



# African Values, Ethics, and Technology

---

Questions, Issues, and Approaches

---

*Edited by*  
Beatrice Dedaa Okyere-Manu

palgrave  
macmillan

Beatrice Dedaa Okyere-Manu  
Editor

# African Values, Ethics, and Technology

Questions, Issues, and Approaches

palgrave  
macmillan

*Editor*

Beatrice Dedaa Okyere-Manu  
School of Religion, Philosophy and Classics  
University of KwaZulu-Natal  
Pietermaritzburg, South Africa

ISBN 978-3-030-70549-7      ISBN 978-3-030-70550-3 (eBook)  
<https://doi.org/10.1007/978-3-030-70550-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use. The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: © Alex Linch shutterstock.com

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# CONTENTS

<b>1</b>	<b>Introduction: Charting an African Perspective of Technological Innovation</b>	<b>1</b>
	Beatrice Dedaa Okyere-Manu	
<b>Part I</b>	<b>The Fourth Industrial Revolution and African Ethics</b>	<b>15</b>
<b>2</b>	<b>The Fourth Industrial Revolution (4IR) and Africa's Future: Reflections from African Ethics</b>	<b>17</b>
	Munamoto Chemhuru	
<b>3</b>	<b>Africa in the Fourth Industrial Revolution: A status quaestionis, from the cultural to the phenomenological</b>	<b>35</b>
	Malesela John Lamola	
<b>Part II</b>	<b>African Values and Technology</b>	<b>53</b>
<b>4</b>	<b>African Reasons Why AI Should Not Maximize Utility</b>	<b>55</b>
	Thaddeus Metz	

<b>5 Values and Technological Development in an African Context</b>	<b>73</b>
Bernard Matolino	
<b>6 African Cultural Values, Practices and Modern Technology</b>	<b>89</b>
Ovett Nwosimiri	
<b>Part III Technology, African Ethics and Sexual Relations</b>	<b>103</b>
<b>7 Shifting Intimate Sexual Relations from Humans to Machines: An African Indigenous Ethical Perspective</b>	<b>105</b>
Beatrice Dedaa Okyere-Manu	
<b>8 The Death of Isintu in Contemporary Technological Era: The Ethics of Sex Robots Among the Ndebele of Matabo</b>	<b>123</b>
Herbert Moyo	
<b>Part IV Technology, African Values and Human Relationship</b>	<b>137</b>
<b>9 The Importance of a Neo-African Communitarianism in Virtual Space: An Ethical Inquiry for the African Teenager</b>	<b>139</b>
Thando Nkohla-Ramunenyiwa	
<b>10 The Ambivalent Role of Technology on Human Relationships: An Afrocentric Exploration</b>	<b>155</b>
Sophia Chirongoma and Lucia Mutsvedu	
<b>11 Interrogating Social Media Group Communication's Integrity: An African, Utilitarian Perspective</b>	<b>173</b>
Elias G. Konyana	

<b>Part V Bioethics, African Values and Technology</b>	<b>187</b>
<b>12 Bioethics and Technology: An African Ethical Perspective</b>	<b>189</b>
Wilfred Lajul	
<b>13 The Use of Sex Selection Reproductive Technology in Traditional African Societies: An Ethical Evaluation and a Case for Its Adaptation</b>	<b>217</b>
Samuel Awuah-Nyamekye and Joseph Oppong	
<b>14 Assisted Reproductive Technologies and Indigenous Akan Ethics: A critical Analysis</b>	<b>229</b>
Stephen Nkansah Morgan	
<b>Part VI African Religious Values and Technology</b>	<b>245</b>
<b>15 The Impact of Technologies on African Religions: A Theological Perspective</b>	<b>247</b>
Nomatter Sande	
<b>16 Technologization of Religion: The Unstoppable Revolution in the Zimbabwean Mainline Churches</b>	<b>263</b>
Martin Mujinga	
<b>Index</b>	<b>281</b>



# African Reasons Why AI Should Not Maximize Utility

*Thaddeus Metz*

## INTRODUCING THE QUESTION OF HOW TO PROGRAMME ARTIFICIALLY INTELLIGENT AUTOMATED SYSTEMS

How should automated systems governed by artificial intelligence (AI) be programmed so as to act in accordance with sound moral norms? For instance, how should one programme a self-driving car that learns in the course of navigating streets? How should one programme a robot that is able to provide nursing to patients or domestic labour to family members, upon adapting to their various idiosyncrasies? How should one programme a military weapon that takes advantage of the way soldiers tend to engage in battle? How should one programme a device that could mine underground on its own, changing its direction and method of extraction upon updating calculations of which kind of ore is likely to be in which place?

In asking how to programme these systems so that they do the right thing, one need not presume that the machines count as moral agents in any robust sense. Instead, it is perfectly sensible to view them as lacking

---

T. Metz (✉)  
University of Pretoria, Pretoria, South Africa  
e-mail: [th.metz@up.ac.za](mailto:th.metz@up.ac.za)

moral agency and instead as being mere tools of their designers, who are human persons and hence moral agents. If a self-driving car is not disposed to stop for pedestrians, it is not the car that has performed a culpable wrong, but the one who made it (or, more carefully, the one who had let it onto the street upon knowing, or having had a duty to know, how it was made). The question is in the first instance about how those making automated systems governed by artificial intelligence should construct them, supposing they know they will be deployed in social contexts (and not merely contained in a laboratory).

It appears that the dominant answer to that question has been a utilitarian one, according to which a machine ought to be programmed so as to do what, in the light of available data and processing power, is expected to maximize what is good for human beings and to minimize what is bad for them in the long run (Shulman et al. 2009; Majot and Yampolskiy 2014; Marwala 2014, 2017; Hibbard 2015; Oesterheld 2015, 2016; Kinjo and Ebina 2017; Bauer 2020<sup>1</sup>). As I indicate below, this approach has the advantages of appearing to capture the nature of rational choice, to be impartial and hence morally attractive, as well as to be amenable to being formalized and coded.

Utilitarianism is a characteristically western conception of moral-practical reason, having been advanced in various of its respects by classic philosophers such as Thomas Hobbes, Blaise Pascal, Jeremy Bentham, Adam Smith, and John Stuart Mill, and continuing to guide much thought about ethics and public policy in the twenty-first-century American-European-Australasian context. In this chapter, I appeal to characteristically African values to question the moral aptness of programming AI automated systems to maximize utility. Although I draw on perspectives that are particularly prominent in contemporary sub-Saharan ethics, and hence (presumably) in the worldviews of the ‘traditional’ peoples and cultures that inform them, the objections to utilitarianism will, at a certain level of abstractness, resonate with those from a variety of moral-philosophical backgrounds, particularly from the Global South. Utilitarianism prescribes a number of immoral actions in the light of some plausible beliefs common

<sup>1</sup>For unusual or tangential applications, there are Grau (2005), who deems utilitarianism apt for robot-robot relations, even if not for robot-human interaction; Gloor (2016), who urges that AI be used primarily to avoid extremely bad outcomes for human beings, and not to produce particularly good ones; and Bonnefon et al. (2015, 2016), who argue that western people are generally utilitarian about how self-driving cars should be programmed.



in African ethical thought, and supposing that moral actions are necessarily rational ones, these criticisms also implicitly cast doubt on the apparent rationality of utilitarianism.

Although there is recent literature on how one might apply AI to resolve problems in Africa and on ethical issues facing AI's application to Africa (e.g. World Wide Web Foundation 2017; Access Partnership 2018; Gwagwa 2019a; Sallstrom et al. 2019; Ormond 2020), there is literally nothing as yet on how one might do so in the light of indigenous African values as distinct from western ones. By 'African' values, I mean ones *salient* in the massive sub-Saharan region of the continent, that is, beliefs about morality found amongst many indigenous black peoples (as opposed to, say, those of Arab descent in the north) over a long span of time and not found amongst many other societies around the world.<sup>2</sup> Recently, one scholar has noted 'the need to define African values and align AI with them' (Gwagwa 2019b), but has not yet sought to meet the need, while another has said that 'African cultural values need to be taken into account when defining a framework for AI on the continent' (Spini 2019), but has not developed such a framework. Here I aim to make some headway when it comes to heeding these calls.

I do not do so for reasons of relativism. It is not my view that the values that should govern technology in a certain society are necessarily those held by most in that society. I believe that majorities can be mistaken about right and wrong action, as nineteenth-century Americans were in respect of slavery. Instead, I draw on under-considered African ethical perspectives in the thought that any long-standing philosophical tradition probably has some insight into the human condition and has something to teach those outside it. Many of the values I identify as 'African' will, upon construing them abstractly, be taken seriously by many moral philosophers, professional ethicists, and the like around the world, especially, but not solely, outside the West.

In the following, I begin by defining my target, saying more about what the nature of utilitarianism is, explaining why theorists have been drawn towards it, and illustrating what it would look like when applied to AI automated systems (section "[Utilitarianism in the Context of AI](#)"). Then, I advance, on African grounds, four major objections to

<sup>2</sup>Such a definition of what 'African' means entails that in order to count, a value need neither be held by all those in Africa nor be held only by those in Africa. For more on what geographical labels plausibly mean, see Metz (2015).

programming AI automated systems to maximize utility, arguing that doing so would fail to respect human dignity (section “[Human Dignity and AI](#)”), inadequately uphold group rights (section “[Group Rights and AI](#)”), violate a principle of family first (section “[Family First and AI](#)”), and counterintuitively forbid certain kinds of self-sacrifice on behalf of others (section “[Self-sacrifice and AI](#)”). Although one might have thought that utilitarianism expects a lot from a moral agent, I argue that it cannot even get that right, for it forbids one from helping others when helping oneself would do the most good impartially construed. I conclude by briefly noting some avenues for future research, such as considering what African grounds there might be to question the other major western moral theory, Kantianism (section “[Conclusion: From Utilitarianism to Kantianism](#)”).

## UTILITARIANISM IN THE CONTEXT OF AI

In this section I tell the reader what I mean by ‘utilitarianism’ and related terms such as ‘maximize utility’, show why it has been taken so seriously by ethical theorists, and give some examples of how it might inform AI automated systems. This section is therefore largely expository, saving critical discussion for later sections.

As will be familiar to many readers, utilitarianism is the doctrine that for any action to be rational and moral, it must be expected to maximize what is good for human beings (and perhaps animals) and to minimize what is bad for them in the long run.<sup>3</sup> By standard utilitarianism, what is good for us is subjective, a matter of either pleasant experiences, satisfied preferences, or positive emotions, and what is bad consists of pain, dissatisfaction, or negativity. Subjective well-being is taken to be the only thing good for its own sake or at least the only sort of good that should be action-guiding for us as moral agents. Everything else on earth, at least in the mineral and vegetable kingdoms (again, potentially excluding the animal kingdom), is at best of merely instrumental value for being of use to foster our subjective well-being.<sup>4</sup>

<sup>3</sup>I therefore address act-utilitarianism in this chapter, setting aside rule-utilitarianism, which has been much less influential in AI circles. I believe that many of the objections to the former apply to the latter, but it takes extra work to demonstrate that. For one advocate of rule-utilitarianism in the context of AI, see Bauer (2020).

<sup>4</sup>There are of course those who pair a consequentialist combinatorial function with an objective account of final value, with an early advocate being Moore (1903) and a more recent one being Railton (1984). I believe the Afro-centric objections made below to classic

Now, on the face of it, everyone's well-being matters equally, such that one would be objectionably biased to leave out anyone's interests when considering how to act. It would also appear irrational to leave out anyone's interests when making decisions, for a given person's subjective well-being has just as much value in itself as anyone else's. Furthermore, it is *prima facie* rational to produce as much of what is good for its own sake as one can; every bit of well-being that one could promote provides an agent reason to promote it, making it irrational to do anything less than the best one is in a position to do. Morality, too, seems to counsel maximizing what is good for us and minimizing what is bad, for surely it is preferable from the moral point of view to do all one can for the sake of humanity.

This consequentialist account of moral-practical reason is compelling, and it is not surprising to find it, or at least various elements of it, invoked in a wide array of western contexts. For instance, it appears to capture the logic of many everyday decisions that at least 'modern' western people make. When deciding whether to use a bicycle or a car to get to work, they naturally attend to the results of the two options, not merely for themselves, but also for others. Riding the bike would be painful and would take longer, and yet it would be beneficial in the long run in respect of one's health and hence one's happiness. Taking the car would be more pleasant and quicker, but it would cost more money and pollute the environment, risking lung cancer to others. Utilitarianism prescribes choosing whichever option would have the most good outcomes with the least bad outcomes in the long term, impartially weighing everyone's interests given available information. People do not always in fact choose in that way, but, for the utilitarian, they should, as doing so would be prescribed by a consistent application of the logic they themselves tend to use to make decisions.

When it comes to state officials, those governing people living in a certain territory, it is common to appeal to what is called in public policy circles 'cost-benefit analysis'. Since all citizens matter equally, those who make and carry out the law should do so in ways that are going to maximize benefits and to minimize costs, taking the good of all citizens impartially into account. Consider that it would be patently unjust for those with political power to use government resources such as jobs and money to

utilitarianism apply with comparable force to objective consequentialism, at least if it includes agent-neutrality (and so is unlike Sen 2000).

benefit themselves, their families, or a certain racial group at the expense of the general welfare.

Still more, at least the maximizing element of utilitarianism has been dominant in western economics for at least three centuries, with actors in market exchanges seeking to obtain the most amount of profit or goods and with the least amount of expense. Of course, in the context of markets, people are motivated by self-interest and not the interests of all. However, Adam Smith has famously argued that even if one's intention is not to benefit others when maximizing profit for oneself, in the long run that practice often enough ends up benefiting society more than other courses of action one could have undertaken, 'as if by an invisible hand'.

Beyond these considerations of the nature of rationality and morality, programmers and those who work with them or otherwise think about programming are drawn towards utilitarianism because it appears quantifiable and hence able to be coded (as noted in Anderson and Anderson 2007: 18; Oesterheld 2015).<sup>5</sup> Clearly, some pleasures are greater than others and some preferences are stronger than others. Utilitarians believe that, in principle, we could assign cardinal values to such states, ascribing real numbers to degrees of subjective well-being and woe. While that is of course difficult for a human being to calculate, artificial intelligence might be in a terrific position to estimate how much pleasure versus pain a given course of action would be expected to produce and to identify the one with the highest net balance (Anderson et al. 2005; Anderson and Anderson 2007: 18).

Applied to, say, a self-driving car, it would be natural for a utilitarian programmer to have it minimize the number of people killed. It would be irrational and immoral, so the argument goes, to favour the interests of, say, the driver as opposed to pedestrians. Instead, everyone's interests count equally from the moral point of view, meaning that the car ought to be directed to do whatever would produce the greatest pleasure and the least pain in the long run, which presumably would come from killing one person instead of three, supposing those were the only options available.

It might seem as though a utilitarian weapon would be nonsensical. After all, utilitarianism is impartial, with a person's nationality making zero difference to her moral standing as capable of subjective well-being, whereas during a time of war one is expected to take sides, that is, to place

<sup>5</sup>For a dissenting perspective, that it would be impossible to programme enough information for a machine to account for long-term results, see Allen, Varner, and Zinser (2000: 256).

the lives of the soldiers from one's country ahead of those of others. However, it is worth considering the point that we should in fact want programmers to develop weapons that would serve only just causes such as rebutting aggression as well as ones that would not treat enemy soldiers, even those fighting for an unjust cause, as though they do not matter at all.

Having expounded utilitarianism and why it has been attractive, one might wonder what in the world is wrong with it. Despite its strengths, the African philosophical tradition provides strong reason for thinking that it has irredeemable weaknesses.

## HUMAN DIGNITY AND AI

Although utilitarianism ascribes a moral status to every individual human being that is alive and either is sentient, has preferences, or exhibits emotions, it does not accord them a dignity. The former amounts to the view that for whichever being is capable of subjective well-being, there is *pro tanto* moral reason to promote its welfare. The latter is the idea that a person has a superlative non-instrumental value that merits respect. To have a moral status, and hence to be owed dutiful treatment directly, does not imply that one is a person who is good for its own sake to a degree higher than anything else in the world. In addition, for there to be moral reason to promote an individual's well-being differs from there being moral reason to avoid degrading a person.

Improving others' well-being does have a place in African thought, but most often insofar as doing so can sometimes be a way of expressing respect for people who have a dignity. Traditionally speaking, sub-Saharanans tend to believe that our dignity is constituted by either vital force or group membership. Vital force is an imperceptible energy that has come from God and permeates everything in the universe in varying amounts and complexities, where, of perceptible beings, human persons have the most (Wiredu 1996: 157–71; Ilesanmi 2001; Deng 2004; Iroegbu 2005; Etieyibo 2017). From this perspective, it is not our quality of life that matters fundamentally for morality, as per utilitarianism, but rather the fact of life itself; our nature is sacred for exhibiting a certain quantity or quality of divine energy.

Other times African thinkers maintain that persons have a dignity by virtue of membership or relationship. One scholar remarks, 'The dignity of human beings emanates from the network of relationships, from being in community; in an African view, it cannot be reduced to a unique,

competitive and free personal ego' (Botman 2000), while another says, '(T)he human person in Africa is from the very beginning in a network of relationships that constitutes his alienable dignity' (Bujo 2001: 88). There are different views on precisely which relationships matter, with some holding that they are ones with a clan or a specific community that has cared for one or that you have cared for (Cobbah 1987; Ikuenobe 2016), and others suggesting the relevant relationships are with the human species as a whole (perhaps Gyekye 2010: sec. 6).

Both views are controversial from a cross-cultural perspective. For example, western philosophers would be inclined to ground our dignity on our capacity for autonomy, while thinkers in the Judeo-Christian-Islamic tradition would maintain that it is constituted by a God-given soul (i.e. a spiritual substance as opposed to an imperceptible force). Although a specific and defensible account of what it is that gives us a dignity would be useful when articulating a comprehensive ethic, it is unnecessary to consider utilitarianism's implications for AI. So long as individual persons have a dignity that merits respectful treatment, regardless of what confers that dignity on us, a moral agent will be forbidden from treating persons merely as a means to the greater good. That is, a typical dignity-based ethic will accord human rights to each person, where to have a human right is for others to have a duty not to subordinate or harm a person that should be upheld even if not doing so would promote a marginally greater amount of value in the world.

It is well known that a utilitarian ethic has difficulty accounting for human rights. Here is a classic example from contemporary Anglo-American ethics (Nozick 1974: 28–30). Imagine that a mob angry about an apparent crime will severely beat and then kill two innocent people unless you, the sheriff in town, frame, corporally punish, and execute one innocent person. From a utilitarian perspective, it would be rational and moral for you to do terrible things to the person, if that were indeed the only way to prevent others from doing terrible things to a larger extent. However, from a human rights perspective, you should not, for doing so would degrade the person. Instead of treating him as though he is valuable for his own sake, it would treat him as though he has a merely instrumental value for the reduction of bad outcomes in the world.

Of course, utilitarians are initially inclined to reply that the case is inaccurately described. They will suggest that the outcomes of a sheriff harming and killing one innocent person would likely be worse in the long run than letting two innocents be harmed and killed by others. After all, upon

the sheriff framing an innocent person, the criminal is still free to commit crime again, and when people discover that the sheriff has done so, the criminal justice system will be weakened, allowing all the more murders to take place down the road.

However, it is not unfair to suppose, for the sake of argument, that there was in fact no crime committed, but that the sheriff was not in a position to convince members of the public of that in time before rioting. In addition, it is reasonable to imagine, also for the sake of argument, that no one would ever discover that the sheriff has framed an innocent party for the non-existent crime. It is intelligible to ask what the right thing to do would be under these conditions, where a human rights perspective will look compelling to many African ethicists and to many other adherents to the idea of human dignity around the world.

What does a human rights framework mean for how to programme AI automated systems? It would involve directing them to treat persons in certain ways that are incompatible with maximizing utility. For example, consider a self-driving car. The most basic version of what is widely known as ‘the trolley problem’, often applied to self-driving cars (for just one influential example, see Awad et al. 2018), appears to tell in favour of utilitarianism.<sup>6</sup> In that case, the car has the two options of either, say, striking two pedestrians if it continues forward or veering in another direction where it would strike one pedestrian. Here, it appears that it would be right for the car to strike the one, which lends *prima facie* support to utilitarianism.

However, upon reflection, the support is weak, as there are many other situations in which it would intuitively be a human rights violation to kill one so as to save two. For example, imagine that the self-driving car happens to have two critically injured passengers in it and that they will die unless they get to the hospital in time. Suppose further that the only way to get them to the hospital in time would be for the self-driving car to mow down one pedestrian. It appears that respect for the person’s dignity forbids killing him in order to save the lives of two others. In short, he has a human right to life that should be observed when programming a self-driving car.

<sup>6</sup>For the view that the trolley problem is importantly distinct from the self-driving car controversy, see Nyholm and Smids (2016).

## GROUP RIGHTS AND AI

Although it is characteristic of African moral thought to ascribe a dignity to individual persons, it is likewise characteristic of it to maintain that certain kinds of wholes merit moral consideration. There are some sub-Saharan theorists who appear to believe only in group rights, such that, say, a family or clan alone has rights and the individual has none, or at least none that could ever conflict with them (e.g. Ake 1987). However, much more common is the view that there are both individual rights and collective rights, which might come into conflict with each other and need to be balanced in some way.

Consider, for example, the African ('Banjul') Charter on Human and Peoples' Rights (Organization of African Unity 1981), which includes many rights of individual human beings alongside those ascribed to a people, which is probably best understood as a nation (or perhaps a country). The latter include rights of a people not to be dominated and to resist domination (Article 20), and rights of a people to natural resources, socio-economic development, and an environment necessary for that (Articles 21, 22, 24).

Similarly, see the Charter for African Cultural Renaissance (African Union 2006). It too accords individuals many rights to access and develop their culture, but alongside them the document speaks of African peoples evolving, being provided resources, and being enriched (Preamble and Articles 3 and 5), the rights of minorities to their cultures (Article 5), and the cultural advancement of African states (Article 18).

It is natural to think of the talk of duties with respect to peoples, minorities, and states as group-based. However, it is a contested matter, amongst philosophers and related thinkers around the world, whether to ascribe moral rights to groups as distinct from their individual members. Some, particularly in the West, maintain that what might appear to be group rights are ultimately the rights of the individuals who compose them. While a well-developed theory of rights to culture, to development, and to resist domination would be welcome, it is not necessary in order to cast doubt on utilitarianism. Supposing that such rights exist, regardless of what ultimately grounds them, we find reason to resist programming AI automated systems in a utilitarian manner.

To see this point, consider the right of a people to resist foreign domination of it and what such a right entails for how to programme a smart automated weapons system. Above I suggested that a utilitarian weapons



system would naturally be deployed to rebut aggression, but that is not always so. Sometimes aggression that is intuitively unjustified would serve to maximize utility in the long run, and sometimes rebutting such aggression but failing to maximize utility in the long run is intuitively justified.

Such seems to be the case when the aggression is undertaken by a much greater majority than the non-aggressing minority being suppressed. Suppose that powers beyond Africa use their might to enslave a small sub-Saharan people, forcing it to extract minerals and related resources to drive foreign economies. Utilitarianism could permit such imperialism and forbid resistance against it, if the numbers who flourished from it were great and the numbers who suffered from it were small. However, if the dominated people rose up against its taskmasters, it would be apt for a smart automated weapons system to assist it, despite the fact that doing so would fail to maximize utility. Whereas a utilitarian weapon would operate on the side of the oppressor, a morally justified weapon would operate on the side of the oppressed, regardless of the numbers involved.

### FAMILY FIRST AND AI

In the previous section I noted that many African ethicists believe in human, that is, individual rights while also accepting some kind of collective rights. Here, I somewhat similarly note that many African ethicists believe that morality has an impartial dimension while also contending that it has a partial one, where the partiality is antithetical to utilitarian reasoning.

Ascribing a dignity to individuals, at least in virtue of their vital force or membership in the human family, entails a certain kind of impartiality, egalitarianism, or cosmopolitanism. That is, everyone matters from the moral point of view, such that a moral agent can have obligations to help others regardless of, say, their religion or nationality. As Kwame Gyekye has pointed out (2010: sec. 6), it is such impartial moral thought that plausibly explains the widespread traditional practice of welcoming visitors to an African village. Hospitality to those one does not know makes good sense if all human beings have a dignity that demands honouring.

However, there is an additional salient part of African ethics that is instead partial, directing one to aid those related to you before aiding others or aiding the former to a greater degree than the latter. ‘Family first’ and ‘charity begins at home’ are widely accepted (e.g. Wiredu 1992: 200; Appiah 1998; Ramose 2003: 385–86; Molefe 2019: 84–86) and

indicate that sub-Saharan morality is well understood as including special obligations of beneficence. If one had to choose between saving the life of one's mother or that of a stranger, a characteristic African view would entail that one would do wrong not to save one's mother. For utilitarianism, in contrast, one should flip a coin to decide whom to save if the outcomes of either option would be the same, or one should save the stranger if the outcomes of doing so would be marginally better than saving one's mother.

Traditionally speaking, it was blood relations that determined who has priority when it comes to aid (on which see Appiah 1998). That, however, is a controversial basis for grounding duties to help others. More common these days is the view that, having related communally with others for a long while or to a strong degree (as with family or colleagues), one owes more of oneself and one's resources to them than to others. Regardless of the precise way in which people must be related to the moral agent in order to be prioritized, so long as priority should be accorded to those related to her in some way or other—as many readers will agree—utilitarianism is suspect.

It follows that when programming AI automated systems, from an African perspective one ought to do so in a way that respects partial duties. If there is a robot in the home, while it should be programmed to be hospitable to visitors, it should also be programmed, at least as a default, to save the life of one's mother before that of a visitor who is not a member of the family or particularly close to it. A utilitarian robot, in contrast, would calculate the expected utility of saving one's mother compared to saving the visitor, and the calculations might not go in favour of mum, particularly if she is old and sick and the visitor is young and healthy.

Similar remarks apply to contexts in which there is relationality, even if to a less intense degree than that encountered in a family. For example, suppose a doctor has had a long-standing relationship with a particular patient and has programmed a robot to nurse him while staying in a hospital. It would be wrong for the doctor instead to allow the robot to disregard this patient's needs, whenever doing so would be necessary to care for a patient of another doctor with marginally greater needs.

### SELF-SACRIFICE AND AI

Here I discuss the last major respect in which values that are African, while also being attractive to readers from a variety of backgrounds, are incompatible with utilitarianism and in a way that would have a bearing on

programming AI automated systems. It concerns the permissibility of making sacrifices of one's own welfare for the sake of others.

It is well known that it is typical for an African ethical perspective to prescribe strong and substantial duties to help other people. That is particularly true in respect of family members, as per the previous section. However, even more impartial interpretations of African morality place heavy demands on a moral agent to improve others' quality of life. For example, Gyekye's influential moderate communitarian ethic prescribes doing all one can to promote the common good, that is, to meet the needs of everyone (albeit in ways that do not violate certain individual rights), without any distinction drawn between what is beyond the call of duty and what is not (1997: 35–76). Kwasi Wiredu's interpretation of African morality in terms of sympathetic impartiality and the Golden Rule similarly expects a lot (1996, 2009: 15–16); putting oneself in others' shoes, one will come to see that they would be glad if one aided them in significant ways.

Such an approach to positive obligations might seem to dovetail neatly with utilitarianism, for it, too, is known for demanding much of a moral agent. However, two Anglo-American philosophers long ago noted respects in which utilitarianism would actually forbid you from doing what is best for others in the long run; that is when the most good in the world would be produced if one instead directed benefits towards oneself (see Nozick 1974: 41 on 'utility monsters' and Slote 1985 on 'agent-sacrificing permissions').

Consider two cases to illustrate the concern (from Metz 2013: 188–89), after which I indicate what it might mean for AI automated systems. For a first case where the most good and least bad in the world would result if subjective well-being were conferred on oneself and woe were placed on others, consider a mother and son who are both hungry. Suppose that maximizing values of nutrition, taste, and the like prescribes mum eating a portion of food that has become available. Even so, it would surely be permissible for the mother to decide to give the food to her son. For a second case, imagine that a vicious canine is chasing you and a young woman, whom you do not know, and that minimizing fear, pain, and missed appointments prescribes outrunning, if not tripping, the woman so that the dog gets to her. Even so, it would be permissible for you instead to draw the dog towards yourself, enabling the woman to escape. In both cases, utilitarianism would forbid you from helping others, but for many

African ethicists, and more generally intuitively, the right thing to do would instead be to help them.

Consider how the permissibility of making sacrifices for others, despite a net loss of subjective well-being in the world, could influence the way to programme smart machines. Basically, any time one is in charge of how such a machine is deployed, one should be able to opt to direct it to save others before oneself. Consider a high-ranking military official who can choose a programme for a weapons system in a battlefield. He might direct it to discount his life relative to the lives of two other soldiers, even if the world would be in marginally better shape, in terms of net welfare, with him in it compared to them. Or if one is able to determine how to direct the life-saving procedures of a care robot in the home, one could set them to save the life of one's sickly child before oneself, in case of, say, a fire. Surely, such programming would be morally permissible, contrary to the dictates of utilitarian reasoning.

### CONCLUSION: FROM UTILITARIANISM TO KANTIANISM

Utilitarianism has been the most common western ethic invoked when considering how to programme artificially intelligent automated systems, while Kantianism (Powers 2006; Ulgen 2017; cf. Allen et al. 2000), or at least some kind of deontological approach informed by western values (Anderson et al. 2005; Anderson and Anderson 2007), has been the runner up. If the project here has been found attractive by many readers, it is worth thinking about how it might be extended. If characteristic African values ground powerful objections to programming smart machines according to the principle of utility, might they have similar implications for doing so according to a principle of respect for autonomy (or an ethic of prima facie duties)?

At this stage, the answer in the light of the specific argumentation in this chapter seems equivocal. Kantianism, after all, accepts the ideas that human dignity grounds human rights (section “[Human Dignity and AI](#)”) and also that one may sacrifice one's happiness in order to help others achieve their ends (section “[Self-sacrifice and AI](#)”). However, it might be that the other two African reasons to reject utility maximization apply with comparable force to Kantianism. That is, if one believes that there are collective rights, it might be that they are not reducible to the dignity of individual persons (section “[Group Rights and AI](#)”). And if one believes that there are special obligations, it might be that they are not a function

of promise making or accepting benefits from a cooperative scheme, the central ways that Kantians account for positive duties to specific individuals (section “[Family First and AI](#)”). It is worth considering in future work whether these African values provide grounds to doubt that AI should act in accordance with Kantian duty or whether there are other resources in the sub-Saharan ethical tradition to ground such doubt.<sup>7</sup>

## REFERENCES

- Access Partnership. (2018). Artificial Intelligence for Africa: An Opportunity for Growth, Development, and Democratisation. Retrieved May 11, 2020, from <https://www.accesspartnership.com/artificial-intelligence-for-africa-an-opportunity-for-growth-development-and-democratisation/>.
- African Union. (2006). Charter for African Cultural Renaissance. Retrieved May 6, 2020, from <https://au.int/en/treaties/charter-african-cultural-renaissance>.
- Ake, C. (1987). The African Context of Human Rights. *Africa Today*, 34, 5–12.
- Allen, C., Varner, G., & Zinser, J. (2000). Prolegomena to Any Future Artificial Moral Agent. *Journal of Experimental and Theoretical Artificial Intelligence*, 12, 251–261.
- Anderson, M., & Anderson, S. L. (2007). Machine Ethics: Creating an Ethical Intelligent Agent. *AI Magazine*, 28, 15–26. Retrieved August 12, 2020, from <https://www.aaai.org/ojs/index.php/aimagazine/article/view/2065>.
- Anderson, M., Anderson, S. L., & Armen, C. (2005). Towards Machine Ethics: Implementing Two Action-Based Ethical Theories. Retrieved July 15, 2020, from <https://www.aaai.org/Papers/Symposia/Fall/2005/FS-05-06/FS05-06-001.pdf>.
- Appiah, A. (1998). Ethical Systems, African. In E. Craig (Ed.), *The Routledge Encyclopedia of Philosophy*. London: Routledge.
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The Moral Machine Experiment. *Nature*, 563, 59–64.
- Bauer, W. (2020). Virtuous vs. Utilitarian Artificial Moral Agents. *AI & Society*, 35, 263–271.
- Bonnefon, J.-F., Shariff, A., and Rahwan, I. (2015). Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars? Retrieved October 4, 2020, from <https://arxiv.org/pdf/1510.03346v1.pdf>.

<sup>7</sup>For comments on talks based on prior drafts of this work, I thank Professor Tshilidzi Marwala and audience members at a debate between myself and Professor Marwala about his essay ‘On Rationality: An Artificial Intelligence Perspective’ that was organized by Johannesburg Institute for Advanced Study in 2019, as well as audience members at the 4th Centre for Artificial Intelligence Research (CAIR) Symposium in 2019.

- Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The Social Dilemma of Autonomous Vehicles. *Science*, 352(6293), 1573–1576.
- Botman, H. R. (2000). The OIKOS in a Global Economic Era. In J. Cochran & B. Klein (Eds.), *Sameness and Difference: Problems and Potentials in South African Civil Society*. Washington, DC: The Council for Research in Values and Philosophy. Retrieved July 20, 2020, from <http://www.crvp.org/publications/Series-II/6-Contents.pdf>.
- Bujo, B. (2001). *Foundations of an African Ethic: Beyond the Universal Claims of Western Morality*. Trans. Brian McNeil. New York: Crossroad Publishers.
- Cobbah, J. (1987). African Values and the Human Rights Debate. *Human Rights Quarterly*, 9, 309–331.
- Deng, F. (2004). Human Rights in the African Context. In K. Wiredu (Ed.), *A Companion to African Philosophy* (pp. 499–508). Oxford: Blackwell.
- Etieyibo, E. (2017). Anthropocentrism, African Metaphysical Worldview, and Animal Practices. *Journal of Animal Ethics*, 7, 145–162.
- Gloor, L. (2016). Suffering-focused AI Safety. Retrieved March 2, 2020, from <https://longtermrisk.org/files/suffering-focused-ai-safety.pdf>.
- Grau, C. (2005). There Is No ‘I’ in ‘Robot’: Robotic Utilitarians and Utilitarian Robots. Retrieved March 2, 2020, from <https://www.aaai.org/Papers/Symposia/Fall/2005/FS-05-06/FS05-06-007.pdf>.
- Gwagwa, A. (2019a). Artificial Intelligence Adoption and Use Cases in Africa. Retrieved March 2, 2020, from [https://www.academia.edu/42070923/Artificial\\_Intelligence\\_adoption\\_and\\_use\\_cases\\_in\\_Africa](https://www.academia.edu/42070923/Artificial_Intelligence_adoption_and_use_cases_in_Africa).
- Gwagwa, A. (2019b). Artificial Intelligence (AI) and Algorithms Ethics in Africa. Retrieved March 2, 2020, from [https://www.academia.edu/41800605/Artificial\\_Intelligence\\_AI\\_and\\_Algorithms\\_Ethics\\_in\\_Africa](https://www.academia.edu/41800605/Artificial_Intelligence_AI_and_Algorithms_Ethics_in_Africa).
- Gyekye, K. (1997). *Tradition and Modernity: Philosophical Reflections on the African Experience*. New York: Oxford University Press.
- Gyekye, K. (2010). African Ethics. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved March 6, 2020, from <http://plato.stanford.edu/archives/fall2010/entries/african-ethics/>.
- Hibbard, B. (2015). *Ethical Artificial Intelligence*. Retrieved March 6, 2020, from <https://arxiv.org/pdf/1411.1373.pdf>.
- Ikuenobe, P. (2016). The Communal Basis for Moral Dignity: An African Perspective. *Philosophical Papers*, 45, 437–469.
- Ilesanmi, S. (2001). Human Rights Discourse in Modern Africa. *Journal of Religious Ethics*, 23, 293–322.
- Iroegbu, P. (2005). Right to Life and the Means to Life: Human Dignity. In P. Iroegbu & A. Echekwube (Eds.), *Kpim of Morality Ethics* (pp. 446–449). Ibadan: Heinemann Educational Books.

- Kinjo, K., & Ebina, T. (2017). Optimal Program for Autonomous Driving under Bentham- and Nash-Type Social Welfare Functions. *Procedia Computer Science*, 112, 61–70.
- Majot, A., & Yampolskiy, R. (2014). AI Safety Engineering Through Introduction of Self-Reference into Felicific Calculus via Artificial Pain and Pleasure. *Proceedings of the 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering*. Chicago, 1–6.
- Marwala, T. (2014). *Artificial Intelligence Techniques for Rational Decision Making*. Cham: Springer.
- Marwala, T. (2017). Rational Choice and Artificial Intelligence. Retrieved June 19, 2020, from <https://arxiv.org/pdf/1703.10098.pdf>.
- Metz, T. (2013). *Meaning in Life: An Analytic Study*. Oxford: Oxford University Press.
- Metz, T. (2015). How the West Was One: The Western as Individualist, the African as Communitarian. *Educational Philosophy and Theory*, 47, 1175–1184.
- Molefe, M. (2019). *An African Philosophy of Personhood, Morality, and Politics*. Cham: Palgrave Macmillan.
- Moore, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
- Nozick, R. (1974). *Anarchy, State and Utopia*. New York: Basic Books.
- Nyholm, S., & Smids, J. (2016). The Ethics of Accident-Algorithms for Self-Driving Cars: An Applied Trolley Problem? *Ethical Theory and Moral Practice*, 19, 1275–1289.
- Oosterheld, C. (2015). Machine Ethics and Preference Utilitarianism. Retrieved June 19, 2020, from <https://reducing-suffering.org/machine-ethics-and-preference-utilitarianism/>.
- Oosterheld, C. (2016). Formalizing Preference Utilitarianism in Physical World Models. *Synthese*, 193, 2747–2759.
- Organization of African Unity. (1981). African (Banjul) Charter on Human and Peoples' Rights. Retrieved June 17, 2020, from [https://www.achpr.org/public/Document/file/English/banjul\\_charter.pdf](https://www.achpr.org/public/Document/file/English/banjul_charter.pdf).
- Ormond, E. (2020). The Ghost in the Machine: The Ethical Risks of AI. *The Thinker*, 83, 4–11. Retrieved June 19, 2020, from [http://www.thethinker.co.za/resources/Thinker\\_83\\_new/83%20ormonde.pdf](http://www.thethinker.co.za/resources/Thinker_83_new/83%20ormonde.pdf).
- Powers, T. (2006). Prospects for a Kantian Machine. *IEEE: Intelligent Systems*, 22, 46–51.
- Railton, P. (1984). Alienation, Consequentialism, and the Demands of Morality. *Philosophy and Public Affairs*, 13, 134–171.
- Ramose, M. (2003). The Ethics of *Ubuntu*. In P. H. Coetzee & A. P. J. Roux (Eds.), *The African Philosophy Reader* (2nd ed., pp. 379–389). Cape Town: Oxford University Press.
- Sallstrom, L., Morris, O., & Mehta, H. (2019). Artificial Intelligence in Africa's Healthcare: Ethical Considerations. *Observer Research Foundation Issue Brief*

- No. 312. Retrieved June 19, 2020, from <https://www.orfonline.org/research/artificial-intelligence-in-africas-healthcare-ethical-considerations-55232/>.
- Sen, A. (2000). Consequential Evaluation and Practical Reason. *The Journal of Philosophy*, 97, 477–502.
- Shulman, C., Tarleton, N., & Jonsson, H. (2009). Which Consequentialism? Machine Ethics and Moral Divergence. In C. Reynolds & A. Cassinelli (Eds.), *Proceedings of AP-CAP 2009: The Fifth Asia-Pacific Computing and Philosophy Conference* (pp. 23–25). AP-CAP. Retrieved August 4, 2020, from <https://intelligence.org/files/WhichConsequentialism.pdf>.
- Slote, M. (1985). *Common Sense Morality and Consequentialism*. London: Routledge and Kegan Paul.
- Spini, L. (2019). Ethics of Artificial Intelligence – Global Considerations and Potential for Africa. Presentation Made at the World Science Forum Held in Budapest. Retrieved September 1, 2020, from <https://worldscienceforum.org/programme/2019-11-21-thematic-sessions-ii-c-ethics-of-artificial-intelligence-global-considerations-and-potential-for-africa-129>.
- Ulgen, O. (2017). Kantian Ethics in the Age of Artificial Intelligence and Robotics. *Questions of International Law* 43: 59–83. Retrieved September 4, 2020, from [http://www.qil-qdi.org/wp-content/uploads/2017/10/04\\_AWS\\_Ulgen\\_FIN.pdf](http://www.qil-qdi.org/wp-content/uploads/2017/10/04_AWS_Ulgen_FIN.pdf).
- Wiredu, K. (1992). The Moral Foundations of an African Culture. In K. Wiredu & K. Gyekye (Eds.), *Person and Community; Ghanaian Philosophical Studies, I* (pp. 193–206). Washington, DC: Council for Research in Values and Philosophy.
- Wiredu, K. (1996). *Cultural Universals and Particulars: An African Perspective*. Bloomington: Indiana University Press.
- Wiredu, K. (2009). An Oral Philosophy of Personhood. *Research in African Literatures*, 40, 8–18.
- World Wide Web Foundation. (2017). Artificial Intelligence: Starting the Policy Dialogue in Africa. Retrieved September 16, from <https://webfoundation.org/research/artificial-intelligence-starting-the-policy-dialogue-in-africa/>.