

O

Janina Loh, Wulf Loh (eds.)

**SOCIAL
ROBOTICS
AND THE GOOD LIFE**

The Normative Side of Forming
Emotional Bonds With Robots

JL

[transcript] Philosophy

Janina Loh, Wulf Loh (eds.)
Social Robotics and the Good Life

Philosophy

Janina Loh (née Sombetzki) is an ethicist (Stabsstelle Ethik) at Stiftung Liebenau in Meckenbeuren on Lake Constance. They have been a university assistant (Post-Doc) at Christian-Albrechts-Universität Kiel (2013-2016) and in the field of philosophy of technology and media at Universität Wien (2016-2021). Their main research interests lie in the field of critical posthumanism, robot ethics, feminist philosophy of technology, responsibility research, Hannah Arendt, theories of judgement, polyamory and poly-ethics as well as ethics in the sciences.

Wulf Loh is an assistant professor at the Int. Center for Ethics in the Sciences and Humanities (IZEW) at Eberhard Karls Universität Tübingen and supervisor of various technology development projects. His areas of expertise are within applied ethics, social philosophy, political philosophy, and philosophy of law.

Janina Loh, Wulf Loh (eds.)

Social Robotics and the Good Life

The Normative Side of Forming Emotional Bonds With Robots

[transcript]

The free availability of the e-book edition of this publication was financed by the Fachinformationsdienst Philosophie.



Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>



This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 (BY-SA) which means that the text may be remixed, build upon and be distributed, provided credit is given to the author and that copies or adaptations of the work are released under the same or similar license. For details go to

<https://creativecommons.org/licenses/by-sa/4.0/>

Creative Commons license terms for re-use do not apply to any content (such as graphs, figures, photos, excerpts, etc.) not original to the Open Access publication and further permission may be required from the rights holder. The obligation to research and clear permission lies solely with the party re-using the material.

First published in 2023 by transcript Verlag, Bielefeld

© Janina Loh, Wulf Loh (eds.)

Cover layout: Kordula Röckenhaus, Bielefeld

Printed by Majuskel Medienproduktion GmbH, Wetzlar

Print-ISBN 978-3-8376-6265-8

PDF-ISBN 978-3-8394-6265-2

<https://doi.org/10.14361/9783839462652>

ISSN of series: 2702-900X

eISSN of series: 2702-9018

Printed on permanent acid-free text paper.

Contents

Introduction – Social Robotics and the Good Life

The Normative Side of Forming Emotional Bonds With Robots <i>Janina Loh, Wulf Loh</i>	7
--	---

Part I – Understanding, Defining, Conceptualizing

Virtues, Robots, and Good Lives: Who Cares?

<i>Charles M. Ess</i>	25
-----------------------------	----

The Relational Turn

Thinking Robots Otherwise <i>David J. Gunkel</i>	55
---	----

From Tool Use to Social Interactions

<i>Anna Strasser</i>	77
----------------------------	----

Reflecting (on) Replika

Can We Have a Good Affective Relationship With a Social Chatbot? <i>Eva Weber-Guskar</i>	103
---	-----

Part II – Design, Imitation, Trust

You Can Love a Robot, But Should You Fight With it?

Perspectives on Expectation, Trust, and the Usefulness of Frustration Expression in Human-Machine Interaction <i>Jacqueline Bellon, Tom Poljanšek</i>	129
---	-----

Empathic Machines?

Ethical Challenges of Affective Computing from a Sustainable Development Perspective

Cordula Brand, Leonie N. Bossert, Thomas Potthast 157

Part III – Care, Love, Sex

Granny and the Sexbots

An Ethical Appraisal of the Use of Sexbots in Residential Care Institutions for Elderly People

Karen Lancaster 181

Alice Does not Care

Or: Why it Matters That Robots “Don’t Give a Damn”

Imke von Maur 209

Emotional Embodiment in Humanoid Sex and Love Robots

Cindy Friedman, Sven Nyholm, Lily Frank 233

Contributors 257

Introduction – Social Robotics and the Good Life

The Normative Side of Forming Emotional Bonds With Robots

Janina Loh, Wulf Loh

Robots and Their Tasks in Society

Robots have existed as a concept that is still common today for a hundred years. Historically, the term “robot” goes back to the Czech word “robota”, which stands for work, hard labor, and forced labor. Karel Čapek first used it in the play R.U.R. Rossum’s Universal Robots (1921) to refer to humanoid contraptions that are at the service of humans. This historical vision of the robot as an artificial slave subsequently was the guiding idea shaping the development of robotics.¹

Without doubt, robots can now be found in many areas of human life. Industry was the first sector they made their entrance into around the middle of the 20th century with the *Unimate*² robot. Particularly in this area, robots are assigned to jobs that are considered “dull, dangerous, and dirty”. Although it is anything but clear which activities fall under this category (Marr 2017), any socio-ethical doubts arising from this have not been able to put a stop to the advancing robotization of industrial production. Today, tasks that are considered boring because they are repetitive and monotonous, as well as dirty

1 By “robot”, we mean an electro-mechanical machine that a) has some form of independent body, b) possesses at least one processor, c) is equipped with sensors that collect information about the world d) as well as with at least one effector that translates signals into mechanical processes. A robot’s behavior e) is or at least appears autonomous, enabling it to f) interact with or influence its environment (cf. Loh 2019a: 7; Misselhorn 2013: 43).

2 A composite of “universal” and “automation”; constructed by Joseph Engelberger in 1961.

and even dangerous, are increasingly being performed by robots on assembly lines, in production halls and warehouses around the globe. The Kuka robots in the automotive industry and the Amazon warehouse robots are but some recent examples of this phenomenon.

Social Robotics and Robots as Social Companions

Robots are increasingly entering the personal sphere of everyday human interaction. This area of so-called *social robotics* is characterized by a large number of everyday activities, that at the same time can be very private and sensitive for various reasons (Breazeal 2002; Duffy 2008, 2004; Fong et al. 2003; Markowitz 2015; Seibt et al. 2016; Fronemann et al. 2022). The main fields in which robots are employed here are therapy and care, specifically in the form of activation and keeping company, but also education and even sexuality, friendship, and love. In those areas, the attribution of the three Ds no longer applies. Even when the robots are primarily designed as auxiliary tools for daily tasks, the typical activities are marked by interdependence and physical and emotional closeness.

From robots for sales assistance such as Paul (who guides customers through the aisles of the electronics retailer Saturn), care robots like Care-O-Bot, entertainment robots such as Pepper, to sex robots (Markowitz 2015: 41), are all examples for the growing group of social robots that are used in close proximity to humans. Depending on their respective tasks and the extent to which they enter into direct interaction with humans, they need to possess social skills in some form. This anthology is dedicated to them and to the question about the relationships they allegedly form or are supposed to be able to form with humans.

In this regard, the authors of this anthology take a closer look at three topics of social robotics in particular. *The first part* revolves around questions of defining and understanding basic elements of human-robot interaction (HRI). Charles Ess, David Gunkel, Anna Strasser, and Eva Weber-Guskar examine the basic anthropological and ontological assumptions of HRI, ask what we mean by “social agent” and, correspondingly, by “artificial social agent”, address whether interaction with robots should be reconsidered in general, and when it is morally appropriate to speak of successful HRI.

The second part deals with questions of the design of robots as social companions, the imitation of emotions, and corresponding or associated reac-

tions such as trust and expectations with regard to the behavior of a robot. Cordula Brand, Leonie N. Bossert, Thomas Potthast, Jacqueline Bellon, and Tom Poljanšek in their contributions cover the ethical implications that arise from the embodied and anthropomorphic design of robots.

Finally, in the *third part*, the authors address whether specific emotional relationships with robots are possible, namely caring, loving, and sexual relationships. Imke von Maur, Lily Frank, Cindy Friedman, Sven Nyholm, and Karen Lancaster are concerned with the possibility and ethical appropriateness of these forms of relationships and its repercussions.

Part I – Understanding, Defining, Conceptualizing: Robots as Social Companions

Already Čapek raises numerous philosophical, ethical, and anthropological questions in his “R.U.R. Rossum’s Universal Robots”. These include, for example, the nature of man, the responsibility of scientists, and what it means to form an emotional bond with another being. For instance, his piece ends with the prospect of a love affair that seems to be developing between two robots. Thus, in the historical understanding of the robot established by Čapek, a broad foundation is already laid for the discussions that were to arise in the decades that followed and that extend into social robotics. Is it possible to live a good, successful life with and around robots, including forming intimate romantic and sexual relationships with them?

In his text “Virtues, Robots, and Good Lives: Who Cares?”, *Charles Ess* deals with these questions, which also concerned Čapek, from a virtue ethics perspective. Maybe a reformulated, “relational” virtue ethics, supplemented by an “ethical pluralism”, can help circumvent the problems of an “ethical relativism and computer-mediated colonization”? By way of applying virtue ethics to sex robotics as an example of emotionally intimate relationships with robots, Ess is skeptical, however, that a sexual human-machine relationship can ever be said to be truly “complete.” On the other hand, it is too quick to simply depreciate robots on the basis of, say, their lack of autonomy, as this ultimately leads to a “reinscribing of traditional patriarchal and racist attitudes”. Therefore, it is necessary to further develop a “pluralistic” ethics of virtue into an “ethics of care” appropriate to the current developments in social robotics.

In order to answer the question whether relationships with robots are desirable, David Gunkel, Eva Weber-Guskar and Anna Strasser, first address the

status of robots in general. Can we understand them as agents with whom genuine relationships are possible or are robots unable to transcend their traditional object status? With these considerations, the authors tie in with a broad tradition in robot ethics, in which basically three currents can be distinguished.

Within the first research area, authors ask to what extent robots can be considered as potential moral agents (Floridi/Sanders 2004; Misselhorn 2013; Moor 2006; Sullins 2006; Wallach/Allen 2009). Accordingly, they consider the degree to which robots are capable of moral action and which competencies they must possess to this end. Depending on the respective understanding of agency, morality, and the competencies to be realized for this purpose, this includes the attribution of freedom and autonomy as a condition for moral action, cognitive competencies (such as thinking, mind, reason, judgment, intelligence, consciousness, perception, and communication), but also empathy and emotions. Defining the “minimal conditions” for understanding robots as “social agents” is also *Anna Strasser’s* concern in her text “From Tool Use to Social Interactions”. For this purpose, she considers the case of “joint actions” as an example for “social interaction” and seeks to establish a more appropriate understanding of HRI or the relationships we enter into with robots.

Within the second area, authors deal with the question whether robots should be regarded as moral patients, i.e. as objects of moral consideration (Damiano/Dumouchel 2018; Darling 2012, 2017; Duffy 2003, 2004, 2008; Gerdes 2017; Johnson 2011; Tavani 2018). These approaches are concerned with how to deal with artificial systems, what kind of moral value they may have, even if they may be incapable of moral agency themselves. Topics include, for example, the formulation of codes of ethics in corporations, the desirability and possibility of relationships with and to robots, the question of exploiting or “enslaving” robots, or the assessment of the use of robots for therapeutic purposes. Some thinkers discuss the possibility of ascribing rudimentary rights to some types of robots. Just as Immanuel Kant in § 17 of the second part of his *Metaphysics of Morals* is opposed to cruelty towards animals, because this would lead to morally questionable attitudes in us humans, Kate Darling, for example, argues in favor of robot rights because people are then more likely to maintain their moral virtues in other interactions as well.

Nonetheless, the fact that robots are regarded merely as moral patients does not preclude the possibility of humans having emotional relationships with them. Against the background of this possibility, *Eva Weber-Guskar* in her

text “Reflecting (on) Replika: Can We Have a Good Affective Relationship With a Social Chatbot?” rejects the possibility of a relationship comparable to that between humans, using the example of the social chatbot Replika. However, Weber-Guskar also emphasizes that “the lack of emotional mutuality” does not justify a general rejection of the possibility of affective relationships with “social or emotional(ized) AI”.

A third strand within robotics ethics transcends the obvious dichotomy between moral agency and patiency. Authors here discuss alternatives to the classical distinction between subjects and objects of moral action. Within the framework of these “inclusive” or “inclusivist” approaches (Loh 2019a, 2020, 2022), the focus is on problems with a traditional conception of the (human) person that underlies the notion of the moral agent. The understanding of the human being as the core of ethical thinking, as the main moral agent, as the pivot of the attribution of abilities, competences, and values is questioned and challenged in these inclusive approaches.

In his text “The Relational Turn: Thinking Robots Otherwise,” *David Gunkel* describes the project of such an inclusive approach in order to “introduce and formulate a meta-ethical theory”. First, he takes a look at the peculiarities of the classical, exclusive ethical positions. In a second step, he then outlines his alternative of a relational Thinking Otherwise, and in a third step meets possible objections. Gunkel is thus concerned with a general new understanding of the possibility of entering into relationships with robots. He shows that questions of moral status of and potential relationships with robots have less to do with the robots themselves and more to do with “us and the limits of who is included in and what comes to be excluded from that first-person plural pronoun, ‘we.’”

Part II – Design, Imitation, Trust: Anthropomorphization and its Function for Social Robotics

Robots as social companions can only carry out their activities in close proximity to humans, if they are accepted by people in their immediate private space. People do not want to be assisted in their personal hygiene by scary or repulsive machines, they do not want to be cared for and touched mentally and physically by cold apparatuses – this seems a pretty straightforward assumption and is therefore at the bottom of almost all social robot design. The latter is quasi unanimously catered to an image of robots that their human

users can identify with and thus engage with more easily. They are, according to Kate Darling, “specifically designed to socially interact with humans” (2012). Such a trust-inspiring design is in many cases anthropomorphic (i.e., human-like), or more rarely, zoomorphic (i.e., animal-like).

The anthropomorphization of non-human entities does not only concern their outer form, but can also refer to their behavior and thus to the attribution of human competencies (Fink 2012: 200). Therefore, in the field of social robotics, a distinction is sometimes made between anthropomorphic design, which primarily comprises externally perceptible criteria such as “shape, speech capabilities, facial expression,” and the like, and anthropomorphic interaction design, which targets the “social phenomenon that emerges from the interaction between a robot and an [sic!] user” and is sometimes called “anthropomorphism” (in the proper sense) (all citations in Lemaignan et al. 2014: 226; cf. Złotowski et al. 2015).

Our capacity for anthropomorphization seems to be a psychological fact and thus primarily a topic of the social sciences, psychology, and (in the context of robots) science and technology studies. However, the anthropomorphic lens through which we often view and evaluate the nonhuman world frequently serves as a vehicle for placing humanized beings in the moral universe (see *Eva Weber-Guskar’s* text in this volume for more on this). After all, the more human we assess a counterpart, the more similar we make it to ourselves, the more we identify with it, the more willing we are also to assign it a (moral) value similar to that of humans. Indeed, we are forced to do so to a certain extent if we do not want to be argumentatively inconsistent. The anthropomorphic gaze gives a moral value to all beings it encounters.

In psychology, anthropomorphization is traditionally viewed in a negative light, “as a bias, a category mistake, an obstacle to the advancement of knowledge, and as a psychological disposition typical of those who are immature and unenlightened, i.e., young children and ‘primitive people’” (Damiano/Dumouchel 2018: 2; cf. Duffy 2003: 180-181). Some authors even go a step further by declaring humanized robots to be a kind of “cheating technology” that is ethically problematic (Turkle 2011: 514). Humanized robots not only deceive us into believing that they possess mental states, but also into the “illusion of relationship” that we can actually only enter into with humans (Turkle 2005: 62; cf. Damiano/Dumouchel 2018: 1; Lin 2012: 11).

This argument of a deceptive technology or a “culture of simulation” (Turkle 2011) can be interpreted in terms of virtue ethics, insofar as the good life importantly also depends on human relationships (Nussbaum 2007). In

this sense, humans commit a moral error when they replace human-human relationships as the genuine form of relationship with a mirage evoked by a robot (see also the text by *Charles Ess* in this volume).³ Accordingly, a “simulated feeling is never feeling, simulated love is never love” (Turkle 2010: 4; cf. Damiano/Dumouchel 2018: 5).

The traditionally negative connotation of anthropomorphism in psychology⁴ is countered by studies in social robotics that contrast these concerns with a positive interpretation of the human capacity for anthropomorphization. For example, Luisa Damiano and Paul Dumouchel outline an optimistic approach that views anthropomorphism not as a “cognitive error” but “as a fundamental tool” (2018: 5) that can support and enhance HRI. They state that what makes social robots special is that they “tend to blur the traditional ontological categories that humans use to describe the world,” most notably the subject-object dichotomy, but also the categories of animate and inanimate, sentient and non-sentient (2018: 4) (on overcoming the subject-object dichotomy in inclusive robot ethics approaches, see also the text by *David Gunkel* in this volume).

Other authors view anthropomorphizing robots only as “desirable where it enhances the function of the technology” (Darling 2017: 174). Humans can and should identify with social robots used in elder care and households, as these machines interact with their owners in an intimate way. Otherwise, they would be unable or unwilling to engage with the artificial system. On the other hand, authors conclude that a robot must “not be too similar to a human being if it is supposed to elicit empathy” (Misselhorn 2009: 117). Often referring to the so-called “uncanny valley” (Mori 1970: 33-35), they argue that otherwise, the inability to clearly categorize the robot adequately according to such important categories as animate/ inanimate will irritate, repel, and scare us.

It is interesting to note, however, that with regard to the question of whether it makes sense to design robots in such a way that people can engage

3 The argument against anthropomorphism can also be spelled out deontologically and utilitarian: deontologically, if I neglect my duties towards other people, for example, because of my emotional attachment to a robot; and utilitarian, if I should, for example, attribute the ability to suffer to robots due to anthropomorphizing them, and therefore include them in the calculation of total utility.

4 There is also a positive interpretation of anthropomorphism in psychology, pointing out that under the anthropomorphic gaze nonhuman beings can become “familiar, explainable, or predictable” (Fink 2012: 200).

and form relationships with them, the focus is primarily on positive emotional forms of relationships such as care, love, and friendship. As discussed earlier, negative affective relationships are mostly used as reasons for rejecting an anthropomorphic design. In their text “You Can Love a Robot, But Should You Fight With it?”, *Jacqueline Bellon* and *Tom Poljanšek* raise the question of whether such “frustrated-related concepts of human emotion” can also be at times conducive to a good life. If this is the case, the simulation of negative emotions should not be excluded in HRI.

In social robotics, the question of appropriate design of robots has so far been limited to direct interaction in close proximity between humans and machines. In their text “Empathic Machines? Ethical Challenges of Affective Computing from a Sustainable Development Perspective”, *Cordula Brand*, *Leonie N. Bossart* and *Thomas Potthast* extend this narrow perspective using a justice-based approach, namely the “Sustainable Development framework”. By doing so, they evaluate whether and how the presumed advantage of “affective computing” can actually be realized for all humans in a just and needs-based way.

Part III – Care, Love, and Sex With Robots as Social Companions

Depending on the status we are willing to ascribe to robots (Part I of this volume) and depending on their design, which can be attractive or repulsive to us (Part II of this volume), some people actually enter into friendly, sexual, caring, or loving relationships with robots. In a very literal sense, it is sex robots that are closest to us.

For most people, sex robotics may sound like pure science fiction. In reality, however, there are already several large international companies that mass-produce and sell sex robots – including two Chinese companies (DS Doll Robotics and Shenshen All Intelligent Technology Co.) and one from the USA (RealDoll with RealBotix). TrueCompany in 2010 was the first company in the world to launch a sex robot called Roxxy. Roxxy had interactive capabilities such as, according to the now offline homepage, “hear what you say, speak, feel your touch, move their bodies, are mobile and have emotions and a personality.” Roxxy was said to be able to develop its own personality (or as many different roles as desired) through interaction with its users. But it was also possible to give her one of five pre-programmed personalities. Besides that, she could be given different hairstyles and hair colors. Aside from the

aforementioned skills, Roxxy should have also been able to “listen, talk, carry on a conversation and feel your touch” and even “have an orgasm” (TrueCompanion 2019).

Other examples include Shenshen all Intelligent Technology’s Emma robot, Matt McMullen’s sex robot Harmony (Realbotix), Samantha that is supposedly equipped with a “moral code”, as well as LumiDoll’s sex robot Kylie (Mlot 2018; Morgan 2017). It is clear that in sex robots highly questionable gender stereotypes are oftentimes upheld and heteronormative, patriarchal, instrumentalizing, and discriminatory power structures are confirmed (see also the text by *Charles Ess* in this volume). As a result, the spectrum of ethical issues is evident (Cheok et al. 2017; Danaher 2017; Danaher/McArthur 2017; Kubes 2019; Levy 2012, 2008; Loh 2019b; Scheutz 2012; Whitby 2012).

In her text “Granny and the Sexbots”, *Karen Lancaster* deals with the possibility and moral desirability of using sex robots in elder care. In doing so, she addresses not one but two taboo topics – aside from sex robotics, the fact that older people may also have a wish to experience a fulfilled sexuality.

At the same time, Lancaster’s text bridges the gap between two fields of social robotics, namely between sex robotics on the one hand and medical, therapeutic and nursing robotics on the other. Here, too, the use of sex robots is discussed. But apart from sex robots, numerous assistance systems are already being used today to support caregivers in the medical, therapy, and nursing sectors in their often extremely physically and mentally demanding work. From lifting and transport systems to companion robots that activate, entertain, and thereby reduce loneliness, to therapy robots that promote communication with patients, a broad spectrum of artificial systems exists for a variety of different tasks in hospitals, therapy, and care facilities. One field that has been developing steadily for a good 15 years, for example, is the robot use in the therapy of children with autism (Richardson et al. 2018; critically Elder 2017).

The artificial seal Paro is a good example of a zoomorphically designed care assistance robot. Paro is modeled after a young harp seal and mainly used in geriatric care and therapy, especially for people with severe dementia. These people particularly tend to isolate themselves from their human caregivers, but often open up to animals. As a robot, Paro cannot be accidentally hurt as it might happen to real animals. Nonetheless, it is said to offer many of the benefits of a regular human-animal-interaction therapy in these application contexts (Shibata/Wada 2011; Wada et al. 2008).

Because of the numerous challenges associated with robots as potential social companions, some thinkers are skeptical about the question of whether it makes sense or is morally desirable to develop robots in such a way that humans want to form relationships with them. In this vein, in her text “Alice Does not Care. Or: Why it Matters That Robots ‘Don’t Give a Damn’, *Imke von Maur* also rejects the use of care robots with the aim “to reduce loneliness”. Since “real care” involves both meanings of the word, robots are according to her not capable of caring in this sense. Even more, people run the risk of “giving up expectations of real care and true relationships” when they get involved with robots.

The debates about the pros and cons of the possibility and especially the moral desirability (in terms of a good life) of emotional relationships with robots are pointedly summarized in the text “Emotional Embodiment in Humanoid Sex and Love Robots” by *Cindy Friedman, Sven Nyholm, and Lily Frank*. By way of three ethical questions, they discuss some of the central approaches to robot ethics, and in doing so bring to the forefront the challenges we face in the theoretical conception (cf. Part I of this volume), the actual design (cf. Part II), and the practical use of robots as potential social companions (cf. Part III).

The authors of this anthology show that the challenges we face regarding the questions of the good life and the possibility of emotional relationships with robots are undoubtedly manifold. But they also show that addressing these questions is worthwhile as we become more sensitive to the complexity of a society in the age of digitalization, automation, and robotization.

References

- Ackermann, Evan (2015): “Care-O-bot 4 Is the Robot Servant We All Want but Probably Can’t Afford”, in: *IEEE Spectrum*; <https://spectrum.ieee.org/automaton/robotics/home-robots/care-o-bot-4-mobile-manipulator>.
- Anderson, Michael/Anderson, Susan Leigh/Armen, Chris (2006a): “MedEthEx. A Prototype Medical Ethics Advisor.” In: *Proceedings of the Eighteenth Conference on Innovative Applications of Artificial Intelligence*, Boston, Massachusetts, pp. 1759-1765.
- Anderson, Michael/Anderson, Susan Leigh/Armen, Chris (2006b): “An Approach to Computing Ethics”, in: *Intelligent Systems IEEE 4*, pp. 2-9.

- Breazeal, Cynthia (2002): *Designing Sociable Robots*, Cambridge MA, London: The MIT Press.
- Cheok, Adrian/Karunanayaka, Kasun/Yann Zhang, Emma (2017): “Lovotics. Human-Robot Love and Sex Relationships.” In: Patrick Lin/Ryan Jenkins/Keith Abney (eds.), *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, New York: Oxford University Press, pp. 193-213.
- Damiano, Luisa/Dumouchel, Paul (2018): “Anthropomorphism in Human-Robot Co-evolution.” In: *Frontiers in Psychology* 9, pp. 1-9; <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00468/full#B54>.
- Danaher, John (2019): “Should We Be Thinking about Robot Sex?” In: John Danaher/Neil MaArthur (eds.), *Robot Sex. Social and Ethical Implications*, Cambridge, Massachusetts, London: The MIT Press, pp. 3-14.
- Danaher, John/McArthur, Neil (eds.) (2017): *Robot Sex. Social and Ethical Implications*, Cambridge, Massachusetts, London: The MIT Press.
- Darling, Kate (2012): “Extending Legal Protection to Social Robots.” In: *IEEE Spectrum*, September 10 2012; <https://spectrum.ieee.org/autaton/robotics/artificial-intelligence/extending-legal-protection-to-social-robots>.
- Darling, Kate (2017): “Who’s Jonny? Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy.” In: Patrick Lin/Ryan Jenkins/Keith Abney (eds.), *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, New York: Oxford University Press, pp. 173-188.
- Duffy, Brian R. (2003): “Anthropomorphism and the social robot.” In: *Robotics and Autonomous Systems* 42, pp. 177-190.
- Duffy, Brian R. (2004): “Social Embodiment in Autonomous Mobile Robotics.” In: *International Journal of Advanced Robotic Systems* 1, pp. 155-170.
- Duffy, Brian R. (2008): “Fundamental Issues In Affective Intelligent Social Machines.” In: *Open Artificial Intelligence Journal* 2, pp. 21-34.
- Elder, Alexis (2017): “Robot Friends for Autistic Children. Monopoly Money.” In: Patrick Lin/Ryan Jenkins/Keith Abney (eds.), *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, New York: Oxford University Press, pp. 113-126.
- Fink, Julia (2012): “Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction.” In: Shuzhi Sam Ge/Oussama Khatib/John-John Cabibihan/Reid Simmons/Mary-Anne Williams (eds.), *Social Robotics. 4th International Conference, ICSR 2012, Chengdu, China, October 2012. Proceedings*, Berlin, Heidelberg: Springer, pp. 199-208.

- Floridi, Luciano/Sanders, J. W. (2004): "On the Morality of Artificial Agents." In: *Minds and Machines* 14, pp. 349-379.
- Fong, Nourbakhsh/Terrence, Illah/Dautenhahn, Kerstin (2003): "A survey of socially interactive robots." In: *Robotics and Autonomous Systems* 42, pp. 143-166.
- Fronemann, Nora/Pollmann, Kathrin/Loh, Wulf (2022): "Should my robot know what's best for me? Human-robot interaction between user experience and ethical design", in: *AI & Society* 37, pp. 517-533.
- Gerdes, Anne (2017): "The Issue of Moral Consideration in Robot Ethics." In: *ACM SIGCAS Computers & Society* 45, pp. 247-279.
- Johnson, Deborah G. (2011): "Computer Systems. Moral Entities but Not Moral Agents." In: Michael Anderson/Susan Leigh Anderson (eds.), *Machine Ethics*, New York, pp. 168-183.
- Kubes, Tanja (2019): *Bypassing the Uncanny Valley. Sex Robots and Robot Sex Beyond Mimicry*. In: Janina Loh/Mark Coeckelbergh (eds.), *Feminist Philosophy of Technology*, Stuttgart: J.B. Metzler, pp. 59-73.
- Lemaignan, Séverin/Fink, Julia/Dillenbourg, Pierre (2014): "The Dynamics of Anthropomorphism in Robotics." In: *Hri 14 Proceedings of 2014 ACM/IEEE International Conference on Human-Robot Interactions*, New York: ACM, pp. 226-227.
- Levy, David (2008): *Love + Sex with Robots. The Evolution of Human-Robot Relationships*, New York, London, Toronto, Sydney, New Delhi, Auckland: Harper Perennial.
- Levy, David (2012): "The Ethics of Robot Prostitutes." In: Patrick Lin/Keith Abney/George Bekey (eds.), *Robot Ethics. The Ethical and Social Implications of Robotics*, Cambridge, Massachusetts, London: Oxford University Press, pp. 223-231.
- Loh, Janina (2019a): "Responsibility and Robotethics: A Critical Overview." In: *Philosophies. Special Issue Philosophy and the Ethics of Technology*, 4/58; <https://www.mdpi.com/2409-9287/4/4/58/htm>.
- Loh, Janina (2019b): "What is Feminist Philosophy of Technology? A Critical Overview and a Plea for a Feminist Technoscientific Utopia." In: Janina Loh/Mark Coeckelbergh (eds.), *Feminist Philosophy of Technology*. Stuttgart: J.B. Metzler, pp. 1-24.
- Loh, Janina (2020): "Ascribing Rights to Robots as Potential Moral Patients." In: John-Stewart Gordon (ed.), *Smart Technologies and Fundamental Rights. Book series Philosophy and Human Rights*, Brill, pp. 101-126.

- Loh, Janina (2022): “Posthumanism and Ethics.” In: Stefan Herbrechter/Ivan Callus/Manuela Rossini/Marija Grech/Megen de Bruin-Molé/Christopher John Müller (eds.), *Palgrave Handbook of Critical Posthumanism*, Palgrave Macmillan, Cham, Online First Publication; https://doi.org/10.1007/978-3-030-42681-1_34-2.
- Loh, Janina (2023, in print): “Are Dating Apps and Sex Robots Feminist Technologies? A Critical-Posthumanist Alternative.” In: Jordi Vallverdú (ed.), *Gender in AI and Robotics. Gender Challenges from an Interdisciplinary Perspective*, Springer.
- Markowitz, Judith A. (ed.) (2015): *Robots that Talk and Listen. Technology and Social Impact*, Berlin, Boston, München: De Gruyter.
- Marr, Bernard (2017): “The 4 Ds Of Robotization: Dull, Dirty, Dangerous and Dear.” In: *Forbes*, October 16, 2017; <https://www.forbes.com/sites/bernardmarr/2017/10/16/the-4-ds-of-robotization-dull-dirty-dangerous-and-dear/?sh=3dfib5a13eod>.
- Marsh, Amy (2010): “Love among the objectum sexuals.” In: *Electronic Journal of Human Sexuality* 13; <http://www.ejhs.org/volume13/ObjSexuals.htm>.
- Misselhorn, Catrin (2009): “Empathy and Dyspathy with Androids: Philosophical, Fictional and (Neuro-) Psychological Perspectives.” In: *Konturen* 2, pp. 101-123.
- Misselhorn, Catrin (2013): “Robots as Moral Agents?” In: Frank Rövekamp/Friederike Bosse (eds.), *Ethics in Science and Society. German and Japanese Views*, München: iudicium, pp. 42-56.
- Mlot, Stephanie (2018): “Sex Robot Samantha Upgraded With Moral Code.” Originally In: *Geek.com*; <https://www.nexusnewsfeed.com/article/science-futures/sex-robot-samantha-upgraded-with-moral-code/>.
- Moor, James H. (2006): “The Nature, Importance, and Difficulty of Machine Ethics.” In: *Intelligent Systems IEEE* 4, pp. 18-21.
- Morgan, Rhian (2015): “Looking for robot love? Here are 5 sexbots you can buy right now.” In: *METRO News*, September 13; <https://metro.co.uk/2017/09/13/looking-for-robot-love-here-are-5-sexbots-you-can-buy-right-now-6891378/>.
- Mori, Masahiro (1970): “Bukimi no tani.” In: *Energy* 7, pp. 33-35; translated by Karl F. MacDorman/Takashi Minato (2005): “On the Uncanny Valley.” In: *Proceedings of the Humanoids-2005 workshop. Views of the Uncanny Valley*, Tsukuba, Japan.

- Nussbaum, Martha: *Frontiers of justice. Disability, nationality, species membership (= The Tanner lectures on human values)*, Cambridge MA: Belknap Press 2007.
- Richardson, Kathleen/Coeckelbergh, Mark/Wakunuma, Kutoma/Billing, Erik/Ziemke, Tom/Gómez, Pablo/Vanderborght, Bram/Belpaeme, Tony (2018): "Robot Enhanced Therapy for Children with Autism (DREAM). A Social Model of Autism." In: *IEEE Technology and Society Magazine* 37, pp. 30-39.
- Scheutz, Michael (2012): "The Inherent Danger of Undirectional Emotional Bonds between Humans and Social Robots." In: Patrick Lin/Keith Abney/George Bekey (eds.), *Robot Ethics. The Ethical and Social Implications of Robotics*, Cambridge, Massachusetts, London: Oxford University Press, pp. 205-221.
- Seibt, Johanna/Noskov, Marco/Schack Andersen, Soren (eds.) (2016): *What Social Robots Can and Should Do. Proceedings of Robophilosophy 2016/TRANSOR 2016*, Amsterdam: IOS Press.
- Shibata, Takanori/Wada, Kazuyoshi (2011): "Robot Therapy. A New Approach for Mental Healthcare of the Elderly – A Mini-Review." In: *Gerontology* 57, pp. 378-386.
- Stasiénko, Jan (2015): "Bizarre marriages. Weddings as a form of legitimization of intimate relations with non-human agents", pp. 80-93; https://www.researchgate.net/publication/308910979_Bizarre_marriages_Weddings_as_a_form_of_legitimization_of_intimate_relations_with_non-human_agents.
- Sullins, John P. (2006): "When Is a Robot a Moral Agent?" In: *International Review of Information Ethics* 6, pp. 23-30.
- Tavani, Herman T. (2018): "Can Social Robots Qualify for Moral Consideration? Reframing the Question about Robot Rights." In: *Information* 9; <https://www.mdpi.com/2078-2489/9/4/73>.
- Terry, Jennifer (2010): "Loving Object." In: *Trans-Humanities* 2, pp. 33-75.
- Turkle, Sherry (2005): "Relational Artifacts/Children/Elders. The Complexities of CyberCompanions." In: *Proceedings of the Cognitive Science Society Workshop on Android Science*, Cambridge MA, pp. 62-73.
- Turkle, Sherry (2007): "Authenticity in the age of digital companions." In: *Interaction Studies* 8, pp. 501-517.
- Turkle, Sherry (2010): "In good company? On the threshold of robotic Companions". In: Yorick Wilks (ed.), *Close Engagements with Artificial Com-*

- panions. Key social, psychological, ethical and design issues, Amsterdam: University of Oxford Press, pp. 3-10.
- Turkle, Sherry (2011): *Alone Together. Why We Expect More from Technology and Less from Each Other*, New York: Basic Books.
- Wada, Kazuyoshi/Shibata, Takanori/Musha, Toshimitsu/Kimura, Shin (2008): "Robot Therapy for Elders Affected by Dementia." In: *IEEE Engineering in Medicine and Biology Magazine* July/August, pp. 53-60.
- Wallach, Wendel/Colin, Allen (2009): *Moral Machines. Teaching Robots Right from Wrong*. Oxford, New York: Oxford University Press.
- Whitby, Blay (2012): "Do You Want a Robot Lover? The Ethics of Caring Technologies." In: Patrick Lin/Keith Abney/George Bekey (eds.), *Robot Ethics. The Ethical and Social Implications of Robotics*, Cambridge, Massachusetts, London: Oxford University Press, pp. 233-247.
- Złotowski, Jakub/Proudfoot, Diane/Yogeeswaran, Kumar/Bartneck, Christoph (2015): "Anthropomorphism. Opportunities and Challenges in Human-Robot Interaction." In: *International Journal of Social Robotics* 7, pp. 347-360.

Part I – Understanding, Defining, Conceptualizing

Virtues, Robots, and Good Lives: Who Cares?

Charles M. Ess

1 Introduction

I begin with a general overview of the primary elements of Virtue Ethics (VE) as a global tradition which in its Western^{*1} developments turns centrally on *phronēsis*, both “practical wisdom” and the capacity for reflective judgment. Starting with Antigone, *phronēsis* grounds what become Western^{*} traditions of civil disobedience and conscientious objection. At the same time, there are strong arguments that *phronēsis* is not computationally tractable (i.e., *phronēsis*’s processes and capabilities cannot be fully replicated by computational technologies and techniques such as Artificial Intelligence (AI)). This will have important consequences for the uptake of VE as an increasingly central framework for Information and Computing Ethics (ICE). This is in part as ICE *qua* global ethics must avoid what I have called “computer-mediated colonization,” i.e., the imposition of Western^{*} norms and values upon the rest of the world. A *pluralistic* VE – as I have developed in terms of a *pro hen* interpretive pluralism – seeks to avoid such colonization as well as the fragmentation and potential violence of a simple ethical relativism: that is, as such relativism abandons efforts to achieve shared norms, practices, etc., cultural differences can be used to justify violence in the name of protecting local identities, norms, practices, etc. against all “Others” who may (appear to) threaten such local identities, etc.² How privacy is understood across diverse

1 Following a convention suggested by Grodzinsky, Miller, and Wolf (2008), the asterisk indicates that this term is deeply fraught, contested, and highly ambiguous – perhaps even no longer meaningful at all in certain ways. There is no space here to explore the possible variations and nuances: the term can be used here only as a shorthand or heuristic. I will also flag *privacy*^{*} in this way for similar reasons.

2 For example, during their visit to Oslo, Norway, in late January, 2022, when confronted by journalists with questions regarding their plans for protecting young girls’ rights to