An International Journal in Philosophy, Religion, Politics, and the Arts

Existenz

Volume 17, No 1, Spring 2022

ISSN 1932-1066

Artificial Systems In-Between Humans and Artifacts

From Autonomous Weapons to Affective Computing

Catrin Misselhorn

University of Goettingen, Germany

catrin.misselhorn@uni-goettingen.de

Abstract: In this essay I advance the thesis that artificial systems form a distinct category positioned in-between humans at one end of the spectrum and artifacts at the other end. I argue that these systems are not mere artifacts, but that they need to be considered as agents. They can even be moral agents in a functional sense, although they fall short of full moral agency as it pertains to humans. This view is elaborated with respect to lethal autonomous weapon systems designed and trained to act as moral agents. Particularly regarding such systems, the question arises as to whether decisions about life and death should be left to machines. I discuss three arguments to the effect that such decisions should not even in war be delegated to artificial moral agents. Henceforth the crucial question arises whether and how humans and machines can cooperate effectively. This brings in a second characteristic which is responsible for the special status of artificial systems in-between humans and artifacts: they are relational artifacts capable of entering emotional and social interactions with humans. Artificial systems cannot really be equal participants in social relationships for they do not have the necessary abilities such as consciousness or intentionality, yet they can simulate them well enough to profoundly challenge established social practices of human relationships.

Keywords: Artificial morality; machine ethics; autonomous weapon systems; *jus in bello*; responsibility gap; emotional AI; affective computing; relational artifacts.

Introduction

The trilogy of books that I have written in the philosophy and ethics of artificial intelligence, namely, *Grundfragen der Maschinenethik*, *Künstliche Intelligenz und Empathie*, and *Künstliche Intelligenz – Das Ende der Kunst?* has one theme in common: it explores the special locus that artificial systems occupy in a spectrum that spans from human beings on one end to artifacts on the other end. In *Grundfragen der Maschinenethik*,[1] I pose the question whether artificial systems can be moral agents. I defend the view that in a basic sense at least some artificial systems can be moral agents, but I agree with James Moor that "they would fall short of being full ethical agents" comparable to human beings.[2] In *Künstliche Intelligenz und Empathie* I discuss the special status of devices that display emotional AI in virtue of them being relational artifacts.[3] While they themselves do not

[1] Catrin Misselhorn, *Grundfragen der Maschinenethik*, Ditzingen, DE: Reclam, 2018. [Henceforth cited as *GM*, all translations are mine]

[2] James H. Moor, "The Nature, Importance, and Difficulty of Machine Ethics," *IEEE Intelligent Systems* 21/4 (July/August 2006), 18-21, p. 21. [Henceforth cited as *NID*]

[3] Catrin Misselhorn, *Künstliche Intelligenz und Empathie. Vom Leben mit Emotionserkennung, Sexrobotern & Co,*

Catrin Misselhorn, "Artificial Systems In-Between Humans and Artifacts: From Autonomous Weapons to Affective Computing," *Existenz* 17/1 (2022), 11-18

First published 2-22-2024

have emotions, they nonetheless can arouse emotions in humans, especially empathy, and enter social relations with them. As the title suggests, in *Künstliche Intelligenz – Das Ende der Kunst?* I discuss the domain of art and show why generative AI can perform certain aesthetic decisions without being able to produce art in the proper sense of the term, as AI systems cannot bear aesthetic responsibility for their products.[4]

Machine Ethics and Artificial Morality

Artificial morality starts from the observation that with increasing autonomy and intelligence machines face situations that require moral decisions. The more complex the areas of application of autonomous systems are, the more challenging become the moral decisions they need to take; for example, in care, in autonomous driving, or in autonomous weapon systems. In each one of these areas, fundamental moral decisions are at stake, including decisions regarding the life and death of humans.

These concerns are novel in the history of ethics. Traditionally, the ethics of technology has to do with questions regarding the moral evaluation of the impact of technologies upon human life, for example, it explores the dangers or benefits of using nuclear power. In this type of deliberation, nuclear power is obviously not considered to be a moral agent. It is only with the advances of digital technology, especially artificial intelligence, in recent decades that it has become conceivable that certain machines can be endowed with the capacity to take moral decisions and to act upon them.

The term "artificial morality" is in this context used analogously to the term "artificial intelligence" as referring to a discipline that is concerned with modelling or simulating human cognitive abilities. Correspondingly, artificial morality aims at modelling or simulating human moral decision-making and agency. Whereas ethics provides a theory of human morality, machine ethics concerns itself with the theoretical and ethical framework for thinking about artificial morality, including its impact on the individual and social life of humans, as well as with the question whether, from a moral point of view, the development of artificial systems with moral capacities is desirable. In contrast to traditional normative ethics which is providing moral standards for human moral behavior, machine ethics is not just developing moral standards, it is about implementing these standards in machines. Machine ethics is therefore an interdisciplinary endeavor that involves philosophy as well as computer science and robotics with the aim of designing so-called artificial moral agents.

As I have argued, machine ethics presupposes a gradual view of agency (*GM* 87-90). I distinguish various types of agents along two dimensions which are taken from philosophical action theory, namely their degree of intelligence and their degree of autonomy. Autonomy should in this context not be understood in the Kantian sense as setting oneself ends but in a more technical sense as operating independently of external intervention. For example, a robotic vacuum cleaner is expected to move around in the house and do its work independently of permanent human supervision and control.

The gradual view of agency includes various types of animals, humans, group agents but also some kinds of artificial systems. According to this view a chess program is an artificial agent, although it is not a moral agent. Such a program can distinguish the information relevant to the game of chess, process it, and take decisions with the objective of winning the game. It is taking into consideration the current position of the pieces on the board and can determine which moves are permissible in a game situation. On this basis, it calculates which alternative is most promising under the given circumstances.

This example arguably refutes an important objection against the assumption that there can be artificial agents. This objection holds that it is not the machine that takes the decisions, but the humans who programmed it. Against this objection one can argue that the greater the advances of artificial intelligence research are, the less can the decisions of artificial systems in a particular situation be attributed to its human programmers or users. Even in the case of a comparably simple chess program, the idea is inadequate that the developers could directly determine the move that system is going to choose in a specific situation.

Support for this point provides the fact that such a program plays chess far better than its programmers, who could certainly not compete with a chess world champion. One can therefore regard

Ditzingen, DE: Reclam, 2021. [Henceforth cited as *KIE*, all translations are mine]

[4] Catrin Misselhorn, *Künstliche Intelligenz – Das Ende der Kunst?*, Ditzingen, DE: Reclam, 2023.

a chess program as being an artificial agent, insofar as its behavior in a specific situation is goal directed and intelligent given that the things that the device does are suitable to reach its objective. Moreover, its behavior is beyond prediction and control of the humans who programmed and use it. One might even say that playing against a chess program would lose its point if one would not ascribe certain agential qualities to the artificial opponent.

This is of course an epistemic point since the way the program acts is as a matter of fact determined by the software although it is not predictable or controllable by the programmers or users once it is running. However, one can argue that the same holds true for human beings in a deterministic world. Still, even if human behavior were fully determined by the laws of nature, it would be beyond prediction and control, at least as things stand today. Determinism is not a reason to deny human beings the status of agents although there might be other arguments against determinism. And one needs to keep in mind that indeterminism does not improve the situation since it only brings in an element of chance which seems to be as problematic for agency as determinism.

The lessons that were drawn from looking at the chess program can be transferred to moral decision-making. The aim of machine ethics is to design so-called explicit moral agents. The actions of explicit moral agents do not only have to conform to what is morally permissible or imperative, they must be the result of moral information processing. The intermediate status of explicit moral agents is again demonstrated by the fact that their moral capacities fall in-between Immanuel Kant's distinction between acting according to duty and acting out of duty. This does not mean that artificial agents can be moral in the strict Kantian sense. Kant would reject this claim not least because they possess neither will nor inclinations that could get in conflict with the moral law.

Explicit moral agents are situated in-between moral subjects in the Kantian sense, who act out of duty, and Kant's example of the prudent merchant whose self-interest happens to coincide with moral duty.[5] In contrast to the prudent merchant, an explicit moral agent must be able to recognize and process morally relevant aspects of a situation as such, and to respond to them in a morally appropriate way. This is what is meant by the requirement that they not only act in accordance with moral duty, but because of moral information processing which results in the application of the relevant moral norm to the situation in question.

Explicit moral agents still fall short of the level of moral agency that humans possess. Full moral agents have further capacities such as consciousness, intentionality, and the capacity for critical reflection and moral deliberation that is to date present only in humans. It is questionable whether machines can ever achieve these capacities and in the foreseeable future machines will be moral agents at most in a functional sense. This sense of functional morality can be defined in terms of information processing. The idea behind this view is that at least some aspects of moral behavior can be captured in terms of information processing in the same way as artificial agency in the context of playing chess.

Functional moral agents are subject to various constraints. First, functional relations refer only to the cognitive aspect of morality. The emotional dimension is only captured insofar as emotions can be functionally modeled independently of their phenomenal quality. Artificial moral agents are so far neither capable to feel compassion for others nor the nagging guilt that torments persons after a moral wrongdoing.

Second, functional relations can be modeled in various ways, as Ned Block has shown, for instance, in terms of relations of meaning, neural connections or by machine-table states.[6] Functional analysis is relative to the type of functionalism on which it is based. The ascription of functional moral agency thus depends on certain human purposes and interpretations, whereas full moral agents possess this status intrinsically.

In *Künstliche Intelligenz und Empathie* I discuss an example of a chatbot that is supposed to recognize when someone is in a mental crisis and subsequently offers help [*KIE* 69-70). I consider this task to be a basic form of functional moral agency. However, if one day human life on Earth were to be extinct while the chatbot is working unperturbed, it would no longer

[5] Immanuel Kant, "Groundwork of the Metaphysics of Morals (1785)," transl. Mary J. Gregor, in *Immanuel Kant, Practical Philosophy*, ed. Mary J. Gregor, Cambridge, UK: Cambridge University Press 1996, pp. 37-108, here pp. 52-3, Ak 4:397.

[6] Ned Block, "Troubles with Functionalism," in *Consciousness, Function, and Representation: Collected Papers, Volume 1*, Cambridge, MA: The MIT Press 2007, pp. 63–101, here pp. 66-70.

have the status of being an artificial moral agent, since there would be no one left to attribute this status to the device, and to interpret the pixels on the display accordingly. In contrast to it, even the last surviving human being would still be a full moral agent.

In *Grundfragen der Maschinenethik* I argue that even if artificial moral agents can be ascribed functional moral agency, this does not imply that they are also morally responsible for their deeds (*GM* 126-8). Since they lack capacities such as consciousness, freedom of will, and intentionality which are required for full moral agency, they do not fulfill the conditions needed for ascribing moral responsibility. I believe that agency without responsibility is the central mark of AI systems and that the decoupling of moral agency and responsibility has fundamental ethical consequences.

Artificial Moral Agents in War

Lethal autonomous weapon systems (LAWS) are enabled to select and attack their targets without human intervention. Machine ethics comes into play to increase the probability that lethal autonomous weapon systems comply with legal and moral norms. Among others, Ronald Arkin has set himself the task to develop an ethics module for a lethal autonomous weapon system. It is programmed to take moral and legal decisions in war situations autonomously.[7]

Guided by international agreements and treaties such as, for example, the Geneva Convention, Arkin's approach is inspired by traditional just war theory, especially the *jus in bello* regulations that outline how war should be fought, once it had begun. His system implements four principles:

* Discrimination: Distinguish legitimate targets from non-legitimate targets. The first category includes combatants and military targets; the second category covers non-combatants and non-military or protected objects.
* Military necessity: Only attack legitimate targets when an attack promises some military advantage.
* Humanity or Unnecessary Suffering: Minimize unnecessary suffering and incidental injury to people, and collateral damage.
* Proportionality: Means must be proportionate to their purposes.

Additionally, Arkin's technical architecture includes a responsibility advisor which is supposed to guarantee that a human being always takes the responsibility for the operations performed by the system (*GLB* 143-54). The advisor informs human operators ahead of time that they are morally responsible for the operation and explicitly asks them to confirm their acceptance of that responsibility by submitting their name and duty number.

The responsibility advisor also updates human operators on changes made to the system and once again asks them to explicitly accept responsibility provided someone wants to stop an operation selected by the system. An intervention by one person suffices to overrule a decision to kill. But removing the system's ethical barrier and ordering a killing operation not in line with the system's decision requires authorization by an identified legitimate second person.

Arkin's main argument for designing lethal autonomous weapon system as artificial moral agents of this type is that a war without people is for him a more humane war. He thinks that lethal autonomous weapon system will be better able to comply with the *jus in bello* regulations, thereby saving lives, particularly civilian lives, and helping to avoid war atrocities.

Three Ethical Objections against Artificial Moral Agents in War

Against Arkin's positive assessment can be objected that lethal autonomous weapon systems might not just make the decision to go to war easier, but that it is intrinsically wrong from a moral point of view to use them. In contrast to pacifism, which strictly condemns the use of violence and the killing of people in war, I am accepting just war theory for the sake of argument. That is, I assume that there is a domain-specific normative ethics that considers the use of violence and the killing of people in war to be morally permissible (or at least excusable) under certain circumstances. Even if these assumptions were true, there are three fundamental moral objections against lethal autonomous weapon system from a moral point of view. These objections are: the argument from responsibility that goes back to Robert Sparrow, the argument from human agency by Alex Leveringhaus, and the argument from moral duty that I developed myself.

[7] Ronald C. Arkin, *Governing Lethal Behavior in Autonomous Robots*, Boca Raton, FL: Chapman & Hall CRC, 2009. [Henceforth cited as *GLB*]

The Responsibility Argument

One important objection against lethal autonomous weapon system is that they make the attribution of responsibility for acts of killing in war impossible. Robert Sparrow, who elaborated on this argument, describes the emergence of responsibility gaps. For him, an act of killing in war is only morally permissible if it fulfills the criteria of *jus in bello*, and there is someone who bears responsibility for the act.[8] Autonomous weapons systems may perhaps comply with *jus in bello* criteria, and arguably do so even better than humans do. Nevertheless, for Sparrow these weapons are morally prohibited if no one can be held responsible for their actions.

The crucial issue is exactly who is responsible for the killings done by lethal autonomous weapon system. Sparrow says that ultimately no one can be held morally responsible for the behavior of such systems, neither the programmers, nor the commanders, nor the operators, not to speak of the machine itself (*KR* 69-71). This is not just a case of collective responsibility where the responsibility is distributed across many persons, so that it might get minimalized to the point of disappearing at the extreme. Although this is a very fundamental problem of war which arises because of it being always a collective undertaking, the argument regarding the responsibility gap does not reduce to the problem of collective responsibility.

A responsibility gap arises if a lethal autonomous weapon system violates the *jus in bello* provisions, even though (a) it was not intentionally programmed or manipulated to violate the ethical or legal norms of warfare; (b) it was not foreseeable that it would do so; and (c) there was no human control over the machine from the start of the operation (*KR* 73). The point of the responsibility argument is that one cannot reasonably ascribe responsibility for the acts of an autonomous weapon system to anyone under these conditions, for the criteria for attributing responsibility are neither met by the humans involved nor by the machine itself. There is no intention, no awareness of the consequences, and no control. But if no one bears responsibility for the act of killing of such a system, then it is morally not permissible to deploy such machines.

However, there seems to be a simple rebuttal to this argument from Arkin's point of view. Since his approach involves a responsibility advisor, it seems that there is always a human being who takes on responsibility for the killings of the lethal autonomous weapon system. Responsibility lies ultimately in the hands of the operator who can, in principle, also revise the machine's course or, if necessary, stop the operation altogether. Arkin's system is designed to be constantly supervised by a person who may intervene. A user may recognize violations of *jus in bello* and can prevent the machine's lethal action. The system is under human control even after the start of the operation. Consequently, there does not appear to be a responsibility gap at first glance.

However, a second glance reveals problems with ascribing responsibility in this manner. It seems unfair that the operator must take full responsibility for the system's actions, whereas the programmers get away with it. At least part of the responsibility should thus lie with them, whose algorithms are decisive for the system's behavior. The operator is responsible only in the sense of not having properly supervised the system and prevented it from doing something inappropriate.

Moreover, to the end of taking a decision in a situation, a human supervisor must rely on information presumably provided by the system, and without having access to independent data the supervisor might not be able to correct the system. Furthermore, given the fact that a system is subject to several quality control procedures during development, this might convince its user that the system's decisions will be superior to one's own.

While effective human control of a system is theoretically possible, it might not be realistic in practice. It is psychologically seen unlikely that a person can perform incessant monitoring, that is, to be attentive for long periods of time and then be ready to take decisions instantly and intervene within seconds if needed. These points raise doubts that moral responsibility can be delegated by the push of a button as Arkin suggests. Even if the user has nominally accepted the responsibility by using the responsibility advisor, the user may actually not be morally responsible unless the conditions for responsibility ascriptions are fulfilled.

The Argument from Moral Agency

The second fundamental argument discussed here against lethal autonomous weapon system goes back to Alex Leveringhaus. This argument is stressing

[8] Robert Sparrow, "Killer Robots," *Journal of Applied Philosophy* 24/1 (February 2007), 62-77, here p. 67. [Henceforth cited as *KR*]

the ethical importance of human agency.[9] It is morally questionable, Leveringhaus says, to leave the decision and execution of a lethal action up to a machine. His reason is that the capacity for guilt, compassion, and mercy enables people to refrain from lethal action sometimes. Borrowing Michael Walzer's naked soldier example, Leveringhaus agrees that soldiers can find it morally inappropriate to kill naked soldiers, despite them being legitimate targets according to jus in bello, because they do not pose a direct threat (*EAW* 92-3). Their vulnerability makes them appear primarily as fellow humans and not as enemies. In contrast, a machine would kill them without hesitation. From a moral point of view, in war, as in all matters of life and death, this human ability to act otherwise has intrinsic value according to Leveringhaus. He rejects lethal autonomous weapon systems as being morally bad.

A fundamental objection against this argument is that the human capacity to act otherwise might not be morally valuable under all circumstances. Leveringhaus brings the example of a bank robber who is taking a hostage. He does not consider it to be morally appropriate to let the hostage-taker get away just because one might feel compassion with the villain. Leveringhaus analyzes this scenario as a case of self-defense that is extended to another person (*EAW* 114-6).

However, there are doubts that being confronted with a hostage-taking bank robber really is a case of self-defense. There is far greater moral pressure to kill the captor if this is the only way of freeing the hostage rather than a situation where this were the only possibility to protect one's own life. To be sure, one may arguably kill the villain to save one's life, but one is not morally obligated to do so. It seems to be morally permissible not to kill somebody in self-defense, as long as no one else is harmed. Not so clear is, whether it is morally equally right to let the captor get away with the hostage taking.

This objection becomes even stronger if one considers cases in which it would not just be morally permissible to kill somebody, but morally obligatory. If there were a moral duty to kill in certain situations, the argument from human agency would no longer be conclusive. It is never morally adequate to deviate from a strict moral duty. The capacity to act otherwise would arguably not have intrinsic value in such situations. The argument from human agency would lose its force as a moral objection to lethal autonomous weapon systems if there were a moral duty to kill in war.

The Argument from Moral Duty

In *Grundfragen der Maschinenethik* I support the argument that there is no moral duty to kill in war (*GM* 180-4). This is the decisive argument against lethal autonomous weapon systems that brings out the grain of truth in the other two arguments. The argument from moral duty does not contradict the other two, but provides grounding for them. If there were indeed a moral duty in certain situations, then the argument of human agency would not be applicable to such cases. Compassion and mercy would be misguided, and it is at least questionable whether responsibility would be an issue in such a situation. In the case of war, the question becomes whether soldiers in war have a moral duty to kill or whether killing is merely morally permissible or excusable.

Arkin seems to believe that there is such a duty. An autonomous weapon may only initiate an act of killing in cases where there is a moral obligation to do so and it is not merely morally permissible (*GLB* 96). The argument of human agency would be rejected if there were such a duty. Clearly, a moral duty to kill does not immediately arise from the law of war. Rather, Arkin's point seems to be that such a duty results if an act of killing in war is morally permissible and militarily necessary. The crucial question is, hence, whether military necessity alone is sufficient to turn moral permissibility into moral obligation. Military necessity is said to be laid down in the Rules of Engagement. Arkin's example is a situation in which it is a military necessity to attack enemy convoys (*GLB* 193).

Of course, such a necessity is anything but strictly defined. Whether a convoy is to be attacked depends, among other things, on whether there is a sufficient probability of success. Yet, firing on a single convoy will hardly decide the outcome of an entire war. According to Arkin, there are even five degrees of necessity. The evaluation of an act of killing puts the respective degree of necessity, for instance, in relation to the expected damage to non-combatants

[9] Alex Leveringhaus, *Ethics and Autonomous Weapons*, Oxford, UK: Palgrave Macmillan 2016, pp. 90-117. [Henceforth cited as *EAW*]

and civilian targets. It is difficult to derive a moral duty to kill from such a gradable view of military necessity since this duty would consequently have to be gradable, too. This contradicts the strictly necessary and unconditional character of duties.

Since there is arguably no moral duty to kill in war, the argument from moral agency remains compelling. The argument from moral duty needs to be seen not as an alternative to the arguments from responsibility and moral agency. It can rather provide a deeper grounding of these arguments. If the argument from moral duty is cogent and there is no moral obligation to kill in war, the argument from moral agency would hold firm.

The argument from moral duty also provides a justification for the claim that someone must take responsibility for killing in war. If an act of killing takes place in a specific situation that makes the act morally permissible then it may be done, but it need not be done. Hence, someone must decide whether to execute it and this person must take responsibility for this decision which was not compulsive. The consequence of all three arguments, the argument of moral duty being the most fundamental one, is that even in war it is morally wrong to delegate the decision regarding human life and death to machines.

One central question for machine ethics becomes, therefore, what it means to and how to ensure that humans exercise meaningful control over the decisions of artificial systems. The task is to develop viable ethical approaches for the cooperation between humans and machines. Since social cooperation involves emotions and empathy, particularly when it comes to morally significant matters, I argue that the perspective regarding machine ethics needs to be broadened to integrate emotional AI.

From Artificial Moral Agents to Affective Computing

In *Grundfragen der Maschinenethik* I explored the special status of artificial agents being placed in-between mere artifacts and human moral agents, and in *Künstliche Intelligenz und Empathie* I turn to the relational nature of artificial intelligence and robotic systems in terms of their emotional and social capacities. In contrast to other types of artifacts, some artificial systems can execute emotional and social interactions with humans without having emotions or other socially relevant mental states.

The concept "relational artifact" was coined in 2001 by Sherry Turkle in the context of a research proposal to the National Science Foundation to the end of labelling machines that simulate an emotional or social interaction with humans;[10] while in 2008 Christopher Scholtz called them "subject simulating machines."[11] Although emotions might also be projected one-sidedly from humans to inanimate objects such as dolls, teddy bears, or cars, relational objects bear a special status. Besides interacting with the users in physical space, they are also proactive and show a variety of behavioral patterns that seem to display internal mental states, and sometimes they even articulate needs and react to the response elicited from the users. It is precisely this set of characteristics that renders relational artifacts more attractive than dolls or toys.

Although relational artifacts are at first glance designed to fulfill the needs and expectations of the users perfectly, they fall short of being genuine partners in life. What is lacking is the dimension of mutual recognition which is crucial for intimate relationships. It is not a genuine relationship merely to project feelings onto an otherwise insentient artifact. Beyond empathy in the narrow sense of the term, it is important in close human relationships that there is someone who perceives who I am, who recognizes and cares for my needs, feelings, and thoughts and *vice versa*. If one takes away recognition by another subject, then one's counterpart is being treated as a mere object, and, so I argue, one also objectifies oneself in that process. One does not recognize one's own feelings, suffering, and thoughts as worthy of consideration by someone who is in fact able to understand them, if one expresses them to be unheard by the deaf ears of an artificial system.

Conclusion

In this essay I developed a thematic thread that spans through my trilogy of books. I argue that artificial

[10] Sherry Turkle, *Relational Artifacts*. Final Report on Proposal to the National Science Foundation SES-01115668, 2004.

[11] Christopher P. Scholtz, *Alltag mit künstlichen Wesen: Theologische Implikationen eines Lebens mit subjektsimulierenden Maschinen am Beispiel des Unterhaltungsroboters Aibo*, Göttingen, DE: Vandenhoeck & Ruprecht 2008, p. 18.

intelligence is fundamentally different from other technologies, for even though it is not on a par with humans, it is getting closer to them and occupies an intermediate space in-between inanimate objects and humans. This intermediate position is decisive for the opportunities as well as for the risks associated with artificial intelligence. It holds for the agential qualities of artificial systems which open-up the possibility that unexpected and surprising things can happen, which may bring a positive effect but also morally questionable consequences, particularly since artificial systems are not responsible for their actions. The same holds for the status of artificial systems as relational artifacts which renders them emotionally engaging but also have ethically problematic implications.