2013a, b: 17).

indeed has a particular property S? And finally (3) how could we define really sharp boundaries between different kinds of entities?¹¹

Coeckelbergh's scepticism about property-based ethical classification leads him to plead for an ecological anthropology as proposed by Tim Ingold (Ingold 2000). In Ingold's "Ecology of Materials", all entities are considered as nodes in a field of relationships, which can only be understood in relational, ecological, and developmental or growth terms (Ingold 2012). Following Ingold, Coeckelbergh interprets moral standing as an expression of active, developing relationships between entities. Instead of asking what property P counts for moral standing S, the new question is: "How should we relate to other beings as human beings who are *already part of the same world* as these non-human beings, who *experience* that world and those other beings and are *already engaged* in that world and stand already in relation to that world?" (Coeckelbergh 2013a, b). Thus, Coeckelbergh considers *relations* as basic conditions for moral standing: "The relational approach suggests that we should not assume that there is a kind of moral backpack attached to the entity in question; instead moral consideration is granted within a dynamic relation between humans and the entity under consideration" (Coeckelbergh 2010). On this account, relations are not to be seen as properties, but rather "as a priori given in which we are already engaged, making possible the ascription of moral status to entities" (Swart 2013).

We will return to reviewing this approach more closely when we come to David Gunkel's relation-based theory of "robot rights". But let us first turn to the juridico-legal discourse on the robots-and-rights topic, just to add this perspective to what we have seen in moral philosophy.

The Juridical Perspective: The "Accountability Gap" Implying a "Responsibility Gap"

Following Jack M. Balkin's¹² clear-cut analysis, there are **three key problems** that robotics and AI agents present for law:

- **Firstly**, there is the **problem of how to deal with the emergence of non-human agents in the social worlds of humans**. How should we "distribute rights and responsibilities among human beings when non-human agents create benefits like artistic works or cause harms like physical injuries"? The difficulty here arises from the "fact that

the behavior of robotic and AI systems is 'emergent'; their actions may not be predictable in advance or constrained by human expectations about proper behavior. Moreover, the programming and algorithms used by robots and AI entities may be the work of many hands, and may employ generative technologies that allow innovation at multiple layers. These features of robotics and AI enhance unpredictability and diffusion of causal responsibility for what robots and AI agents do" (Balkin 2015: 46).¹³

- **Secondly**, there is the **problem of the "substitution effect"**. What we already see now will become even clearer in the future: "People will substitute robots and AI agents for living things—and especially for humans. But they will do so only in certain ways and only for certain purposes. In other words, people tend to treat robots and AI agents as special-purpose animals or special-purpose human beings. This substitution is likely to be incomplete, contextual, unstable, and often opportunistic. People may treat the robot as a person (or animal) for some purposes and as an object for others" (Balkin 2015: 46).
- **Thirdly**, we are not dealing here with a static configuration of challenges to the law. Rather, **we are faced with a steadily evolving dynamic field of often disruptive factors**. As Balkin put it: "We should not think of essential characteristics of technology independent of how people use technology in their lives and in their social relations with others. [...] Innovation in technology is not just innovation of tools and techniques; it may also involve innovation of economic, social and legal relations. As we innovate socially and economically, what appears most salient and important about our technologies may also change" (Balkin 2015: 48f).

The most obvious legal problem with robots that can potentially harm people's physical integrity or property is an "**accountability gap**" implying a "**responsibility gap**". There are at least two important legal levels at which this issue creates problems: **criminal law** and **civil law** (Keßler 2019).¹⁴

In **criminal law** only natural persons—in the sense of the law, real, living people—can be held responsible for their actions. At civil law level, on the other hand, legal persons, such as companies, can be included. Robots fall after the current state in neither of the two fields.

¹¹For further discussion see Swart (2013).

¹²Jack M. Balkin is Knight Professor of Constitutional Law and the First Amendment at Yale Law School. He is the founder and director of Yale's Information Society Project, an interdisciplinary center that studies law and new information technologies.

¹³Balkin also reminds us on Lawrence Lessig's famous dictum that "Code is Law", meaning "that combinations of computer hardware and software, like other modalities of regulation, could constrain and direct human behavior. Robotics and AI present the converse problem. Instead of code as a law that regulates humans, robotics and AI feature emergent behavior that escapes human planning and expectations. Code is lawless" (Balkin 2015: 52).

¹⁴The following brief exposition relies on the kind advice of my Würzburg colleague Christian Haagen (Forschungsstelle Robotrecht, University of Würzburg).

Manufacturers of robots can be held responsible if they have been proven to use the wrong materials in construction or have made mistakes in programming: for example, if a self-propelled forklift in a high-level warehouse drops a pallet and is clearly related to faulty programming. For this damage then the manufacturer of the truck should be liable. However, proving such errors is likely to become even more difficult in the future.

The robots of the future will learn independently and will make decisions based on their past experiences. Programmers can and must provide a faultless framework for this. However, if the robots are wrong in their experience-based decisions, manufacturers cannot be held so easily accountable. In general, one could say: as long as the manufacturers have acted with all possible and reasonable care and have made no mistakes, they cannot be guilty according to our criminal understanding.

At the *civil law* level, there is still the possibility of liability, which does not depend on fault. The manufacturer could then be prosecuted if the damage is due to a fault on the robot due to the manufacturer. Otherwise, it will be difficult to hold someone liable.

If the robots violate their duties and cause damage, they would be responsible, for example, financially. In principle, robots act independently with their experience-based decisions. They just cannot be punished like humans. For example, imprisonment means nothing to robots. Robots do not have their own assets. Maybe one will fund up, with whose funds the penalties for the machines are paid. Or, ultimately, it is the manufacturers who are responsible for the mistakes of their robots. You need to watch the devices, detect and fix errors. In the worst case, they have to call back the robots.

Recent Juridical Tendencies Towards Advocating Legal Personality for Robots

Some legal scholars like Gunther Teubner¹⁵ argue in favour of granting rights and legal personality to AI systems, depending on the degree of independence that AI systems are bestowed with. Teubner holds that personification of non-humans is “best understood as a strategy of dealing with the uncertainty about the identity of the other, which moves the attribution scheme from causation to double contingency and opens the space for presupposing the others’ self-referentiality” (Teubner 2006: 497); hence Teubner does not recognize any “compelling reason to restrict the attribution of action exclusively to humans and to social systems, as

Luhmann argues. Personifying other non-humans is a social reality today and a political necessity for the future. The admission of actors does not take place ... into one and only one collective. Rather, the properties of new actors differ extremely according to the multiplicity of different sites of the political ecology” (Teubner 2006: 497; see also Calerco 2008; Campbell 2011). On Teubner’s account, granting legal personality to AI systems would fill the Accountability Gap, thereby maintaining the integrity of the legal system as a whole and advancing the practical interests of humans (Teubner 2006).

Going beyond Teubner’s view, US legal scholar Shawn J. Bayern¹⁶ has argued that, under US law of limited liability companies (LLCs), legal personality could be bestowed on any type of autonomous systems. By means of a special transactional technique, legal entities (mainly LLCs) in the US could be governed entirely by autonomous systems or other software, without any ongoing necessary legal oversight or other involvement by human owners or members (Bayern 2019). For the time being, however, it seems clear that any autonomous system would probably “lack the basic acumen necessary to take many business decisions” (Turner 2019: 177). Thus, it seems unclear if Bayern’s point of view would be shared by the courts (Turner 2019).

In a recent groundbreaking book called “Robot Rules” (Turner 2019), legal scholar Jacob Turner¹⁷ deals with the Bermuda triangle of AI-related legal problems: Who is responsible for harms as well as for benefits caused by AI? Should AI have rights? And last but not least: How should ethical and juridical rules for AI be set and implemented?

Rather than literally formulating “robot rules” See also Coeckelbergh (2014), Turner’s book sets out to “provide a blueprint for institutions and mechanisms capable of fulfilling this role” (Turner 2019: 372). On Turner’s account, there are four reasons for protecting the rights of others that could be applied to at least some types of AI and robots:

1. The ability to suffer
2. Compassion
3. The value of something or somebody to others
4. “Situations where humans and AI are combined” (Turner 2019: 145)

Regarding the “**Argument from the Ability to Suffer**”, it is clear that some sort or degree of “artificial consciousness”¹⁸ would be a necessary precondition for a claim to legal protection based on the ability to suffer. “Pain”, in this

¹⁵Gunther Teubner has been Professor of Private Law and [Legal Sociology](#) at the [University of Frankfurt](#) until 2007; in 2011 he has taken up an “ad personam” Jean Monnet Chair at the [International University College of Turin](#).

¹⁶Shawn J. Bayern is Larry and Joyce Beltz Professor of Torts at the Florida State University College of Law.

¹⁷Jacob Turner is a barrister at Fountain Court Chambers (UK) after having been judicial assistant to Lord Mance at the UK Supreme Court.

¹⁸Turner explains: “For an entity to be conscious, it must be capable of (1) sensing stimuli, (2) perceiving sensations and (3) having a sense of

context, could be understood as “just a signal which encourages an entity to avoid something undesirable”; thus defined, it would not be difficult in Turner’s eyes to “acknowledge that robots can experience it” (Turner 2019: 152). Turner’s conclusion here is that “if an AI system was to acquire this quality, *then* it should qualify for some moral rights” (Turner 2019: 146).

The “**Argument from Compassion**” works on the premise that we obviously tend to protect certain entities because (and as far as) we have “an emotional reaction to them being harmed” (Turner 2019: 155). In the case of robots, it may suffice that they look as if they were conscious and had the ability to suffer to trigger our psychological tendency to develop feelings and compassion for them.

The “**Argument from Value to Humanity**” draws on a different observation. It is based on the fact that a whole range of things can be protected by law not because these things have a particular definable use, but “rather for a panoply of cultural, aesthetic and historical reasons” (Turner 2019: 165). These reasons can be deemed to constitute an “inherent value” of the objects at stake. To exemplify this idea, Turner refers to art. 20a of the German Constitution (Grundgesetz) which says: “Mindful also of its responsibility towards future generations, the state shall protect the natural foundations of life and animals”.

Last but not least, there is an “**Argument from Posthumanism**” referring to hybrid organisms, Cyborgs and “electronic brains”. This argument draws on the fact that in technically enhanced human bodies augmented by AI technology and human minds are not always really “separate”. Rather humans and AI seem to combine to become a symbiotic entity—thus developing into something “greater than just the sum of their parts” (Turner 2019: 167). Accordingly, in these cases the strict distinction between what is human and what is artificial may become increasingly fluid or even obsolete (cf. Turner 2019: 169).

With humans augmented by AI, boundary issues can arise: “when, if ever, might a human lose their protected status? [...] What about if 20%, 50% or 80% of their mental functioning was the result of computer processing powers?” (Turner 2019: 168). Probably we could suppose a broad consensus here that augmentation or replacement of human organs and physical functions with artificial substitutes “does not render someone less deserving of rights” (Turner 2019: 168).

While rights in general are social constructions, legal personality is, more specifically, a juridical fiction, created in and through our legal systems. On these grounds, it is up to us to decide to what exactly it should apply and to define its precise content (cf. Turner 2019: 175).

self, namely a conception of its own existence in space and time” (Turner 2019: 147).

As Turner points out, legal personality—instead of being a single notion—is “a technical label for a bundle of rights and responsibilities” (Turner 2019: 175). It is a juridical “artifice” designed to make sure that “legal people need not possess all the same rights and obligations, even with the same system”. In the case that one country grants (or plans to grant) legal personality to AI systems, this could certainly have a domino effect on other nations (cf. Turner 2019: 180).

On the other hand, arguments are brought forward against granting legal personality to AI systems. One draws on the idea of “Android Fallacy”, meaning the mistaken conflation of the concept of personality tout court with “humanity” as such (cf. Turner 2019: 189).

Another point of departure for rejecting legal personality for AI systems is the fear that robots with e-personality could be (mis-)used and exploited as liability shields by human actors (or whole companies) for selfish motives.

Furthermore, one could argue that robots should not be given rights or even legal personality because they as themselves are unaccountable rights violators (Turner 2019: 193).

David J. Gunkel’s “Robot Rights” (2018)

Unsatisfied by traditional moral theorizing on human-machine relations (Gunkel 2007, 2012), philosopher David J. Gunkel¹⁹ has made a strong case for “robot rights” as different from human rights. The latest and systematically most accomplished versions of his approach are his books “Robot Rights” (Gunkel 2018) and “How to Survive a Robot Invasion: Rights, Responsibility, and AI” (Gunkel 2019). Gunkel’s approach is based on a critique of all traditional moral theorizing in which ontological reflection actually precedes the ethical one: First you ask (and try to clarify) what a certain entity “is”; then you can proceed to the ethical question of whether or not this entity can or should be attributed a certain moral value (Gunkel 2018: 159). Gunkel pleads for thinking otherwise. He wants to “deconstruct” the aforementioned “conceptual configuration” and remix the parameters involved.

Gunkel’s argument draws heavily on Emmanuel Levinas’ assertion that ethics precedes ontology, not the other way around. For Levinas, intersubjective responsibility originates in face-to-face encounters (Lévinas 1990).²⁰ On this account,

¹⁹David J. Gunkel is Presidential Teaching Professor of [Communication Studies](#) at [Northern Illinois University](#). Currently he is seen as the most prominent philosophical author on robot rights issues.

²⁰“L’éthique, déjà par elle-même, est une ‘optique’.” Cf. *ibid.* 215: “La relation avec l’autre en tant que visage guérit de l’allergie. Elle est désir, enseignement reçu et opposition pacifique du discours. [...] Voilà la situation que nous appelons accueil du visage. L’idée de l’infini se produit dans l’opposition du discours, dans la socialité. Le rapport avec le visage avec l’autre absolument autre que je ne saurais contenir, avec l’autre, dans ce sens, infini, est cependant mon Idée, un commerce. Mais

“intersubjective experience proves ‘ethical’ in the simple sense that an ‘I’ discovers its own particularity when it is singled out by the gaze of the other. This gaze is interrogative and imperative. It says ‘do not kill me’. It also implores the ‘I’, who eludes it only with difficulty, although this request may have actually no discursive content. This command and supplication occurs because human faces impact us as affective moments or, what Levinas calls ‘interruptions’. The face of the other is firstly expressiveness. It could be compared to a force” (Campbell 2001).

On Gunkel’s account, this means that “it is the axiological aspect, the *ought* or *should* dimension, that comes first, in terms of both temporal sequence and status, and the ontological aspect (the *is* or *can*) follows from this decision” (Gunkel 2018: 159).

If one follows the thesis of “the underivability of ethics from ‘ontology’” (Duncan 2006: 277), encountering others and otherness changes its meaning as regards the sequence of challenges associated with it. On Gunkel’s view, “we are initially confronted with a mess of anonymous others who intrude on us and to whom we are obliged to respond even before we know anything at all about them and their inner workings” (Gunkel 2018: 159f). On these grounds, Gunkel advocates proposals to grant robots “with a face” some basic moral rights to be respected by humans in their common social worlds (Gunkel 2018: 171–175).

Gunkel interprets his approach as “applied Levinasian philosophy” (Gunkel 2018: 170), knowing that Levinas never wrote about robots, technology or robotics. As it seems, Gunkel’s “applied Levinasian philosophy” is also inspired by Silvia Benso’s book on the “face of things” (Benso 2000). Accordingly, Gunkel admits some difficulties arising from the application of Levinas’s philosophy to the study of robot rights. He is aware of the fact that Levinas’ ethics exclusively concerns relationships between humans and between humans and god. Therefore, as Katarzyna Ginszt notes in her comment on Gunkel’s hermeneutics, applying Levinas’ thought to robots would “require reading Levinas beyond the anthropocentric restrictions of the ‘Other’ that is presupposed to be a human being” (Ginszt 2019: 30). Moreover, Gunkel would have to make sure that this kind of “broadening the boundaries” could not easily be misunderstood as relativistic but could be clearly recognized as “relational” in Coeckelbergh’s sense (Ginszt 2019: 30). Closely read, Gunkel’s formulations meet these requirements.

The EPSRC Paper on “Principles of Robotics”

Nevertheless, Gunkel’s plea for robot rights is generally perceived as bold proposal and has earned much criticism. In the eyes of many robo-ethicists Gunkel goes too far in his account of the ethical implications of having robots in our society (cf. Bryson 2019). Those who do not want to take the “relational turn” in roboethics insist that “Robots are simply not people. They are pieces of technology” (Boden et al. 2017: 126). From this point of view, it seems indeed counterintuitive to attribute rights to robots. Rather, it is stringent to emphasize that “the responsibility of making sure they behave well must always lie with human beings. Accordingly, rules for real robots in real life must be transformed into rules advising those who design, sell and use robots about how they should act” (Boden et al. 2017: 125; see also Pasquale 2018).

How this responsibility might be practically addressed is outlined in a paper by the British Engineering and Physical Research Council (EPSRC) called “Principles of Robotics. Regulating robots in the real world” (Boden 2011): “For example, one way forward would be a licence and register (just as there is for cars) that records who is responsible for any robot. This might apply to all or only operate where that ownership is not obvious (e.g. for a robot that might roam outside a house or operate in a public institution such as a school or hospital). Alternately, every robot could be released with a searchable online licence which records the name of the designer/manufacture and the responsible human who acquired it [...]. [...] Importantly, it should still remain possible for legal liability to be shared or transferred e.g. both designer and user might share fault where a robot malfunctions during use due to a mixture of design problems and user modifications. In such circumstances, legal rules already exist to allocate liability (although we might wish to clarify these, or require insurance). But a register would always allow an aggrieved person a place to start, by finding out who was, on first principles, responsible for the robot in question” (Boden 2011: 128).

Robots in the Japanese *koseki* System: Colin P.A. Jones’s Family Law Approach to Robotic Identity and Soft-Law Based Robot Regulation

A completely different, and probably more culture-relative, soft law model for robotic identity and robot registration is based on Japanese family law. This might sound counterintuitive in Euro-American or African contexts where robots are frequently seen as unnatural and threatening. Yet Japan, being in the vanguard of human–robot communication, is different. In Japanese popular culture robots are routinely

la relation se maintient sans violence dans la paix avec cette altérité absolue. La ‘résistance’ de l’Autre ne me fait pas violence, n’agit pas négativement; elle a une structure positive: éthique.”

depicted as “an everyday part of the natural world that coexists with humans in familial contexts” (Yamaguchi 2019: 135). Moreover, there is a political dimension to it. Since 2007, with Prime Minister Shinzo Abe’s “Innovation 2025” proposal, followed by the “New Robot Strategy” of 2015 and the subsequent blueprint for a super-smart “*Society 5.0*”, also the Japanese government has been eagerly promoting “the virtues of a robot-dependent society and lifestyle” (Robertson 2014: 571). In Abe’s vision—which is futuristic and nostalgic at the same time—robots can help actualize the historicist cultural conception of “beautiful Japan” (Robertson et al. 2019: 34). Robots are seen as a “dream solution to various social problems, ranging from the country’s low birth rate, an insufficient labor force, and a need for foreign migrant workers, to disability, personal safety, and security concerns” (Yamaguchi 2019: 135).

As anthropologist Jennifer Robertson²¹ reports, nationwide surveys even suggest “that Japanese citizens are more comfortable sharing living and working environments with robots than with foreign caretakers and migrant workers. As their population continues to shrink and age faster than in other postindustrial nation-states, Japanese are banking on the robotics industry to reinvigorate the economy and to preserve the country’s alleged ethnic homogeneity. These initiatives are paralleled by a growing support among some roboticists and politicians to confer citizenship on robots” (Robertson 2014: 571). On this basis, the somewhat odd idea of robots acquiring Japanese civil status became a reality. Japanese family law provided the institutional framework for this.

In Japan, every legally significant transition in a citizen’s life—birth, death, marriage, divorce, adoption, even change of gender—is supposed to be registered in a *koseki* (戸籍), the registry of a Japanese household’s (*ie* [家]) members. In fact, it is this registration (which is historically a part of the foundation of civil law and government infrastructure in Japan) that gives legal effect to the above-mentioned events. An extract of a citizen’s *koseki* serves as the official document that confirms basic details about their Japanese identity and status (Chapman 2008).

Law scholar Colin Jones²² sees a basic analogy between Japanese family law and institutional requirements for robot regulation. The crucial point of reference is family law’s concern with parental liability for the interests, torts and crimes of minors. On Jones’ account, many issues of robot law might be “amenable to an approach that sees robots treated analogously to ‘perpetual children’. The provisions on parental liability for harm caused by children ... might provide as useful a model for allocating responsibility for

robots as anything in products liability or criminal law – if we could just figure out who the ‘parents’ are” (ibid.: 410).²³

Obviously, this was not too much of a problem for Japanese authorities in the paradigmatic case of *Paro*,²⁴ a therapeutic “mental commitment robot” with the body of an artificial baby harp seal manufactured in Nanto City, Japan. On November 7th 2010, as Jennifer Robertson notes, “*Paro* was granted its own *koseki*, or household registry, from the mayor of Nanto City, Toyama Prefecture. Shibata Takanori, *Paro*’s inventor, is listed as the robot’s father ... and a ‘birth date’ of 17 September 2004 is recorded. Media coverage of *Paro*’s *koseki* was favorable. [...] this prototypical *Paro*’s *koseki* can be construed as a branch of Shibata’s ... household, which is located in Nanto City. Thus, the ‘special family registry’ is for one particular *Paro*, and not for all of the seal-bots collectively” (Robertson 2014: 590f).

As mentioned earlier, the *koseki* conflates family, nationality and citizenship. In the case of *Caro*, that means that, by virtue of “having a Japanese father, *Paro* is entitled to a *koseki*, which confirms the robot’s Japanese citizenship” (Robertson 2014: 591). Oddly, the fact that *Paro* “is a robot—and not even a humanoid—would appear to be less relevant here than the robot’s ‘ethnic-nationality’” (ibid.). Accordingly, not *Sophia* (a social humanoid robot that became a Saudi Arabian citizen in October 2017) but *Paro* was the first robot ever to be granted citizenship.

The fact that robots can be legally adopted as members of a Japanese household definitely is an inspiration for robot law theory. Colin Jones has outlined what a “Robot *Koseki*” would look like if it were systematically contoured as a (Japanese) law model for regulating autonomous machines. His approach to “practical robot law” on providing “definitions that can be used ... to establish a framework for robotic identity. Hard or soft laws defining what is and is not a robot would be—should be—the starting point for either applying existing rules to those definitions or developing new rules” (Jones 2019: 418).

That means that a “Robot *Koseki*” would be essentially informational, yet in a robot law perspective. It would start with differentiating registered Robots from unregistered robotic AI systems, the latter thus remaining “robots without capital R” (W.S.). Unregistered AI systems may have many attributes commonly associated with “robots”. Yet they “would not be Robots for purposes of the registration system, or the rules and regulations tied to it” (Ibid.). That means that, in order to be eligible for *koseki* registration, “Robots

²¹Jennifer Robertson is Professor of Anthropology and the History of Art at the University of Michigan, Ann Arbor.

²²Colin P.A. Jones is Professor of Law at Doshisha Law School, Kyoto.

²³As Jones admits, children are “not the only area of family law that may be a useful reference. The field also deals with responsibility for adults with diminished capacity, those judicially declared incompetent or subject to guardianship or conservatorships” (Jones 2019: 411).

²⁴Its name comes from the Japanese pronunciation of “personal robot” (*pāsonaru robotto*) (cf. Robertson 2014: 590).

with capital R” (W.S.) would have to meet certain technical and normative criteria, e.g. technical specifications, safety, possible liability nexuses and so forth. In this way a “Robot Koseki” would provide third parties with “assurances” that the registered Robots satisfy certain minimum standards on which “hard and soft law requirements as well as technical rules and regulations” could then be built by governments and private actors (Jones 2019: 453).²⁵

Robots registered in the “Robot Koseki” would also be assigned “unique identifying codes or numbers that would become a key part of its identity. Codes identifying members of the same series or production line of robots could also be used. Robot Identification Numbers could even serve as taxpayer identification numbers if the Robot is accorded legal personality and the ability to engage in revenue-producing activities” (Ibid.: 455). In the Internet of Things, the “Robot Koseki” would need to work “in a way so that the current registration details of each Robot were accessible to other technology systems . . . interacting with it”, either in a centralized or distributed database system (ibid.: 464).

As Jones insists, the “Robot Koseki” would not only entail data about machines. Some of the key registration parameters should also “provide information about people involved in the creation and ongoing existence of the Robot, people who through the system will effectively become a part of the Robot’s identity”, i.e. “the maker (or manufacturer), programmer, owner, and user” of the Robot (ibid.: 465). In Jones’ eyes, this is where the Japanese *koseki* system provides “a particularly useful model, since it involves the registration of a single unit (the family) that is comprised of multiple constituents. If we are to develop robot law from family law analogies and attempt to regulate Robots as a form of ‘perpetual children’, then the *koseki* system will make

it possible to identify who is analogous to their parent(s)” (Ibid.).

This amounts to suggesting a soft law basis for hard law robot regulation. According to Jones, *koseki*-style Robot registry would not immediately call for governmental legislation but could be organized primarily by industry action. The registry could start as “a creature of code, of soft law and technical standards”; thus, it would first be driven by “industry players, professional associations or open standards organization comparable to the Internet Engineering Task Force, which has developed many of the rules and standards governing the technical aspects of the Internet” (Ibid.: 461). Based on these standards, then, both hard and soft law requirements as well as technical rules and regulations could be built by governments and private actors (ibid.: 453).

Overall, Jones’ layout of a “Robot Koseki” shows two things: (a) how to solve the problem of “robotic identity” both in a legally compatible and legally effective way without requiring the concept of “robot rights” (in Gunkel’s sense) as a basis²⁶; (b) how to provide a helpful soft-law basis for hard-law robot regulation.

Yet apart from soft law: Which *ethical* key distinctions should a hard-law robot regulation be based on?

Joanna J. Bryson’s Outline of Roboethics

In a recent paper, AI ethicist Joanna J. Bryson²⁷ (who was one of the most influential co-authors of the aforementioned EPSRC paper) has marked two central normative criteria for integrating AI-embedding systems like robots in our society: coherence and a lack of social disruption (Bryson 2018: 15). Bryson’s premises here are that “the core of all ethics is a negotiated or discovered equilibrium that creates and perpetuates a society” and “that integrating a new capacity like artificial intelligence (AI) into our moral systems is an act of normative, not descriptive, ethics” (Bryson 2018: 15). Moreover, Bryson emphasizes that “there is no necessary or predetermined position for AI in our society. This is because both AI and ethical frameworks are artefacts of our societies, and therefore subject to human control” (Bryson 2018: 15).

Bryson’s main thesis on the ethics of AI-embedding system is: While constructing such systems as either moral agents or patients is indeed possible, neither is desirable. In particular, Bryson argues “that we are unlikely to construct a coherent ethics in which it is ethical to afford AI moral

²⁵Jones envisages that “Robot Koseki” parameters would include “requirements and specifications such as those relating to: (1) the method the Robot uses to interact with other technology systems (WiFi, USB, QR codes, Bluetooth, RFID, etc.); (2) basic safety parameters as to size, speed of motility, etc.; (3) location (e.g. incorporation of GPS; compatibility with geo-fencing systems, etc.); (4) cybersecurity requirements (anti-malware/requirements, etc.); (5) access requirements (i.e. if the Robot Koseki system requires Robots to submit to software updates for various purposes, the Robot will have to be set to accept such updates regularly); (6) privacy protection (e.g. mandatory data encryption and access restrictions for video, voice, and other data recorded by the Robot); (7) operating system; (8) override capability (e.g. a kill switch that can be used remotely to shut the Robot down remotely when necessary in emergency situations); (9) sensory capabilities for perceiving the world (video, sound, motion sensors, facial recognition technology, etc.); and (10) a ‘black box’ that records all that is happening inside the Robot (software updates, a log of what and how the robot may have ‘learned’ to do things, etc.), and which can be used for forensic purposes, if necessary. Further mechanisms may be necessary to (for example) address the safety, integrity and rights (or denial) of access to the vast amount of data robots may be able to record and store. Roboticians will doubtless have other suggestions as to what technological parameters should be included.

²⁶The extent to which entitlement and defense rights for robots would come with robot citizenship as in the case of *Paro* and *Sophia* needs further discussion.

²⁷Joanna J. Bryson was associate professor in the department of computer science at the University of Bath and is now Professor of Ethics and Technology at the Hertie School of Governance, Berlin.

subjectivity. We are therefore obliged not to build AI we are obliged to” (Bryson 2018: 15).

So Bryson’s recommendations are as follows: “First, robots should not have deceptive appearance—they should not fool people into thinking they are similar to empathy-deserving moral patients. Second, their AI workings should be ‘transparent’ [...]. This implies that clear, generally-comprehensible descriptions of an artefact’s goals and intelligence should be available to any owner, operator, or other concerned party. [...]. The goal is that most healthy adult citizens should be able to make correctly-informed decisions about emotional and financial investment. As with fictional characters and plush toys [...], we should be able to both experience beneficial emotional engagement, *and* to maintain explicit knowledge of an artefact’s lack of moral subjectivity” (Bryson 2018: 23).

In my eyes, Bryson’s position is plausible and convincing. Quite obviously, even the most human-like behaving robot will not lose its ontological machine character merely by being open to “humanizing” interpretations. Rather, robots are, and will probably remain, more or less perfect simulations of humans and their agency.

But even if they do not really present an anthropological challenge (cf. Wolfe 1993), they certainly present an ethical one. I endorse the view that both AI and ethical frameworks are artefacts of our societies—and therefore subject to human choice and human control (Bryson 2018). The latter holds for the moral status of robots and other AI systems too. This status is in no way logically or ontologically set; rather, it is, and remains, a choice, not a necessity: “We can choose the types and properties of artefacts that are legal to manufacture and sell, and we can write the legislation that determines the legal rights and duties of any agent capable of knowing those rights and carrying out those duties” (Bryson 2018: 16). To this adds that self-disclosing AI would “help people match the right approach to the right entities, treating humans like humans, and machines like machines” (Bowles 2018: 188; Macrorie et al. 2019).

Conclusion

However, also the relational model sketched by Coeckelbergh and Gunkel 2020 seems to be helpful when it comes to avoiding incoherency and social disruption in ethics systems (see Van Wynsberghe 2012, 2013; van Wynsberghe and Robbins 2014; Wong and Simon 2020) Wong forthcoming. If the claim of ethics is uncompromising in the sense that it concerns action and agency as such (and not only a limited range hereof), one can argue for a *normative basic ethical continuum* in the behaviour of humans, meaning that there should be no context of action where a complete absence of human respect for the integrity of other beings (natural

or artificial) would be morally allowed or even encouraged. This might also help to minimize the risk of being morally deskilled by using technology (Coeckelbergh 2012b; Wong 2012; Vallor 2015).

With that in mind, we could consider AI-embedding machines at least as awe-inspiring. Facing them, we encounter the work and latest evolutionary product of our own intelligence: a culmination of human creativity, we are at the cutting edge of human creativity. And we are allowed to be astonished, in the sense of the *thaumázein* in Plato’s Theaitetos.²⁸ Along these lines one could think of a roboethics based upon what we owe to ourselves as creators and users of such sophisticated technology. Avoiding disrespectful treatment of robots is ultimately for the sake of the humans, not for the sake of the robots (Darling and Hauert 2013).

Maybe this insight can contribute to inspire an “overlapping consensus” (Rawls 1993: 133–172) in further discussions on responsibly coordinating HRI. Re- or paraphrasing Rawls in this perspective could start with a three part argument: (a) Rather than being dispensable, it seems reasonable to maintain the aforementioned normative basic ethical continuum in the behaviour of humans (some would add: and their artefacts); accordingly, we should (b) look for ways to stabilize this continuum, facing up (c) to the plurality of reasonable though conflicting ethical frameworks in which robots and rights can be discussed. If mere coordination of diversity is the maximum that seems achievable here, nothing more can be hoped for than a *modus vivendi* in the AI and roboethics community. Yet there might also be some “common ground” on which to build something more stable. In a Rawlsian type of “overlapping consensus” for instance, diverse and conflicting reasonable frameworks of AI and roboethics would endorse the basic ethical continuum cited above, each from its own point of view. If this configuration works out, both could be achievable: a reasonable non-eliminative hybridity of the ethical discourse on robots and rights—and the intellectual infrastructure for developing this discourse with due diligence.

²⁸This is in line with how *autómata* were viewed in antiquity (Schürmann 2006). *Autómata* were not used as industrial machinery in production. Rather, they were almost always aimed at creating amazement in the spectators, who could hardly comprehend the mysteriously autonomously moving artefacts they were seeing. Archytas of Tarentum, for instance, is said to have constructed a mechanical dove which looked deceptively real and was perhaps also able to take flight using some form of pneumatic propulsion (Gell. NA 10,12,8–10). In a similar way, an artificial snail created by Demetrius of Phalerum (308 BC; Pol. 12,13,12) and a statue of Nysa seem to have been hugely awe-inspiring; they were displayed in the procession of Ptolemaeus II Philadelphus (c. 270 BC; Ath. 5198 s.Co) for the purpose of demonstrating the power of Hellenistic rulers and legitimizing their position in relation to the people.

References

- Anderson, M., & Anderson, S. L. (Eds.). (2018). *Machine ethics*. Cambridge: Cambridge UP.
- Andreotta, A. J. (2020). The hard problem of AI rights. *AI & Society*. <https://doi.org/10.1007/s00146-020-00997-x>.
- Balkin, J. M. (2015). The path of robotics law. *California Law Review*, 6, 45–60.
- Basl, J. (2019). *The death of the ethic of life*. Oxford: Oxford UP.
- Bayern, S. J. (2019). Are autonomous entities possible? *Northwestern University Law Review Online*, 114, 23–47.
- Behdadi, D., & Munthe, C. (2020). A normative approach to artificial moral agency. *Minds and Machines*, 30, 195. <https://doi.org/10.1007/s11023-020-09525-8>.
- Bennett, B., & Daly, A. (2020). Recognising rights for robots: Can we? Will we? Should we? *Law, Innovation and Technology*, 12(1), 60–80.
- Benso, S. (2000). *The face of things: A different side of ethics*. Albany, NY: SUNY Press.
- Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge UP.
- Birch, T. H. (1993). Moral considerability and universal consideration. *Environmental Ethics*, 15, 313–332.
- Boden, M. (2011). *Principles of robotics*. Available via The United Kingdom's Engineering and Physical Sciences Research Council (EPSRC). Retrieved April 2011, from <https://www.epsrc.ac.uk/research/ourportfolio/themes/engineering/activities/principlesofrobotics/>
- Boden, M., Bryson, J. J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., Newman, P., Parry, V., Pegman, G., Rodden, T., Sorrell, T., Wallis, M., Whitby, B., & Winfield, A. (2017). Principles of robotics: Regulating robots in the real world. *Connection Science*, 29(2), 124–129.
- Bowles, C. (2018). *Future ethics*. Hove: Now Text Press.
- Bremner, P., Dennis, L. A., Fisher, M., & Winfield, A. F. (2019). On proactive, transparent, and verifiable ethical reasoning for robots. *Proceedings of the IEEE*, 107(3), 541–561.
- Brockman, J. (Ed.). (2019). *Possible minds: 25 ways of looking at AI*. New York: Penguin Press.
- Bryson, J. J. (2018). Patience is not a virtue: The design of intelligent systems and systems of ethics. *Ethics and Information Technology*, 20, 15–26.
- Bryson, J. J. (2019). *The past decade and future of AI's impact on society*. Available via BBVA. Retrieved from <https://www.bbvaopenmind.com/en/articles/the-past-decade-and-future-of-ais-impact-on-society/>
- Bryson, J. J., Diamantis, M. E., & Grant, T. D. (2017). Of, for, and by the people: The legal lacuna of synthetic persons. *Artificial Intelligence and Law*, 25(3), 273–291.
- Calerco, M. (2008). *Zoographies: The question of the animal from Heidegger to Derrida*. New York: Columbia UP.
- Campbell, K. (2001). Legal rights. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Winter 2017 Edition). Available via Stanford Encyclopedia of Philosophy. Retrieved February 19, 2020, from <https://plato.stanford.edu/entries/legal-rights/>
- Campbell, T. C. (2011). *Improper life: Technology and biopolitics from Heidegger to Agamben*. Minneapolis: University of Minnesota Press.
- Chapman, D. (2008). Sealing Japanese identity. *Critical Asian Studies*, 40(3), 423–443.
- Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology*, 12(3), 209–221.
- Coeckelbergh, M. (2011). Is ethics of robotics about robots? Philosophy of robotics beyond realism and individualism. *Law, Innovation and Technology*, 3(2), 241–250.
- Coeckelbergh, M. (2012a). *Growing moral relations: Critique of moral status ascription*. Basingstoke, NY: Palgrave Macmillan.
- Coeckelbergh, M. (2012b). Technology as skill and activity: Revisiting the problem of alienation. *Techné*, 16(3), 208–230.
- Coeckelbergh, M. (2013a). David J. Gunkel: The machine question: Critical perspectives on AI, robots, and ethics. *Ethics and Information Technology*, 15, 235–238.
- Coeckelbergh, M. (2013b). *Human being @ risk: Enhancement, technology, and the evaluation of vulnerability transformations*. Philosophy of engineering and technology (Vol. 12). Dordrecht: Springer.
- Coeckelbergh, M. (2014). The moral standing of machines: Towards a relational and non-Cartesian moral hermeneutics. *Philosophy & Technology*, 27(1), 61–77. <https://doi.org/10.1007/s13347-013-0133-8>.
- Coyne, R. (1999). *Technoromanticism: Digital narrative, holism, and the romance of the real*. Cambridge: MIT Press.
- Danaher, J. (2017a). *Robot sex: Social and ethical implications*. Cambridge: MIT Press.
- Danaher, J. (2017b). *Should robots have rights? Four perspectives*. Available via Philosophical Disquisitions. Retrieved from <https://philosophicaldisquisitions.blogspot.com/2017/10/should-robots-have-rights-four.html>
- Danaher, J. (2019). *Automation and utopia: Human flourishing in a world without work*. Cambridge: Harvard UP.
- Darling, K., & Hauert, S. (2013). *Giving rights to robots*. Available via Robohub. Retrieved from <http://robohub.org/robots-giving-rights-to-robots>. Accessed.
- Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Cham: Springer Nature.
- Dubber, M., Pasquale, F., & Das, S. (Eds.). (2020). *Oxford handbook of artificial intelligence*. Oxford: Oxford UP.
- Duncan, R. (2006). Emmanuel Levinas: Non-intentional consciousness and the status of representational thinking. In A.-T. Tymieniecka (Ed.), *Logos of phenomenology and phenomenology of the logos, Book 3. Analecta Husserliana: The yearbook of phenomenological research* (Vol. 90, pp. 271–281). Dordrecht: Springer.
- Ekbja, H. R. (2008). *Artificial dreams: The quest for non-biological intelligence*. New York: Cambridge UP.
- Floridi, L. (2013). *The ethics of information*. Oxford: Oxford UP.
- Floridi, L. (2017). Robots, jobs, taxes, and responsibilities. *Philosophy & Technology*, 30(1), 1–4.
- Ginszt, K. (2019). The status of robots in moral and legal systems: Review of David J. Gunkel (2018). *Robot rights*. Cambridge, MA: MIT Press. *Ethics in Progress*, 10(2), 27–32. <https://doi.org/10.14746/eip.2019.2.3>.
- Gunkel, D. J. (2007). *Thinking otherwise: Philosophy, communication, technology*. West Lafayette: Purdue UP.
- Gunkel, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. Cambridge: MIT Press.
- Gunkel, D. J. (2018). *Robot rights*. Cambridge: MIT Press.
- Gunkel, D. J. (2019). *How to survive a robot invasion: Rights, responsibility, and AI*. London: Routledge.
- Gunkel, D. J. (2020). *An introduction to communication and artificial intelligence*. Cambridge: Wiley Polity.
- Haraway, D. J. (2008). *When species meet*. Minneapolis: University of Minnesota Press.
- Hart, E., Timmis, J., Mitchell, P., Nakano, T., & Dabiri, F. (Eds.). (2012). *Bio-inspired models of network, information, and computing systems*. Heidelberg/Dordrecht/London/New York: Springer.
- Hohfeld, W. N. (1913). Some fundamental legal conceptions as applied in juridical reasoning. *The Yale Law Journal*, 23, 16–59.
- Ingold, T. (2000). *The perception of the environment: Essays on livelihood, dwelling and skill*. London: Routledge.
- Ingold, T. (2012). Toward an ecology of materials. *Annual Review of Anthropology*, 41, 427–442.
- Jones, J.P. (2019). The Robot Koseki: A Japanese Law Model for Regulating Autonomous Machines. *J. Bus. & Tech. L.*, 14, 403–467.
- Keßler, F. (2019). *Wie verklage ich einen Roboter?* Available via Spiegel. Retrieved February 17, 2020, from <https://www.spiegel.de/karriere/kuenstliche-intelligenz-so-koennten-roboter-haften-wenn-sie-fehler-machen-a-1263974.html>

- Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication Theory*, 15(2), 127–147.
- Latour, B. (2015). *Face à Gaïa: Huit Conférences sur le Nouveau Régime Climatique*. Paris: Éditions La Découverte.
- Lévinas, E. (1990). *Lévinas*. Giessen: Focus.
- Lin, P., Jenkins, R. K., & Bekey, G. A. (Eds.). (2014). *Robot ethics 2.0: From autonomous cars to artificial intelligence*. Cambridge: MIT Press.
- Loh, J. (2019). *Roboterethik: Eine Einführung*. Berlin: Suhrkamp.
- Macrorie, R., Marvin, S., & While, A. (2019). Robotics and automation in the city: A research agenda. *Urban Geography*. <https://doi.org/10.1080/02723638.2019.1698868>.
- Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133), 20180089. <https://doi.org/10.1098/rsta.2018.0089>.
- Nourbakhsh, I. (2013). *Robot futures*. Cambridge: MIT Press.
- Nyholm, S. (2020). *Humans and robots: Ethics, agency, and anthropomorphism*. London: Rowman & Littlefield International.
- Pasquale, F. (2018). A rule of persons, not machines: The limits of legal automation. *The George Washington Law Review*, 87(1), 1–55.
- Prescott, T. J. (2017). Robots are not just tools. *Connection Science*, 29(2), 142–149.
- Rawls, J. (1993). *Political liberalism*. New York: Columbia University Press.
- Robertson, J. (2014). Human rights vs. robot rights: Forecasts from Japan. *Critical Asian Studies*, 46(4), 571–598.
- Robertson, L. J., Abbas, R., Alici, G., Munoz, A., & Michael, K. (2019). Engineering-based design methodology for embedding ethics in autonomous robots. *Proceedings of the IEEE*, 107(3), 582–599.
- Rosenthal-von der Pütten, A. M., Krämer, N. C., & Herrmann, J. (2018). The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *International Journal of Social Robotics*, 10, 569–582.
- Sandini, G., & Sciutti, A. (2018). Humane robots—From robots with a humanoid body to robots with an anthropomorphic mind. *ACM Transactions on Human-Robot Interaction*, 7(1), 1–7. <https://doi.org/10.1145/3208954>.
- Scharre, P. (2018). *Army of none: Autonomous weapons and the future of war*. New York: W.W. Norton.
- Schröder, W. M. (2012). Natur- und Vernunftrecht. In G. Lohmann & A. Pollmann (Eds.), *Menschenrechte: Ein interdisziplinäres Handbuch* (pp. 179–185). Stuttgart: Metzler Verlag.
- Schröder, W. M., Gollmer, K. U., Schmidt, M., & Wartha, M. (2021). *Kompass Künstliche Intelligenz: Ein Plädoyer für einen aufgeklärten Umgang*. Würzburg: Würzburg University Press.
- Schürmann, A. (2006). Automata. In H. Cancik & H. Schneider (Eds.), *Brill's new Pauly*. Available via BRILL. https://doi.org/10.1163/1574-9347_bnp_e210220.
- Siewert, P., Ameling, W., Jansen-Winkel, K., Robbins, E., & Klose, D. (2006). Nomos. In: H. Cancik & H. Schneider (Eds.), *Brill's new Pauly*. Available via BRILL. https://doi.org/10.1163/1574-9347_bnp_e210220.
- Soh, C., & Connolly, D. (2020). New frontiers of profit and risk: The fourth industrial Revolution's impact on business and human rights. *New Political Economy*. <https://doi.org/10.1080/13563467.2020.1723514>.
- Spiekermann, S. (2019). *Digitale Ethik: Ein Wertesystem für das 21. Jahrhundert*. Droemer & Knaur, München: Jahrhundert.
- Swart, J. A. A. (2013). Growing moral relations: Critique of moral status ascription. *Journal of Agricultural and Environmental Ethics*, 26(6), 1241–1245.
- Teubner, G. (2006). Rights of non-humans? Electronic agents and animals as new actors in politics and law. *Journal of Law and Society*, 33(4), 497–521.
- Torrance, S. (2012). *The centrality of machine consciousness to machine ethics*. Paper presented at the symposium 'The machine question: AI, ethics, and moral responsibility', AISB/IACAP world congress 2012, Birmingham, 4 July 2012.
- Turner, J. (2019). *Robot rules: Regulating artificial intelligence*. Cham: Palgrave Macmillan.
- Vallor, S. (2015). Moral deskilling and upskilling in a new machine age: Reflections on the ambiguous future of character. *Philosophy & Technology*, 28, 107–124.
- van Wynsberghe, A. (2012). Designing robots for care: Care centered value-sensitive design. *Science and Engineering Ethics*, 19(2), 407–433.
- van Wynsberghe, A. (2013). A method for integrating ethics into the design of robots. *Industrial Robot*, 40(5), 433–440.
- van Wynsberghe, A., & Robbins, S. (2014). Ethicist as designer: A pragmatic approach to ethics in the lab. *Science and Engineering Ethics*, 20(4), 947–961.
- Walch, K. (2019). *Ethical concerns of AI*. Available via Forbes. Retrieved February 19, 2020, from <https://www.forbes.com/sites/cognitiveworld/2020/12/29/ethical-concerns-of-ai/#10a9affd23a8>
- Wallach, W. (2007). Implementing moral decision-making faculties in computers and robots. *AI & Society*, 22(4), 463–475.
- Wallach, W. (2010). Robot minds and human ethics: The need for a comprehensive model of moral decision-making. *Ethics and Information Technology*, 12(3), 243–250.
- Wallach, W., & Allen, C. (2010). *Moral machines: Teaching robots right from wrong*. New York: Oxford UP.
- Wenar, L. (2015). Rights. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2015 Edition). Available via Stanford Encyclopedia of Philosophy. Retrieved from <https://plato.stanford.edu/archives/fall2015/entries/rights/>. Accessed. 9/25/20.
- Winfield, A. (2007). *The rights of robot*. Available via Alan Winfield's Web Log. Retrieved from <https://alanwinfield.blogspot.com/search?q=rights+of+robots>. Accessed. 9/25/20.
- Winfield, A. (2012). *Robotics: A very short introduction*. Oxford: Oxford UP.
- Wolfe, A. (1993). *The human difference: Animals, computers, and the necessity of social science*. Berkeley/Los Angeles/London: University of California Press.
- Wong, P. H. (2012). Dao, harmony and personhood: Towards a Confucian ethics of technology. *Philosophy & Technology*, 25(1), 67–86.
- Wong, P. H. (forthcoming). *Global engineering ethics*. In D. Michelfelder & N. Doorn (Eds.), *Routledge handbook of philosophy of engineering*. Wong 2020 is still in press; it has not yet been published.
- Wong, P. H., & Simon, J. (2020). Thinking about 'ethics' in the ethics of AI. *IDEES*, 48.
- Yamaguchi, T. (2019) Japan's robotic future. *Critical Asian Studies*, 51(1), 134–140.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

