



HAL
open science

A New Challenge for Machine Ethics Regarding Decision-Making in Manufacturing Systems

Esmail Kondori, Rui Neves-Silva

► **To cite this version:**

Esmail Kondori, Rui Neves-Silva. A New Challenge for Machine Ethics Regarding Decision-Making in Manufacturing Systems. 12th Doctoral Conference on Computing, Electrical and Industrial Systems (DoCEIS), Jul 2021, Costa de Caparica, Portugal. pp.161-172, 10.1007/978-3-030-78288-7_16 . hal-03685938

HAL Id: hal-03685938

<https://inria.hal.science/hal-03685938>

Submitted on 2 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A New Challenge for Machine Ethics Regarding Decision-Making in Manufacturing Systems

Esmaeil Kondori and Rui Neves-Silva

NOVA University of Lisbon, School of Science and Technology, 2829-516, Caparica,
Portugal

kondori@campus.fct.unl.pt, rns@fct.unl.pt

Abstract. In order to deal with increasingly complex manufacturing systems, we need to make sophisticated decisions. A new challenge emerges when dealing with presenting a new decision – making model merged by an off-line (production data including; staffs, machinery, materials) and on-line (sensors, actuators) data to render a shared responsibility of decision’s consequences between machine and human through giving weight to the taken decisions. Undoubtedly, to make an accurate fair prediction, this presented model should follow the ethical rules. However, the mostly past research works about relationship between machine and ethics mainly have focused on human and his responsibility in applying of technology and only humans have engaged in ethical issues. In light of the digital era especially applying AI and Machine learning, necessarily a new approach should be applied to the interplay between the machine, ethics, and human by adding an ethical dimension to those machines which involve with decision making.

Keywords: Sliding decision - making, Machine ethics, Smart manufacturing systems, Artificial Intelligence.

1 Introduction

At the early stage of the fourth industrial revolution, ‘Industrie 4.0’, machines (through current paper machine means ordinary physical machines, autonomous robots, as well as purely algorithmic systems) increasingly penetrate many aspects of our routine life, and gradually are finding their key positions especially in manufacturing systems by making intelligent interactions between physical and autonomous components and expected to be a growing factor in the future superiorities [1]. This means in the age of machines [2], despite all warnings, the interaction between human and machine is entering in the new environment, and systems, with the focus on manufacturing systems, will experience a new integration of humans - automation model via diverse elements such as informatics, robotics, mobile devices and sensors [3]. Simultaneously, the concerns connected to ethical decision-making, which has been challenging issues, significantly has drawn both public and experts’ attention [4], [5]. Since many aspects of machines’ behavior are ethically relevant [6], so in this field, most of attempts have focused, how to design, employ, and then to evaluate a machine which could act

ethically in its performances and also in relation with other agents and humans with minimum cause harm [7].

Despite serious warnings from eminent thinkers about the risk of applying these technologies (because applying automating ethical decision-making operates the risk of taking humans entirely out of the ethical decision-making process), seemingly humans inevitably have been convicted to employ these complex technologies, however, these warnings stand in contrast to the current innovation of technologies [8]. Till now, several works have been executed on machines through complex manufacturing systems [9] and exploring the notions of the use of them, but ethic of machine is still under study. As a pioneer, Kantian [10] proposed his ethical theory by putting the ‘‘ethic’’ into ethical machines.

Dealing with this issue is so crucial, because by clarifying the aspects of machine ethics and then ethical decision making, in a calculated fair approach, we can share the responsibility between human and machine during decision process.

The rest of this paper is structured as follows. Section 2 renders the relationship between our work and the scope of the conference. Section 3 discusses about autonomy and its role in ethical decision making, its importance, and justification of used terminology throughout the paper. Section 3 presents a framework regarding machine ethics, its important, and main challenges in front of it. Section 4 talks about machine responsibility and type of responsibilities. Section 5 present the research approach to this work and Section 6 presents conclusion.

2 Relationship to Technological Innovation for Applied AI Systems

Human’s decision situations have been studied by several research communities such as artificial intelligence, psychology, cognitive science, and decision science [11]. As a pioneer of the field of decision support systems, Herbert Simon made key contributions to enhance our understanding of the decision-making process. He suggested for the first time that the decision-making model of human beings be based on three phases: 1) Intelligence phase, 2) Design phase, and 3) Choice phase [12]. On the other hand, some efforts from other authors led to various models which have been developed to simulate decision behaviors, where [13] classified them into three major categories including: 1) economical approach, 2) psychological approach, and 3) synthetic engineering-based approach.

Among them, some research concentrated simulation-supported planning of manufacturing systems. But this sort of simulation (manufacturing processes) is mainly focused on machinery, facilities and the material which flow into production line [13], [14]. The main weak point of these efforts is that they only consider an operation-based approach [11]. In other words, solely the operative functions are modelled with no attention (or minimum attention) to the role of the staffs’ decisions. In fact, by neglecting the influence of human decision making in manufacturing operations, it changes its role to a passive one [14], [15].

Nowadays, in accordance with emerging new manufacturing paradigm, staffs’ decisions at shop floor level are gradually finding its actual position and considered as an active component during decision situations. Considering the human in an active role [16], there is a huge insight gain about decisions and leads the process towards a measured action as a final target. To satisfy this, we must support the decision situations

by calculated prediction, extracted knowledge, and different scenarios about handling the manufacturing systems.

In this way, by merging both past (off-line) and real-time (on-line) data collected by embedded sensors from manufacturing systems [17], [18]. Getting support from digital twins as a powerful solution which paves the way into a high-level prediction of future events, as a testbed of adopting decision, our machine (AI core) will be fed by collected data then can be trained.

Aligning with the scope of the conference, in this work we are going to propose a decision-making model with the capability of sharing the responsibility of taken decisions ethically. This model will benefit digital twins in order to provide real – time (on-line) data through the installed sensors and at the same time will be supported by past (off-line) data collected from manufacturing system, e.g., production data including staffs, machinery, materials ... to balance the responsibility of decision's consequences between machine and human. In fact, the proposed model through this work is an AI-based decision model as it benefits merged collected data in order to make prediction.

Developing this notion, allows us to do virtual what-if experiments before we try things out for real [14].

3 Autonomy Role in Ethical Decision Making

Autonomy is a vital structure in defining human-machine interaction and broadly changes through machine platforms. Levels of machine autonomy, ranging from teleoperation to fully autonomous systems (e.g. in a fully advanced manufacturing system), have affect the way in which humans and machine interact with each another [19].

According to one of the most famous definition, a machine is a task-oriented device which uses sensors and other information, to be able to physically change its environment, move, and have both the energy and capability to autonomously take decision regarding how to carry out its tasks. In fact, one pivotal characteristic of an autonomous machine is the ability to reply properly to a wide range of situations but based on some defined ethical principles. The main sign of such an ethical agent is demonstrating the autonomy (in different levels as mentioned) in decision making process. Regarding industrial systems, it is expected that future manufacturing systems will still count on collaborations between humans and machine albeit in totally renewed arrangements and in higher levels of autonomy to improve their competitive advantage in order to operate without external interventions [3], [6].

A key component in this new paradigm will be intelligent autonomy and fortified human-machine collaboration.

Entering digital era, and applying new independent advanced machines, many ethical issues are raised because of capability of these machines, in taking independent decision and acting.

However, the most important bottleneck we usually encounter is ethical issues related to design as well as to run of such systems. Whether in practice or in theory, there are many difficulties in front of attributing ethical decision making to machines independently of human intention or will. However, significant improvement in the intelligent automation is promising that in the near future autonomous small-scale

machines and safe large-scale machines will be applied in a structured collaboration framework with humans as part of advanced manufacturing systems [20], [6].

4 What is Machine Ethics?

Nowadays, emerging field of machine ethics has attracted a lot of attentions [10] and many believe, assigning control over vital decisions to machines requires machines do actions in an appropriate ethical way [21].

The destination of machine ethics might be one of the following argued expressions including capability to design machines which are eligible as “artificial moral agents” [22]; which can follow ethical rules; which can make ethical decision; and which have “an ethical aspect [23]. On top of this issue, a fundamental question that must be answered is how a machine achieves at its ethical conclusions during ethical decision-making process [6]. Regarding this and during decision process, two distinctive approaches have been considered: ‘top-down’ and ‘bottom-up’ methods [24]. Top-down approach assumes humans have collected adequate insight on a specific subject so decision making algorithms will be codified by the machine programmer to generate the predictable results: in fact, through this approach, the machine is trained with accordance what a human perceives as an ethical behaviour, which one needs to be determined and the appropriate time to be applied. According to bottom-up approach, the machine designer builds an open-ended system that is capable to gather needed data from its surroundings, to be able to anticipate the outputs of its measures, to opt from among existing alternatives and, more crucially, has the ability to learn from its prior experiences [24].

The final goal of machine ethics, (we believe and, in the following), is that a machine can apply an ideal ethical principle or set of principles [25] to be able to create a shared responsibility of decision’s consequences with human.

4.1 Why Machine Ethic is Important?

Designing then building ethical machines is a practicable, not a theoretical attempt. Acting to design, will ensure that growingly autonomous machines will not a reason of harm to humans and other entities qualified of ethical consideration [26]. Autonomous machines are increasingly taking over human tasks in any levels of decision process in manufacturing systems, while in the beginning, machines were replaced with workers only in assembly lines and workshop floors to carry out uncomplicated and restricted assignments.

With the advent of new technologies especially those which were employed in manufacturing systems, not only simple jobs that were previously carried out purely by human, but complicated ones gradually are being transferred to machines and robots. Due to this notion, Picard [27] says, “the greater the freedom of a machine, the more it will need moral standards.”

For this reason, recently *Machine Ethics* have drawn much attentions and many researchers have been showing a great interest to this issue .

As discussed, in light of digital era, machines in a serious way and increasingly have imposed themselves to human in various domains, from agriculture to industry. Through the interactions, machines take over many types of decisions and activities

which before performed by humans. Therefore, by replacing humans with machines, performing more interactions between this combination, meaningfully will decrease human control over decisions and from another side will lead to an increased number of decisions which are made by machines .

In this sense, the essence of ethics plays a crucial role in decision making to consider substantial problems in sharing of responsibility [28] between machine and human. Clearly, the consequences of taken decisions should reduce the possibility of negative results for humans and/or to mitigate the negative effects machines can cause [21].

Coming back to the topic, why the field of machine ethics important? The following reasons can be brought to answer this question.

First, to abandon this aspect of machine could have serious negative consequences through the decision process especially in manufacturing systems during human – machine interactions. Second, as previously discussed, the main concern is whether these machines will demonstrate ethical behaviour or not. So, the future of machine ethics, especially AI may be at risk [25]. An integral concern connected to AI is different as the huge amount of data must be applied to train AI algorithms. In fact, “AI is fueled by data;” hence, it encounters ethical challenges related to data collection, including consent, IP, and privacy [29], and especially GDPR (General Data Protection Regulation).

4.2 Challenges In front of Machine Ethics

As prior argued, with the advent of new field of machine ethics, the fundamental concern is how to be added an ethical aspect to the machines [25].

By entering to the era of the fourth industrial revolution ‘Industry 4.0’, this concern accelerated as companies try to digitalize the industrial processes, especially those that are driven by AI and intelligent machine learning to make efficient the manufacturing systems, because according to [30], automating ethical decision-making can situate a direct risk to a user’s ethical autonomy.

Due to the traditional mindset in manufacturing systems, there is a historic mistrust proximity to machine on the shop floor among industrial staffs and they psychologically feel a strong segregation in this relation. This segregation has made a big problem in human – machine collaborations and caused to reduce workflow efficiency [31].

In short, as [32] says, the potential downsides to machine ethics include technical safety, transparency and privacy, fundamental values, bias in data, training sets, unemployment / lack of purpose & meaning, growing social unfairness, environmental effects, automating ethics, human rights, effects on the human spirit and etc. As discussed in *Introduction* sector, ethic of machine is still under study. But as an initial definition, the concept of ethics is to bear the consequences of taken decisions in order to maximize the benefits and minimize the risks to a society (manufacturing system) uses the machine. However, it is still a challenging issue to ethical theory to be applied in the mentioned field.

5 Can Machines Be Responsible?

One of the fundamental subjects that must be considered and find a transparent answer

for is human-machine interactions in manufacturing systems, in which concerns are growing; “When a machine makes mistakes, who is responsible?” How can we share the responsibility in human-machine model?

Since in near future machines, according to their capacity, they will be assigned to undertake more tasks especially those are related to human’s safety, health, and his live, therefore human will have to make confidence whether those are appropriately following ethical principles and also consider if there is a risk that could cause harm.

In this regard, namely, to be able to propose a reliable model that can bring us a solution which can solve aforementioned problem, we need to develop new concept based on dynamic models that sensitively reacts to any changes, even a minor, just like the real thing is something new, a digital twin can take this responsibility.

To have accurate predictions or to make an accurate estimation as the basis of robust decision-making process during a manufacturing system, accessibility to a timely real data from real system is crucial. In fact, the models need to be fed by real – time data to be able to help us in making accurate predictions. Since all discussed decision-making models benefiting from static models, consequently it is not expected to access to exact predictions. In this regard, we need to detect the fundamental KPIs, those deeply effect on our proposed decision model (sliding decision making model) through a manufacturing system.

Clearly, by employing digital twins, as a fashion method, this work tries to overcome this pitfall

In the meantime, a few researchers [33] claim that, in the predictable future of technology, “machines will take zero responsibility” for their actions, and humans will be the ultimate responsible for any outcome .

As earlier discussed, some academic reports, for instance [34] believe that it would be threatening if be given a full responsibility to machines for their actions, as they might go out of control even from the programmers [6].

To answer this issue, two prevalent approaches about a machine’s ethics responsibility should be addressed: classical approach and pragmatic approach.

Based on classical approach there is no responsibility to machines for the actions under any situation. This approach views them as mechanical facilities or servants. But in contrast, pragmatic approach assumes ‘artificial ethics’ can be attributed to machines and then machines can be viewed as responsible for their decisions and actions [35].

5.1 Types of Human Ethical Responsibility, can be extended to Machines?

To be able to share ethical responsibility through decision making process in any situation including manufacturing systems between human and machine, firstly we need to know types of responsibilities which are toward human and secondly, whether such types of responsibilities can be extended to machines. Therefore, according to [36], human ethical responsibility can be split into three types: causal responsibility, role responsibility, and liability responsibility .

In causal responsibility, an individual will be responsible for everything caused by him/her, namely his/her actions. In role responsibility, based on a person’s role in a certain area, in short, the task is the responsibility. But liability responsibility identifies who is eligible to be “admired or blamed” for especial measures or conclusions. According to [35], causal responsibility can be dedicated to non-humans e.g., machines.

Since machines come in different varieties then not all types have the ability to bear responsibilities, in fact taking responsibility among them directly depends on how autonomous they are. For example, a thoroughly advanced autonomous machine in a production line can be considered to bear some causal responsibility as it has ability to make decision that leads to outputs in an extensive range of environments while the responsibility of a machine with low autonomy, typically is on the human that controls it [6]. In the meantime, the autonomy level of machine that comes from its training rate, and already defined related manufacturing systems' KPIs, then machine's responsibility during decision process can be determined. In other word, when actions are carried out autonomously by a machine, the consequences of the actions are assessed by the AI-based system. In this regard, through a human - machine interaction, to be able to share the accurate responsibility, first the limitations of a human and the machine should be understood.

Usually under certain conditions, human makes poor predictions. Humans often overweight crucial information and are careless for statistical properties. Machines often show better outcomes than humans in complex interactions where there are different indicators, especially dealing with big data. Hence, based on our proposed decision-making model, obviously it is expected to have a more responsive machine in the aforementioned situations. Using big data, the machine training rate will increase, and it means we will have a machine with a high-level autonomy. As we face more growing dimensions of interaction, the capability of a human to make an accurate prediction will diminish, especially comparison to machines, and it means machine should bear more responsibility than before.

6 Research Approach

Since this paper is a part of a PhD thesis, we expect at the end of this work, combining off-line and on-line data collected during our research, then training the machine (AI core), and applying the collected data in a real case study in a manufacturing system, we can develop our model and propose a decision-making model that ethically balance the responsibility through human – machine interaction.

Regarding to the thesis topic, we are going to answer to a part of thesis main research question in this work.

How to model a decision-making process within a manufacturing system that capable to balance then share a reasoned ethical responsibility between human and machine?

It was expected to be discussed the results of case study, but as this model is in its early development stage, naturally it is being discussed theoretically, however, as one our main target, we are going to apply this model in the next phase of its development on a real case study in a manufacturing system. Thus, we already have put this part as the pillar of our PhD work that will be considered.

6.1 Sliding Decision Making Model

In this section, we would like to propose sliding decision making as a possible solution.

As said, solely having the capacity for ethical performing never guarantee ethically aligned decision making. In this regard, a machine which is trained by both off-line (production data including staffs, machinery, materials) and on-line (installed sensors and actuators) accurate data within a manufacturing system can bring us a strong insight in the modeling of ethical-based decision making.

Dealing with the suggested ethical decision-making model, some steps and features need to be considered. We also will debate how these features of autonomy and choice could have consequences for attributing ethical responsibility to machines.

To be able to enrich our database (by collecting data from the mentioned sources) and in order to train our machine, a 6-stage process is suggested as following:

- 1 Collecting data from installed sensors to provide real-time data
- 2 Collecting data from multi off-line sources, e.g. production line, staffs, equipment, materials, ... to provide past data
- 3 Analyzing the collected data by categorizing and integrating
- 4 Extracting the patterns according to the collected data from its experience
- 5 Enriching the database by gathering huge (big) data and training our machine (AI core) via collected data
- 6 Making decision aligned with ethical principles for doing appropriate prediction.

Completing aforementioned steps, by applying the past data from off-line sources, and also receiving real-time data from embedded sensors (applying digital twins), now the collected data can be analyzed, and machine is capable to detect and extract the repeated behaviors as the structured patterns based on its experiences to take high-quality decisions that are compatible with ethical principles. Through a manufacturing system, these machines will make gradually complicated and take vital decisions related to humans and other agents, so we need to know whether their decisions are reliable and ethically defensible or not. According to its autonomy level, sliding decision making model must assess the ethical implications of their probable proceedings in order to fair sharing the responsibility within human-machine interplay.

To reach this goal in our sliding decision-making model, quantifying the uncertainty and the risk connected to the decision process is vital in determining the amount of responsibility to be transferred to the machine. Applying already detected manufacturing system's KPIs and at the same time by merging them with collected data, now the proposed model has this chance to be trained by appropriate data.

However, designing such machines is not so easy. Hence, we will need to keep the machines responsive about the taken measures and decisions, because the proposed decision-making model needs that the taken decisions be transparent and logically be strong.

To reach to what argued above, generally we should perform two actions :

Firstly, enhance machine ethics in a decision process by taking the two following ways to be able to add an ethical aspect of a machine:

- a. Enhancing the core behavior of the machine by training and in result, increasing its autonomy level, to follow ethical performances
- b. Enhancing the behavior of decision-makers, who uses the machine.

And in the second measure, we should enhance the behavior of machine or human (decision makers) by taking the following measures:

- i) By improving the decisions even at individual level,
- ii) By improving how decision-makers fit within ethical decisions, as a broader social system [36].

6.2 Methodology

In terms of making ethical decisions, sliding decision making model will act based on the autonomy level that the machine is codified. In fact, in its specific autonomy level, the trained machine will benefit from freedom degree assigned to that level to make decision. Therefore, the proposed model will evaluate the implications of taken decisions with accordance to the characteristics of that level, its limitations, and its degree of freedom. So, in our discussion about ethically sharing responsibility between human and machine in a manufacturing system, sliding decision model will act to share the responsibility according to the degree of machine autonomy. More transparent, the more formulated and codified by human and his trace is more visible, the lower autonomy level then the weight of responsibility for those probable mistakes lies squarely with the machines' designer or owner. In contrast, if the measures that are conducted by machine have arrived from its experiences during its training by collected data with minimum human interferences ('bottom-up' approach), means higher autonomy level and clearly the burden of responsibility for the probable mistakes and ethical consequences is more obviously on the machine's shoulders.

Training machines and make them more intelligent means, augmentation in their level of autonomy and independencies. With this, practically we increase their capability in decision making and by advances in their autonomy level, machines will follow ethical principles in sharing and balancing the responsibility with human, as they aware of the consequences of their actions, therefore will take responsibility about already taken decisions without human presence. In fact, as humans does not play a significant role in their decision-making consequences, hence in this situation, based on the proposed decision model, machines have to bear more responsibilities during humans-machine interactions.

7 Conclusion

The future of machine ethics debate has a strong dependency on improvements in both technology and ethical theory from human-machine collaboration point of view.

Once autonomous machines can and will be ethical agents that have capability to make ethical decisions by accepting their responsibility in a balanced human-machine interaction. Here, the fundamental question would be: who will be responsible for the actions performed by autonomous ethical machines? In our proposed decision-making model, i.e. sliding decision making model, the responsibility in decision making has a direct correlation with machine autonomy level. According to this model, three scenarios will be raised:

Low-autonomous machines, Semi-autonomous machines, and High-autonomous machines.

In first scenario, as long as machines perform based on designs and programs

invented by humans, the designer is responsible for any consequences. In second scenario, things will undoubtedly get complicated. What can be said here is to get rid of this risky situation, strongly recommended to be increased the autonomy level of machine to be changed to high-autonomous machine.

In third and last scenario, a trained machine which learns from its experiences while is capable to boost its own decision making, can bring us a higher autonomy level that with this specification, responsibility may shift entirely from human to the machine while is needless of the direct effect of programs, designer, or humans.

All already discussed aforementioned scenarios come back to the machine autonomy levels and its degree of freedom (by quantifying the uncertainty and the risk involved in the decision-making process) the risk in making decisions by sharing responsibility within human-machine interactions.

At the end as previously argued, since the ethics in this work mean taking responsibility regarding made decisions using obtained autonomy, the proposed decision-making model will deal with already collected merged data in our case study in a manufacturing system, and in correspondence with training rate, the proposed model most likely will act aligned with ethical performances, however this phase is still under study and reaching to this, is one of the main goal in our PhD work.

Acknowledgements. This work benefited from partial support from Fundação para a Ciência e Tecnologia through the program UIDB/00066/2020 (CTS – Center of Technology and Systems).

References

1. Sara Lumbreras, The Limits of Machine Ethics, Institute for Research in Technology, Universidad Pontificia Comillas, Madrid 28001, Spain
2. Veruggio, Gianmarco, Fiorella Operto, and George Bekey. Roboethics: Social and ethical implications. In Springer Handbook of Robotics. Berlin and Heidelberg: Springer, pp. 2135–60, (2016)
3. S.R Fletcher and P. Webb, Industrial Robot Ethics: The Challenges of Closer Human Collaboration in Future Manufacturing Systems, In: "A World with Robots: International Conference on Robot Ethics: ICRE 2015", pages 159-169, DOI: 10.1007/978-3-319-46667-5_12 (2017)
4. Yuliang Shi, Chenfei Sun, Qingzhong Li, Lizhen Cui, Han Yu, and Chunyan Miao. A fraud resilient medical insurance claim system. In AAAI, pages 4393–4394, (2016)
5. Yongqing Zheng, Han Yu, Lizhen Cui, ChunyanMiao, Cyril Leung, and Qiang Yang. Smarths: An ai platform
6. Fahad Alaieri and Andr e Vellino, Ethical Decision Making in Robots: Autonomy, Trust and Responsibility, DOI: 10.1007/978-3-319-47437-3_16 (2016)
7. Jason Millar Department of Philosophy, An Ethics Evaluation Tool for Automating Ethical Decision-Making in Robots and Self-Driving Cars, Carleton University, Ottawa, Canada, APPLIED ARTIFICIAL INTELLIGENCE, VOL. 30, NO. 8, 787–809, (2016)
8. Jim Torresen, A Review of Future and Ethical Perspectives of Robotics and AI, Robotics and Intelligent Systems Group, Department of Informatics, University of Oslo, Norway, <https://doi.org/10.3389/frobt.2017.00075>, (2018)
9. Luxton, David D, Recommendations for the ethical use and design of artificial intelligent care

- providers. *Artificial Intelligence in Medicine* 62: 1–10, (2014)
10. Ryan Tonkens, A Challenge for Machine Ethics, *Minds & Machines*, DOI 10.1007/s11023-009-9159-1, (2009)
 11. Robinson, S., et al., Modelling and Improving Human Decision Making with Simulation. Winter Simulation Conference B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, eds (2001)
 12. Rizun, N., Taranenko, Y., Simulation Models of Human Decision-Making Processes. *Management Dynamics in the Knowledge Economy*, Volume 2 no. 2, pp. 241-264; www.managementdynamics.ro (2014)
 13. Kim, N., et al. Human Behavioral Simulation Using Affordance-Based Agent Model. 14th International Conference, HCI International 2011, Orlando, FL, USA (2011)
 14. Zülch, G., Modelling and Simulation of Human Decision -Making in Manufacturing Systems. 1-4244-0501-7/06/, IEEE (2006)
 15. Lin, T., Modelling of Human Decision Making via Direct and Optimization-based Methods for Semi-Autonomous Systems, University of California, Berkeley (2015)
 16. Aerts, D., Sassoli de Bianchi, M., Sozzo, S., Veloz, T.: Modeling Human Decision-making: An Overview of the Brussels Quantum Approach. arXiv:1807.11036v1 [q-bio.NC] (2018)
 17. Szipka, K.,: Modelling and Management of Uncertainty in Production Systems - from Measurement to Decision. KTH Royal Institute of Technology Department of Production Engineering Manufacturing and Metrology Systems (2018)
 18. Hoc, Jean-M., From human – machine interaction to human – machine cooperation. *Journal Ergonomics*, Volume 43 (2010)
 19. Jenay M. Beer, Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction, *Journal of Human-Robot Interaction*, DOI: 10.5898/JHRI.3.2.Beer, (2014)
 20. Iordanis Kavathatzopoulos and Ryoko Asai, Can Machines Make Ethical Decisions?, Department of IT – VII, Uppsala University, Sweden, AIAI 2013, IFIP AICT 412, pp. 693–699, (2013)
 21. Suzanne Tolmeijer, Markus Kneer, Cristina Sarasua, Markus Christen, and Abraham Bernstein, Implementations in Machine Ethics, A Survey, *ACM Comput. Surv.* 53, 6, Article 132 (December 2020), 38 pages. <https://doi.org/10.1145/3419633> (2020)
 22. S. L. Anderson, “Machine metaethics,” in *Machine Ethics*, M. Anderson and S. Anderson, Eds. New York, NY, USA: Cambridge Univ. Press, pp. 21–27 (2011)
 23. M. Anderson and S. L. Anderson, “Guest editors’ introduction: Machine ethics,” *IEEE Intell. Syst.*, vol. 21, no. 4, pp. 10–11, Jul. (2006)
 24. Allen, C., Smit, I., Wallach, W.: Artificial morality: top-down, bottom-up, and hybrid approaches. *Ethics Inf. Technol.* 7(3), 149–155 (2005).
 25. Michael Anderson and Susan Leigh Anderson, *Machine Ethics: Creating an Ethical Intelligent Agent*, , *AI Magazine* Volume 28 Number 4, (2007)
 26. Brent Mittelstadt, *AI Ethics – Too Principled to Fail?*,
 27. Rosalind Wright Picard et al. 1995. *Affective Computing*. MIT Press, Cambridge, MA.
 28. Robert Sparrow, Killer robots. *J. Appl. Philos.* 24, 1 (2007), 62–77 (2007)
 29. M. Taddeo and L. Floridi, "How AI Can Be a Force for Good," *Science*, vol. 361, no. 6404, pp. 751-752, (2018)
 30. Millar, J. 2015. Technology as moral proxy: Autonomy and paternalism by design. *IEEE Technology and Society Magazine* 34 (2):47–55. doi:10.1109/MTS.2015.2425612.
 31. Santis, A., & Siciliano, B. Safety issues for human–robot cooperation in manufacturing systems. *Tools and Perspectives in Virtual Manufacturing*, (2008)
 32. *Artificial Intelligence and Ethics: Sixteen Challenges and Opportunities*, Brian Patrick Green, director of Technology Ethics at the Markkula Center for Applied Ethics.
 33. Asaro, P.: Robots and responsibility from a legal perspective. In: *IEEE International Conference on Robotics and Automation*, Rome, Italy (2007)
 34. Johnson, D.G.: Computer systems: moral entities but not moral agents. *Ethics Inf. Technol.*

- 8(4), 195–204 (2006)
35. Dodig-Crnkovic, G., Persson, D.: Sharing moral responsibility with robots: a pragmatic approach. In: Tenth Scandinavian Conference on Artificial Intelligence, SCAI, 2008, pp. 165–168 (2008)
 36. Jarvik, M.: How to understand moral responsibility? *TRAMES: J. Humanit. Soc. Sci.* 7(3), 147–163 (2003).
 37. A Review of Future and Ethical Perspectives of Robotics and AI, Jim Torresen, *Robot. AI*, doi.org/10.3389/frobt.2017.00075, (2018)
 38. S. L. Anderson, “How machines might help us achieve breakthroughs in ethical theory and inspire us to behave better,” in *Machine Ethics*, M. Anderson and S. Anderson, Eds. New York, NY, USA: Cambridge Univ. Press, pp. 524–530, (2011)