Ethics and Information Technology
https://doi.org/10.1007/s10676-020-09570-y

ORIGINAL PAPER

Critically engaging the ethics of AI for a global audience

Samuel T. Segun[1]

© Springer Nature B.V. 2020

Abstract
This article introduces readers to the special issue on *Selected Issues in the Ethics of Artificial Intelligence.* In this paper, I make a case for a wider outlook on the ethics of AI. So far, much of the engagements with the subject have come from Euro-American scholars with obvious influences from Western epistemic traditions. I demonstrate that socio-cultural features influence our conceptions of ethics and in this case the ethics of AI. The goal of this special issue is to entertain more diverse views, particularly those from Africa; it brings together six articles addressing pertinent issues in the ethics of AI. These articles address topics around artificial moral agency, patiency, personhood, social robotics, and the principle of explicability. These works offer unique contributions for and from an African perspective. I contend that a wider engagement with the ethics of AI is worthwhile as we anticipate a global deployment of artificial intelligence systems.

Keywords Ethics of AI · Africa · Artificial intelligence · Afro-ethics · Utilitarianism · Kantianism

Introduction

With a focus on selected issues in the ethics of AI, I contend that one of such issues, and a very pressing one at that, is the marginalisation of non-western knowledge systems in the study of AI ethics. As our world becomes more interconnected, it is evident that in every corner of world knowledge systems exist that may differ, sometimes significantly, from the predominant epistemic tradition. Unfortunately, little has been done to bridge this gap or take into consideration culture and context thereby perpetuating epistemic injustice. This injustice becomes palpable when we see discourses around digital and AI ethics skewed toward presenting Western ideals, problems, and solutions as prima facie disposition of the field. Although there is some basis to this as much of the advancements made in the field of AI have come from the West. However, with a possible global adoption of this technology, it becomes expedient to have a wider representation of ethics that accounts for diverse ethos and contexts.

Value systems differ across cultures. Birhane (2020) notes that "…Certain matters that are considered critical problems in some societies may not be considered so in other societies" (395). For this reason, an intercultural approach to the ethics of AI should inform the formation of policies and guidelines to regulate the design and use of AI. To this end, there have been recent calls for a more inclusive and intercultural look on ethics for its use in artificial intelligence systems (AIS). This is because of the growing awareness of the need for a global perspective on ethics if we intend to deploy AIS globally. One such call was made by the IEEE's Global Initiative on Ethics of Autonomous Intelligent Systems. In its call, it suggested exploring "…established ethics systems, addressing both scientific and religious approaches, including secular philosophical traditions such as utilitarianism, virtue ethics, and deontological ethics and religious and-culture-based ethical systems arising from Buddhism, Confucianism, African Ubuntu traditions, and Japanese Shinto influences toward an address of human morality in the digital age" (2017, p. 193). Another is UNESCO's (2019) decision to work on a recommendation for a global instrument on the ethics of AI, which would also serve as guidelines for practitioners, governments and policy-makers.

This issue offers a road map for carrying out scholastic intercultural dialogue in the ethics of AI. A refusal to have a broader conception of the ethics of AI would amount to what Fricker (2007) calls testimonial injustice,[1] the tendency to

✉ Samuel T. Segun
ssegun@uj.ac.za

[1] Department of Philosophy, University of Johannesburg, Johannesburg, South Africa

[1] With testimonial injustice, we attribute less credibility to a proposition, opinion, statement or knowledge systems on grounds of prejudice about the speaker's gender, race, ethnicity, sexuality or accent, etc. The harm caused by testimonial injustice is that it deprives the

Published online: 27 November 2020

attribute less credibility to an idea, proposition or account based on prejudices. By offering comparative engagements around critical issues in the ethics of AI, we get one step closer to having a more balanced and robust conception of what should constitute policies, and guidelines for the development of AI systems that are sensitive to human values.

This paper is divided into three sections. In the first, I make a justification for why we need a diverse approach to the ethics of AI, showing a significant disparity in the conception of values, ethos, and moral judgement among cultures. In the second, I offer a roadmap on how we can better engage with the ethics of AI and the development of guidelines that have far-reaching effects. In the third section, I introduce readers to the six timely articles in this special issue.

Why context matters: globally engaging the ethics of AI

Demonstrating the disparity among Western conceptions of ethics and those of other cultures such as African, South-East Asian, Middle Eastern or South American is a daunting task this paper cannot sufficiently cover; instead, I show the glaring differences in value systems, using common themes in Western, African and Chinese ethical systems. I reckon that attaching geographical labels to social, cultural and ethical features is a very problematic way of ascribing homogeneity to a group as not all peoples in that geographical space share those features or values. However, there is a way to construe geographical labels around predominant features of groups. In speaking of African moral theories, Metz[2] (2007, 2017) argues that there are recurrent salient features that can be found in many sub-Saharan cultures that are not found (in the same way) elsewhere in the world. This does not mean these features cannot be found in other cultures, it just means that they are more recurrent in Africa. The same can be said of Western, Middle Eastern and South-East Asian cultures.

In many ways, African ethics, sometimes dubbed "Afro-ethics", as opposed to Western ethics, is collectivist. A similar notion is held about Confucian ethics, a Chinese ethical system. In general, both Afro-ethical and Confucian ethical systems share similarities as collectivists systems and their normative principles, especially construal of what makes a right or wrong action, rest heavily on a collectivist disposition. Unlike Western ideals, which are predominantly built on advancing individualism, Afro-ethical and Confucian moral values share principles that advance collective progress, harmony and group cohesion.

What does all this mean for the ethics of AI? The answer is simple. It profoundly shapes the nature of contributions to the field on issues of data privacy, social robotics, conceptions of artificial moral agency, moral status and patiency, autonomous weapons systems, big data and the likes. Research shows that cultures that are uniquely individualist emphasized individual rights and autonomy while cultures that emphasized relationships and harmony were often collectivist in disposition (Hofstede 2001). This means that decision-making, negotiations and conflict management are often based on group needs to maintain group cohesion rather than individual needs or preferences. The most recent evidence to buttress this point is MIT's Moral Machine project, which significantly highlighted ethical preferences across cultural and geographical spectrums.[3]

In the moral machine experiment, the researchers gathered data from millions of respondents in about 233 countries and territories, who all made a total of 40 million ethical decisions. The paper by Awad et al. (2018) showed a substantial variation in ethical preferences across cultural lines, which, as I have argued above, correlates with geographical mappings, highlighting individualist and collectivist preferences. There were consistent disparities between collectivist ethics in Asian, Middle Eastern and African nations and individualist ethics like you would find among Euro-American nations. Respondents from collectivist cultures showed little care about saving high-profiled individuals in this experiment when compared to Euro-American respondents. This could be attributed to the disparity between individualist and collectivist cultures mentioned above. In the former, the emphasis is often placed on the value of each individual as compared to promoting group interest. So, it is very easy to understand why the value of

Footnote 1 (continued)

speaker or knower a level-playing field to express their ideas, which many times are embraced if another speaker expresses them but does not share the conditions that made the prejudices appear in the first place. Examples abound in history where knowledge systems of indigenous people are jettisoned only to be celebrated when appropriated by European or North American anthropologists.

[2] Metz argues that there are unique interpretations of ethics found amongst the people of sub-Saharan Africa. This does not mean that they lay exclusive claims to these types of ideas. Rather, "… it means merely that certain properties have been recurrent amongst many of those societies for a long span of time in a way they have tended not to be elsewhere around the globe" (2017, p. 62).

[3] The moral machine experiment used a trolley problem-like scenario to gather responses on a variety of ethical decisions. The overall results of the survey showed a few shared principles among respondents regardless of cultural-influenced ethical preferences, such as choosing to save many over few; however, it was clear that this was done in varying degrees across geographical mapping. For instance, among South-East Asians (Japan, China) and Middle Eastern (Saudi Arabia) respondents, and unlike Western respondents, the preference to save younger characters in the scenario were less distinct. This correlates with the notion of respect for elders found among people of collectivist cultures.

an individual would be factored into the decision-making process of a person from an individualist culture.

In individualist societies, the ties between individuals and groups are loose with everyone looking after themselves and their immediate families. On the other hand, a collectivist society ensures that from birth people are integrated into cohesive in-groups that prioritise harmony, loyalty and mutual respect (Realo 1998). From anthropological studies, Hofstede (2011) suggests that much of Western societies are considered individualist and African, Latin American and Asian cultures collectivist. Collectivist cultures like those found in Africa and South-East Asia, are built on the notion of strong ties with the family and society such that it greatly influences the framing of moral obligations and values. Western values on the other hand, and by "Western" I mean Euro-American societies, are often built on individualism which promotes different sets of values that are characteristically distinct from those in a collectivist society.

The close ties between individuals and their communities in a collectivist society make guidelines to issues like privacy differ significantly from approaches adopted in the West. Among Western cultures, great emphasis is placed on data privacy and its protection. The EU's General Data Protection Regulation enacted in 2016, which extensively covers aspects of the use, transfer and protection of individual data is a good example. In contrast, in some collectivist cultures, for example in China, the concept of privacy takes a whole new meaning. Rather than see privacy in a positive light, as a way to guarantee the protection and respect of individualism by regulating an individual's interaction with the world, privacy is often dismissed and given a negative connotation (McDougall and Hansson 2002). Tam (2018) notes that privacy is considered "si yen", which translates to mean seclusion or secrecy.

As interaction with the West continues, China's conception of privacy has evolved. Even though there are new conceptions of privacy that aligns with Western individualism, it is evident that traces of the traditional conception of privacy still surround much of the conversations on data privacy and individual freedoms. Yao-Huai notes that "…at least in the relevant discussion of many contemporary Chinese scholars, the concept of privacy is no longer been limited to the earlier, narrow sense of, but now includes all personal information (i.e., whether shameful or not) that people do not want others to know" (2005, p. 8). Taking into consideration these polarised views on privacy, it behoves international organisations and corporations seeking to extend privacy laws to consider contextual peculiarities.

Another particular theme that runs through all conceptions of collectivist ethics is relationships. In some variations of collectivist ethics, at least in Confucian (Feuchtwang 2016) and Ubuntu (Shutte 2001) ethics, promoting communal relations is considered the highest good. One conception of this form of relational ethics dictates that advancing communal relations is morally good in itself since it shows respect for others based on their capacity to be human. The major proponent of this version of relational morality, and a scholar in Afro-ethics, is Thaddeus Metz. In his seminal work "Toward an African Moral Theory", Metz (2007) insists that there is a very appealing interpretation of African ethics which takes a relational framework. Hence, "An act is right if it prizes other persons in virtue of their natural capacity to relate harmoniously; otherwise, an act is wrong, and especially insofar as it prizes discordance (2016, p. 178). From this interpretation, communal relations are good in themselves not because they help us actualise ourselves or maintain some sort of peace among members of society, even though these things are seen as appurtenances. This construal offers us the platform to ground individual rights, a vital matter in the ethics of AI.

The overall nature of a relational model of ethics renders it characteristically different from Kantian and utilitarian views is its appeal to collectivist values. The moral proclivity of an agent following a relational model is to ensure the good of others and having the community take an important place in the decision-making matrix. On the other hand, Kantian and utilitarian ethics puts the agent at the centre of all decisions, making them individualistic ethics. In addition, the expected object of contemplation for the Kantian and the utilitarian are individuals. An example comes in handy here. Suppose we are to make an ethical decision on whether we ought to lie to achieve a certain goal, say get a promotion at work. We can appeal to Kantian ethics through the categorical imperative by one, contemplating on what a universalisation of that act would mean if others had to lie to achieve their goals (Korsgaard 1996). Two, we will consider if lying in this instance is using others as a means to achieve an end (O'Neill 1975). For the utilitarian, if lying would help us maximise utility while minimising pain, then we are making the right decision to lie and get promoted. Following the relational model of Confucian or African ethics, the consideration would be if lying would affect relations with others. If we had to lie to get promoted, we are likely to be considered deceitful, dishonest, mendacious and untruthful because we have used others as a means to achieve our selfish ends. Remember that in the Metzian interpretation, we consider an act wrong if it promotes selfish ends and causes discordance. In this case, the object of our attention isn't whether the action can be universalised or whether it helps us maximise utility, instead, our focus is on how it affects communal relations.

Another good example to show how these ethical systems can be dichotomised is to consider a capacity-based argument. On a strictly capacity-based argument, these ethical theories show remarkable differences in how we ought to positively relate with others. To positively relate with others,

for a Kantian, we have to appeal to their intrinsic capacity for autonomy and by extension rationality; and for the utilitarian, we must appeal to their intrinsic capacity to feel pleasure or pain. However, from both the Confucian and Afro-ethical perspective, we are enjoined to treat others rightly simply by virtue of them being humans and having the capacity to be part of a communal relationship or a recipient of such relationship (Metz 2013; Bell and Metz 2011). Even though these are all capacity-based views, the intrinsic capacities we appeal to differ; what the collectivist view affords us is a wider spectrum to accommodate other species outside humans and animals into our moral circle. Since inclusion into our moral circle is dependent on if an entity is being identified with or exhibited solidarity towards, as Metz suggests, we can consider certain artificial intelligence systems like care robots, sex robots, robot nannies as morally significant beings.

Another difference between a collectivist and individualist outlook on ethics is with the principle of autonomy. Autonomy as self-governance and in the Kantian sense of an agent having the capacity to act with reference to objective morality rather than under the influence of desires does not exist in this way in collectivist cultures as decision-making is expected to be influenced by group dynamics and needs. Each member of society is expected to be an active member willing to put aside their needs for the good of society. For the utilitarian, one becomes autonomous if they can make moral judgements that help them maximise utility. In the collectivist ethics of Confucianism and Ubuntu (a moral theory from Africa), true autonomy involves having "rich social relations" (Bell and Metz 2011, p. 83) with others. Even though these two ethical systems vary in their interpretation of how one ought to relate with others, it is clear that the community plays a central role in ethical decision-making.

Looking again at the relational model of ethics found in collectivist cultures, it is clear that these cultures prize unity, cohesion, harmony and selflessness over individuality, personal autonomy and the likes (Hofstede 1991). Also, in collectivist societies, the emphasis moves from individual rights to family and community rights as individual actions have consequences not just for the individuals but also for the community; for this reason, individuals are implored to act right for the good of all. In individualist societies, rules are often made to protect the rights of individual members of society and their freedom to express those rights. The goal of rules in collectivist societies, as opposed to individualist societies, is to foster support, cooperation, and a sense of community among members of society. A simple way to perpetuate this stands is by insisting that the needs/right of all outweighs the need of one. This conception would have an impact on how we conceive of ethics by design and ethics for design.[4]

In studying Confucian and Afro-ethics, it is easy to see certain constants that reflect features of collectivism. The same can be said of individualism when you look at utilitarian and Kantian conceptions of right or wrong actions. For Kant, autonomy is central to moral decision-making. Also, the interest of the agent is keenly factored into the decision-making process. If we take a look at Kant's recommendation in the categorical imperative, for us to visualise universalising our actions to appraise if they be right or wrong actions, the agent has to ask, "Would I want this action to become universalised given these circumstances?" The answer would suggest, at least in that context, what action should be taken. In taking this introspective journey, the agent would have to hypothesise been on the other end of the stick if that action is meted on them. For the utilitarian, right and wrong acts are constructs that can either be identified by hedonic variables, preferences, or the lack of pain. This aligns very much with an individualist way to conceiving what is right and wrong. The ability to take action on grounds of its ability to provide some sort of satisfaction is heavily individualistic. On the other hand, Afro-ethics dictates that the needs of the group should be our criterion for telling the right actions apart from wrong actions. The implication would be that individuals may be unable to maximize their satisfaction as utilitarianism enjoins us to. They would more likely have to relinquish some level of satisfaction for the good of the community.

In this section, I have shown how cultural peculiarities across geographical mappings could influence ethical preferences. It is noteworthy that we consider this when engaging in debates on the ethics of AI and recommendations for a global guideline is proposed. Also, accepting contextualised contributions is an evident way to give a global outlook to the ethics of AI. In the next section, I proposed a roadmap that will allow for global participation in discourses around the ethics of AI.

[4] Dignum (2018) offers an insightful way to understand the ethics of AI across three categories. First, ethics by design, which focuses on the technical capacity and integration required to develop autonomous systems that have ethical reasoning capabilities. Second, ethics in design, which addresses governing guidelines and engineering methodologies that are required to analyse the ethical implications of autonomous intelligent systems as they become more ubiquitous. And third, ethics for design, which refers to codes of conducts, guidelines, standards and certifications that guarantee the quality and veracity of engineers, developers and users who engage with research and deployment of AI systems.

Roadmap to global engagement with the ethics of AI

Having acknowledged and shown that ethics is often influenced by cultural and contextual peculiarities, I succinctly offer three recommendations that serve as roadmaps to having an appreciable engagement with the ethics of AI. I reckon that my suggestions in this section are not exhaustive. However, these counsels offer practical ways designers, engineers, developers and researchers in AI can engage ethics at a global level. More so, it addresses the three subareas of the field according to Segun's (2020) classification—robot ethics, machine ethics and computational ethics.

One way to ensure that we have global participation with the ethics of AI in such a way that we have varying voices is to insist that international bodies that seek to create guidelines and recommendations around the use and deployment of AI technology do so having had a wide range of participation from representative groups. With only a few exceptions like the UNESCO's first global standard-setting instrument on the ethics of artificial intelligence in the form of a recommendation and the IEEE's Global Initiative on Ethics of Autonomous and Intelligent Systems, very few efforts have been made to aggregate a global perspective when drafting recommendations and guidelines for the use and regulation of artificial intelligence systems. This, it can be argued, is the tendency to assume that contexts do not matter when we consider ethical issues as we expect that what constitutes right or wrong actions are universals and not particulars. I contend that one way to encourage global participation in the ethics of AI is to insist that committees and bodies that seek to develop guidelines and recommendations should take cognizance of group dynamics, context and peculiarities. In addition, drafts of such proposals should be open to contributions from scholars of varying backgrounds. This in many ways would ensure that a more robust framing of solutions is proposed.

Another viable way to encourage global participation, which inevitably would lead to having a diverse pool of experts to address ethical issues in AI, is to promote scholarship from non-Western perspectives. This special issue is one of such. Deliberate efforts need to be put to encourage an intercultural and comparative take on serious debates in the ethics of AI. The challenge to this approach has been the continued marginalisation of non-Western epistemic traditions. The continuous display of testimonial injustice from publishers, journal editors and reviewers stifle efforts to have this type of dialogue. Most times, reviewers, judging from a Western epistemic grounding, assume that alternative perspectives do not offer any worthwhile contributions. This sort of thinking is exacerbated by the fact that much of what counts as international publishing avenues (journals and publishers) are resident in Europe and America, which continue to perpetuate and expand Western ideals and influences (Chimakonam 2017). The uneven power distribution makes it difficult for wider participation on discussions around ethics, especially with respect to AI. Hence, I suggest a deliberate call to have a global conversation on the ethics of AI.

The third approach is to have funders of research into the ethics of AI support projects that offer diverse views on ethical preferences especially as they apply to artificial intelligence systems. MIT's moral machine project offers us a good look at what diverse opinions on ethical issues may look like. With more funding for projects promoting ethical contributions from marginalised epistemic traditions, we are likely to have more research and published works that showcase prima facie attractive ethical features from non-Western perspectives. Funders should approach entertaining diverse views as an opportunity to support projects that might offer alternative and sometimes better solutions to problems in the field.

By acknowledging that problems in the field are often not evenly distributed and solutions are not either, it becomes expedient to clamour for global participation in discussions around the ethics of AI. I aver that beyond culturally influenced geographical mappings of ethical systems, we must also consider socio-economic factors that may affect conceptions of the role of AI. A strong reason to take this concern rather seriously is that we are a few steps closer to having AI systems deployed across the world.

An outlook on the content of this special issue

In their article, "Applying a principle of explicability to AI research in Africa: should we do it?" Mary Carman and Benjamin Rosman emphasised the need to contextualise AI in and for Africa. The authors propose a different way to conceive of the principle of explicability such that decision-making systems that are powered by AI are fair and just and their recommendations or decisions intelligible for the context they are applied to. This means guarding against biases including racial biases in training data and building systems that are sensitive to African interests and needs.

With a complicated history with the West, Africa continues to experience imposition of ideas, values and technologies that are often not context-driven, the authors warn against uncritically adopting principles under AI for social good proposal, with one of such principle being explicability. By proposing an epistemological and ethical look at the principle of explicability when applied to core areas

of machine learning, in other words answering the questions, "how does it work and who is responsible for how it works?", Mary Carman and Benjamin Rosman offer us unique ways to contextualise the purpose of an AI system to better suit a society. Having considered that Africa plays little role in the development of AI technology, developers, engineers and big tech corporations who intend to have their products used within the continent are enjoined to ensure that values with those of the society are aligned by consulting subject matter experts and knowledgeable stakeholders.

In the article, "An Ontic-ontological Theory for Ethics of Designing Social Robots: A Case of Black African Women and Humanoids", John Lamola offers us a phenomenological theory that should motivate ethical consideration by robot designers and engineers, suggesting that the socio-aesthetic state relating to black women can be used as remediation of design decisions. Lamola develops what he calls "a novel appreciation of the aesthetic and phenomenal ontology of humanlike socially-situated robots, which serves to account for their affective potentialities". He contends that the socio-political issues of race and gender play central roles in design decisions, making it crucial for designers of humanlike social robots to be conscious of their prejudices, ethnic and socio-cultural commitments while designing.

In pursuit of having socially-conscious roboticists, Lamola advocates for value sensitivity in the design of humanlike robots. Value sensitive design is a theoretically-based approach to technology design that takes into account human values throughout the process of design. Within the ethics of AI, value-sensitive design is a delicate matter requiring utmost sensitivity in the appropriation and representation of cultural and ethical values. Lamola's article calls for a wider dialogue with innovators, designers and researchers on robotics to consider the role race and gender play in the design of social and assistive robots.

Is it ever possible to consider AI systems as persons? In his article, "Artificial intelligence and African conceptions of personhood", Christopher Wareham answers this question offering a characteristically African ethical perspective. He avers, following an African account of personhood which he admits is generally anthropocentric, that we can, in fact, accommodate artificial intelligence systems as persons. The paper explores the criteria and circumstances on which personhood in a moral sense can be conferred upon AI. By leaning on the relational model of ethics found among African normative theories, Wareham notes that conceptions of personhood in Africa, unlike Western notions, are partial with a potentiality to exclude AI from our moral circle.

A relational model of ethics, which is representative of many collectivist views on ethics, suggests that the capacity to partake in communal relations would depend on how one is perceived by members of the society. Against this backdrop, Wareham proposes two ways we can conceive of AI personhood from an African perspective. First, we can think of AI as subjects of communal relations or, second, as objects of that relationship. As subjects, we would accord AI rights and duties as moral agents giving their capacity to partake in communal relations; and as objects, we would ascribe moral status to these systems as entities worthy of our moral benevolence.

In the literature, several works defend or refute the claim that AIS can be considered moral agents. One criterion that runs through much of the published works on the subject is that of computational rationality. In his article, "Computationally rational agents can be moral agents", Bongani Mabaso makes a compelling argument for why AI with computational rationality should be considered moral agents. Mabaso offers us a clinical approach to consider AI as moral agents by referring to computational rationality, an attempt at optimising decision-making process using computational resources.

Mabaso develops a logical, philosophical and computationally consistent model for building artificial moral agents and extending the debate from a purely theoretical domain to a practical one, which also accounts for artificial morality. Using a bounded-optimal computational framework to argue for the possibility of artificial moral agency, Mabaso avers that key features agency such as consciousness and autonomy are computationally possible, at least functionally. He notes that, even though complex, the framework created for computational rationality can serve as a basis to develop computational morality.

Can we have an account of moral patiency that is not typically anthropocentric? Fabio Tollon thinks this is possible in his article, "The Artificial View: Toward a non-anthropocentric account of moral patiency". In it, he provides a critique of Torrance's organic view of ethical status, a view that dictates that only moral patients can be moral agents and because AI lack sentience they cannot be the subject of moral consideration. Tollon argues that not only is this view primarily anthropocentric, it also uses this intuition to define who gets ascribed sentience and ultimately what counts as a moral patient. This would mean narrowing our moral circle and effectively alienating artificial intelligence systems who lack sentience but may even be behaviourally indistinguishable from other sentient beings.

Tollon uncovers conceptual and epistemic challenges with the organic view and advances a combination of socio-relational and an intentional stance as criteria for the consideration of moral patiency. According to this view, behavioural cues are taken into account as opposed to having maximal criteria that depend on artificial intelligence systems possessing real qualitative states. With behavioural cues, Tollon believes we can have an idea of an entity's internal state and possibly its capacity for affective states. Under this view, we ascribe patiency based on immediacy and proximate value of

an artificial intelligence system to existing members of our moral circle. We consider their extrinsic capacities for social relationships and not the intrinsic capacities like sentience.

Danielle Swanepoel in her paper, "The possibility of deliberate norm-adherence in AI", questions if we should grant moral status to artificial intelligence systems. She contends that to grant AI moral status they must meet, at least minimally, deliberate norm-adherence, which they currently cannot. By deliberate norm-adherence, Swanepoel refers to acting out of reverence for the norm. Considering the need to proffer, however minimally, a framework to evaluate if we should accommodate AI into our moral circle, Swanepoel makes two critical distinctions; one, between norm-compliance and norm-endorsement and two, between deliberate norm-adherence and deliberate norm-violation.

The paper insists that we should not extend the status of moral agency to AI and if we must do so it must be with certain limits. One key reason for Swanepoel's position is that we have still not been able to answer the question of moral accountability for AI. Even if we succeed in building systems with actions that appear to adhere to norms, Swanepoel contends their inability to endorse those norms or deliberately choose to violate those norms make their adherence to the norm pretentious.

There are quite many issues to be addressed in the ethics of AI and lots of work published about them. A special issue dedicated to addressing questions around the role contexts, cultures and non-conventional ethical preferences play in the ethics of AI is a step in the right direction. I am convinced that this special issue would stir up dialogue around developing more robust and context-driven projects in the ethics of AI.

References

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., et al. (2018). The moral machine experiment. *Nature, 563*(7729), 59–64. https://doi.org/10.1038/s41586-018-0637-6

Bell, D. A., & Metz, T. (2011). Confucianism and Ubuntu: Reflections on a dialogue between Chinese and African traditions. *Journal of Chinese Philosophy, 38,* 78–95.

Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed*. https://doi.org/10.2966/scrip.170220.389

Chimakonam, J. O. (2017). African philosophy and global epistemic injustice. *Journal of Global Ethics, 13*(2), 120–137. https://doi.org/10.1080/17449626.2017.1364660

Dignum, V. (2018). Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology, 20,* 1–3. https://doi.org/10.1007/s10676-018-9450-z

Feuchtwang, S. (2016). Chinese religions. In L. Woodhead, H. Kawanami, & C. H. Partridge (Eds.), *Religions in the modern world: Traditions and transformations* (pp. 143–172). London: Routledge.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London: McGraw-Hill.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks: Sage publications.

Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture, 2*(1), 2307.

Korsgaard, C. (1996). *Creating the Kingdom of Ends*. New York: Cambridge University Press.

McDougall, B. S., & Hansson, A. (2002). *Chinese concepts of privacy* (Vol. 55). Leiden: Brill Academic Pub.

Metz, T. (2007). Toward an African moral theory. *The Journal of Political Philosophy, 15*(3), 321–341.

Metz, T. (2013). The virtues of African ethics. In S. Van Hooft (Ed.), *The handbook of virtue ethics* (pp. 276–284). Durham: Acumen Publishers.

Metz, T. (2016). An African theory of social justice: Relationship as the ground of rights, resources and recognition. In C. Boisen & M. C. Murray (Eds.), *Distributive justice debates in political and social thought: Perspectives on finding a fair share* (pp. 171–190). New York: Routledge.

Metz, T. (2017). Toward an african moral theory. *Themes, issues and problems in African philosophy* (pp. 97–119). Cham: Palgrave Macmillan.

O'Neill, O. (1975). *Acting on principle*. New York: Columbia University Press.

Realo, A. (1998). Collectivism in an individualist culture: The case of Estonia. *Trames, 2*(52/47), 19–39.

Segun, S. T. (2020). From machine ethics to computational ethics. *AI & Society*. https://doi.org/10.1007/s00146-020-01010-1

Tam, L. (2018). Why privacy is an alien concept in Chinese culture. Retrieved from https://scmp.com/news/hong-kong/article/2139946/why-privacyalien-concept-chinese-culture.

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2017). Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems, version 2. IEEE. Retrieved from http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.

Shutte, A. (2001). *Ubuntu: An ethic for the new South Africa*. Cape Town: Cluster Publications.

UNESCO. (2019). Elaboration of a Recommendation on the ethics of artificial intelligence. Retrieved from https://en.unesco.org/artificial-intelligence/ethics.

Yao-Huai, L. (2005). Privacy and data privacy issues in contemporary China. *Ethics and Information Technology, 7*(1), 7–15.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.