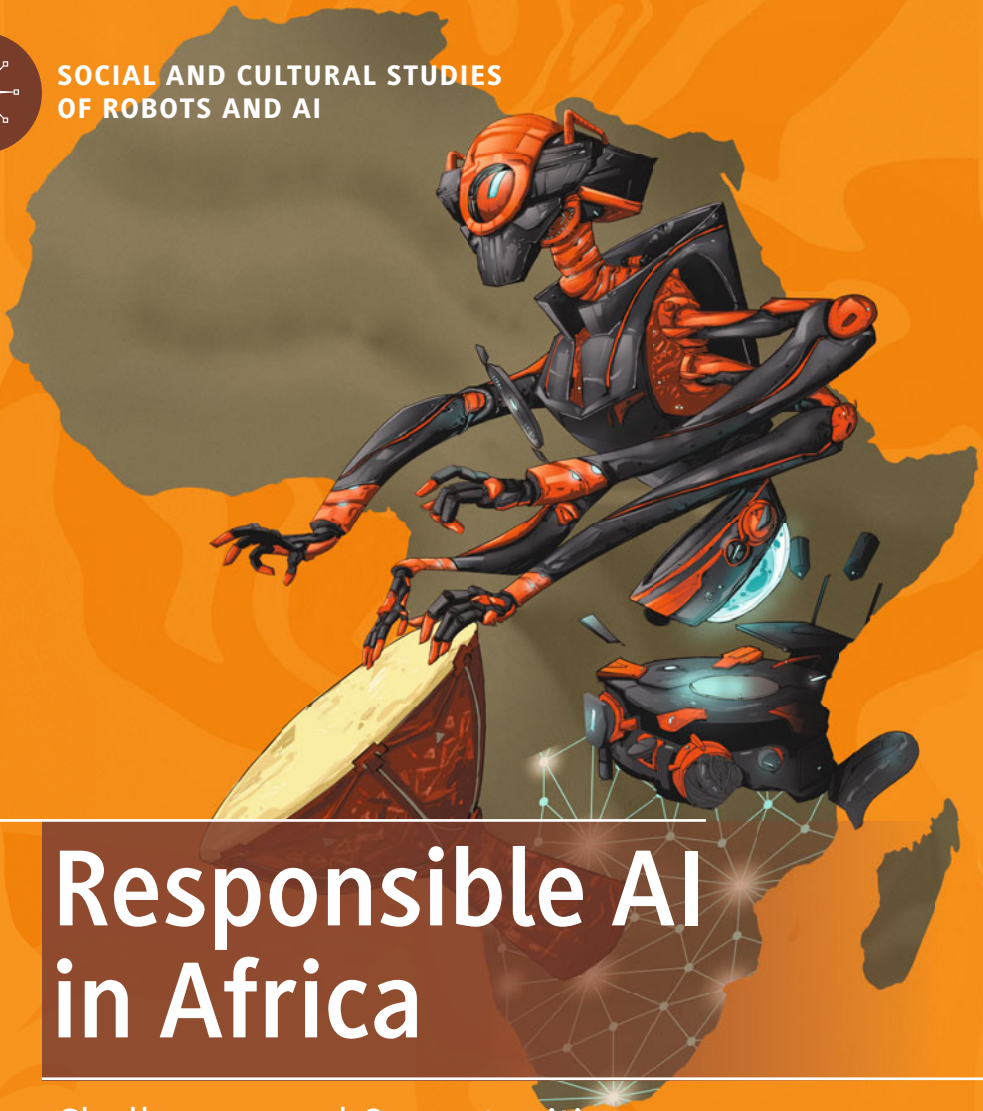
SOCIAL AND CULTURAL STUDIES OF ROBOTS AND AI



Responsible AI in Africa

Challenges and Opportunities

Edited by
Damian Okaibedi Eke
Kutoma Wakunuma
Simisola Akintoye

OPEN ACCESS

palgrave
macmillan

Social and Cultural Studies of Robots and AI

This is a groundbreaking series that investigates the ways in which the "robot revolution" is shifting our understanding of what it means to be human. With robots filling a variety of roles in society—from soldiers to loving companions—we can see that the second machine age is already here. This raises questions about the future of labor, war, our environment, and even human-to-human relationships.

Damian Okaibedi Eke · Kutoma Wakunuma ·
Simisola Akintoye
Editors

Responsible AI in Africa

Challenges and Opportunities

palgrave
macmillan

Editors
Damian Okaibedi Eke
De Montfort University
Leicester, UK

Kutoma Wakunuma
De Montfort University
Leicester, UK

Simisola Akintoye
Department of Law
De Montfort University
Leicester, UK

ISSN 2523-8523 ISSN 2523-8531 (electronic)
Social and Cultural Studies of Robots and AI
ISBN 978-3-031-08214-6 ISBN 978-3-031-08215-3 (eBook)
https://doi.org/10.1007/978-3-031-08215-3

Cover illustration: © Oluwatimilehin E. Makinde (Ajani)

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

The ethics of Artificial Intelligence (AI) is a popular topic. But Africa is usually not on the radar when it comes to academic discussions about AI ethics and AI policy, not even when it comes to "global" and intercultural approaches. This "forgetting" is likely due to biases and stereotypes about Africa on the part of Western interlocutors. To those who believe that Africa has little to do with high tech and innovation, a title such as "Responsible AI in Africa" sounds almost like an oxymoron, and at best comes across as a marginal topic.

For many decades, if not centuries, Africa has been synonymous with *problems* in the Western media and imagination. I remember how in the 1980s TV images of starving children in Africa flooded our living room. This was the first thing I heard about Africa as a child. Africa became synonymous with poverty, droughts, war and (other) human rights violations. Underdevelopment. Challenges. Disaster. Today, American or European mainstream media hardly report about Africa. Unless you specifically search for it, you don't hear about it at all. Consequently, for most people in the West, including policymakers, Africa is a kind of *terra incognita*. We don't know about it, and as the lucky but also somewhat tragic heirs of the Enlightenment, we risk to see it as a place where ignorance thrives. Many Westerners do not know *any* African philosopher or scientist. Africa is often absent on Western maps of the global knowledge and innovation landscape. Development and deployment of high tech such as AI are supposed to happen elsewhere: in

Silicon Valley, in Asia, sometimes in Europe. And in terms of global ethics, all humanities people seem to talk about are colonialism and neocolonialism. Again: Africa is connected with problems. At worst, it is the sad setting of horrible narratives about slavery, colonialism and exploitation. At best, there is the *hope* that Africa, like a sleeping beauty, may awaken and develop in the far future—perhaps kissed by the prince of foreign investment.

While there is a grain of truth in this picture of Africa, it is also a very one-sided and distorted one. Africa is also a place where AI is used and developed: in social networks, in businesses, in healthcare, in agriculture, in education and academic contexts. Of course, as several chapters in this volume show, there are challenges. There is sometimes insufficient digital literacy, lack of data infrastructure, inadequate policies and lack of enough funds. And like in Western countries, there are risks related to bias and discrimination, responsibility, the future of work and climate change (Coeckelbergh 2020). But AI is becoming more popular and, like everywhere, AI is already changing the world—also in Africa. There are start-ups, there is research, and there is innovation. Africa is also a place where the future is made.

One country that has understood this is China, which since years has been pouring loans and investments into Africa, benefiting manufacturing, agro industry, telecommunication and infrastructures. Soft power. When the West will realise that Africa is also a market and a trade partner, other players will already have benefited—as do the Africans themselves already, for example via local businesses and employment. Although here there is of course the danger of neocolonialism and empire, with China taking over the role of European ex colonial powers. More generally, Ethiopian-born cognitive scientist Abeba Birhane (2020) has rightly warned for what she calls "the algorithmic colonization of Africa": the "invasion" of AI may echo colonialism by neglecting local interests and disadvantaging minority groups. And Africa remains a key provider of natural resources to the rest of the world, which often involves exploitation and violence.

But like other high tech, AI is there and is there to stay, and it also offers chances. Partly, this remains a matter of foreign investment. In April 2019 Google opened its first African AI research centre in Ghana. Africa is also a potential, an opportunity. One may even worry that the world may miss out on *talent* from Africa, also for the development of new technologies. That's a different picture of Africa: Africa as a chance, an

opportunity and potential laboratory for creating the future: in the first place for and by Africans themselves, but also for the benefit of other countries and humanity. AI can potentially empower Africa and Africans and lead to more prosperity for all. Yet this assumes that the right kind of socio-economic conditions are created, that the technology is employed in a socially just and democratic way.

Moreover, there are also opportunities to reach new users who simply want entertainment, as the Indian anthropologist Payal Arora has shown in her book *The Next Billion Users* (2019): not all narratives about the Global South have to be about "development" and "economic growth", digital technologies can also be used for entertainment *and that's OK*. Africans and others in the Global South also deserve to have fun! This means, again, that Africa is a business opportunity, for example for AI, but also a place where people want the same things *everyone else wants on the planet*: freedom and integrity, well-being and social security, inclusion, justice and democracy; but also: play, dance, relaxation and simply having a good time. AI can contribute to *that*, too, for example in the context of games.

Yet Africa is not a country; it is a continent (in fact the second largest on earth), and a very diverse one. Consider just the cheerful linguistic diversity. There are an estimated 1500 to 2000 African languages, ranging from Arabic and Berber languages in Algeria and Morocco to Yoruba, Swahili, Lingala, Wolof and Kiswahili in other parts of Africa, and the colonial Afrikaans, French, English and Portuguese. And within countries there are also several languages; many Africans are multilingual. Today, only native anglophones, supported by the global postcolonial and empire-style dominance of English, can afford the luxury of monolingualism. This is a convenience but also a problem. Africa is not only rich in resources, but also in culture. This raises the question: what does its diversity in terms of language, ethnicity, culture, etc., mean for thinking about the ethics of AI?

As suggested earlier, ethics is partly a universal matter. We are all human beings and share similar needs and aspirations. A global ethics can refer to universal principles, for example universal human rights. This is also the case in ethics and policy of AI. However, it is equally important to cater for diversity and respect cultural differences. And preferably more than respect, which always keeps a distance, does not necessarily engage: it is even better to *learn* from other cultures. With regard to AI, this means that African people are invited to explore what they can

do with Western ethical principles and policy frameworks (e.g. the AI ethics frameworks developed by the High-Level Expert Group on AI of the European Commission (2019), which I had the honour to contribute to), but *also* that non-Africans in America, Europe, Asia, etc., should try to *learn from African ethics and African approaches to policy*. From the non-African side, this can be done in order to pay lip service to in-group political correctness or to support domestic identity politics, but it can also be motivated by a genuine interest in African ethics and its values. Perhaps African ethics can even be one of the sources for the renewal of Western ethics. For example, one could argue that the latter is too individualistic and not social enough, and that Westerners do not sufficiently respect family, truth and old age—just to name some values that are often identified as African. There may also be interesting links between character-oriented African ethics and virtue ethics. And of course there are different kinds of ethics in different cultures and different parts of Africa. African ethics is not one thing; it is about people, communities and places. AI ethics in and for Africa should engage with that cultural diversity and explore what AI ethics and robot ethics means in those specific cultural contexts. For example, in this volume, Kwanya writes about the concept of work in the Kenyan context and Dignum suggests that AI ethics can benefit from, for example, Ubuntu philosophy. Finally, it may also be helpful to research how tech innovation and use of digital technologies actually works in African contexts, including cases in which that might *already* be responsible. There are examples of good strategies, inside and outside Africa. Ethics should not only be about what could go wrong, but also about what already goes right and why, and about how we can shape a good common future.

These comments also raise the question: *who* should develop responsible AI in Africa? The answer must be: in the first place Africans themselves. Neocolonialism is game over, and AI ethics should not be the place to perpetuate or renew it. Unfortunately, as Stahl et al. remind us in this volume, AI strategies are often dominated by powerful global actors. Such structural problems need to be addressed. Yet it should not be an excuse to refrain from developing one's own policy ecosystem with regard to AI and similar technologies. This is about creating opportunities for responsible innovation. This is about harnessing the benefits of AI for citizens. This is about taking responsibility for one's own technological future.

Luckily, many Africans did not wait for Western voices or money to do this. But as the editors of this volume suggest, there is a need for more strategy and policy in this area, also with regard to ethical AI. More attention needs to be paid to the ethically and societally responsible development and deployment of AI. Many contributors to this volume help us to think about what that means and how it might be achieved in Africa. AI should not just be about money and business. It is also about justice and solidarity. It is about gender equity. It is about cultural engagement. It is about human rights. It is about good government. It is about communities. It is about what kind of world we want to live in. Ensuring this is in the first place a task and responsibility for African citizens and their leaders in civil society, business and politics. But non-Africans (and those who have a foot in both worlds!) can help with this, as equal partners. Business and trade partners, but also conversation partners in academia and elsewhere: fellow humans looking for the good life and the good society in local and global contexts, given new technological possibilities and challenges. Being from the non-African part of the world, and as a philosopher specialised in the use of words and aware of the danger of neocolonial interventions and hegemonic discourse, I can only say: let's get together and talk about this. Let's talk about what African ethics of AI and global ethics of AI means and could mean. Let's try to learn from each other's experiences and exchange thoughts. This book is hospitable to such dialogues and offers valuable and pioneering material for those much needed discussions, imaginations, strategies, policies and perhaps even *innovations* in both African and global AI ethics and policy.

January 2022

Mark Coeckelbergh
University of Vienna
Vienna, Austria

References

Arora, Payal. (2019). *The Next Billion Users: Digital Life Beyond the West*. Cambridge, MA: Harvard University Press.

Birhane, Abeba. (2020). "Algorithmic Colonization of Africa." *SCRIPTed: A Journal of Law, Technology, & Society* 17(2). Available at: https://script-ed.org/article/algorithmic-colonization-of-africa/

Coeckelbergh, Mark. (2020). *AI Ethics*. Cambridge, MA: MIT Press.

High-Level Expert Group on Artificial Intelligence. (2019). *Ethics Guidelines for Trustworthy AI*. Brussels: European Commission. Available at: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Mark Coeckelbergh is Professor of Philosophy of Media and Technology and Vice Dean of the Faculty for Philosophy and Education at the University of Vienna. Previously he was President of the international Society for Philosophy and Technology and member of the High-Level Expert Group on AI of the European Commission. He is the author of numerous books including *AI Ethics* (MIT Press), *Introduction to Philosophy of Technology* (Oxford University Press), and *The Political Philosophy of AI* (Polity Press).

Acknowledgements

We would like to thank everyone who has provided support and assistance towards the successful publication of this book. Particular thanks are due to colleagues at De Montfort University, UK, particularly members of the Centre for Computing and Social Responsibility, who provided stimulating conversations that helped the formulation of the initial ideas of this book. We are also indebted to our colleagues and collaborators from the EU *Human Brain Project*. We also owe special thanks to Sarah Hills and Rachael Ballard who expertly supported this work as handling editor and series editor, respectively.

Finally, **Damian Okaibedi Eke** thanks his family; Wife—Linda and Children; Chizitere, Chidimma, Chizurumoke and Nnabuihe who provided Daddy the conducive environment to work on this book.

Kutoma Wakunuma thanks her family, particularly her beloved daughter Angelica who is always loving, patient and understanding of Mummy's sometimes busy schedule.

Simisola Akintoye thanks her husband for the love and support provided during this process. She is also thankful to her three boys—Enoch, Ethan and Elias who have been nothing but loving boys to mum during this time.

This research received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under Grant Agreements No. 945539 (Human Brain Project SGA3).

Contents

Notes on Contributors

Prof. Ade-Ibijola Abejide holds a Ph.D. in Computer Science (Artificial Intelligence) from the University of the Witwatersrand, Johannesburg, South Africa. He is currently an Associate Professor of Artificial Intelligence and Applications in the Department of Applied Information Systems at the University of Johannesburg. He is rated by the National Research Foundation (NRF) of South Africa and mainly interested in Problems and Algorithms in AI, and Innovations for driving the 4th Industrial Revolution.

Dr. Akintoye Simisola is the Data Protection Officer for the European Union Future and Emerging Technologies Human Brain Project where she is the task lead for compliance, data protection, data governance and data management. She is an international privacy practitioner, data protection consultant and senior lecturer in law at De Montfort University Centre for Law, Justice and Society. She has broad research interests in legal and ethical compliance of organisations with applicable laws. Her work involves continuous, up-to-date collaborative research in privacy, ethics and responsibility. As a legal expert, she regularly sits on panels involved in critical dialogues at both national and international levels that explores issues around balancing the competing interests of research and innovation with privacy and data protection. Her current research investigates the changing role of Africa's emerging digital transformation. She is currently a co-chair of AI Policy in Africa at De Montfort University Centre for Computing and Social Responsibility.

Aruleba Kehinde holds a Ph.D. in Computer Science from the University of the Witwatersrand, Johannesburg, South Africa. He is currently a lecturer at the University of Leicester, United Kingdom. Before that, he was a postdoctoral fellow at Walter Sisulu University, South Africa. His expertise extends beyond computer science. He actively collaborates with researchers across disciplines. His research interest is in machine learning, data ethics and ICT4D.

Borokini Favour is an AI and tech policy researcher. She currently works as a Data and Digital Rights Researcher with Pollicy, an award-winning Ugandan feminist collective and civic technology organisation, where she researches questions and problems posed by AI and other technologies on social justice and inequality in Africa, with an emphasis on African women. She also researches with the Africa Data Governance Initiative (ADGI), Responsible Data Governance for Artificial Intelligence (REDGOVAI), and is a content writer for Ethical Intelligence, a UK technology ethics consulting firm.

Chintu Schmidt Shilukobo Ph.D. is a Zambian born Mathematician and Computing Scientist who researches in Artificial Intelligence, Quantum Computing, Vehicular Networks, the Braess's paradox and Multi-Objective Optimisation (Vector Optimisation). He earned his Ph.D. in the Department of Computing and Mathematical Sciences at the University of Greenwich in the United Kingdom. He was until recently the regional head of the British Computer Society in Southern and Eastern Africa. He previously held a position as one of the directors of the Institute for the Management of Information Systems in the UK where he worked closely with Professor Simon Rogerson who was the Institute's chairperson. He is the C.E.O. and head tutor at Cyber Academy Zambia a college he founded in 2000.

Dignum Virginia is Professor of Responsible Artificial Intelligence at Umeå University, Sweden, and associated with the TU Delft in The Netherlands. She is the director of WASP-HS, the Wallenberg Program on Humanities and Society for AI, Autonomous Systems and Software, the largest Swedish national research programme on fundamental multidisciplinary research on the societal and human impact of AI. She is a member of the Royal Swedish Academy of Engineering Sciences, and a Fellow of the European Artificial Intelligence Association (EURAI). Her current research focus is on the specification, verification and monitoring

of ethical and societal principles for intelligent autonomous systems. She is committed to policy and awareness efforts towards the responsible development and use of AI, as member of the European Commission High-Level Expert Group on Artificial Intelligence, the working group on Responsible AI of the Global Partnership on AI (GPAI), the World Economic Forum's Global Artificial Intelligence Council, lead for UNICEF's guidance for AI and children, the Executive Committee of the IEEE Initiative on Ethically Aligned Design, and as founding member of ALLAI, the Dutch AI Alliance. Her book *Responsible Artificial Intelligence: Developing and Using AI in a Responsible Way* was published by Springer-Nature in 2019. She studied at the University of Lisbon and the Free University of Amsterdam and has a Ph.D. in Artificial Intelligence from Utrecht University in 2004. In 2006, she was awarded the prestigious Veni grant by the NWO (Dutch Organization for Scientific Research). She is a well-known speaker on the social and ethical impacts of Artificial Intelligence, is member of the reviewing boards for all major journals and conferences in AI and has published over 250 peer-reviewed papers.

Eke Damian Okaibedi is a Research Fellow at the Centre for Computing and Social Responsibility at De Montfort University, UK. He is a Philosopher and has a Ph.D. in Computer Ethics. He is the Data Governance coordinator for the EU Human Brain Project and leads the International Brain Initiative (IBI) Data Sharing and Standards Working Group Task Force on International Data Governance. His research interests cover critical philosophical issues at the intersection of Technology, Information and Society including; Data Governance, Ethics of Emerging Technologies, Responsible Innovation and ICT4D.

Kwanya Tom is a professor of Knowledge Management at the Technical University of Kenya. He is currently also serving as the Director of the School of Information and Communication Studies. Prior to joining academics fulltime in 2013, he worked as a consultant on public information and knowledge management. He is the author of several peer-reviewed journal articles and book chapters, two university-level books, several conference papers and four edited books. He has also supervised and examined several postgraduate dissertations and theses. He has served as an external examiner of information science in six universities in Kenya. He sits in the Information Science Services Technical Committee of the Kenya Bureau of Standards (KEBS).

Leach Tonii is a Research Assistant at the Centre for Computing and Social Responsibility, De Montfort University, Leicester, UK. Her current role is in the Human Brain Project (HBP)—an EU funded Future Emerging Technology flagship. Her research interests cover a wide range of topics including perceptions and narratives of AI, ethics of emerging technologies and responsible research and innovation.

Obaido George is a Postdoctoral Scholar in the Department of Computer Science and Engineering, Jacobs School of Engineering, University of California, San Diego, United States. He received his Ph.D. in Computer Science at the University of the Witwatersrand, Johannesburg, South Africa. He also obtained his M.Sc. in Computer Science from the same University. His research interests lie at the intersection of utilising machine learning, deep learning and formal methods to solve many societal problems.

Ogoh George is a Research Fellow at the Centre for Computing and Social Responsibility, De Montfort University. His current role is in the Human Brain Project (HBP)—an EU funded Future Emerging Technology flagship. His research interests cover a wide range of topics in emerging technology ethics including responsible innovation, data governance and data protection.

Okolo Chinasa T. is a Ph.D. Candidate in the Department of Computer Science at Cornell University. Her research interests include AI ethics, computer vision, global health and information & communication technologies for development. Her research leverages ethnographic methods to understand how frontline healthcare workers in low-resource regions perceive and value artificial intelligence (AI). She also examines FATE (Fairness, Accountability, Transparency, and Ethics) in AI-enabled technologies deployed throughout the Global South, with a focus on health care.

Dr. Okonkwo Chinedu holds a Ph.D. in Computer Science from the North-West University, Potchefstroom Campus in South Africa, and is currently a Research Fellow at the University of Johannesburg in South Africa. His research interests include Applications of Artificial Intelligence in Education, Computer Science Education and Technology Adoption. He has actively collaborated with researchers from a variety of other disciplines on Artificial Intelligence.

Oyeniji Oluyinka is a Doctoral Researcher at the Center for Computing and Social Responsibility, De Montfort University. His area of research interests include ethics of emerging technologies, neuroscience and neurotechnologies, artificial intelligence and cloud computing.

Ruttkamp-Bloem Emma is head of the Department of Philosophy at the University of Pretoria. She is the coordinator of the Ethics of AI Research Group at the Centre for AI Research (CAIR) in South Africa and co-chair of the Steering Committee for the Southern African Conference for AI Research (SACAIR). She is the current rapporteur of the UNESCO World Commission for Ethics of Scientific Knowledge and Technology (COMEST). She chaired the UNESCO Ad Hoc Expert Group that prepared the initial draft of a Recommendation for a global instrument on the ethics of AI in 2020 that was adopted by member states in November 2021. She is the South African representative at the Responsible AI Network Africa (RAIN) and the South African country advisor at the International Group of AI (IGOAI). She is a member of various advisory boards such as the Wallenberg AI, Autonomous Systems and Software Programme (Human Sciences) hosted by Umeå University in Sweden, the Global AI Ethics Institute and the *Z-Inspection*® Advisor Board. She has been the South African representative at the International Union of the History and Philosophy of Science and Technology (IUHPST) since 2014. She is an associate editor of *Science and Engineering Ethics* and a member of the editorial board of Springer's *Synthese Library Book Series* and the editorial board of *Acta Baltica: Historiae et Philosophiae Scientiarum*. She has a Ph.D. in Philosophy in the domains of mathematical logic and the philosophy of science. Currently, in AI ethics, she works in machine ethics, the ethics of social robotics, data ethics and the philosophy of human-technology relations. In terms of policymaking, she focuses on generating culturally sensitive policies for trustworthy AI technologies while aiming for global regulation. In the philosophy of science, her work is centred on debates in scientific realism, the structure of scientific theories and the status of machine learning-based methodologies in the discovery/justification debate.

Stahl Bernd Carsten is Professor of Critical Research in Technology and Director of the Centre for Computing and Social Responsibility at De Montfort University, Leicester, UK. His interests cover philosophical

issues arising from the intersections of business, technology and information. This includes ethical questions of current and emerging information and communication technologies, critical approaches to information systems and issues related to responsible research and innovation.

Dr. Wakunuma Kutoma is Associate Professor Research and Teaching in Information Systems at De Montfort University where she works within the Centre for Computing and Social Responsibility. Her research interests are in understanding the social and ethical implications of ICTs and the role that emerging technologies like Artificial Intelligence play in both the developed and developing world. In particular, her research work has focussed on Responsible Innovation, computing ethics, digital innovation and plastic waste, ICT for development and gender. In addition to projects funded by other agencies, she has also been involved in several EU-funded projects focussed on emerging technologies, ethics and Responsible Innovation. She also acts as a European Commission Ethics Expert and Project Evaluator. She is also Subject Group Leader for the Information Systems group and Programme Leader for M.Sc. Computing within the School of Computer Science and Informatics. Her teaching involves the area of ICT for Development, Computing Ethics as well as Responsible Innovation in ICT. She also supervises a number of Ph.D. students and has examined several Ph.D. thesis.

List of Tables

Introducing Responsible AI in Africa

Damian Okaibedi Eke, Kutoma Wakunuma, and Simisola Akintoye

Background

In the last few years, a growing and thriving AI ecosystem has emerged in Africa. Within this ecosystem, there are local tech spaces as well as a number of internationally driven technology hubs and centres established by big tech companies such as Twitter, Google, Facebook, Alibaba Group, Huawei, Amazon and Microsoft have significantly increased the development and deployment of AI systems in Africa. While these tech spaces and hubs are focused on using AI to meet local challenges (e.g. poverty, illiteracy, famine, corruption, environmental disasters, terrorism and health crisis), the ethical, legal and socio-cultural implications of AI in Africa have largely been ignored. To ensure that Africans benefit from

D. O. Eke (✉) · K. Wakunuma · S. Akintoye
De Montfort University, Leicester, UK
e-mail: damian.eke@dmu.ac.uk

K. Wakunuma
e-mail: kutoma@dmu.ac.uk

S. Akintoye
e-mail: simi.akintoye@dmu.ac.uk

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_1

the attendant gains of AI, ethical, legal and socio-cultural impacts of AI need to be robustly considered and mitigated.

On the global level, a number of national, regional and international bodies, think-tanks, research institutions and private companies have developed or are in the process of developing ethical principles and guidelines for AI (Jobin et al., 2019; Ulnicane et al., 2021). These emerging principles such as transparency, justice and fairness, non-maleficence, responsibility and privacy that shape global AI ethics discourse are informed by ethical perspectives and traditions from Western Europe, North America and East Asia (Gupta and Heath, 2020). Ethical narratives, perceptions and principles from the Global South, particularly Africa, are glaringly missing from the global discussion of AI ethics. There is a general belief that socio-cultural and political contexts shape expectations of AI and the challenges and risks it poses. It is therefore safe to assume as Hargety and Rubinov (2019) suggested that AI ethics concepts such as 'bias', 'human rights', 'privacy', 'justice', 'solidarity', 'trust', 'transparency', 'openness' and 'fairness' mean different things to different people. The meaning and scope of these concepts emerge from cultural contexts in which they are discussed. Citing the example of Nordic AI policies, Robinson (2020) notes the fundamental influence cultural values have on the way these concepts are conceptualised in national and regional policies. As he pointed out, cultural values contribute to value-laden technology policies in a way that can address societal concerns and interests that are different in different places. This is at the heart of responsible AI—the idea of developing AI systems that will not only be compliant to laws (including human rights provisions) but that are socially/culturally sensitive and acceptable as well as be ethically responsible.

Indeed, embedding cultural values and beliefs into the development and implementation of AI policies and strategies is an imperative for both AI developers and policymakers. People's contextual understanding of reality must be represented in the design and implementation of the technology to improve acceptability. AI development and use in Africa needs to be sensitive to African cultural values, beliefs and ethical principles which are currently lacking in the global discussion on AI ethics and guidelines. As AI continues to grow in Africa, there is a risk of alienating those for whom these services are meant for, if principles and values from the Global North are imposed on them. Whereas African societies cannot be described as monolithic, African people share common values

steeped in rich ethical traditions, described differently yet similar in many communities, that can shape AI development and governance.

To contribute to this discussion, this book presents cutting-edge research and insights on the current challenges and prospects of developing a Responsible AI in Africa from both African and world-leading scholars in AI ethics. The contributions evaluate the importance of contextual values and principles on the development of effective AI and its ethics, governance and strategy for Africa. The book offers a much-needed African AI narrative that is missing in the global AI ethics and governance discourse. It contains original contributions on the current-state-of-the-art, challenges, prospects, the meaning and scope of Responsible AI in Africa. The book succeeds in advancing our understanding of some specific challenges and concerns AI raises in Africa and provides insights on the African ethical foundations that can help in mitigating specific AI concerns in Africa as well as ensure that AI is developed to meet societal hopes, expectations and needs.

AI in Africa

As noted by Schwab (2016), AI is a significant component of the fourth industrial revolution that will lead to fundamental changes in the way we live, work and relate to one another. PricewaterhouseCoopers AI sizing the price report estimates that by 2030, AI technologies could increase the global economy by $15.7 trillion (14%), with increased productivity level of about $6.6 trillion and consumption side-effects of $9.1 trillion (PricewaterhouseCoopers, 2017). This PWC report also showed that although AI is at its early development, the AI market in Europe, North America and China is more advanced than other regions. To put this into proper perspective, the financial gains for the markets in Africa, Oceania and low-income Asian markets are estimated to be around $1.2 trillion while China is about $7.0 trillion, $3.7 trillion for North America and $1.8 trillion for Northern Europe. These figures indicate that in Africa, AI development and deployment are still at its early stages and face a number of challenges towards being a transformative force in society.

However, the nature of AI promises to bring about fundamentally socio-cultural changes in Africa including in areas such as political activities, poverty, environmental sustainability, transportation, agriculture, health care, education, financial transactions and religious and traditional belief systems. Many of these AI systems are no longer described

as dreams but are becoming a reality in Africa but mainly driven by companies with roots in the Global North. In addition to the big technology companies establishing operations in Africa, home grown experts are increasingly establishing technology spaces similar to the US Silicon Valley and Silicon Wadi in Israel. Most of these tech spaces are aptly named as 'Silicon Savannah' in Kenya, 'Sheba Valley' in Ethiopia and 'Yabacon Valley' in Nigeria. These tech spaces and many African networks (such as DeepLearning Indaba, Responsible AI network—Africa), local AI start-ups and local stakeholders (including centres of higher education, governments and broader AI community) are fostering a growing ecosystem aimed at developing AI systems that are sensitive to African interests, concerns and culture.

Therefore, AI as a tool or system that performs a specific intelligent task (or otherwise known as artificial narrow intelligence—ANI) is growing and thriving in Africa. However, despite the great benefits these AI systems promise for Africa, there is an appreciation that it is critical to ensure the values and needs of Africa are considered in the design and implementation of these systems. There are also substantial sociocultural and organisational challenges that undermine the adoption and implementation of AI across the continent. This includes lack of digital infrastructure, education, inadequate data, public policies and funding (Kiemde and Kora, 2021, 2020). Thus, for Africa to begin to capitalise on the opportunities of AI, there needs to be cooperation between African stakeholders as well as the establishment of an enabling environment for AI to thrive. This includes structural reform to support innovation, development of effective policies and regulations for digital growth.

Responsible AI

Responsibility denotes accountability and having control and authority for or over something. It is an important aspect that needs to be taken seriously in any technology design, development, implementation and eventual main-stream buy-in. As such, adoption, adaptation, access and use should also be accounted for in Responsible AI. To this end, there ought to be responsibility in AI, more so when it comes to how it is contextualised and applied in an African context. As such, the question needs to be asked with respect to what Responsible AI means in the Global South, with particular respect to the African context. This is especially in the light of the fact that as is known, AI, like many other

technologies, is a western import, designed and developed with mainly western values, yet the technology is expected to be adopted and used in much similar ways as in the Global North. This is despite research showing that AI is not neutral, therefore suggesting that use as well as social and ethical considerations will differ depending on geographical locale, cultural, social and political norms as well as economic standing. Because technologies like AI bring about their own challenges that call for a considerable amount of responsibility, it becomes imperative to understand how Africa is addressing the social and ethical challenges that are brought to light due to the application of AI or the desire to apply AI. To this end, it becomes necessary to understand what Responsible AI means for Africa and how is this considered within the context of ethical challenges that result from its potential adoption and use or is Responsible AI considered at all? In the first instance, Wakunuma et al., (2021) call for the reconceptualisation of the notion of responsible innovation which covers terms like Responsible AI because they have been developed in the Global North with little reference to what these may mean in the Global South. The authors argue that RI should take into account diverse RI practices that may be dependent on community initiatives as well as indigenous knowledge and cultural values. In much similar terms, Carman and Rosman (2021) posit that AI and therefore its ethical considerations should be compatible with the societal values within which they operate from. This is a clear recognition of the fact that as societal value systems differ, so too will the ethical concerns and solutions thereof that may pertain to different societies.

However, solutions to ethical concerns will be challenging to come by if as Gwagwa et al. (2021) have noted that Africa still scores very lowly in its AI readiness and will therefore need to depend on continued support from international partners and technology firms. This dependency does not foster confidence in finding ethical solutions when technology accessibility and implementation are premised on support from others. The dependency can lead to challenges in developing appropriate policies that should speak to Responsible African AI rather than risk the embedding of values developed by the very international technology firms that Africa depends on. Perhaps that is why there are very few African countries with AI policies that robustly address ethical concerns. Further, the fact that there are few African AI experts on the international AI stage also shows a shortage of AI skills—a lack of diversity among those with skills and a lack of financial resources which are much needed to accelerate

the development of AI which does not help with a robust understanding and awareness of ethical concerns of AI that can enable responsible AI in Africa.

Global AI Ethics: African Perspectives

In their comprehensive and robust review of 84 guidelines on Ethical AI published around the world, Jobin et al., (2019) identified 11 overarching ethical principles: transparency, justice and fairness, non-maleficence, responsibility, privacy, beneficence, freedom and autonomy, trust, sustainability, dignity and solidarity. However, the authors admitted that a further thematic analysis revealed "significant semantic and conceptual divergences in both how the eleven ethical principles are interpreted and the specific recommendations or areas of concern derived from each" (Ibid, p. 7). It is important to note that none of the guidelines reviewed was developed in or for African contexts. This points to the fact that the global AI ethics debate is being shaped without Africa in mind. The underlying moral traditions behind the ethical principles shaping this debate therefore emerge from non-African contexts, while the AI applications and tools will potentially be used in African contexts. Since Africa does not lack moral traditions or ethical principles worthy of being considered in the global AI ethics debate, the continued global discourse on Responsible or Ethical AI without perspectives from Africa amounts to *epistemic injustice*. *Epistemic injustice* is a concept that defines unfairly discriminating against one's capacity as a knower (Byskov, 2021). Africa has well-established philosophical and cultural traditions that can provide unique perspectives on identified ethical principles for the design, development and application of AI.

While the rationale behind the lack of African perspectives in the global AI ethics debate is not the focus here, this book serves as a counter to this epistemic injustice and makes a case for why and how African voices, ethics, interests, visions, concerns, expectations and fears should become part of the increasing global discussion on Responsible AI. For AI to be sensitive to African socio-cultural contexts, it is important to consider African perspectives. AI systems are designed to solve problems within contexts. The values, interests and moral traditions of these contexts need to be factored into the design and deployment of any AI technology. Therefore, global Responsible AI frameworks that can make AI align with

diverse societal needs, concerns and interests are needed. African perspectives are and should be critical components of these frameworks. There are two major implications of this. First, it will mean the development of AI applications that respond to African needs, expectations, interests, values and beliefs. Second, it will contribute to epistemic justice in the global AI ethics discourse.

Structure of Volume

In eight main chapters, this book explores the concept of Responsible AI in Africa. In this introductory chapter, the editors introduce the concept of Responsible AI in the African context. It highlights the lack of Africa contexts, voices, interests and values in the global discussion of AI ethics and calls for the reconsideration of the Responsible AI landscape. Following this, Ruttkamp-Bloem starts the book off by making a case for actionable AI ethics in Africa that is driven by dynamic and epistemic just ethical systems. She focuses on the AI ethics policy environment in Africa and concludes that the fast-changing nature of AI technology requires a dynamic AI ethics policy ecosystem characterised by engagement with diverse stakeholders from different backgrounds, interests and values. This contribution highlights the importance of considering African contexts and values (particularly the communitarian concepts of personhood and interconnectedness within a community) in developing actionable AI ethics that ensures trust, social acceptability and cultural sensitivity. It suggests that culture should be the global calculus for AI ethics with regard to being the source for AI ethics and translating into more relatable contexts for communities.

Responsible AI is, however, necessitated by identifiable challenges to AI design and implementation. These challenges differ in different contexts and disciplines. In the third chapter, Okolo et al. highlight the challenges and opportunities AI presents to Africa. Their contribution provides detailed information on principles of Responsible AI and empirically sound evidence of the landscape of AI in Africa. This chapter also raises concern regarding the increasingly aggressive presence of big tech companies in Africa, particularly Chinese companies, which touches on the power imbalance in the AI ecosystem between the Global North and the Global South. It also provides recommendations for improving responsible AI in Africa.

Chapter four focuses on identification of specific ethical perspectives around the possible deployment of an AI system in Kenya, Africa. Kwanya perspectives on ethical concerns related to possible integration of co-bots in workplaces. This chapter explored the perceptions of data scientists in Kenya on ethical issues that can affect the acceptability of co-bots in workplaces. Kwanya's contribution highlights specific socio-cultural concerns, fears and expectations of AI in African societal contexts albeit in Kenya.

In chapter five, Abejide Ade-Ibijola and Chinedu Okonkwo unpack the emerging challenges facing the design and adoption of AI in Africa. The design and adoption of AI face a number of challenges in Africa due to our unique political and socio-cultural contexts. Their contribution not only highlights challenges to widespread design and adoption of AI in Africa such as lack of structured data ecosystem, skills acquisition, relevant policies and ethics, insufficient infrastructure, it provides recommendations for addressing these challenges.

Chapter six touches on the issues of AI and gender. Borokini et al. present critical perspectives on the use of gendered chatbots in commercial banks in Nigeria. Through an analysis of identifiable features of chatbots used by financial institutions in Nigeria, the chapter shows that the majority of available conversational agents are gendered to appear female. The anthropomorphic project of human features/characteristics on these AI applications reinforce gender stereotypes with critical implications for human behaviour. The authors also point out that the increasing use of chatbots raises crucial concerns not only related to gender equality and possible biases against women, but also to the future of work in a society where the unemployment rate is high. Most importantly, this chapter provides recommendations for AI designers on how chatbots can be designed in a way to subvert stereotypes and for policymakers on how to develop policies for gender-inclusive AI designs in Nigeria in particular and African in general.

In Chapter seven, Stahl et al. present AI policy as a response to the need for AI Ethics in Africa. Their contribution is based on the analysis of AI strategies and initiatives available in North Africa. The authors explored how ethical issues are framed and addressed in North African AI strategies. This chapter also highlights the gaps and opportunities in connection with the current AI strategy landscape and suggests how AI policies in Africa should address ethical issues in line with African socio-cultural values.

Further to this, Eke et al., provide critical perspectives on how the future of Responsible AI in Africa can be shaped. This chapter previews the current and future landscape of AI design and deployment in Africa and highlights the unfair neglect of African socio-cultural contexts in the global AI ethics discourse despite the concerns AI raises in African contexts. To achieve a globally just, fair and transparent AI therefore, the authors identified the need to integrate African contexts, interests, values, fears, hopes, expectations and aspirations into AI. They went further to map out what Africans need to do to achieve Responsible AI within and outside of Africa.

Finally, Virginia Dignum discusses the rationale, the scope, nature and limitations of current global efforts in Responsible AI and governance. The chapter begins with a reasoned conceptual clarification and demystification of AI following the well-documented hype that surrounds it. It provides insights into the lessons Responsible AI in Africa can learn from various national and regional AI governance initiatives. Most importantly, Dignum suggests that Responsible AI can benefit from social perspectives embedded in African philosophies such as Ubuntu philosophy. In her words, these social perspectives can ''complement the currently predominant individualistic view of AI systems, to one that acknowledges and incorporates the collective, societal''.

Conclusion

Overall, this book highlights the need for increased discussion and application of African contextual narratives, especially ethics, into AI. It serves as a call for local and international AI stakeholders and professionals to be aware of African narratives and to consider them in the design, development and deployment of the AI applications in Africa. African moral traditions can inform decision-making in AI applications (especially the ones deployed in Africa). To achieve fair, just and transparent AI, it is high time these moral traditions and contexts were considered in the global AI ethics discourse. Ethical AI cannot be achieved in Africa and around the world without consideration of African socio-cultural and ethical contexts or the inclusion of African voices into the global Responsible AI discourse. Failing to consider African contexts and narratives will only lead to ineffective or misguided applications that will neither benefit African societies nor promote human flourishing in Africa. This book, therefore, introduces responsible AI in Africa discourse in a way that will

facilitate culturally sensitive and inclusive AI systems that can improve rather than worsening African societal situations.

References

Byskov, M.F., 2021. What Makes Epistemic Injustice an "Injustice"? J. Soc. Philos. 52, 114–131.https://doi.org/10.1111/josp.12348

Carman, M., Rosman, B., 2021. Applying a principle of explicability to AI research in Africa: should we do it? Ethics Inf. Technol. 23, 107–117.

Gupta, A., Heath, V., 2020. AI ethics groups are repeating one of society's classic mistakes. MIT Technol. Rev. URL https://www.technologyreview.com/2020/09/14/1008323/ai-ethics-representation-artificial-intelligence-opinion/ (accessed 6.29.21).

Gwagwa, A., Kachidza, P., Siminyu, K., Smith, M., 2021. Responsible artificial intelligence in Sub-Saharan Africa: landscape and general state of play.

Hagerty, A., & Rubinov, I. (2019). Global AI ethics: A review of the social impacts and ethical implications of artificial intelligence. ArXiv Prepr. ArXiv190707892.

Jobin, A., Ienca, M., Vayena, E., 2019. The global landscape of AI ethics guidelines. Nat. Mach. Intell. 1, 389–399.https://doi.org/10.1038/s42256-019-0088-2

Kiemde, S.M.A., Kora, A.D., 2021. Towards an ethics of AI in Africa: rule of education. AI Ethics.https://doi.org/10.1007/s43681-021-00106-8

Kiemde, S. M. A., & Kora, A. D. (2020). The challenges facing the development of AI in Africa. In *2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI)*, pp. 1–6. IEEE.

PricewaterhouseCoopers. (2017). PwC's Global Artificial Intelligence Study: Sizing the prize [WWW Document]. PwC. URL https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html (accessed 2.25.22).

Robinson, S.C., 2020. Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). Technol. Soc. 63, 101421. https://doi.org/10.1016/j.techsoc.2020.101421

Schwab, K. (2016). The fourth industrial revolution: What it means, how to respond. World Econ. Forum. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Ulnicane, I., Eke, D.O., Knight, W., Ogoh, G., Stahl, B.C., 2021. Good governance as a response to discontents? Déjà vu, or lessons for AI from other emerging technologies. Interdiscip. Sci. Rev. 46, 71–93.https://doi.org/10.1080/03080188.2020.1840220

Wakunuma, K., Castro, F. de, Jiya, T., Inigo, E.A., Blok, V., Bryce, V., 2021. Reconceptualising responsible research and innovation from a Global South perspective. J. Responsible Innov. 8, 267–291.

Epistemic Just and Dynamic AI Ethics in Africa

Emma Ruttkamp-Bloem

Introduction

I argue that the road to responsible AI ethics in Africa should be paved by regulations driven by epistemic just and dynamic AI ethics systems, rather than only by good intentions. Not only does Africa need dynamic and adaptive AI ethics systems as critical enabler of progress in fast-moving AI technologies, but it also needs such systems to ensure Africa is included—and their contribution understood—in the global conversations needed to speak to the transnational nature of Big Tech Companies and their combined potential threat to all of humanity.

Some of the main motivations for this argument include: The fast-changing nature of AI technologies means that regulation always seems out of step. I thus plead for the acknowledgement of the dynamic role of AI ethics to alert humanity to possible harm from AI technologies and to flag where legal protection is needed. AI ethics as a system that has its ear to the real world perhaps more immediately and intimately than

E. Ruttkamp-Bloem (✉)
University of Pretoria, Pretoria, South Africa
e-mail: emma.ruttkamp-bloem@up.ac.za

Centre for AI Research, Council for Scientific and Industrial Research, Pretoria, South Africa

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_2

International Law does seems eminently suited for this role. Of course the general debate about the priority of ethics vs. the law is an old and complex one, and I regret that I cannot do more here than just place a stake for AI ethics in this debate.

Secondly, a motivation speaking to epistemic justice is that Africa has become the ethical dumping ground of the main players on the AI technology scene. Because of weak regulations and other factors, African states are vulnerable to exploitation by transnational companies of members of poor communities in various ways, from labour exploitation to not being able to determine which data represents Africa and not owning our own data.

A third motivation straddling both the plea to recognise the dynamic nature of AI ethics and the plea for an epistemic just AI ethics system is that full adoption and trust in AI technologies and realisation of the benefits of AI technologies are in a tight reciprocal necessary and sufficient relationship, and in Africa, the potential benefits of AI technologies can introduce real change for many of its inhabitants. Both these reasons speak to the need for AI ethics awareness, sensitivity and literacy. In order to address this need, it is necessary to link AI ethics reflection to the lived world of Africa's inhabitants and to meet them in their own context.

A final essential motivation in terms of epistemic justice is that Africa has been more or less excluded from AI ethics debates globally and Western terminology and approaches dominate the domain. I argue here that there is a real need to be sensitive to the communal nature of African ethics, to rethink typical Western style formulation of AI ethics principles (which does not imply weakening the protection they are intended to offer at all), to take context as well as culture into account and to authentically and effectively integrate ethics with technological innovation in Africa.

In the good company of thinkers such as Aristotle, Stuart Mill, Jeremy Bentham, Paul Ricoeur, Michael Sandel, Robert Nozick, Amartya Sen and Martha Nussbaum, I subscribe to the view that ethics is the medium through which decisions are made about what the right thing to do is (morality). Thus, a moral problem relates to what the right thing to do is, while an ethical problem relates to how best to consider a solution to a moral problem. This is of course a complex debate in moral philosophy, but not one which I can delve into here. But even so, it should be obvious that on this kind of account of ethics there are different ethical systems to choose from in order to engage with moral dilemmas and decisions,

not only in terms of the school of thought a system belongs to (deontological, consequentialist, virtue approaches, communitarian approaches, religious/spiritual approaches, etc.), but also in terms of the values that inform these systems. Ultimately, the choice of ethical system is heavily influenced by culture and also by context.

The particular relevance of these two factors when considering the AI ethics domain in Africa is explained and argued for in Sects. 2 and 3. Obstacles in the way of actionable AI ethics range over different dimensions, some relating to political, economic, social or educational contexts (Sect. 2), while others relate to culture (Sect. 3). By building on arguments in previous sections, I conclude this chapter by suggesting the fast-changing nature of AI technology and the fluidity of the AI readiness[1] of any country, together with the role that culture and context play in adoption of and openness to AI ethics regulations, imply actionable AI ethics in Africa has to be realised in epistemic just and dynamic, adaptive, agile systems.

The African Context and Actionable AI Ethics

When considering factors impacting on adherence to AI ethics policy, both the *context* within which policies are formulated and the context in which they are to be actionalised are relevant for very practical reasons. Factors such as Internet penetration, civil and political stability, quality education and factors as mundane as access to electricity, among many others, all impact on both the implementation of AI ethics frameworks and the appetite for adoption of AI technologies and sensitivity to AI ethics concerns. It is clear that such factors range over a "continuum of scientific, technological, economic, educational, legal, regulatory, infrastructural, societal, and other dimensions" (UNESCO, 2020). These factors essentially determine the "AI readiness" of a country in UNESCO terms (ibid.). It also relates to a novel notion of "national AI capital" (NAIC), suggested by Momčilović (2021) as a country's "capacity to apply and develop, and cope with the challenges of various artificial intelligence systems, in order to increase the country's social and economic well-being and competitiveness" (ibid.). This leads to a definition of a third related concept, the AI ethics capital (AIEC) of a country, as "the

[1] https://unesdoc.unesco.org/ark:/48223/pf0000374266

state of the art multi-disciplinary knowledge, skills, and competencies of individual AI actors, which drive individual AI actors' ethical habits and inform a country's AI ethics guidelines; which as such, in their turn, facilitate the creation of personal, social and economic wellbeing as a result of the potential of harmonious and ethical co-existence of humans with technology thus created" (Ruttkamp-Bloem, 2020).

These three notions—AI readiness, National AI Capital and AI Ethics Capital—all imply that AI ethics regulation should be scalable, as all three concepts indicate dynamic statutes. This points to a need to engage with the context in which AI ethics regulation is formulated and will be applied via instruments such as AI ethics impact assessment and readiness methodologies and above all being sensitive to the fact that there is a difference of context between the West, the East and the Global South in many ways, which necessitates being sensitive to what can be expected in each context, while both acknowledging existing AI infrastructure and governance policies and assisting where necessary in culturally respectful manners.

While it is always risky to speak of the African continent as if it is a homogeneous whole, it allows me to make some general observations. Africa has the youngest population of all continents—a median age of 19.7 years at present[2]—and it is clear that while this status brings immense social and economic opportunities, it also brings with it urgent social, moral, legal and economic responsibilities. Furthermore, in terms of ethical concerns around AI technology, we should not make the mistake to think that this young population necessarily implies a high level of either AI ethics literacy or protection via AI ethics regulation. It is imperative that it is acknowledged that while "generation Z" is a technology savvy generation, this brings its own concerns: We do not yet know what the full scope of the impact is of growing up in a technologically driven world in the way that this generation has (see, e.g., UNESCO, 2020), and given the socio-economic and political instability in some African states, this impact has to be closely monitored so that it does not exacerbate existing inequality or other potential forms of social harm.

In addition, the potential for good that AI technology holds for Africa is huge, but can only be fully realised if AI technology is trusted, as full

[2] https://www.statista.com/statistics/1226158/median-age-of-the-population-of-africa/

adoption of technology only happens in contexts where there is trust in technology, and only full adoption guarantees economically successful and socially supportive AI ecosystems, as alluded to already. In this sense, African states seem to face a serious moral dilemma with regard to their AI approach and AI governance approach: "Do they go all out and become a global role player with the eye on economic gains that AI offers? Or do they take time to stop and think about the social and ethical impact on vulnerable groups in their communities?" (Ruttkamp-Bloem, 2021). To some extent however, I suggest this is a false dichotomy as the moral challenge is more complex. It is necessary to have participative inclusion (UNESCO, 2020) during all phases of the AI system[3] lifecycle[4] (Ibid.) of as many different cultural, ethnic, social-economic, age, gender and other groups as possible such that the mechanisms, products and benefits of AI technologies belong to all. This means that states can in fact not secure or guarantee economic gain from AI technologies without concerning themselves with all groups in their societies, and ensuring members of all groups are both protected against possible harm from AI systems (which includes addressing issues of structural harm that may potentially be amplified by data-driven AI technologies) and are enabled to actively participate in AI ecosystem building in the particular state.

Some specific important challenges in the context of AI governance include the issue of the exclusion of Africa from so-called global AI debates; the level of AI adoption and successful implementation of AI governance policies (impacted on by many factors that are not uniform across the continent, e.g., Internet penetration on the continent,[5] effectiveness and agility of legal processes, etc.); AI ethics, information and communication literacy (e.g., access to equal STEAM—STEM plus the

[3] "AI systems are information-processing technologies that embody models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments. AI systems are designed to operate with varying degrees of autonomy by means of knowledge modelling and representation and by exploiting data and calculating correlations" (UNESCO, 2020).

[4] The AI system lifecycle is fully defined by UNESCO (2020) as ranging "from research, design, and development to deployment and use, including maintenance, operation, trade, financing, monitoring and evaluation, validation, end-of-use, disassembly, and termination".

[5] Internet penetration in Africa ranges from 4.6% to 85.2% (https://www.statista.com/statistics/1124283/internet-penetration-in-africa-by-country/).

Arts—education for all, etc.); and lastly, the collectivist/individualist clash between African and Western ethical traditions.

All of these merit further discussion, but here I will just highlight a few issues. First, when we consider the issue of exclusion of Africa from "global" AI debates,[6] it should be clearly understood that we deal here with issues of epistemic injustice that cut across both hermeneutic and testimonial injustice in Miranda Fricker's (2007) sense. What is interesting here is not just that the exclusion of African academics and AI practitioners is a form of epistemic injustice, but that it in fact may be contributing to general harm in terms of fairness and bias concerns as this exclusion is the product of deeper structural injustice in the West and globally. In particular, given that a person's social positioning, influenced by factors such as race, gender and class, determines what knowledge they have access to as well as how the mechanisms with which they ascribe social meaning and gain knowledge are developed,[7] I think it is probable that epistemic injustice in the sense of the exclusion of Africa and Africans from global debates on AI ethics on the one hand contributes to what Kate Crawford (2017) calls allocation and representational harms, and on the other hand, is exacerbated as a result of such harms, in a concerning kind of feedback loop.

Testimonial injustice occurs when a hearer "awards a speaker's claims less credibility than it deserves because of a prejudice that the hearer holds towards the speaker based on operations of power that come about as a result of given social identities" (ibid.). In its turn, hermeneutic injustice is the "withholding from a certain social group the proper tools with which to make sense of or articulate social experiences which prohibits [members of such groups] from functioning adequately as equal agents in society" (ibid.). If one considers that representational harm from AI systems is a cultural and social harm, which occurs when "systems reinforce the subordination of some groups along the lines of identity; so that's race, class, gender, etc." (Crawford, 2017), and that the primarily economic and transactional harm of allocation harm occurs "when a system allocates or withhold[s] certain groups an opportunity or resource" (ibid.),

[6] Of course, I don't deny that much is being done in order to address this exclusion in certain contexts such as within UNESCO and the OECD. However, this is far from the norm in these debates.

[7] See, e.g., Mohanty (1993), Spivak (2010), Mason (2011), and Medina (2011).

one can perhaps get a glimpse of the scope of the harm to Africa and its people brought about by excluding them from global AI conversations.

I claim that a social systems analysis of the kind advocated for by Crawford (e.g., ibid., Crawford & Calo, 2016) and others such as Campolo et al., (2017), may very well show that there is a feedback loop specifically between testimonial injustice and representational harm and between hermeneutic injustice and allocation harm (especially in the latter case if allocation harm is also seen in terms of access to equal quality education). It is because of being excluded, of epistemic injustice on a grand scale and of resulting exclusionary practices in the tech community, that identity prejudice feeds so easily into the harms Crawford (2017) identifies. The point I am making here is that excluding Africa from global discussions specifically in AI, given the potential of data-driven AI for amplifying structural bias, unfairness and exclusion, does far more harm than simply ensuring AI technology stays in the hands of the North. It is necessary that this is acknowledged and actively combatted by advocating for inclusive, international and diverse tech teams, ensuring travel between Africa and the North becomes easier for AI practitioners, inviting speakers from Africa to global AI forums and acknowledging the existing skills and expertise in Africa, among other initiatives.

One of the biggest contributing factors relating to the above is the fact that work done in Africa is not recognised adequately. Of course there are unique challenges in Africa, but nevertheless there is a lot to be excited about in terms of AI development in Africa. This is illustrated by many initiatives such as the courses offered at the African Institute for Mathematical Science[8] and various evening and weekend classes, AI boot camps and innovation hubs sponsored from industry in many African universities. Then there are initiatives such as, to name just a few, Data Science Africa,[9] the Deep Learning Indaba,[10] the Masakhane Natural Language Processing community,[11] the IBRO-SIMONS Computational Neuroscience group Imbizo,[12] the Sisonke-Biotek grassroots-focused research

[8] https://aims.ac.za/.

[9] http://www.datascienceafrica.org/.

[10] https://www.2030vision.com/projects/deep-learning-indaba.

[11] https://www.masakhane.io/.

[12] https://imbizo.africa/.

initiative at the interface of machine learning and healthcare,[13] and also government sponsored programmes such as Rwanda's digital ambassador programme.[14] There are many private sector technology actors in Africa too—two examples of big AI technology exhibitions are East Africa COM[15] and AI EXPO Africa.[16] Then, there are academic initiatives such as the Technical University Munich and the Ghanaian Kwame Nkrumah University of Science and Technology's Responsible AI Network,[17] and the Division for Science and Innovation's Centre for AI Research in South Africa,[18] among many others.

Second, the concern around full AI adoption and successful implementation of AI governance policies relate, among other factors, to the practice of "ethics-dumping" which is a term introduced by the European Commission (Ruttkamp-Bloem, 2021), and first applied in the AI ethics context by Floridi (2019). Basically, this refers to the practice of transnational tech companies moving core operations to countries with weak AI governance regulation. The reason why this is such a strong threat is that few African states have national AI strategies. Key components of such strategies should include at least improved telecommunications infrastructure in Africa to increase Internet penetration, adequate AI regulation and, perhaps most importantly, the establishment of enabling and collaborative AI environments (Pillay, 2020). Even fewer countries have national strategies to address social and ethical concerns around AI technology. Key obstacles in formulating such strategies include lack of sufficient research into re- and up-skilling so as to offset potential job losses; lack of equal access to quality education focused on STEAM teaching, that is, STEM teaching with at least AI ethics, information and communication literacy included; reaching every member of society, across socio-economic and age and gender divides; actively ensuring inclusion in global debates; and both ensuring the quality of data and addressing the "data desert" in Africa (ibid.). In the case of the latter,

[13] https://www.sisonkebiotik.africa/.

[14] https://www.minict.gov.rw/projects/digital-ambassadors-programme.

[15] https://tmt.knect365.com/eastafricacom/.

[16] https://aiexpoafrica.com/.

[17] https://ieai.mcts.tum.de/responsible-ai-in-africa-network/.

[18] https://www.cair.org.za.

the private sector is playing an ever more crucial role in many African AI ecosystems (see the UNECA Africa Data Revolution Report, 2016).

Furthermore, in summary, in their recent AI Needs Assessment Survey in Africa 2021 (Sibhal et al., 2021), UNESCO states that "Member States have requested UNESCO's support for standard setting, policy advice, capacity building, network development and for addressing gender equality-related concerns in the development and use of AI" (ibid.): 32 countries requested UNESCO's support for building human and institutional capacities in AI-related domains in its fields of competence; 26 countries requested policy advice for the development of aspects of AI policy concerning education, sciences, culture and communication and information; 21 countries requested support from UNESCO in terms of setting standards; 27 countries requested support in building partnerships for the development and use of AI to help them achieve their developmental priorities; and 17 countries requested support for addressing gender equality-related concerns in the development and use of AI (ibid.).

The need for enhancing capacities for AI development and implementation of ethical AI governance is thus widely recognised, rather than downplayed in Africa. The need for AI literacy is also widely recognised, if not always immediately or uniformly addressed. One success story of addressing this issue is the Rwandan Digital Ambassador's Programme,[19] which is a government-funded and government-driven initiative for bringing digital literacy to rural areas by sending out graduate students and young entrepreneurs to communities to provide digital literacy training in local languages and focusing on locally relevant digital content and services (http://www.hsrc.ac.za/en/news/impact-centre/African-AI). Data Science Nigeria has developed an elementary school textbook on the nature and role of data in African lives, and focused on children being taught skills in contexts they know. These examples illustrate the importance of intentional capacity building and public interest-driven AI approaches in Africa (e.g., AI for social good and AI for development initiatives focused on food security in agriculture, or on distribution of medical supplies[20]) combined with the need for understanding existing infrastructure and being realistic about what is possible.

[19] https://www.minict.gov.rw/projects/digital-ambassadors-programme

[20] Think here of the famous Zipline initiative that originated in Rwanda and Ghana (https://flyzipline.com/).

Being focused on African problems and needs also feeds back into the discussion above on trust and adoption of AI technology being key for the success of this technology.

There are many factors in Africa that impact on the possibility of establishing this trust I referred to in the Introduction as an essential ingredient for successful realisation of the benefits of AI technologies for all. From the above, it is clear that these fall into different categories: (i) socio-political factors ranging from epistemic injustice and being at the receiving end of structurally biased and non-transparent AI systems, to religious and cultural concerns; (ii) literacy and digital inequality concerns which may lead to feelings of helplessness or even despair, especially in cases where there are other hardships to face such as political instability, access to clean water, to electricity, to education, etc.; and (iii) concerns around labour exploitation, job loss and access to up-skilling. Overall, there is also a shared concern about the implications of AI technology for cultural diversity, which brings me to the next section.

Towards an African AI Ethics

More than in any other culture perhaps, African ethics is deeply seated in the societal beliefs about what is morally right and wrong on the one hand, and in the behaviour society deems appropriate to bring about social justice and harmony (see, e.g., Gyekye, 2011). In the West, ethics is not as entangled with societal thinking, as the approach is individualistic rather than collectivist. This difference in approach to the role and nature of ethical systems should be acknowledged, if one wants to speak at all of any kind of global AI ethics regulation, but also in particular, if one wants to establish an effective AI ethics paradigm or domain in Africa.

Furthermore, the typical human rights approach to AI ethics dominant in the West may not sit naturally in Africa, as much of African political thought focuses on duties (responsibilities) rather than rights, or at the least, recognises that the rights vs. duties debate is central to African political thinking. This debate rests on the African notion of personhood, and this is a complex moral-political concept in African philosophy, which needs to be taken into account when thinking of AI ethics regulation in Africa. Gyekye (1997, p. 2) differentiates between metaphysical and moral perspectives on the notion of personhood in relation to social structuring of societies. The metaphysical perspective concerns questions such as whether a person is an "atomic, self-sufficient individual" (ibid.,), and

thus the "ontological priority" of an individual over the community vs. her communal nature (ibid.). Moral questions focus on the nature and status of individual rights, the place and role of duties, and the nature and role of a sense of "shared life or common (collective) good" (ibid.).

There is a contrast between Western acknowledgement of values such as autonomy, freedom and dignity belonging to individuals and the African ethical tradition according to which individuals depend on society for their very status of personhood and their general well-being (ibid., pp. 1–2). This difference also carries over to the notion of individual rights in the different traditions (see also, e.g., Ake, 1987; Deng, 2004; Metz, 2011; Molefe, 2019). In the African tradition, the relation between an individual and society is determined by a community of people (Gyekye, 1997, p. 2). Thus, the "communal structure" (ibid., p. 3) of African societies is the core characteristic of the social structure of African cultures (see also, e.g., Masolo, 2004; Tshivhase, 2015; Matolino, 2018; Metz, 2018). In this sense, one's status as an individual, one's very uniqueness, is only a secondary quality, as one is "first and foremost … several people's relatives and several people's contemporary" (Gyekye, 1997, 2).[21]

In the AI ethics context, this communal aspect of African societies should be taken seriously, as there are many (e.g., Raso, 2018, Latonero, 2018) who view AI ethics regulation through the lens of human rights and the communal structure of African societies implies that this individualistic approach may not make sense in Africa. Let me be clear, I am certainly not implying that human rights are not respected in Africa, or should not be, or cannot be, but I am cautioning that there needs to be sensitivity to how the concept of individual (and by implication human) rights is interpreted in Africa and therefore to how AI ethics regulations are formulated. In the global context, the actionability of AI ethics regulation depends on a number of factors such as the divide between the goals of members of the tech world and abstract ethical guidelines among others (see, e.g., Mittlestadt, 2019, Jobin et al., 2019; Hagendorff, 2020), and certainly, those factors apply in the African context too. But, I urge here for the additional acknowledgement of the potential

[21] I acknowledge that there are variations on this view, between moderate (e.g., Gyekey, 1997) and more strict (e.g., Menkiti, 1984) African communitarianism, but the communal core remains across different nuances.

impact of recognition of cultural language and traditions on AI ethics and actionability in general and in Africa in particular.

The naturalist approach to rights depicts rights as being "held simply by virtue of being a person (human being)" (Donnelly, 1982b, p. 391) and is based on individualistic moral and political frameworks (Gyekye, 1997, p. 33). But if it is taken into account that the concept of personhood is at the heart of the debate on rights vs. duties in African communitarian theories (Molefe, 2019, p. 147), it becomes clear that more nuance is necessary (e.g., Donnelly, 1982a).[22] In other words, the "thick" view of rights according to which rights are prior to duties (e.g., Griffin, 2009) is not necessarily naturally the African view, as on the latter account, rights should not naturally trump cultural, moral and political grounds for action (Molefe, 2019, p. 147), but are rather related to human dignity in more complex causal relationships (ibid, p. 152). There is in fact a continuum of views on rights in African literature, with persons such as Ake (1987), claiming that there are no individual rights in African moral and political thought, only communal duties, on one side of the continuum, more moderate views such as Gyekye's (1997) in which rights and duties are not in a one-to-one correlative relationship, but are nevertheless on equal moral footing and in a mutually dependent relationship, to African scholars accepting individual rights into African political theories to varying degrees (e.g., Wiredu, 1997; Metz, 2011; Matolino (2018)). I will here briefly consider Gyekye's moderate communitarian approach[23] to this debate, with some reference to other views, such as Molefe's. It is impossible to do justice here to the richness of this discussion in African literature and I can do no more than alert readers here to its overall importance, and specifically its importance in the context of actionable AI ethics for Africa.[24]

[22] I am not implying that there is no pushback against this naturalised account of rights in Western communitarian views, as of course there is, thinking of writers such as Taylor, McIntyre and Sandel among others, but the focus here is on the African view of rights, based as it is on African notions of personhood.

[23] Supporters of radical communitarianism include writers such as Mbiti (1970, 1990) and Menketi (1984).

[24] The general debate on the role of culture in global AI ethics regulation is a very complex one (see, e.g., Kesserwan, 2018, Hongladarom, 2020, Goffi et al., 2021) and can obviously in itself also not be done justice to here.

On Gyekye's moderate communitarian view, the community and the individual are ascribed "equal moral status" (Ibid., p. 9) as the one cannot exist without the other. Such a moderate view ties to a specific interpretation of the notion of the common good, which differs in important ways from the Western or individualistic interpretation of this notion. The Western understanding of the common good as "the aggregate of the particular goods of individual persons, which, like individual rights, ought to be respected" (Ibid, p. 13), implies not only the prioritisation of the individual, but also that the individual's value system, does not depend on and may be totally different from that of their community. On this view "the pursuit of a common good in an individualistic society will do violence to the autonomy and freedom of the individual and fetter her ability to choose her own good and life plans. But not only that: … the pursuit of the common good will result in intolerance of other conceptions of the good and inappropriate use of political power to realise the common good" (Ibid.).

If one now considers that even Gyekye's moderate view of communitarianism still implies that "communal life is not optional" (Ibid, p. 5) for an individual and that her personhood is "constituted by the social relationships she finds herself in" (Ibid), because of her "natural sociality" (ibid.), it becomes clear that the Western view of the common good and ultimately of rights cannot just be adopted as it is and should not just be "domesticated" (Molefe, 2019, p. 148) to fit the African context. Rather, in-depth reflection is needed to consider how these notions are interpreted and applied in African moral and political theories. On Gyekye's account, the individual cannot be ontologically prior to society, as this would imply not only that the individual's choice to join a community is optional, but also that the forming of communities is contingent on such decisions, and this is in opposition to the notion of our natural sociality (Ibid, pp. 5–6). It must furthermore be understood that cultural values are inherited from the social structure of the community and cannot be generated by the individual (Ibid., p. 7) as the individual can only realise their full potential as a member of a society.

On Gyekye's (ibid, p. 14) view then, in contrast to the Western view, the common good "literally and seriously means a good that is common to individual human beings—at least those embraced within a community … It is linked, … to the concept of our common humanity and, thus, cannot consist of, or be derived from, the goods or preferences of particular individuals; thus, the common good is not a surrogate for

the sum of the different individual goods [as in the West]", because if it were, it would suggest the that "common" can only be contingently realised (ibid.). Rather, the common good includes moral or political values that are "embracive of fundamental or essential goods" (ibid.) such as dignity, peace and respect. "The common good can, thus, be regarded as that which inspires the creation of a moral, social, or political system for enhancing the well-being of people in a community generally" (ibid., p. 15).

In this context, Gyekye (ibid., p. 34) formulates his moderate view of rights as belonging to individuals and contributing to their self-realisation, even though the way in which individuals express themselves as the result of these rights is still best done within a social framework within which communal values such as compassion are more important than individual rights. Gyekye (ibid., p. 36) writes: "… even though rights belong primarily to individuals, insofar as their exercise will often, directly or indirectly, be valuable to the larger society, their status and roles will nevertheless (have to) be recognized by communitarian theory". To deny this would be "sawing off the branch" on which communitarianism sits (ibid.). Society acknowledges social values such as "peace, harmony, stability, solidarity, and mutual reciprocities and sympathies" (ibid, p. 37) and no individual may exercise their rights in a way that compromises these values (ibid.) and thus the claims of "individuality and community ought to be equally morally acknowledged" (ibid, p. 38).

Thus, individual rights are only valid, or only have meaning, within human society. This means that such rights come with social "responsibility" (duties) (ibid, p. 38)—where "responsibility" refers to "a caring attitude or conduct that one feels one ought to adopt with respect to the wellbeing of another person or other persons" (ibid, p. 39), and thus, Gyekye is defending a duty-based view of rights on which rights and duties ("responsibilities") are closely related (ibid, p. 38). This is the case as the "relational character of the individual by virtue of her natural sociality immediately makes her naturally oriented to other persons with whom she must live. Living in relation to others directly involves an individual in social and moral roles, obligations, commitments, and responsibilities, which the individual must fulfil" (ibid., p. 39) as the communitarian notion of the common good implies that the individual should always do what is best for the community. This implies that moral duties/ responsibilities are "elevated" to the level of rights (ibid, p. 40).

If we briefly turn to consider the view of Molefe, he points out that the African idea of personhood is "grounded on a different ethical sensibility than that which informs the discourse on rights" (Molefe, 2017) and that acknowledging this is core to understanding the debate on rights vs. duties in African communitarian theories (Molefe, 2019 p. 147). "The idea of personhood, … envisages an other-regarding morality of duties. It is these (other-regarding) duties [virtues], I submit, that take priority even over rights" (ibid). This implies the moral currency in the African context is different from the West: Rather than making each individual member of a community a moral concern, the community is the moral concern in African communitarianism.

On the Western "minimalist" account of rights (e.g., Griffin, 2009), human rights "… function to protect our normative agency, which requires autonomy (the ability to choose one's own ends), liberty (freedom from coercion and manipulation) and welfare (provision of basic needs to be able to lead a human life like education)" (Molefe, 2019, p. 157). Molefe's "maximalist" duty-based view of rights is in direct contrast to this minimalist account of rights, as it is based at its core on the "ineliminable residue of human dependency" (Wiredu, 1998, p. 293), which relates to the recognition of our shared humanity rather than on differences between individuals (Molefe, 2019, pp. 158–160). The maximalist conception of personhood is based on a social or relational ethics (ibid, p. 160), which is "an ethics motivated by the needs and interests of others" (ibid), and in this sense echoes Gyekye's (1997) insistence on an ethics of sensitivity or care towards others and their needs. "The basis of these other-regarding virtues is the spontaneous human capacity to recognise human needs" (Molefe, 2019, p. 160). The kind of needs at issue here are basic needs in order to live a life of dignity. Such a life is only possible if an individual can attain personhood, which is dependent on living in a society in which the conditions for pursuing and attaining personhood are good (ibid, p. 163). Masolo (2004, p. 494) speaks of this duty to provide the basic goods for all in terms of the "economy of affection".

Given this brief—and necessarily superficial due to the wider scope of this article—introduction into African thinking on the rights vs. duties debate, at least two considerations to take into account when considering the nature and actionability of AI ethics in Africa become clear. Firstly, some of the most influential views of African ethics claim that it has a social and duty-based character that is different from the individualistic

rights-based character of Western ethics. Secondly, it should be clear that the manner in which AI ethics regulations are phrased should take into account the cultural embeddedness of any action that is required.

In terms of AI ethics culture matters, because for a global (or any other) AI ethics policy to be successful, it should belong to every role-player[25] as an active participant (Ruttkamp-Bloem, 2020). Among other things, this means the terms in which such policies are formulated must be familiar and acceptable to everyone at its receiving end, whether as user or researcher or developer or deployer of the technology at issue. Implementation and negotiation can only happen if role-players can understand each other and if every role-player feels heard and recognised as a credible and valid participant.[26]

Returning to our conversation on Africa, Hagerty and Rubinov (2019) write that, "AI is likely to have markedly different social impacts depending on geographical setting. Likewise, perceptions and understandings of AI [and addressing its disruption and successfully mitigating its potential harm] are likely to be profoundly shaped by local cultural and social context" (Hagerty & Rubinov, 2019). In conclusion of this section, I now briefly consider concretely the difference between Western and African ethical values and principles that may credibly drive AI ethics regulation to illustrate the importance of cultural nuance.

In terms of general African ethical values, Gyekye (1997, p. 40) identifies values such as peace, dignity, compassion, solidarity, reciprocity, cooperation, interdependence and social well-being as principles of communitarian morality, which impose "responsibilities [duties] on the individual with respect to the community and its members" (ibid). These values are for instance clearly taken up by the Masakhane natural language processing organisation, which is a South African grassroots organisation whose mission is "to strengthen and spur NLP

[25] With 'role-players' I mean AI actors in the sense of the UNESCO (2020) draft Recommendation: "AI actors can be defined as any actor involved in at least one stage of the AI system life cycle, and can refer both to natural and legal persons, such as researchers, programmers, engineers, data scientists, end-users, business enterprises, universities, public and private entities, among others" (UNESCO, 2020).

[26] While the focus here is on epistemic just approaches to AI ethics formulation, which will bring Africa to the table and assist in enabling responsible AI in Africa, it is worth pointing out that openness to inter-, trans-, and multi-disciplinary disciplinary research is another essential ingredient of successful participation for all, as well as for successful implementation of AI ethics regulations and negotiation (ibid.).

research in African languages, for Africans, by Africans" (https://www.masakhane.io/). Their list of principles includes: "Umuntu Ngumuntu Ngabantu"—"loosely translated from isiZulu and meaning 'a person is a person through another person' or 'I am because you are'. This principle [of Ubuntu] proposes relationality over individualism for stronger social cohesion towards sustainable communities" (ibid.). Other principles include African-centricity, ownership (of NLP research processes), openness, multidisciplinarity, kindness (towards members of the NLP community), responsibility (taking ethical impact of research seriously), data sovereignty ("Africans should be able to decide what data represents our communities globally, retain ultimate ownership of that data, and know how it is used" (ibid)), reproducibility and sustainability.

If we now compare the Western context, where individualism, autonomy and human rights of individuals dominate as general values, we can clearly see how culture impacts on the AI ethics context. In general, AI ethics values and principles in documents generated in the West include some mention of human rights and human dignity, inclusiveness, flourishing of individuals and societies, autonomy, explainability, transparency, fairness and non-discrimination, awareness and literacy, responsibility, accountability, good governance, sustainability, robustness, privacy, solidarity and trust (see, e.g., Jobin et al., 2019).

Even though the Masakhane project is focused on NLP and the values listed for the West are from broad AI ethics documents, it is clear that there is some overlap of values, if not directly, e.g., sustainability being mentioned from within different cultures; then by implication in terms of at least some shared sentiment, e.g., solidarity and Ubuntu, and openness and inclusivity. It may very well be that there may be values that stand alone as very context-dependent, but which are nonetheless universally comprehensible, such as African-centricity; and others that may overlap in terms of sentiment, but in which the description has a very contextual slant, such as the interpretation of data sovereignty in the Masakhane case. But, in essence, it should be clear that taking culture seriously in the formulation of AI ethics values and principles does not need to make for incomprehensible differences. Rather, it makes for understanding—and respecting—the community that is on the receiving end of AI ethics regulation better and also ensures better potential adherence as regulation is formulated in familiar terms. This is why care should be taken in terms of concrete formulation of AI ethics policy that in principle it would be possible for different communities to find affinity with the manner in

which values and principles taken up in AI ethics regulation have been expressed.

As a brief example, in the current version of the UNESCO Recommendation, there is a value called "Living in peaceful, just and interconnected societies" (UNESCO, 2021). In the first draft of the Recommendation, this value was referred to as "Living in harmony" (UNESCO, 2020) and the idea behind it was to incorporate principles from the African philosophy of Ubuntu and Eastern philosophies such as Buddhism and Taoism, in order to demonstrate concretely the global character of the Recommendation. This value was not simply focused on inclusion and diversity, or on respecting human rights, and even though it incorporates the essence of what is meant by solidarity, it is still different and needed to be expressed in its own terms. It is clear if one looks at the current version (UNESCO, 2021) that while a solid compromise was reached after negotiation, a fair measure of cultural content has unfortunately been lost in translation.

This means we have a long way to go to learn how to practice cultural respect without somehow imagining that not expressing a value in Western terms would necessarily mean transgressing International Law, or that it would not be powerful enough somehow. Embracing culture does not mean that the meaning or interpretation of values and principles is open or relative to absurd degrees and it certainly does not mean International Law is not complied with, it simply means that there has to be engagement in epistemic just circumstances with every member of every community on the receiving end of AI ethics regulation to ensure that the meaning ascribed to values is articulated clearly and communicated as much as possible in terms that have synergy with the culture at issue. This will obviously improve sensitivity to AI ethics and adherence to resulting regulation. This brings us to some concluding remarks.

Conclusion

I conclude that in order to ensure the actionability of AI ethics regulation in Africa, AI ethics should be realised in epistemic just and dynamic systems driving AI policymaking. Reasons for advocating for the dynamic nature of AI ethics as a system include the fast-changing nature of AI technology, the need to take inter-, trans- and multi-disciplinary research seriously, the dynamic status and difference of contexts of application in terms of concepts such as AI (ethics) readiness and AI Ethics National

Capital, and also the fact that this system should be accessible to various cultures. In its turn, reasons for defending an epistemic just AI ethics system lie in the simple fact that culture matters in conversations on ethics. Listing un-interpreted values as abstract concepts alienates the members of the tech community, but expressing such concepts from within one dominant culture in fact alienates entire countries or even continents. There should be *engagement* with different interpretations such that there can be articulation and communication of ascribed meaning (always within the confines of International Law).

The way to meet this challenge is not to turn to hegemonic forms of discourse based on epistemic and other injustices as the easy way out; but to work to have cultural difference actualise the dynamics of ethics for a shared human goal.[27] There is thus an urgent need for a protocol for cultural engagement in AI ethics discussions. While human rights may be the lens through which many in the West consider AI ethics, culture should perhaps be the global calculus for AI ethics in the sense of being the source for interpreting AI ethics and translating it into familiar terms for each community. This may contribute to trust in technology becoming tangible because in such a scenario, trust would be constructed from the bottom up. The point made here is that cross-cultural understanding and collaboration on the one hand, and respect for socio-economic context on the other may ensure that the "responsibility of solidarity with the least advanced to ensure that the benefits of AI technologies are shared" (UNESCO, 2021) that lies with the most technologically advanced countries can be successfully taken up. Interestingly enough, the African philosophy of Ubuntu may be the very golden threat needed to knit the world into a more equal future and allow for a sustainable global AI ethics narrative, focusing such narrative on living in harmony, on shared human values and on the ultimate interconnectedness of all humans.[28]

[27] The role of 'technology ambassadors' (e.g., https://techmonitor.ai/leadership/innovation/tech-ambassadors) may be far more valuable than has been acknowledged up to now to enable such engagement. Rather than being deployed only to Silicon Valley to represent their countries' interests in liaising with Big Tech Companies, they could also be deployed to inter-governmental discussions on AI ethics regulation and play the role of cultural interlocutors, rather than focusing only on political interests.

[28] See http://www.hsrc.ac.za/en/news/impact-centre/African-AI.

Bibliography

Ake, C., 1987. The African context of human rights. Afr. Today 34, 5–12.
Campolo, A., Sanfilippo, M.R., Whittaker, M., Crawford, K., 2017. AI Now 2017 report.
Crawford, K., 2017. The trouble with Bias,‖ in 31th Conference on Neural Information Processing Systems (NIPS). Long Beach, CA, USA.
Crawford, K., Calo, R., 2016. There is a blind spot in AI research. Nat. News 538, 311.
Deng, F., 2004. Human rights in the African context, in: Wiredu, K. (Ed.), Companion to African Philosophy. Blackwell, Oxford, pp. 499–508.
Donnelly, J., 1982a. Human rights and human dignity: An analytic critique of non-Western conceptions of human rights. Am. Polit. Sci. Rev. 76, 303–316.
Donnelly, J., 1982b. Human rights as natural rights. Hum Rts Q 4, 391.
Floridi, L., 2019. Translating principles into practices of digital ethics: Five risks of being unethical. Philos. Technol. 32, 185–193.
Fricker, M., 2007. Epistemic injustice: Power and the ethics of knowing. Oxford University Press.
Goffi, E.R., Colin, L., Belouali, S., 2021. Ethical assessment of AI cannot ignore cultural pluralism: A call for broader perspective on AI ethic. Arribat-Int. J. Hum. Rights Publ. CNDH Moroc. 1, 0–0.
Griffin, J., 2009. On human rights. OUP Oxford.
Gyekye, K., 2011. African ethics. Stanf. Encycl. Philos.
Gyekye, K., 1997. Tradition and modernity: Philosophical reflections on the African experience. Oxford University Press.
Hagendorff, T., 2020. The ethics of AI ethics: An evaluation of guidelines. Minds Mach. 30, 99–120.
Hagerty, A., Rubinov, I., 2019. Global AI ethics: A review of the social impacts and ethical implications of artificial intelligence. ArXiv Prepr. ArXiv190707892.
Hongladarom, S., 2020. The ethics of AI and robotics: A buddhist viewpoint. Lexington Books.
Jobin, A., Ienca, M., Vayena, E., 2019. The global landscape of AI ethics guidelines. Nat. Mach. Intell. 1, 389–399.https://doi.org/10.1038/s42256-019-0088-2
Kesserwan, K., 2018. How can indigenous knowledge shape our view of AI. Policy Options Polit.
Latonero, M., 2018. Governing artificial intelligence: Upholding human rights & dignity. Data Soc. 1–37.
Masolo, D.A., 2004. Western and African communitarianism: A comparison.
Mason, R., 2011. Two kinds of unknowing. Hypatia 26, 294–307.
Matolino, B., 2018. Restating rights in African communitarianism. Theoria 65, 57–77.

Mbiti, J.S., 1990. African religions & philosophy. Heinemann.
Mbiti, J.S., 1970. Christianity and traditional religions in Africa. Int. Rev. Mission 59, 430–440.https://doi.org/10.1111/j.1758-6631.1970.tb00979.x
Medina, J., 2011. Toward a foucaultian epistemology of resistance: Counter-memory, Epistemic friction, and *Guerrilla* pluralism. Foucault Stud. 9–35. https://doi.org/10.22439/fs.v0i12.3335
Menkiti, I., 1984. Person and community in African traditional thought, in: Wright, R.A. (Ed.), African philosophy: An introduction. University Press of America, Lanham, pp. 171–181.
Metz, T., 2018. What Is the essence of an essence? Comparing Afro-relational and western-individualist ontologies. Synth. Philos. 65, 209–224.
Metz, T., 2011. Ubuntu as a moral theory and human rights in South Africa. Afr. Hum. Rights Law J. 11, 532–559.
Mittelstadt, B., 2019. Principles alone cannot guarantee ethical AI. Nat. Mach. Intell. 2019. https://doi.org/10.1038/s42256-019-0114-4
Mohanty, S., 1993. The epistemic status of cultural identity: On 'Beloved' and the postcolonial condition. Cult. Crit. 24, 41–80.
Molefe, M., 2019. An African philosophy of personhood, morality, and politics. Springer.
Molefe Motsamai, 2017. Critical comments on Afro-communitarianism: The community versus the individual. Filos. Theor. J. Afr. Philos. Cult. Relig. 6, 1–22.https://doi.org/10.4314/ft.v6i1.1
Momčilović, A., 2021. NAIC Foundations: Is human capital the only thing (becoming and) remaining important? By Aco Momcilovic, EMBA. URL https://www.ssbm.ch/blog/naic-foundations-is-human-capital-the-only-thing-becoming-and-remaining-important-by-aco-momcilovic-emba/ (accessed 9.10.21).
Pillay, N., 2020. Artificial intelligence for Africa: An opportunity for growth, development, and democratization. University of Pretoria, viewed.
Raso, F.A., Hilligoss, H., Krishnamurthy, V., Bavitz, C., Kim, L., 2018. Artificial intelligence & human rights: Opportunities & risks (SSRN Scholarly Paper No. ID 3259344). Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3259344
Ruttkamp-Bloem, E., 2021. Artificial intelligence presents a moral dilemma. Mail Guard.
Ruttkamp-Bloem, E., 2020. The quest for actionable AI ethics, in: Gerber, A. (Ed.), Artificial intelligence research, communications in computer and information science. Springer International Publishing, Cham, pp. 34–50. https://doi.org/10.1007/978-3-030-66151-9_3
Sibal, P., Neupane, B., Orlic, D., 2021. Artificial intelligence needs assessment survey in Africa. UNESCO, France.

Spivak, G.C., 2010. Can the subaltern speak?, in: Morris, R.C. (Ed.), Can the subaltern speak? Reflections on the history of an idea. Columbia University Press, New York, pp. 21–78.

Tshivhase, M., 2015. On the possibility of authentic self-expression. Communicatio 41, 374–387.https://doi.org/10.1080/02500167.2015.1093329

UNECA, 2016. Africa data revolution report, https://repository.uneca.org/handle/10855/23716.

UNESCO, 2020. Preliminary report on the first draft of the recommendation on the ethics of Artificial Intelligence, https://unesdoc.unesco.org/ark:/48223/pf0000374266.

UNESCO, 2021. Draft text of the recommendation on the ethics of Artificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000378426?posInSet=5&queryId=1edccad9-dc25-4319-ba46-f3bb300fd2f2.

Wiredu, K., 1997. Cultural universals and particulars: An African perspective. Indiana University Press, Indiana.

Wiredu, K., 1998. Moral foundations of an African culture, in: Coetzee, P.H., Roux, A. (Eds.), The African philosophy reader. Routledge Taylor and Francis Group, New York.

Responsible AI in Africa—Challenges and Opportunities

Chinasa T. Okolo, Kehinde Aruleba, and George Obaido

Introduction

Since its inception , the development and integration of artificial intelligence have been mostly concentrated within the Global North. This concentration of power has direct ties to the colonial history of resource extraction from the Global South that deprived nations in this region of autonomy and means to industrialise. This disparity has limited the ability of artificial intelligence applications to be effective, meaning that these tools are able to operate in a functional manner that doesn't compound existing inequities, within such contexts. Effective AI adoption and implementation of artificial intelligence are dependent on a variety of factors, such as having a local workforce with the required training to develop these solutions, sufficient infrastructural capacity to

C. T. Okolo
Cornell University, Ithaca, NY, USA

K. Aruleba
University of Leicester, Leicester, UK

G. Obaido (✉)
University of California, Berkeley, CA, USA
e-mail: rabeshi.george@gmail.com

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_3

handle the computationally-heavy training of algorithms, representative datasets, governmental support and regulation to govern the appropriate fair use of these technologies, independent and civil institutions and policymakers that safeguard from harmful applications and reinforce responsibility and accountability. Within the African continent, we find these factors sorely lacking and a major contributor to the dearth of solutions implemented that incorporate artificial intelligence. In this chapter, we present a number of challenges to effective AI adoption and implementation in Africa. We examine topics such as digital literacy, infrastructure and government support then lead into an analysis of the AI startup and research landscape within the continent. The chapter then defines what responsible AI looks like for Africa and provides actionable recommendations for improving its progress.

What Signifies Responsible AI

The concept of responsible AI has garnered significant attention in recent years (Boden et al., 2017; Gwagwa et al., 2020; Neri et al., 2020; Arrieta et al. 2020). According to Boden et al. (2017), responsible AI is the tendency of behaving in a positive, desirable, or socially acceptable manner. To formulate a central theme for responsible AI, several organisations created frameworks for building AI responsibly. These organisations formed an organisation termed the Partnership on Artificial Intelligence (PAI).[1] The partnership included organisations such as Amazon, Facebook, Google, Microsoft and IBM. PAI documented best practices to bring diverse organisations together to build AI systems responsibly to benefit the people and society. PAI described responsible AI as an approach geared towards ethical, social consequences that must be considered towards the development and deployment of AI systems.

Succinctly, the goals of the partnership on AI are fourfold:

- First, to develop and share best-practice methods and approaches in the research, development, testing and fielding of AI technologies;
- Second, to advance public understanding of AI across varied constituencies, including on core technologies, potential benefits and costs;

[1] https://www.partnershiponai.org/

- Third, to provide an open and inclusive platform for discussion and engagement on the future of AI, and to ensure that key stakeholders have the knowledge, resources and overall capacity to participate fully in these important conversations; and
- Fourth, to identify and foster aspirational efforts in AI for socially benevolent applications.

AI enables society to automate more tasks and automate to a larger extent than before, but it is important to understand who or what is responsible for the benefits and harms of using this technology? And, if this problem should be tackled pro-actively in the domains of technology and policy, what does the development of "responsible AI" mean? (Coeckelbergh, 2020) To address this question within an African context, we describe the different tenets of responsible AI as mentioned by Nyabola (2016) and Shearlaw (2016), such as accountability, transparency, explicability, transparency and bias.

Principles of Responsible AI

Accountability

One of the common principles of responsible AI discourse is accountability (Arrieta et al., 2020; Raji et al., 2020a; Rakova et al., 2021). According to Raji et al., (2020a), accountability is the state of being answerable or responsible towards a system with its underlying behaviour and its likely impacts. Their work further emphasised that algorithms cannot be held accountable. They are not legal entities; the organisations using and deploying these algorithms should be held liable through governance structures.

Within the tenets of the law, the notion of responsibility is often coupled with liability and punishment for misdeeds, with accountability viewed as a review, oversight and enforcement (Kohli, 2018). In their work, Doshi-Velez (2017) noted that creating accountable AI systems is crucial because accountability is essential for good public and private governance. Ensuring accountability in AI systems requires guiding actions and providing explanations in line with social values and norms. To ensure accountability, Africa needs stringent policies to govern AI usage to ensure rights preserving and ethics in its formative

design (Gwagwa et al., 2020). While only a few African countries have established data protection laws, many remain apathetic about it.

Transparency

Transparency means different things to different people, a term poignantly described by Weller [22]. Towards an AI discourse, Hollanek (2020) noted that a transparent system would provide information of what it is doing and why, and it must be permissible to be audited. Current AI algorithms are black boxes. The unboxing of AI algorithms has shown to be an engineering challenge, requiring clarity and explanations for end-users and regulators. This level of opacity is seen as surreptitious, incorporated by complex data processing, purported as a matter of deliberate practice (Hollanek, 2020).

Within contemporary African settings, the lack of AI transparency is visibly apparent (Gwagwa, 2020). One notable example is surveillance technology, where many African states are deploying these systems to monitor citizens (Mudongo, 2021). Most foreign organisations predominantly run these systems with the sole purpose of achieving political agendas or silencing critics. Such lack of transparency infringes user's privacy and leads to data exploitation. According to Hollanek (2020), to assess AI transparency, one must recognise that the developer has always been a trickster, applying masking techniques to achieve a result. The gateway to transparency is to ensure that all aspects of algorithm design, politics and morals have to be considered, especially to win trust.

Explicability

We used explicability and explainability interchangeably. The concept of explainability is viewed as the notion of explanation of steps taken by an AI model, in an attempt to ensure transparency, such that the result produced must be clearly understood by a human expert (Neri et al., 2020; Arrieta et al., 2020). For example, in precision medicine, medical practitioners require much more information from an AI model about a medical diagnosis rather than just some binary predictions. Other notable application areas that require explanations include autonomous vehicles, security, finance, among others.

According to Carman and Rosman (2021), to develop AI models sensitive to African interests and values, it is pertinent to adopt the principle of

explicability relevant in an African research context. The authors echoed that for an African adoption, AI systems must be just, fair and intelligible. The study further recommended that AI frameworks be designed to be applied to an African context for transparency. Regrettably, as a continent with diverse cultures and values, it is widely understood that African interests are not considered during AI designs (Wareham, 2021).

Bias Evaluation

AI bias occurs when an algorithm's output becomes prejudiced due to false assumptions based on the data fed into it (Silberg, 2019). Silberg argued that the extent to which AI is used for prediction and decision-making will always be subject to bias challenges. Roselli et al. (2019) argued that AI biases stem from diverse sources, such as the chosen algorithm, input attributes and training data used. Several AI biases have been seen in diverse applications, such as facial recognition systems, autonomous vehicles, health systems, criminal justice systems and recruiting systems (Mehrabi et al., 2021; Perkowitz, 2021; Gebru, 2020; Jo & Gebru, 2020). A typical example of AI bias in the medical field might be that an algorithm may wrongly recognise doctors as male and not female or exclude minorities. In some cases, AI systems may falsely misclassify a black person as a criminal element before standing trial or even likely to re-offend. Such systems often exclude traditionally marginalised groups and result in many diversity issues.

There has been growing use of AI in many sectors in Africa, especially in an unequal society, such as South Africa. Most banks use such technologies to make loan decisions (Adams et al., 2020). However, the metric used for allocating loans through demographics or loan history is widely unavailable. Such systems may be used to disadvantage certain races or genders. Moosajee (2021) attributes this issue to biased results caused by biased data.

Challenges to Effective AI Adoption and Implementation in Africa

In this section, we present the challenges towards AI adoption and implementation in Africa.

Digital Literacy

Digital skills literacy is a significant barrier to the adoption and implementation of artificial intelligence in Africa. Out of all world regions, sub-Saharan Africa has the lowest percentage of citizens equipped with digital skills, equalling to about half of the average level of digital skills adoption seen globally (Madden & Kanos, 2021). The Future of Work in Africa 2021, a report from the World Bank, shows that, on average, citizens in Nigeria, Kenya and South Africa possess a higher level of digital skills compared to the rest of sub-Saharan Africa (Choi et. al 2020). Inspired by the Sustainable Development Goals formed by the UN, the World Bank has formed the Digital Economy for Africa (DE4A) initiative, to digitally enable every African individual, business and government by 2030 (World Bank, 2021). Creating building blocks for a digital economy within the African continent shows promise in enhancing economic growth and alleviating poverty by encouraging entrepreneurship among young adults, increasing farming productivity and yields, and balancing the labour workforce by creating pathways for women to access more jobs (World Bank, 2019a, b). While some of the statistics presented earlier in this section seem dismal, it is important to note the historical and structural issues that have led to such outcomes. Historically, the lack of investment by African governments into infrastructure necessary for supporting digital economies has hampered the growth of digital literacy. With the help of international finance institutions and intergovernmental organisations, countries such as Mozambique and Rwanda have actively begun to develop action plans towards achieving digital transformation (World Bank, 2021). Over the past few years, large technology companies have begun to realise the importance of training local workforces in digital skills. In May 2021, Microsoft announced a partnership with the Nigerian government to significantly build their digital economy (Microsoft, 2021). This collaboration plans to speed Nigeria's transition to becoming a digital economy by making significant investments in internet infrastructure, equipping 5 million people across the country with digital skills,

developing cloud-based tools to fight corruption and leveraging artificial intelligence to preserve cultural heritage. With companies such as Microsoft, Twitter, IBM, Facebook and Google already having such a large presence within the African continent, commitments such as the one seen by Microsoft can help subvert the colonial narratives and dynamics seen in other industries like agriculture and mining. However, their presence is not without great scrutiny or concern. Big Tech is not the saviour Africa needs to look up to and their presence in Africa is driven primarily by profits, monopoly and a rush to grab power more than anything else.

Infrastructure

Over the past decade, internet penetration within the African continent has risen rapidly from an estimated 10% in 2010 to 28% in 2019 (ITU, 2021). The sore state of internet penetration across the African continent can be blamed due to infrastructure issues associated with the lack of access to electricity and low investment into internet infrastructure such as fibre-optic cables, cell towers and base stations. According to the World Bank, 80% of the urban population in sub-Saharan Africa has access to electricity compared to 28% in rural sub-Saharan Africa[2]. The World Bank estimates that reaching the 100 million Africans living in remote regions inaccessible to cellular mobile networks will require an investment of at least $100 billion[3]. While an extremely large number, tech companies such as Google, Facebook and Microsoft have lent their expertise and vast financial resources to improve internet infrastructure across Africa and the Global South to varying levels of success. Alphabet's (Google's parent company) Loon, a project developed in 2011 to bring high-speed internet to remote regions through fleets of balloons, operated in regions such as Sri Lanka, Puerto Rico, Mexico, Brazil, Chile, Argentina and Kenya (Loon, 2017). While this project was disbanded in early 2021, Alphabet has pledged $10 million to support companies and organisations focused on internet connectivity, education and entrepreneurship within Kenya (Teller, 2021). In partnership with Samsung, Ericsson, MediaTek, Opera, Nokia and Qualcomm, Facebook

[2] https://data.worldbank.org/indicator/EG.ELC.ACCS.RU.ZS?locations=ZG&name_desc=false

[3] https://www.worldbank.org/en/news/press-release/2019/10/17/achieving-broadband-access-for-all-in-africa-comes-with-a-100-billion-price-tag

launched Free Basics (also known as Internet.org), an initiative to provide free internet services to underdeveloped countries in 2013[4]. This service provides free internet access to websites containing job ads, weather and health information and full access to the entire internet for those who pay. This two-tiered system has been criticised for harming net neutrality and cannibalising the services of local internet cafes, leading to its eventual shutdown in India (Prasad, 2018). To this date, 32 African countries have participated in this initiative but over the past decade, internet shutdowns have become a common censure tactic for African governments (Killander & Ilori, 2020). The close relationship between Free Basics and the telecommunication companies providing these services, many of which are fully or partially state-owned, is a troubling issue that should be examined more closely. In 2020, Facebook announced "2Africa", their billion dollar project to build an undersea cable that will interconnect 23 countries in Africa, tripling the continent's existing network capacity and providing support for 4G, 5G and broadband access (Facebook Engineering, 2020). Microsoft Airband, launched in 2017 to bring internet connectivity to rural regions, currently has projects in 8 African countries (Ghana, Nigeria, Kenya, South Africa, Democratic Republic of Congo, Zambia, Tanzania and Rwanda)[5]. While Africa continues to be the focus point of internet-related initiatives, there is reasonable scrutiny of these initiatives and their capabilities to exacerbate existing censorship of citizens by African governments and introduce new methods of surveillance to the continent (Shearlaw, 2016).

Price Barriers

The Global System for Mobile Communications estimates that 45% of the population in sub-Saharan Africa subscribes to mobile services (GSMA, 2020). Affordability is a large barrier to preventing the adoption of both mobile devices and services, which is the primary way users within the African continent access the internet. Another factor that impacts affordability is the high telecommunications taxation rates in countries such as the Democratic Republic of Congo, Mozambique, Sierra Leone and Tanzania, which are above the world average (GSMA, 2016). The

[4] https://connectivity.fb.com/

[5] https://www.microsoft.com/en-us/corporate-responsibility/airband

majority of low- and middle-income countries have failed to meet affordability targets set by the Alliance for Affordable Internet (A4AI) and adopted by the UN Broadband Commission for Sustainable Development (Policy, 2018). When represented as a percentage of average per capita Gross National Income (GNI), the average price of one gigabyte of mobile broadband varies between 0.84% in North America and 17.49% in Africa (Policy, 2018). With an average cost of $3.30 per gigabyte of mobile internet—a price that is higher than anywhere in the world except for North America[6]—African consumers are being priced out of access to the internet and stonewalled from improving their livelihoods through digital means. African governments have to be proactive in regulating large, multinational telecom companies in setting fair prices for consumers and providing the necessary investments in both electrical and internet infrastructure to accelerate the adoption of digital skills, which will hopefully have a significant impact on the state of AI development within the continent.

Lack of Local AI Talent

While the technology ecosystem within Africa has grown significantly, there is still a large gap between the pace of software development and AI development within the continent. Fortunately, the interest of outside entities like Google, Microsoft and IBM has led to the establishment of AI research labs on the continent and the local startup ecosystem has also begun to grow. Something we find extremely promising is the emergence of local AI practitioners and research groups that have formed to address local problems in agriculture, healthcare, education and more. Initiatives, such as Bhala, a smart keyboard that is the first mobile application to support spell-checking Ndebele, Shona, Swati, Swahili, Xhosa, and Zulu, is an example of home-grown technology that meets the needs of local populations and fills a gap overlooked by larger players in this space[7]. A subsequent section in this chapter (survey of the landscape of AI in the continent) provides a deeper look into the current state of artificial intelligence within the African continent. We analyse over 100 startups and

[6] https://www.dw.com/en/why-mobile-internet-is-so-expensive-in-some-african-nations/a-55483976

[7] https://www.bha.la/about.html

organisations dedicated to providing AI products, services and education and note trends that show great promise or bring cause for concern.

Artificial intelligence has become nearly ubiquitous in many societies and many tech companies have begun to recognise the importance of democratising the development of artificially intelligent systems and providing equitable access to regions traditionally overlooked in AI development. Large tech companies such as Google, Microsoft, IBM, Facebook and Amazon have made strides to move into the Global South, establishing research labs, development centres, customer support centres, or data centres within this region. The openings of these establishments may initially appear to be beneficial for local ecosystems, but the talent needed to fill these highly specialised roles may not exist locally. This presents room for displacement of local workforces by those who have had the privilege to access relevant training and mirrors systems of colonialism.

Fortunately, the past five years has given rise to grassroots efforts focused on training local communities in artificial intelligence and related technologies such as natural language processing (NLP), computer vision and machine learning. Black in AI[8], a nonprofit organisation founded in 2017, has provided hundreds of Black students and professionals the opportunity to attend top tier machine learning conferences such as Conference on Neural Information Processing Systems (NeurIPS), the International Conference on Learning Representations (ICLR), International Conference on Machine Learning (ICML) and much more. Additionally, the organisation has begun mentorship programmes to guide prospective applicants to graduate programmes in computer science, support current PhD students in their journeys towards tenure track positions in academia and provide resources for entrepreneurs of African descent to build successful AI startups. Other prominent initiatives such as Data Science Africa[9], Masakhane[10], Ghana NLP[11], AI Saturdays Lagos[12] and Deep Learning Indaba[13] have similar missions, contributing greatly to the representation of African scholars at AI/ML conference venues,

[8] https://blackinai.github.io/#/

[9] http://www.datascienceafrica.org/

[10] https://www.masakhane.io/

[11] https://ghananlp.org/

[12] https://aisaturdayslagos.github.io/

[13] https://deeplearningindaba.com/2021/

increasing the number of publications focusing on AI and its applications to local problems and improving access to AI education. Another step to improving AI adoption within the African continent has focused on building institutions to formally train students in the concepts needed to pursue successful careers within this field. The African Institute for Mathematical Sciences (AIMS) was launched in 2003 in South Africa to teach specialised topics in the mathematical sciences such as applied mathematics, bioinformatics, scientific computing, artificial intelligence and more[14]. Since then, AIMS has expanded to Senegal, Ghana, Cameroon, Tanzania and Rwanda, graduating nearly 2000 students. To meet the demand for artificial intelligence practitioners within the continent, AIMS launched the African Masters in Machine Intelligence (AMMI) with sponsorship support from Facebook and Google[15]. DeepMind, a subsidiary of Google that develops AI systems to advance scientific discovery, has recently funded scholarships for students to pursue Master degrees in Computer Science with specialisations in AI, ML and data science at Makerere University in Uganda (Mwamai, 2021)[16]. We believe that industry-academic partnerships between local African institutions are a tangible step in bridging the AI-talent gap within the continent and will help build sustainable pathways to encourage future growth. While governments should be taking on the primary responsibility of funding AI education and entrepreneurship, the support of large industry players has helped fill this gap.

Datasets

Another issue plaguing the effective adoption of artificial intelligence in Africa is the lack of data accessible to African researchers and the relevance of this data to African problems in domains such as agriculture, health care and voice/text recognition. Machine learning relies on vast amounts of data to train algorithms, and if this data is sparse and unrepresentative, the resulting algorithms will be less effective and could cause harm to the vulnerable populations. Within Western countries like the

[14] https://nexteinstein.org/

[15] https://aimsammi.org/about-ammi-2/

[16] http://cs.mak.ac.ug/news/view/18

United States, issues regarding dataset representation of minority populations like Black people and women have gained prominence over the past few years (Buolamwini & Gebru, 2018). However, this conversation has continued to stay focused on dataset bias in the context of Western issues, centring the gaze of these problems on the Global North. In regions where the social construct of race is not present, focusing solely on the lack of racial representation in datasets limits how people address other facets of dataset underrepresentation in the Global South. We find that expanding issues of dataset bias to factors like ethnicity, tribal affiliations and other cultural nuances will help datasets becomes truly inclusive and relevant to solving African challenges. Open-source platforms like Kaggle[17], openAFRICA[18] and Zindi[19] have been supportive avenues for African researchers and AI practitioners to curate and share their datasets, helping to address the lack of datasets within the African continent. Initiatives like the Inclusive Images Challenge from Google[20] have aimed to improve representation of imagery from the Global South, but haven't fully represented the vast diversity within the African continent. This stresses the importance of local communities within the African continent being involved in the creation, sharing and use of datasets. More notably, we find that the formation of grassroots efforts throughout the continent has helped make significant strides in the types of data representing a variety of cultures, languages and regions throughout the continent. Collaborations between entities like Zindi and AI for Development (AI4D) led to the creation of the AI4D Africa Language Challenge in 2020 where over 400 data scientists enrolled to contribute their expertise to build novel datasets[21]. The winners of the challenge submitted datasets encompassing a variety of African languages like Wolof, Igbo, Hausa, Fongbe, Ewe, Kabiye, Kiswahili and Chichewa, many of which aren't present on popular translation services provided by Apple and Google. Other initiatives like the Lacuna Fund[22], which was founded to provide researchers and scientists in low-income countries resources to

[17] https://www.kaggle.com/tags/africa

[18] https://africaopendata.org/dataset

[19] https://zindi.africa/

[20] https://www.kaggle.com/c/inclusive-images-challenge

[21] https://zindi.africa/competitions/ai4d-african-language-dataset-challenge

[22] https://lacunafund.org/about/

produce labelled datasets, have helped improve dataset representation in agriculture, health and languages. More notably, the Lacuna Fund stresses adherence to practices in ethics and privacy, ensuring that the datasets will be owned by their respective creators and openly accessible to the international community. We believe that it is imperative for African researchers to maintain agency over the data they collect and have input on how this data should be shared. Data sharing practices within the African context have been understudied, but a recent paper titled "Narratives and Counternarratives on Data Sharing in Africa" provided much-needed insight on local practitioners involved in these efforts and contributes tangible suggestions towards making datasets context-aware and ensuring the process of collection and sharing is trustworthy (Abebe et al. 2021). While there is still a long way to go in improving the quality and accessibility of datasets representing the African continent, significant progress has been made thus far.

Government Support

Over the past few years, governments have raced to develop legislation that will govern the use and implementation of artificial intelligence for personal and commercial use. However, African governments lag heavily behind those in North America and the European Union. In the 2019 Government AI Readiness Index published by Oxford Insights, Africa is the worst performing region, with no countries listed in the top 50 spots and only 12 in the top 100 (Readiness, 2019). The top five countries who are represented in the top 100 (Kenya, Tunisia, Mauritius, South Africa and Ghana) already have significantly developed tech ecosystems. This brings cause for concern to smaller economies within the African continent who have not developed legislation but could still be impacted by the effects of artificial intelligence. While it is unclear how many African countries have formally instituted regulation on AI, countries like Senegal, Kenya and South Africa have launched regulatory frameworks, data protection laws and acts regulating automated decision-making (Adams, 2021). The Index of Regulation of Artificial Intelligence has monitored the AI policy and regulatory landscape around the world, reporting eight African countries (Ghana, Kenya, Nigeria, Sierra Leone, South Africa, Uganda, Zambia and Zimbabwe) as making strides towards regulating AI (Goitom, 2019). Additionally, the lack of AI legislation proposed by African governments is mainly due to policymakers with

scant technology expertise, and insufficient expertise in AI and related emerging technologies. As the local AI workforce grows within the African continent, it will be important that governments provide opportunities for highly-experienced professionals to contribute to AI legislation by serving on technical advisory panels or being placed in government positions specifically developed to leverage their expertise. The development of artificial intelligence in Western countries has been fuelled by local startups and a similar model could prove successful in Africa. Larger companies like Google, Apple and Amazon have made dozens of acquisitions of AI startups over the past few years, but this model may not be efficient for the African context since there is so little AI activity occurring in comparison with the West. Thus, government support for AI startups and research hubs is crucial. This will ensure that local interests and not those of multinational corporations are prioritised when it comes to solving issues with AI and that the continent doesn't experience an "AI brain drain" seen in other fields such as medicine and engineering. Efforts like the Artificial Intelligence Hub at the University of Lagos run by Data Science Nigeria, which is the first of its kind in Nigeria and likely in the continent, provide free AI courses and research labs for aspiring AI practitioners (Ndiomewese, 2018). While it is not clear how much governmental support this initiative received, it is a good model for governments to follow in establishing nationwide AI hubs. Again, we stress the importance of AI development being "for Africans by Africans" to ensure that colonial cycles of extraction by Western entities and historical dependence on foreign aid don't impede what could be a viable pathway towards economic freedom (Chan et al., 2021).

Survey of the Landscape of AI in the Continent

The AI startup and research organisational landscape in Africa has rapidly increased over the past decade and continues to grow. We collected data on 102 African startups and research organisations that are either in operation, defunct, or with an unknown operating status across 11 African countries. Startups and research organisations were included if they operated within any of the 54 countries on the African continent and if their core product or business offering focused on artificial intelligence and/or its respective applications across a variety of domains. As some startups had either expired or broken links to their respective websites, we still included them but listed their operating status as either Unknown (U) or

No (N). Our search for African companies and organisations working with AI was conducted through Google, LinkedIn, Twitter and startup market intelligence platforms like Crunchbase, Pitchbook, Venture Capital for Africa (VC4A) and Tracxn. We also relied heavily on curated lists from sources such as Briter Bridges[23], a business intelligence company focusing on markets in the Global South and AI Expo Africa. This list is not exhaustive but represents a significant number of companies developing artificial intelligence solutions across the African continent.

Companies and organisations operating on the continent span a wide range of industries and domains. From the startups we listed, we found 30 different industries they operate in. Finance, health care, agriculture and research were the largest segments, accounting for nearly 50% of all the companies. Within many startups on the continent, the financial sector has been a priority, and as the shift to AI-powered solutions has begun, there is no surprise that this trend has moved towards finance as well. However, despite the goal of many fintech startups on the continent to improve financial access, they might in fact not be improving the lives of people at the bottom of society. We also find similar trends in health care, where AI has claimed to either match or even exceed the diagnostic capabilities of doctors (Liu et al., 2019). However, this claim only holds true in the high-resourced environments these technologies are developed and tested in, which mostly happen to be in Western societies. As infrastructure within the African continent has scaled rapidly, we expect to see AI being leveraged further across a variety of domains. While AI being applied to fields such as agriculture, health care and finance may help improve overall access to these vital services, if not developed properly, they could indeed exacerbate existing inequities.

A significant number of companies have sprung up as consultancies to help larger businesses incorporate AI strategy into their current respective solutions. These companies were also classified as "AI" companies for the sake of simplicity. Another growing trend we see is the establishment of research groups and initiatives to train AI researchers on the continent and to tackle gaps within AI development that fail to include African users. Groups such as Masakhane[9] have conducted novel research to build datasets and machine translation tools for African languages while

[23] https://briterbridges.com/

other initiatives such as Datascience Nigeria[24] have democratised access to AI education by hosting bootcamps, summits and online competitions. A small number of AI developers on the continent have begun to develop their own libraries to improve AI and data science methods, with some like DeepQuest AI[25] by brothers Moses and John Olafenwa having thousands of users around the world.

The startups in our analysis operate across a total of 11 African countries, with a majority of them based in Nigeria (25%) and South Africa (33%). Compared to other regions within the African continent, Western Africa and Southern Africa have 35 and 33 organisations respectively focusing on AI. Most notably, we find that countries from Francophone Africa (Benin, Burkina-Faso, Cape Verde, Côte D'Ivoire, Democratic Republic of the Congo, Republic of Guinea, Madagascar, Mali, Mauritania, Niger, Senegal, Togo and Tunisia) are missing from the AI ecosystem. While the overall startup and research landscape in AI within the African continent are promising, it is essential that AI services built on the continent include Africans from a diverse set of backgrounds and regions. We find that the artificial intelligence startup ecosystem in Africa is relatively young. All of these startups and organisations were founded between 2010 and 2020, with 72 (70%) of them founded within the past five years. Additionally, almost 80% of the startups in our analysis are early-stage, showing significant room for these companies to grow. While artificial intelligence has been around for decades, we presume that many of the companies in our analysis that were founded in the early 2010s may have adjusted their respective strategies to incorporate AI methodologies.

A majority of the AI solutions developed by startups within the African continent cater to businesses that are aiming to improve their respective AI offerings or introduce AI into their operational systems. As financial technology (also known as fintech) becomes increasingly popular, we find that a significant number of AI startups are developing technology to improve banking processes or make financing decisions for customers based on existing data. Agricultural businesses and healthcare facilities are also popular options for AI startups to provide services for these industries that are rapidly digitising. Our analysis notes a growing trend of AI being

[24] https://www.datasciencenigeria.org/

[25] https://deepquestai.com/

introduced into warehouse operations and for manufacturing capabilities which could prove positive in scaling growth in these sectors and having a positive economic impact.

Out of 102 startups, 22 of them (21%) had either all white or majority white/non-African founding teams. Many of these companies were based in South Africa, which we find particularly concerning due to the types of technologies such as facial recognition being developed by these startups which would disproportionately affect a major part of the respective population. In the United States, research has shown how facial recognition systems from companies such as Microsoft, IBM, Amazon and others are biased against subjects of darker skin tones and those with female characteristics (Buolamwini, 2018). These companies have either chosen to improve the gender and racial makeup of the datasets used to train these systems or abandon facial recognition technology altogether (Heilweil, 2020). Surveillance technologies like facial recognition are being deployed within large cities throughout the continent and have received major backlash from citizens in countries like South Africa and Zimbabwe, but continue to stay in use (Chutel, 2018; Hawkins, 2018). This again raises concern due to policing systems in countries such as South Africa that are significantly biased towards Black South Africans who make up a considerable portion of the population.

From the 102 startups we analysed, we found that there is a lack of information and transparency on what exactly most startups are doing: what methods they are using to build their predictive models, where their data is sourced, and how well their models perform. These issues have been characterised as fairness, accountability, transparency and ethics (FATE) and dozens of startups, mainly in the United States and Europe, have formed to address these issues of data and model observability. As this field continues to widen, it will be important for African startups working with AI technologies to construct AI observability platforms of their own or work with other local startups in this space. More significantly, it is hard to conclusively say if there is a "true" AI element within many of the startups we ed. However, this is an issue that is not relegated solely to startups within the African continent. In 2019, London venture capital firm MMC studied almost three-thousand "AI startups" across the EU, concluding that 40 per cent of these companies do not incorporate artificial intelligence in their products (MMC, 2019). The overuse of the phrase "artificial intelligence" has led to unrealistic exaggerations of technology and excessive trust on what AI can do. Academic research labs and

AI startups have purported to build AI systems that can detect emotions, gender, sexuality and even political orientation; however, these tasks are nearly impossible for a human to accomplish (Birhane, 2021; Heckman, 2020; Wojcik & Remy, 2019). This work has led to calls for the EU to ban these tools (Asher-Schapiro, 2021) and we hope that policymakers within the African continent recognise these harms and actively begin to introduce regulation banning these tools. With the business models of many of these types of companies primarily being motivated by maximising profit at any cost, the autonomy and well-being of everyday citizens should not be disregarded. Current AI technologies are already believed to cause harm to marginalised people around the globe, and as AI grows within the African continent, there could be repercussions if its development isn't well-governed.

Emerging Trends and Concerns in AI Deployments in Africa

China is making a push for AI leadership and doubling down on its soft power initiatives in Africa as part of China's Grand Strategy to tap emerging markets, shape global governance norms and expand its influence (HDI, 2021; Nantulya, 2018). In his report to the 19th Party Congress in October 2017, Chinese President Xi Jinping outlined his vision for China becoming a global science and technology leader by 2050 (Shepherd & Qiu, 2017). A growing consensus singles out China as a major driver and influencer of authoritarian tech (Feldstein, 2019a, b). Several experts have claimed that Chinese governments are working closely with Chinese companies to export authoritarian tech to like-minded governments to promote an alternative governance model (Polyakova & Meserole, 2019; Sharma, 2020; Mozur et al., 2019). These reports can be validated with the recent increase in the export of Chinese tech, which are gross human rights violators to countries such as Zimbabwe, Uganda and Ethiopia (Feldstein, 2019a, b). Most of the countries in Africa rely on Chinese companies for their digital and telecom services. For example, the Ethiopian government uses the services and infrastructures of ZTE, a Chinese telecom to monitor its citizens' communications, Hikvision, the world's leading surveillance camera manufacturer, recently opened an office in Johannesburg, CloudWalk Technology, a startup based in Guangzhou, recently signed a deal with the

Zimbabwean government (Hawkins, 2018), and Transsion a Shenzhen-based company that has never sold a handset in its native China, but its brands iTel, Tecno and Infinix sell more smartphones than any other produces in Africa (Bayes, 2019; Bloomberg, 2018).

As reported in the AI Global Surveillance (AIGS) Index, Chinese companies led by Huawei are leading the supply of AI technologies with much focus on surveillance systems around the world (Feldstein, 2019a, b). Huawei technology has been linked to more countries in the South than any other company. It is aggressively infiltrating into the sub-Saharan Africa technology market by providing not only equipment and technological tools but also offering operation, management and support to set up these tools. To date, at least 12 African countries are using Huawei digital surveillance technology (Olander, 2019; Jili, 2020). An investigative report in The Wall (2019) highlighted that Huawei employees, technicians, provide other services to Uganda and Zambia governments that are not disclosed publicly. They helped the Uganda government spy on their political opponents by using cell data to track their locations and intercept their encrypted social media and communications. Also, in Zambia, the employees assisted the government in gaining access to Facebook pages and phones of bloggers critical of the president so they could be tracked and arrested. In Uganda, the government splashed $126 million on CCTV from Huawei (Reuters, 2019; Woodhams, 2020). The police say the new CCTV system will help reduce violent crime; however, the opposition leaders and civil society leaders believe that the law enforcement bodies are overburdened and too corrupt to identify criminals using footage. They believe the cameras with facial recognition technology will target and identify demonstrators in violent clampdowns as the election approaches in 2021.

Facial recognition technology has become increasingly pervasive around the world today, with the rising concerns about privacy, potential abuses, security, bias and freedom (Zeng et al., 2019); this has led to cities such as San Francisco ban its usage (Conger, 2019). In 2018, the government of Zimbabwe employed a surveillance network developed by CloudWalk to provide a mass facial recognition programme. In exchange for the technology, Zimbabwe sends images of its citizens, which give China an edge in AI technologies compared to other Western countries (Mind Matters News, 2020; Techzim, 2018). Beyond the human rights concerns, the deal pointed to another angle to the China-Africa tech story: the quest for technological advantage. As one local outlet put it:

"the Zimbabwe Government is sending our faces to China so China's AI can learn to see black faces". Existing AI facial recognition technologies are principally trained on white and East Asian datasets; the Zimbabwe deal offered Cloudwalk valuable data for improving its recognition of other ethnicities—thereby strengthening the arsenal of surveillance tools available to authoritarian governments (Bayes, 2019).

> Its terms require Harare [the capital of Zimbabwe] to send images of its inhabitants—a rich data set, given that Zimbabwe has absorbed migration flows from all across sub-Saharan Africa—back to CloudWalk's Chinese offices, allowing the company to fine-tune its software's ability to recognize dark-skinned faces, which have previously proved tricky for its algorithms.—Andersen (2020a, b)

By expanding into African markets, China's tech companies are gaining access to that sought-after commodity, data. Yes, these companies are playing a positive role in connecting African citizens, consumers and businesses. Nevertheless, they also have another role in helping Beijing promote its model of the internet as a controlled space and as a data-driven instrument of social and political control.

As reported by Tilouine and Kadiri (2018), the Africa Union (AU) building headquarters at Addis Ababa gifted to AU by China was serving the Chinese government more purpose than initially assumed. In January 2018, AU officials accused China of hacking its headquarters computer every night (when no one was in the office, but daylight had broken in Shanghai) for five years and downloading confidential data. Beijing had funded the building in Ethiopia, and a Chinese state-owned company built it. Chinese workers still maintain the building to this day, and even its elevator symbols are written in Chinese. This is bothersome. That one of the most prominent political organisations in the continent had been unknowingly sending all of their confidential data directly to the Chinese state certainly raises concerns about the implications of China's growing influence in the technological infrastructure of Africa. The overt Chinese presence on the continent in construction, technology and business has been attributed to the availability of generous loans with affordable interest rates and willing partnership for development (Future Africa,2018).

Recommendations for Improving Responsible AI Progress in Africa

According to Feldstein (2019a, b), local governments have the right to undertake surveillance systems that are unbiased and rooted in limiting their citizens' freedom and enforcing political repression. For example, tracking technologies play a crucial role in preventing terrorism. They give the government the ability to monitor threats and act accordingly. Governments can also use face recognition tools in finding missing people and victims of human trafficking. However, technology and power struggles have changed the nature of how governments use surveillance systems and what they intend to monitor. Generally, the legal standards required to utilise surveillance systems legitimately are high, and governments find it challenging to meet them. Countries with weak legal enforcement or authoritarian systems "*routinely neglect these obligations*" (Feldstein, 2019a, b).

Several authors have investigated the implication of racial bias surveillance and facial recognition algorithms (Cavazos et al., 2020; Bacchini & Lorusso, 2019; Raji et al., 2020b; Seutloali, 2015). Despite all the identified implications and recommendations made by these authors, it is important to state that no algorithm improvement is safe from the risk of contributing to racial discrimination if the social context in which it unfolds mainly consists of racial prejudice. A key recommendation is to stop seeing face recognition technologies as tools that do not see race and focus on working hard to monitor and get rid of racist from these tools. We need software and systems trained on datasets that are equally made up of faces representing all races. Until a solution to bias is found, algorithms need to be tested regularly for racially biased error rates.

Awareness is also a key precondition. We cannot hope that any of these recommendations are ever achieved unless we become definitively aware that facial recognition technology is doomed to be racially biased—at least until racism is permanently erased. Of course, the most effective recipe for a racism-free face recognition technology is to struggle for a racism-free society.

It is crucial to align AI initiatives and the data used for training AI models to local communities in Africa and the Global South in general. Engaging these communities, offering training solutions, understanding local issues and their unique needs will help create an avenue for developing more inclusive AI technology.

Conclusion

As AI development extends across the world and begins to make significant progress within Africa, it is imperative that local development of AI is encouraged and actively supported by governments, international agencies and large tech companies who have already begun to expand their global footprint throughout the continent. With current discourse and development of artificial intelligence focused on the West and China, there is little work that understands the nuances and sociotechnical implications of AI development in Africa. Combined with the lack of policy regulating the use of AI in many African countries, some of which has been problematic, Africa remains ripe for continued exploitation through new avenues presented by AI.

Although it will take considerable effort and expense to grow African countries into "AI superpowers", leveraging existing strengths in the software development and AI research communities while investing in infrastructure are viable steps towards this goal. AI has strong potential to transform livelihoods within Africa, but it is up to the continent to diligently focus on ensuring the potential risks and harms of AI don't outweigh the benefits.

References

Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S.L. and Sadagopan, S., 2021, March. Narratives and counternarratives on data sharing in Africa. In Proceedings of the 2021 *ACM Conference on Fairness, Accountability, and Transparency* (pp. 329–341).

Abiodun, F, (2018). 'Developing Public Sector Leaders for Africa: So China is Spying on the African Union? Who is Surprised?', *Future Africa*, 30 January. Available at: https://futureafrica.net/so-china-is-spying-on-the-african-union-who-is-surprised/ (Accessed: 11 June 2021).

Adams, R., Fourie, W., Marivate, V. and Plantinga, P., 2020. Introducing the series: can AI and data support a more inclusive and equitable South Africa? *Human Sciences Research Council*.

Adams, N.-R. (2021) 'AI law and Afrofuturism—how to regulate AI's future in Africa', *Michalsons,* 11 February. Available at: https://www.michalsons.com/blog/ai-law-and-afrofuturism-how-to-regulate-ais-future-in-africa/47558 (Accessed: 11 June 2021).

Aliya, R. (2019) 'Europe's AI start-ups often do not use AI, study finds', *Financial Times,* 4 March. Available at: https://www.ft.com/content/21b19010-3e9f-11e9-b896-fc36ec32aece (Accessed: 11 June 2021).

Alliance for Affordable Internet. (2018a) 'UN Broadband Commission Adopts A4AI "1 for 2" Affordability Target', *News Archives—Alliance for Affordable Internet*, 23 January. Available at: https://a4ai.org/un-broadband-commission-adopts-a4ai-1-for-2-affordability-target/ (Accessed: 10 June 2021).

Alliance for Affordable Internet. (2018b). *Affordable Internet Access: The cost challenge*. Available at: https://a4ai.org/affordable-internet-access-the-cost-challenge/ (Accessed: 11 June 2021).

Andersen, R. (2020a) 'Chinese AI Is Creating an Axis of Autocracy', *The Atlantic, 1 September*. Available at: https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/ (Accessed: 11 June 2021).

Andersen, R. (2020b). The panopticon is already here. Available at: https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/ (Accessed: 10 June 2021).

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, pp. 82–115.

Asher-Schapiro, A. (2021) 'EU urged to ban "discriminatory" AI tools that detect gender or sexuality', *Reuters*, 16 April. Available at: https://www.reuters.com/article/eu-tech-lgbt/eu-urged-to-ban-discriminatory-ai-tools-that-detect-gender-or-sexuality-idUSL8N2M86XH (Accessed: 11 June 2021).

GSM Association. (2016) *Connected Society Consumer Barriers to Mobile Internet Adoption in Africa*. Available at: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2016/07/Consumer-Barriers-to-mobile-internet-adoption-in-Africa.pdf (Accessed: 10 June 2021).

GSM Association. (2020) *Sub-Saharan Africa 2020 The Mobile Economy*. Available at: https://www.gsma.com/mobileeconomy/wp-content/uploads/2020/09/GSMA_MobileEconomy2020_SSA_Eng.pdf (Accessed: 10 June 2021).

Bacchini, F. and Lorusso, L. (2019), Race, again: how face recognition technology reinforces racial discrimination, *Journal of Information, Communication and Ethics in Society, (17)3*, pp. 321–335. https://doi.org/10.1108/JICES-05-2018-0050

Bayes, T., (2019) 'African networks, smartphones - and surveillance', *Merics,* 18 June. Available at: https://merics.org/en/analysis/african-networks-smartphones-and-surveillance (Accessed: 11 June 2021).

Birhane, A., (2021). The Impossibility of Automating Ambiguity. *Artificial Life*, pp.1–18.

Biryabarema, E. (2019) 'Uganda's cash-strapped cops spend $126 million on CCTV from Huawei', *Reuters*, 15 August. Available at: https://www.reuters.com/article/us-uganda-crime-idUSKCN1V50RF (Accessed: 11 June 2021).

Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., Newman, P., Parry, V., Pegman, G., Rodden, T. and Sorrell, T., 2017. Principles of robotics: regulating robots in the real world. *Connection Science, 29(2)*, pp. 124–129.

Buolamwini, J. and Gebru, T., (2018), January. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on Fairness, Accountability and Transparency* (pp. 77–91). PMLR.

Buolamwini, J. (2018) *Gender Shades*. Available at: http://gendershades.org/ (Accessed: 11 June 2021).

Carman, M. and Rosman, B., 2021. Applying a principle of explicability to AI research in Africa: should we do it?. *Ethics and Information Technology, 23(2)*, pp. 107–117.

Cavazos, J.G., Phillips, P.J., Castillo, C.D. and O'Toole, A.J., 2020. Accuracy comparison across face recognition algorithms: Where are we on measuring race bias?. *IEEE transactions on biometrics, behavior, and identity science, 3(1)*, pp.101–111.

Chan, A., Okolo, C.T., Terner, Z. and Wang, A., 2021. *The Limits of Global Inclusion in AI Development*. arXiv preprint arXiv:2102.01265.

Chen, L. Y. (2018) 'Are you a robot?', *Bloomberg,* 28 March. Available at: https://www.bloomberg.com/news/articles/2018-03-28/this-chinese-phone-maker-has-taken-over-africa-for-better-and-worse (Accessed: 11 June 2021).

Choi, J., Dutz, M.A. and Usman, Z. eds., 2020. *The Future of Work in Africa: Harnessing the Potential of Digital Technologies for All*. World Bank Publications.

Chutel, L. (2018) 'China is exporting facial recognition software to Africa, expanding its vast database', *Quartz Africa*, 25 May. Available at: https://qz.com/africa/1287675/china-is-exporting-facial-recognition-to-africa-ensuring-ai-dominance-through-diversity/ (Accessed: 11 June 2021).

Coeckelbergh, M., 2020. Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics, 26(4)*, pp. 2051–2068.

Conger, K., Richard, F., and Kovaleski, S. (2019) 'San Francisco bans facial recognition technology', *The New York Times*, 14 May. Available at: https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html (Accessed: 08 June 2021).

Ding, J. (2018) 'Deciphering China's AI dream', *Future of Humanity Institute,* 14 March. Available at. https://www.fhi.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf (Accessed 11 June 2021).

Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Scott, K., Schieber, S., Waldo, J., Weinberger, D. and Weller, A., 2017. Accountability of AI under the law: The role of explanation. arXiv preprint arXiv:1711.01134.

Dzoma, G. (2018) 'Zimbabwe Government Is Sending Our Faces To China So China's Artificial Intelligence System Can Learn To See Black Faces', *Techzim,* 8 November. Available at: https://www.techzim.co.zw/2018/11/zimbabwe-government-is-sending-our-faces-to-china-so-chinas-artificial-intelligence-system-can-learn-to-see-black-faces/ (Accessed: 11 June 2021).

Economy, E.C., 2018. China's new revolution: The reign of Xi Jinping. *Foreign Affairs, HeinOnline Journal*. pp. 60–61.

Ehl, D. and Grün, G-C. (2020) 'Why mobile internet is so expensive in some African nations', *DW*, 03 November. Available at: https://www.dw.com/en/why-mobile-internet-is-so-expensive-in-some-african-nations/a-55483976 (Accessed: 11 June 2021).

Facebook Engineering. (2020) 'Building Africa, a transformative subsea cable to better connect Africa', *Meta Engineering 13 May,* Available at: https://engineering.fb.com/2020/05/13/connectivity/2africa/ (Accessed: 10 June 2021).

Feldstein, S., 2019a. The global expansion of AI surveillance (Vol. 17). Washington, DC: *Carnegie Endowment for International Peace*.

Feldstein, S., 2019b. Can a UN Report Help Rein in Expansive and Abusive Digital Surveillance?. *World Politics Review*.

Gebru, T., 2020. Race and Gender. *The Oxford handbook of ethics of AI*, pp. 251–269.

Goitom, H., (2019), 'Sub-Saharan Africa', Regulation of Artificial Intelligence in Selected Jurisdictions. pp. 129–131.

Greitens, S.C., 2020. Dealing with Demand for China's Global Surveillance Exports. *Brookings Institution Global China Report*.

Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I. and De Beer, J., 2020. Artificial intelligence (AI) deployments in Africa: benefits, challenges and policy dimensions. *The African Journal of Information and Communication*, 26, pp. 1–28.

Gwagwa, A., 2020. Recommendations on the inclusion of sub-Saharan Africa in Global AI Ethics. *Research ICT Africa*. pp. 1–13.

Hawkins, A. (2018) 'Beijing's Big Brother Tech Needs African Faces', *Foreign Policy,* 24 July. Available at: https://foreignpolicy.com/2018/07/24/beijings-big-brother-tech-needs-african-faces/ (Accessed: 11 June 2021).

HDI. (2021) *China is Ready to Be come a Technological Superpower*. Available at: https://www.hdi.global/infocenter/insights/2021/china-technological-superpower/ (Accessed 11 June 2021).

Heckman, C. (2020) 'AI can now read emotions – should it?', *The Conversation*, 8 January. Available at: https://theconversation.com/ai-can-now-read-emotions-should-it-128988 (Accessed: 11 June 2021).

Heilweil, R. (2020) 'Big tech companies back away from selling facial recognition to police. That's Progress', *Vox,* 10 June. Available at: https://www.vox.com/recode/2020/6/10/21287194/amazon-microsoft-ibm-facial-recognition-moratorium-police (Accessed: 11 June 2021).

Hollanek, T., 2020. AI transparency: A matter of reconciling design with critique. *AI & SOCIETY*, pp. 1–9.

Oxford Insights. (2019). 'Government Artificial Intelligence Readiness Index 2019'. Available at: https://www.oxfordinsights.com/ai-readiness2019 (Accessed: 11 June 2021).

ITU. (2021) *Measuring Digital Development: Facts and Figures 2020*. Available at: https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx (Accessed: 9 June 2021).

Jili, B. (2020). Chinese Surveillance Tools in Africa. Available at: https://cld.web.ox.ac.uk/files/finaljilipdf (Accessed: 11 June 2021).

Jo, E.S. and Gebru, T., 2020, January. Lessons from archives: Strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 306–316).

Killander, M. and Ilori, T. (2020) 'Internet shutdowns in Africa threaten democracy and development', *The Conversation,* 26 July. Available at: https://theconversation.com/internet-shutdowns-in-africa-threaten-democracy-and-development-142868 (Accessed: 11 June 2021).

Kohli, N., Barreto, R. and Kroll, J.A., 2018, February. Translation tutorial: a shared lexicon for research and practice in human-centered software systems. In the *1st Conference on Fairness, Accountability, and Transparency*. New York, NY, USA (Vol. 7).

Liu, X., Faes, L., Kale, A.U., Wagner, S.K., Fu, D.J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C. and Ledsam, J.R., 2019. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet Digital Health, 1(6)*, pp.e271–e297.

Madden P. and Kanos D., (2021) *Figures of the week: Digital skills and the future of work in Africa*. Available at: https://www.brookings.edu/blog/africa-in-focus/2020/07/22/figures-of-the-week-digital-skills-and-the-future-of-work-in-africa/ (Accessed: 9 June 2021).

Makerere University. (2021) *DeepMind Supports Master's Scholarships to Study AI at Makerere University*. Available at: https://www.mak.ac.ug/slide/deepmind-supports-master%E2%80%99s-scholarships-study-ai-mak (Accessed: 11 June 2021).

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. and Galstyan, A., 2021. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR), 54(6)*, pp. 1–35.

Microsoft. (2021) 'Microsoft collaborates with the Nigerian government to accelerate digital transformation in the country', *Microsoft News Center,* 03 May. Available at: https://news.microsoft.com/en-xm/2021/05/03/microsoft-collaborates-with-the-nigerian-government-to-accelerate-digital-transformation-in-the-country/ (Accessed: 9 June 2021).

MMC. (2019). The State of AI Divergence. Available at: https://mmc.vc/wp-content/uploads/2019/02/The-State-of-AI-2019-Divergence.pdf (Accessed: 10 June 2021)

Moosajee, N. (2021) 'Fix AI's racist, sexist bias', *The Mail & Guardian*, 14 March. Available at: https://mg.co.za/article/2019-03-14-fix-ais-racist-sexist-bias/ (Accessed: 11 June 2021).

Mozur, P., Kessel, J.M., and Chan, M. (2019). 'Made in China, exported to the world: the surveillance state', *The New York Times,* 24 April. Available at: https://www.nytimes.com/2019/04/24/technology/ecuador-surveillance-cameras-police-government.html (Accessed: 11 June 2021).

Mudongo, O., 2021. Africa's Expansion of AI Surveillance-Regional Gaps and Key Trends. *Research ICT Africa*. pp. 1–6.

Nantulya, P., 2018. Grand strategy and China's soft power push in Africa. Africa Center for Strategic Studies, pp. 1–3.

Ndiomewese, I. (2018) 'This Is Nigeria's First Ever Artificial Intelligence Hub', *Techpoint Africa,* 11 June. Available at: https://techpoint.africa/2018/06/11/nigerias-first-ever-artificial-intelligence-hub/ (Accessed: 11 June 2021).

Neri, E., Coppola, F., Miele, V., Bibbolino, C. and Grassi, R., 2020. Artificial intelligence: Who is responsible for the diagnosis?. *La Radiologia Medica, 125(6)*, pp. 517–521.

Nyabola, N. (2016) 'Facebook's Free Basics Is an African Dictator's Dream', *Foreign Policy*, 27 October. Available at: https://foreignpolicy.com/2016/10/27/facebooks-plan-to-wire-africa-is-a-dictators-dream-come-true-free-basics-internet/ (Accessed: 10 June 2021).

Olander, E. (2019) 'China And The Rise Of AI Surveillance Technology In Africa', *The China Africa Project,* 19 September. Available at: https://chinaafricaproject.com/analysis/china-and-the-rise-of-ai-surveillance-technology-in-africa/ (Accessed: 10 June 2021).

Parkinson, J., Bariyo N., and Chin J. (2019) 'Huawei Technicians Helped African governments spy on political opponents', *The Wall Street Journal*, 15 August. Available at: https://www.wsj.com/articles/huawei-technicians-helped-african-governments-spy-on-political-opponents-11565793017 (Accessed: 11 June 2021).

Perkowitz, S., 2021. The Bias in the Machine: Facial Recognition Technology and Racial Disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing*.

Polyakova, A. and Meserole, C., 2019. Exporting digital authoritarianism: The Russian and Chinese models. *Policy Brief, Democracy and Disorder Series* (Washington, DC: Brookings, 2019), pp. 1–22.

Prasad, R., 2018. Ascendant India, digital India: how net neutrality advocates defeated Facebook's Free Basics. *Media, Culture & Society, 40(3)*, pp. 415–431.

Raji, I.D., Smart, A., White, R.N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D. and Barnes, P., 2020a, January. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020a Conference on Fairness, Accountability, and Transparency* (pp. 33–44).

Raji, I.D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J. and Denton, E., 2020b, February. Saving face: Investigating the ethical concerns of facial recognition auditing. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 145–151).

Rakova, B., Yang, J., Cramer, H. and Chowdhury, R., 2021. Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), pp. 1–23.

Roselli, D., Matthews, J. and Talagala, N., 2019, May. Managing bias in AI. In *Companion Proceedings of The 2019 World Wide Web Conference* (pp. 539–544).

Seutloali, G.M., 2015. Own-race bias in facial recognition amongst black, coloured and white participants. *PhD diss., University of Cape Town*.

Sharma, I., 2020. China's Neocolonialism in the Political Economy of AI Surveillance. *Cornell International Affairs Review, 13(2)*, pp. 94–154. https://doi.org/10.37513/ciar.v13i2.554

Shepherd, C. and Qiu, S. 2017. China's Xi lays out vision for 'new era' led by 'still stronger' Communist Party. Available at: https://www.reuters.com/article/us-china-congress-idUSKBN1CM35L

Maeve Shearlaw. (2016) 'Facebook lures Africa with free internet—but what is the hidden cost?', *The Guardian,* 01 August. Available at: https://www.theguardian.com/world/2016/aug/01/facebook-free-basics-internet-africa-mark-zuckerberg (Accessed: 10 June 2021).

Silberg, J. and Manyika, J., (2019) *Tackling bias in Artificial Intelligence (and in Humans*. Available at: https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans (Accessed 11 June 2021).

Teller, A. (2021) 'Loon's final flight', *Medium*, 21 January'. Available at: https://blog.x.company/loons-final-flight-e9d699123a96 (Accessed: 10 June 2021).

Tilouine, J. and Kadiri, G. (2018) 'A Addis-Abeba, le siège de l'Union africaine espionné par Pékin', *Le Monde.fr, 26 January*. Available at: https://www.lemonde.fr/afrique/article/2018/01/26/a-addis-abeba-le-siege-de-l-union-africaine-espionne-par-les-chinois_5247521_3212.html (Accessed: 11 June 2021).

Wareham, C.S., 2021. Artificial intelligence and African conceptions of personhood. *Ethics and Information Technology, 23(2)*, pp. 127–136.

Wojcik, S. and Remy, E., (2019) 'The Challenges of Using Machine Learning to Identify Gender in Images', *Pew Research Center*, 5 September. Available at: https://www.pewresearch.org/internet/2019/09/05/the-challenges-of-using-machine-learning-to-identify-gender-in-images/ (Accessed: 11 June 2021).

Woodhams, S. 2020. China, Africa, and the Private Surveillance Industry. *Georgetown Journal of International Affairs*, 21, pp. 158–165.

World Bank. (2019a) *TICAD Seminar Series: DE4A initiative*. Available at: https://pubdocs.worldbank.org/en/312571561424182864/062519-digital-economy-from-africa-initiative-Tim-Kelly.pdf (Accessed: 9 June 2021).

World Bank. (2019b) *Achieving Broadband Access for All in Africa Comes with a $100 Billion Price Tag*. Available at: https://www.worldbank.org/en/news/press-release/2019b/10/17/achieving-broadband-access-for-all-in-africa-comes-with-a-100-billion-price-tag (Accessed: 9 June 2021).

World Bank. (2020) *Digital Skills for all African Students: What Will it Take?* 28 Feb. Available at: https://www.worldbank.org/en/news/feature/2020/04/14/digital-skills-for-all-african-students-what-will-it-take (Accessed: 9 June 2021).

World Bank. (2021). *Digital Economy for Africa Initiative*. Available at: https://www.worldbank.org/en/programs/all-africa-digital-transformation (Accessed: 9 June 2021).

X, the moonshot factory. (2017) *Loon, Expanding Internet Connectivity with Stratospheric Balloons*. Available at: https://loon.com/ (Accessed: 10 June 2021).

Zeiger, H. (2020) 'China's Eyes Are Watching Africa Closely.' *Mind Matters News* Available at: https://mindmatters.ai/2020/11/chinas-eyes-are-watching-africa-closely/ (Accessed: 09 May 2021).

Zeng, Y., Lu, E., Sun, Y. and Tian, R., 2019. Responsible facial recognition and beyond. *arXiv preprint* arXiv:1909.12935.

Working with Robots as Colleagues: Kenyan Perspectives of Ethical Concerns on Possible Integration of Co-bots in Workplaces

Tom Kwanya

Introduction

In 2013, a group of US soldiers deployed in Iraq shocked the world by holding an elaborate funeral for a robot, named Boomer, which was destroyed in combat. Grieving over the fallen comrade in arms, the soldiers not only honoured the robot with a 21-gun salute but also with two prestigious medals, the Purple Heart and Bronze Star Medal (Carpenter 2013). The Purple Heart is a decoration awarded in the name of the President to soldiers either killed or wounded in service. The Bronze Star Medal is awarded to soldiers for heroic service in a combat zone. Boomer was considered as male and was recognised "posthumously" for heroic exploits and saving the lives of his comrades in arms (Nyholm and Smids 2020). This funeral and the awards demonstrated unique facts about the type of relationships people can develop with the machines they work closely with. Typically, people can humanise the machines to the extent that they regard them as colleagues and treat them in the same way they would treat their human associates. Similarly,

T. Kwanya (✉)
The Technical University of Kenya, Nairobi, Kenya
e-mail: tkwanya@tukenya.ac.ke

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_4

they can develop emotional "brotherhood" bonds with the machines in a process that Renzullo (2019) describes as the "anthropomorphisation" of devices. Nijssen et al. (2019, p. 42) explain that "interactions with machines may give rise to emotional attachment and a humanized perception of them to the point where we start considering them deserving of moral care - something that is usually only reserved for other humans".

Boomer's funeral was not the only incident in which human beings have assigned human characteristics to machines. In 2017, workers at the Canadian Broadcasting Corporation feted "retiring" mail robots with a retirement party (Isa 2018). The party, complete with beautiful balloons and a cake, was held to celebrate the tireless, selfless and colourful career of the robots in the corporation having served faithfully for a quarter of a century (Hooker and Kim 2019). During the party, members of staff narrated stories of their experiences with the robots and how they would be missed in the company. Some of the employees expressed the fear that it may be difficult adjusting to the new work experience without the robots. By organising and participating in such a party, the workers demonstrated that human beings have the capacity to develop unique bonds with machines which is akin to relationships colleagues develop with people they work closely with in work spaces. Indeed, staff members narrated how they had become so accustomed to the robots to the extent that they were part and parcel of the workspace just like human colleagues.

Other recent anthropomorphisation cases include the granting of citizenship of a humanoid robot known as Sophia by Saudi Arabia in 2017 (Retto 2017). Sophia who became the first humanoid to be granted citizenship by any country is an advanced robot which is capable of interacting closely with human beings. Sophia is linguistically advanced and capable of expressing feelings such as anger, joy, sadness, amazement, annoyance or fear (Weller 2017; Yu 2020). Importantly, she is able to learn from her interactions with human beings thereby improving her knowledge and experiences. Thus, she is able to fit better in her context by demonstrating familiarity with the culture, emotions and linguistic expressions of the people she interacts with. She is even capable of searching for new information using platforms such as Google. She has an Instagram account with 160 k followers and 500 posts as at October 2021. As a Saudi citizen, Sophia enjoys the rights of a legal person similar to other legal persons, including human beings (Pagallo 2018; Parviainen and Coeckelbergh 2020). In fact, she is the first robot to be appointed

an ambassador by the United Nations where she serves as the United Nations Development Programme's innovation champion. Sophia now has siblings such as Japanese Erica and Chinese Jia Jia (Riaz et al. 2020).

From the foregoing, it is evident that robots are moving closer to human beings, particularly, in workspaces. This will inevitably expose human workers to a different working environment laden with myriad benefits, challenges and concerns thereby raising a number of ethical questions: How should human workers perceive and interact with robots working alongside them? Are these robots supposed to be treated as tools or colleagues by the humans working with them? Can human workers truly trust the robots they work with in the same way they can trust their human colleagues? What factors influence the acceptance of robots as co-workers by human workers? This chapter explores these ethical concerns in the Kenyan context.

Emerging Frontiers for Robots

According to Calo et al. (2016), a robot can simply be perceived as a machine which has some means of sensing the environment or receiving instruction; an algorithm or programme which enables it to decipher the sensed data or instruction; and can take appropriate action based on the sensed instruction. Gates (2007, p. 65) argues that robots and related machines will soon be found in all spheres of human life, including in homes. He explains that "as these devices become increasingly affordable to consumers, they could have just as profound an impact on the way we work, communicate, learn and entertain ourselves as the PC has had over the past 30 years". Kaur et al. (2021) explain that robots are getting more advanced and becoming more helpful in diverse spheres of human life. These trends exemplify an earlier opinion of Gates (2007) who projected that the power and potential of robots in everyday life are bound to increase because the devices have capacity to be networked. Therefore, it is possible to have groups of networked robots which are able to work together to accomplish tasks which may not be performed easily or comfortably by human beings.

Gunkel (2018, p. ix) opines that human beings are currently "in the midst of a robot invasion. The machines are now everywhere doing virtually everything". Iqbal et al. (2017) explain that what was hitherto imagined in science fiction is finally being realised. Indeed, the International Federation for Robotics (IFR) (2015) projected a gradual but

steady increase of robots in nearly all spheres of human life in what was described as "conquering of the world" by the machines. Cookson (2015) also predicted a ubiquitous presence of diverse types of robots. In the service industry, for instance, he envisioned robots entertaining guests, taking care of the elderly, cooking and serving food in restaurants, and milking cows on farms. Sun (2016) stated that several robots were already being used on farms for irrigation, harvesting and processing of crops. Solaiman (2017) reported that there were already 12 million service robots in operation and that this number would grow exponentially in the future. Demianova (2018) predicted an annual 12% growth in the application of robots in diverse sectors of the economy.

Clabaugh and Matarić (2018) explained that new technological advancements have facilitated the development of advanced robots which are able to stretch the limits of human–machine symbiosis to levels that have hitherto not been experienced. Currently, there exist robots for many human endeavours. Pagallo (2013, p. 47) reported that there are artificial agents with the ability to "send bids, accept offers, request quotes, negotiate deals and make contracts". Other latest entrants into the robotics universe include machines which are able to "protect and improve the quality of air, water, and soil; safeguard species biodiversity; and effectively manage natural resources" (Mazzolai et al. 2021); diverse categories of social robots used for educational purposes (Konijn et al. 2020; Xia and LeTendre 2021); automation of diverse forms of industrial processes (Stein and Kaivo-Oja 2020); and those performing assistive roles in homes and medical facilities (Nomura 2017), among other roles.

Many challenges hamper advancement of robotics. According to Gates (2007, p. 60), "the robotics industry faces many of the same challenges that the personal computer business faced 30 years ago. Because of a lack of common standards and platforms, designers usually have to start from scratch when building their machines". Dautenhahn et al. (2005) conducted a study in the United Kingdom to understand the perception of robots among 28 adults. The study found that a large majority of the respondents preferred having a robot as a machine, assistant or servant but not as a companion or friend. Most of the respondents also preferred to assign robots household chores not involving the care of children or animals. Scopelliti et al. (2004) conducted a study which revealed that whereas young people demonstrate an openness towards robots in social set-ups, including in workplaces and homes, the elderly are hesitant and are actually frightened of the prospect.

Cookson (2015) argued that although industrial robots have for many years dominated robotics, the balance has tilted in favour of service robots. These robots are more than just machines. Demianova (2018) explains that many workers fear that the increasing use of robots as an affordable source of efficient labour will render many workers jobless leading to new dynamics in the labour markets globally. Sharkey and Sharkey (2012) argue against the use of robots in a way that totally replaces human beings particularly in situations which require emotional support which robots, being machines, cannot give. They give the example of providing care support to the elderly persons who would feel neglected and isolated if the use of robots extensively reduces their interaction with human beings. Pransky (2001) argues that letting children to interact extensively with robots may lead them to have less interaction with other children and humans thereby leading them to think that human–robot interaction is actually the norm. Fridin (2014) holds the view that exposing young children extensively to robots may lead to poor emotional and social development. This may lead to difficulties in engaging with other humans or being confused about the abilities of humans and robots. Interacting with robots more than human beings may also affect children's moral consciousness. She calls for a balance in the levels of exposure of children to robots and other human beings. Several scholars echo this call (Bertolini and Aiello 2018; de Graaf 2016; Fiske et al. 2019; Haring et al. 2019; Lin et al. 2011; Yulianto 2019).

According to Solaiman (2017), several discourses are ongoing about the personality of robots. One of these is the view that robots should acquire their own legal personality so that they, and not their manufacturers or owners, are held responsible for their own mistakes. Gunkel (2018) poses: Can robots be held responsible for the consequences of the mistakes they make in their line of duty? Do robots, being the machines that they are, expect some level of respect from human beings? Can they be treated with dignity reserved for human beings and other living things? Can robots have any rights?

Contextual Information

Robots have been used in diverse sectors of the economy for decades. The robots, which were largely industrial, worked in controlled spaces not shared with human workers (Fryman and Matthias 2012). Industrial robots supported heavy manufacturing processes. They were intimidating

huge mechanical machines operating autonomously in industrial plants (Dauth et al. 2017; Singh et al. 2013). With technological advancement, new types of robots, called collaborative robots or co-bots, have emerged. As opposed to industrial robots, co-bots are designed and deployed to work closely with human workers in less controlled workspaces (Fast-Berglund et al. 2016; Galin and Mamchenko 2020; Kildal et al. 2018; Simões et al. 2020; Veloso et al. 2015). According to Vojić (2020), co-bots have been largely deployed to perform manual duties such as packing, picking, welding or assembling parts of products. There are also co-bots in the service industry offering user support and customer care. Marvel and Norcross (2017) report that the integration of co-bots in workspaces has been on the increase. For instance, Cohen et al. (2019) projected that the US alone will spend 12 billion dollars on co-bots in 2025, up from 710 million dollars in 2017. It is also estimated that more than 5 million co-bots are already sharing human workspaces globally. It is further estimated that about 400,000 co-bots will be joining the human workspaces yearly in the near future. Malik and Bilberg (2019) opine that co-bot deployment results in the automation of up to 70% of the workload making processes and production to become more efficient and profitable.

Kenya is one of the most technologically advanced countries in sub-Saharan Africa (Kwanya 2021). Its capital, Nairobi, is the technological hub of the East and Central African region earning it the title "Silicon Savannah" (Kwanya et al. 2021). The country is reputed for having one of the best Internet connections in Africa due to the number of undersea cables which land in it (Bramann 2017). Kenya also boasts of the use of the latest technologies in its economy; a dominant presence of multinational ICT companies implying access to latest technologies; a growing population of young people amenable to technological developments; a relatively well-educated population (adult literacy is about 78%); and a history of technological innovation. Indeed, Nairobi is the home of globally-celebrated technological innovations such as mobile money transfer platform, M-pesa, among others (Kwanya 2021).

Kenya's short-term development blueprint is anchored on the "Big 4 Agenda" which is aimed at improving the national economy by enhancing manufacturing, improving food security and nutrition, attaining universal healthcare coverage and providing affordable housing to the citizens (Macharia 2019; Musundi et al. 2021). The Government of Kenya has committed to use emerging innovative technologies to attain the "Big 4

Agenda" targets. It is specifically seeking to mainstream the use of artificial intelligence, machine learning and robotics to achieve its development agenda (Mvurya 2020). Consequently, the government and other stakeholders are already experimenting with co-bots in various sectors of the economy. The latest initiative was the deployment of a medical co-bot, known as Robodoc, to support the country's response to the COVID-19 pandemic. It is expected that the country will experience an influx of co-bots in the next few years (Kwanya 2021).

Literature Review

The category of robots which work and interact closely with human beings is known as collaborative robots or co-bots (Peshkin and Colgate 1999; Kwanya 2021). According to Castillo et al. (2021), co-bots are designed for close physical and functional collaboration with human workers. Colgate et al. (1996) explained that co-bots are robotic devices which perform assigned tasks in collaboration with human workers. They work in close proximity with human beings and share their workload. Their close physical proximity and sharing of work imply that co-bots are constantly in contact with the human beings they work with. Peshkin and Colgate (1999) further explained that co-bots work collaboratively but fairly independently from human intervention. Compared to industrial robots, co-bots are more flexible, adaptable and safe (Fast-Berglund et al. 2016). Table 1 summarises the characteristics of co-bots and compares the same with industrial robots.

According to Yilma et al. (2020), co-bots possess human-like characteristics making them able to collaborate with them on a day-to-day basis. Pazienza et al. (2019) explained that co-bots are specially designed for close and constant interaction with human beings in ordinary workspaces. Cohen et al. (2021) argued that co-bots are critical for the realisation of the 4th Industrial Revolution. Co-bots blur the boundaries between the digital and physical work environments as well as the distinction between human workers and machines. Sladić et al. (2021) also explain that co-bots are designed to collaborate with other robots and humans in performing tasks. Adriaensen et al. (2021) emphasise that while most of the jobs assigned to robots were those which are ergonomically or psychologically challenging for humans to do, co-bots work on the same tasks in the same workspace with human beings. According to Margherita and Braccini (2021), co-bots utilise a human-worker-centric approach in

which the robot does not entirely replace human labour but enriches it. According to Campbell (2021), the greatest benefit of co-bots lies in the fact that they provide an alternative to the use of costly and complex traditional robots. Sony et al. (2021) assert that co-bots will drive the realisation of the fourth industrial revolution by synergising human labour with robotic power. Gjeldum et al. (2021) emphasise that it is the ability of co-bots to directly mingle with and share tasks with humans that make them more versatile in the Industry 4.0 workspaces.

A number of models of interaction between humans and robots in the human workspaces have been identified. The first model of interaction is co-existence. According to Wang et al. (2019), co-existence is a situation where the robot and human share space but perform different tasks. The second model is cooperation. Gjeldum et al. (2021) explain that in cooperation, the robot and the human perform different components of a task but sequentially. The subtasks performed by the robot and the human are mutually independent. The third model is collaboration. According to Vicentini (2020), collaboration involves the robot and human working on the same task at the same time to yield a mutual result. Collaboration brings the robot and human being together to work on the same assignments simultaneously in a relationship that Gupta et al. (2021) describe as mutually-inclusive and contagious. Rossato et al. (2021) explain that collaboration enables organisations to combine the strengths of robots (such as tirelessness, speed and accuracy) with those of the human worker (such as flexibility and dexterity). McQuillen (2021) asserts that when this happens, robots and humans maximise their strengths while also compensating for their individual weaknesses.

Co-bots are now quite common in several sectors of the economy (Fast-Berglund et al. 2016). Many benefits are derived from co-bot deployment in human workspaces. These include precision in job performance (Pazienza et al. 2019; Zhu et al. 2020); improved productivity due to enhanced human–robot collaboration (Gobinath 2021; Sowa et al. 2021); production efficiency due to processing optimisation (Castillo et al. 2021); low production costs (Castillo et al. 2021; Gisginis 2021; Kadir et al. 2018); better quality of products or job outputs (Galin et al. 2020); job enrichment and fulfilment through improved workflows and role distribution (Kadir et al. 2018; Margherita and Braccini 2021); improved capacity to handle complex assignments (Simoes et al. 2019; Zhu et al. 2020); improved safety of human workers when robots take up risky and repetitive duties (Fast-Berglund et al. 2016; Zhu et al. 2020);

quick performance of tedious and sophisticated tasks (Gisginis 2021); as well as improved customisation of goods and services (Kopp et al. 2021).

Many factors have influenced the adoption and use of co-bots. Sladić et al. (2021) identify lack of requisite technical skills to work with robots as one of the factors limiting the deployment of co-bots. Pinto et al. (2021) argue that many human workers, including engineers, do not wholly trust robots and perceive them as being volatile, uncertain, complex and ambiguous. Lambrechts et al. (2021) argue that resistance to change, organisational culture and leadership are among the factors which limit the adoption and use of co-bots. Other factors which hinder widespread adoption and use of co-bots include occupational safety concerns (Kopp et al. 2021). McQuillen (2021) suggests that giving some level of control over the robot to the human coworker enhances the acceptability of the robot by the human. Lambrechts et al. (2021) suggest that reskilling and upskilling human workers to collaborate with co-bots would enhance acceptability and effectiveness of co-bots into their workspaces. Giacometti and Larsson (2017) argued that addressing safety concerns of workers in spaces and tasks shared with robots would also enhance their acceptability and operational gains.

According to Calitz et al. (2017), there is a great opportunity to deploy co-bots in sub-Saharan Africa. Although only a few manufacturing companies and businesses in Africa have deployed co-bots, the machines are acknowledged as having a great potential to drive economic growth of the region. Calitz et al. (2017) also identify the factors contributing to the relative slow adoption of co-bots in sub-Saharan Africa to include perceived high costs of purchase, installation, operation and maintenance; low cost of labour in the region; lack of requisite technical capacity and infrastructure; and lowly educated but highly unionised workforce who fear job losses. According to Chigbu and Nekhwevha (2021), workers in Africa need to acquire new skills to fit the needs of new job tasks which involve collaborating with co-bots. Keet (2021) also suggests that developers of co-bots need to demonstrate cultural awareness when developing co-bots for use in Africa. Dunn (2021) recommends the development and use of policies which facilitate the development of human skills which are relevant for the fourth industrial revolution.

According to Anderson and Anderson (2010), autonomous machines like robots are bound to play a critical role in human life. They further argue that the big question is whether they will do this ethically. Ethical concerns about the widespread use of robots have been on the table for

many years. Indeed, as early as 1941, Isaac Asimov, an American professor, proposed the following three laws of robotics (Asimov 1941):

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm;
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law; and
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Anderson (2008) argues that the above laws, though a pointer to the need for ethical machines, were unsatisfactory, thereby necessitating the continued dialogue on ethics in robotics. Indeed, Maddahi et al. (2021) explain that robotics has advanced over the years and have outgrown these laws. For example, there are now robots, such as Boomer which was mentioned earlier, used in warfare and are definitely designed to harm enemy human beings.

Maddahi et al. (2021) argue that three main ethical issues arise from the use of co-bots. These include privacy and safety of data either generated or stored by the robot; the impact of robot use on the common good of the society in which they are applied; and the safety of human beings working with the robots or operating in their work environment. Ménissier (2020) explains that the Montreal Declaration for the Responsible Development of Artificial Intelligence aimed to provide a framework for ethical application of artificial intelligence in society. A number of ethical principles can be derived from the declaration. These include respect for autonomy, caution, diversity inclusion, responsibility, democratic participation as well as protection of privacy and intimacy. Maddahi et al. (2021) explain that these principles are human-centric and do not have adequate provisions for advanced robotic machines.

Other ethical dilemmas relate to the possibility and nature of bonding relationships between robots and their human co-workers as well as the socioeconomic effects of widespread automation of tasks (Maddahi et al. 2021). Anderson (2008) wonders whether autonomous machines can have a moral standing. Tolksdorf et al. (2021) explain that in cases where robots are trusted to take care of vulnerable persons, such as children and the elderly, these machines may not be expected to uphold values such as human morals. Therefore, ethical dilemmas arise regarding

the moral safety of the persons under their care. Again, there are also concerns about human beings absconding from their duty of taking care of their offspring and parents. Leaving these fundamental responsibilities to robots alters the social structure of the society and may lead to children growing up with "machine" minds which may affect their capacity to relate well with other human beings when they grow up (de Graaf 2016). Indeed, Anderson et al. (2005) explained that scientists find it difficult to make ethics computable. Robots are disruptive technologies which also stimulate social changes in their environments. If not handled carefully, such changes may have outstanding impact on the humanness of society as we know it today. Tan et al. (2021) point out that these disruptive changes have more serious and long-term social safety implications that can be comprehensively deciphered or anticipated. Liang et al. (2021) explain these ethical dilemmas relating to robots will persist and have consequences on human rights, morals, values, justice and equity.

Anderson (2008) citing the arguments of Asimov (1976) explains that some people hold the view that autonomous machines should have rights to act independently without having to work as slaves for human beings. Chomanski (2021) also advocates for the freedom of artificial persons arguing that they should be treated just like human beings and should not be enslaved or exploited. In this regard, Peeters and Haselager (2021) suggest that robots should be designed with features that enable them to consent to assignments or other forms of interaction with their human operators or co-workers. According to Bennett and Daly (2020), these debates led to discussions on the status of robots as legal persons. These discussions have not been conclusive. Nonetheless, they explain that the rights assigned to a robot may depend on many factors about the robot, including its assumed gender. Nonetheless, other scholars (Kelley et al. 2010; Calverley 2006) hold the view that perhaps it would be better to treat robots in the same way we treat domestic animals or pets. In this case, they would not have exactly the same rights or moral expectations as human beings. The diversity of views on the rights of robots points to the fact that the subject is complex and non-conclusive.

According to Boada et al. (2021), current debates on ethical issues relating to ubiquitous robot use in society are fragmented and conceptually disordered. This limits the perception of ethical risks emanating from integrating robots into the social spheres of human life and the interventions therein. It is also difficult to prescribe a set of ethical principles for machines because ethics is relative. Similarly, ethical dilemmas exist which

make it difficult to prescribe actions in all imaginable decision situations in human life.

Although the sub-Sahara is lagging behind the developed countries on the adoption and use of robots, the situation is changing fast. According to Rapanyane and Sethole (2020), the rise of the use of artificial intelligence and robots in the sub-Sahara is driven by the need to facilitate the fourth industrial revolution in the region. Indeed, Mayer (2018) and Naudé (2017) argued that meaningful industrialisation on the continent will be realised fast by embracing artificial intelligence and robotics in manufacturing. In Kenya, Banga and te Velde (2018) argued that the country can make fast progress in manufacturing by using intelligent systems and equipment. These views echoed positions taken by Kibor and Obwoge (2014), Arunda (2020) and Anitah et al. (2019). Although statistics about the actual level of adoption of robots in Kenya's workspaces are blurry, it can be deduced that this situation will not hold for long. It is just a matter of time before there are a sizable number of robots working alongside human beings in diverse sectors of the economy in Kenya. This will definitely lead to ethical concerns and dilemmas. It is evident from this literature review that Kenyan perspectives to the ongoing debates on ethical issues emerging from the entrance of robots in the job market are lacking. This chapter explores this subject as a means of contributing Kenyan perspectives to this important global debate. The degree to which these salient ethical issues are identified, discussed and addressed will determine the success of robots in Kenyan workspaces.

Methodology of Study

This chapter is founded on a qualitative study conducted to investigate the ethical concerns of Kenyan workers on the integration of co-bots in their workspaces. According to Kahlke (2014), qualitative studies enable researchers to understand issues under investigation from the perspectives of the respondents. Dongre et al. (2010) argued that applying qualitative methods of data collection and analysis helps researchers to explore research issues in-depth and to adequately integrate the opinions of the respondents. This chapter discusses the perceptions and attitudes of Kenyan workers towards collaborative robots in their workspaces. Therefore, it relied heavily on the views of the respondents. A qualitative approach was, therefore, deemed most appropriate to collect data for the study.

Data was collected from 20 purposively selected information scientists in Kenya. According to Malterud et al. (2016), the sample size of populations in qualitative studies depends on the objectives and context of each study. However, they point out that the "information power" or level of expertise of the respondents is critical in determining their number. Thus, the more expert the respondents, the smaller the sample size and vice versa. According to Braun and Clarke (2021), it is possible to reach information saturation with 20–24 interviews. In the study anchoring this chapter, the respondents were experts in information science who were selected through information-oriented purposive sampling. Information scientists were selected because they are likely to encounter robots in non-industrial environments. A sample size of 20 respondents was also deemed as adequate to provide an exhaustive view of the qualitative issues under study. Qualitative data was collected from the selected respondents through telephone interviews. This data collection technique was considered appropriate because of restrictions on physical meetings due to COVID-19 pandemic. A semi-structured interview schedule was used to guide the interviews. The key questions asked included:

1. Are you aware of co-bots?
2. In your opinion, what extent are they deployed in Kenya?
3. Have you encountered a co-bot in your professional work?
4. In your opinion, what are the benefits of co-bots?
5. Are you willing to accept a co-bot as a work colleague? Please, explain your answer.
6. Are you willing to share your workspace with a co-bot? Please, explain your answer.
7. In your opinion, is it ethical to treat a co-bot a colleague in the workplace?
8. If your response to 7 above is yes, what ethical concerns would you have about treating a co-bot as a colleague?
9. How can the concerns identified in 8 above be addressed to enhance acceptability of co-bots in workspaces?

Findings of the Study

The collected data was analysed thematically based on the questions above. The findings of the study are presented here according to the themes. The themes are awareness of collaborative robots, extent of their deployment in Kenya, past experience with co-bots, benefits of integrating

co-bots in human workspaces, willingness to work in close proximity with co-bots, ethical concerns about sharing workspaces with machines such as co-bots, as well as strategies to enhance acceptance of co-bots in human workspaces.

Awareness of Co-bots

All the respondents were aware of the existence of co-bots. When asked to explain what they knew about co-bots, the overall response was that co-bots are robots which have been specially designed and developed to operate in ordinary spaces alongside human beings. Some verbatim responses are as reported hereunder:

> *Co-bots are small robots which interact directly with human beings in ordinary life spaces such as offices, markets and homes.* **R11**
>
> *Co-bots are robotic machines which help human beings to perform normal tasks in homes and offices efficiently.* **R07**
>
> *Cheap, simplified and adaptable robots which are able to operate in normal spaces used by human beings to perform a wide array of tasks to make human life more convenient and comfortable.* **R13**
>
> *A category of robots which can work alone but also together with people and other robots in completing shared job tasks.* **R10**
>
> *Co-bots are advanced human-friendly robots which can operate in normal job locations and can work closely with human beings.* **R3**

It can be deduced from the responses that information workers in Kenya have a good understanding of what co-bots are. They describe co-bots as a category of robots which are designed to work in close physical interactions with human beings and less controlled and human-friendly environments. They also perceive co-bots as representing an advancement in robotics aimed at making robots pleasant and safe to work in close proximity with human beings. This description tallies with the definitions of co-bots found in the reviewed literature (Castillo et al. 2021; Kwanya 2021; Peshkin and Colgate 1999).

These findings demonstrate that people in sub-Saharan Africa are abreast with emerging developments in robotics and associated technologies. Indeed, Calitz et al. (2017) explain that most business enterprises in Africa know about collaborative robots and their potential role in facilitating the realisation of the fourth industrial revolution. Naudé (2017) explained that although the level of awareness of co-bots in Africa is

growing rapidly, more needs to be done to harness their potential. Isa (2018) pointed out that even with relatively cheaper labour, Africa will continue to lag behind developed countries if technologies which provide efficient labour, such as co-bots, are ignored. She emphasises that the advantage associated with cheap labour force in Africa will be eroded incrementally by assistive technologies like collaborative robots. Chinyamurindi and Mey (2017) assert that there is need, therefore, for strategies which will turn the awareness into programmes for adopting co-bots in the workspaces. According to Chigbu and Nekhwevha (2021), Africa will continue being a technological desert if no comprehensive actions are taken to transform technological awareness into reality. Rapanyane and Sethole (2020) assert that assistive technologies are inevitable. The earlier Africa embraces the technologies the better for the continent. In this regard, this level of awareness of co-bots is positive and should be encouraged.

Extent of Deployment of Co-bots in Kenyan Information Workspaces

All the respondents indicated that they were not aware of any co-bot currently deployed in the information workspaces in Kenya. They acknowledged the presence of diverse automation systems in the country but stated that none of these meets the descriptions of co-bots given above. They also stated that there is a limited number of industrial robots in Kenya's manufacturing sector. The findings of the current study concur with Magachi et al. (2017) who investigated the use of industrial robots by listed manufacturing companies in Kenya. They found a low application of industrial robots by the companies. They attributed this low usage to high costs of acquiring and deploying robots as well as inadequate technical skills to operate them. Nganga (2020) also reports low usage of robots in Kenya and attributes this to inadequate skills, infrastructure and policies. Nonetheless, in the wake of health challenges occasioned by the COVID-19 outbreak, Kenya is one of the countries which turned to robots to reduce the spread of the disease. In partnership with the United Nations Development Programme (UNDP) and Japan International Cooperation Agency (JICA), the country's Ministry of Health deployed three robots—*Jasiri*, *Shujaa* and *Tumaini*—at the Jomo Kenyatta International Airport and Kenyatta National Hospital in February 2021. The robots handle high-exposure duties such as temperature screening, automatic disinfection and fumigation, as well as identifying those who are

either standing too close together or not wearing masks. The robots, all given the male gender, also collect other critical health data from high-traffic locations for decision-making and intervention. A number of drones have also been deployed in the country to support telecommunication, security/military, agricultural, health, educational, tourism and customer service operations, among many others (Achieng et al. 2020; Arunda 2020; Banga and te Velde 2018; Forbes et al. 2020; Mvurya 2020; Steer 2017).

It is evident from the findings that the positive level of awareness of co-bots in Kenya is not followed with a commensurate extent of deployment in information workspaces. It is noteworthy, however, that the number and diversity of robots in other workspaces in the country are growing, albeit gradually. It can be concluded, therefore, that it is just a matter of time before more robots enter into Kenya's information workspaces in libraries, mass media institutions, archives, museums, records centres and publishing houses, among others. This situation is not entirely unique to Kenya. In sub-Saharan Africa, only South Africa has co-bots in information workspaces. Ocholla and Ocholla (2020) reported that "Libby", a humanoid robot deployed in the University of Pretoria library in 2019, is the first and only humanoid librarian in sub-Saharan Africa. Libby, originally made in China, is considered as a library employee and offers services alongside other librarians in the institution. They add that although Libby is way ahead of her time, it is a harbinger of more robots to be deployed in information centres in other countries in sub-Saharan Africa. Tella (2020) also argued that the fact that there are few co-bots in information centres in sub-Saharan Africa is by no means a permanent situation. He predicted that robots will soon be part and parcel of sub-Saharan information centres in the near future.

Merits and Demerits of Co-bots in Information Workspaces

Deployment of co-bots in information workspaces can result in several benefits. The respondents explained that the key benefits revolve around helping information workers to perform routine tasks, offering non-stop services, taking up jobs considered unhealthy or risky for human beings, and offering consistent services efficiently. In libraries, for instance, co-bots can shelve books, conduct library orientation for new students, take stock of and label library resources, as well as check-in and check-out information materials (Kwanya et al. 2014). These routine duties take

up a large portion of librarians' time leaving no space for creativity and innovation. By freeing this time, co-bots will enable librarians to spend this valuable time in other roles. In record centres and archives, co-bots can easily work with dusty materials and artefacts which are considered a health hazard for human beings. They can also ensure accurate classification and filing of documents to facilitate prompt searching and retrieval (Kwanya 2021). This will not save the time of the users but also improve their decision-making process ultimately resulting in improved productivity. Co-bots can also be relied on to offer uninterrupted services since they do not have to take leave, fall sick or become moody (Abok and Kwanya 2016). They overcome limitations to consistent services experienced by human workers. This consistency results in steady and superior services throughout the day. In this age where information users seek services on a 24-h, everyday basis, co-bots will aid information centres to expand the reach of their services and reduce barriers to access (Gichora and Kwanya 2015; Nakitare et al. 2020). Some of the verbatim responses are as reported hereunder:

A robot concentrates fully on the job assigned to it since it does not have other responsibilities in life like taking care of sick relatives, attending burials or weddings, dropping or picking children from school. **R07**

Co-bots can perform heavy-duty roles in information centres such as lifting files and artefacts. These tasks would require the effort of many people at a time. In this manner, co-bots will help information centres to cut costs while offering efficient services. **R03**

Robots do not get tired and therefore do not need time off to rejuvenate. As long there is work to be done, they will do it. **R09**

The knowledge held by a co-bot is readily available in the organisation since it is largely explicit. The organisation faces no risk of knowledge loss with staff turnover. In fact, turnover itself is greatly reduced. Co-bots serve faithfully and loyally. They are not looking out for greener pastures. Similarly, they cannot get disgruntled or experience burnout. **R12**

Robots can store and retrieve vast volumes and diversity of data much faster than human beings. This capacity can be used to personalise information services and products to the needs of individual users. This helps to improve the relevance and impact of services. **R10**

Co-bots can keep organisational secrets. They can effectively manage confidential records since they do not gossip or feel the obligation of extending favours. They have no favourites. **R08**

Co-bots can offer transparent and accountable services. They are not corrupt…they cannot take bribes or show impartiality. They can help information centres to overcome challenges associated with negative ethnicity and gender biases which are common in Sub-Saharan Africa workspaces. **R01**

The respondents also explained that despite the many benefits of co-bots, there are also a number of disadvantages of having the machines in information workspaces. Generally, human workers would find co-bots as cold, inflexible and dangerous. Some of the verbatim responses are as hereunder:

What if a co-bot goes berserk while on duty? The consequences would be disastrous. **R11**

Although there are co-bots which can tell stories and share jokes with colleagues, this is quite limited. They have no clue about current affairs. Therefore, their stories and jokes will largely be stuck in time. They can also not tailor their stories or jokes to the mood of the day or personal interests of the colleagues. Working with a co-bot would the most boring part of any worker's life. **R13**

A co-bot does not eat. Therefore, it has not packed lunch to share with a colleague. It lacks the attributes of brotherhood in the workspace. **R04**

Some information materials and artefacts are fragile. It is unimaginable them surviving for posterity with the roughness of robots. **R05**

When faced with unique challenges requiring discretion, co-bots would not be able to reason beyond the programme they have. There are cases, especially in customer services, when workers have to make decisions contrary to the established protocols or policies. The context of the case determines and justifies the decision. **R15**

Robots have no feelings or emotions; they cannot be motivated. They also have no stake in the organisation. They don't care whether the company closes or thrives. They have no capacity to appreciate success or failure. **R17**

Co-bots in Information Workspaces in Kenya

In spite of the demerits discussed above, all the respondents were willing to share their workspaces with co-bots. However, there was divided opinion on whether they would consider such co-bots as colleagues or tools. Those who would consider co-bots as colleagues argued that given that they would share space and tasks, the contribution of the co-bot is much more than that of a tool. They further explained that a tool is

operated but co-bots work independently with minimum human intervention. However, they emphasised that they would remain senior to the co-bot and where needed rather than taking instructions from the co-bot. Therefore, they would remain superior in the workspace and exercise authority over co-bots they work with. Some of the verbatim responses are as reported hereunder:

> *The close proximity in which co-bots and humans work as well as the sharing of tasks qualify them to be considered as colleagues. They can be cataloguing books together. Maybe the human worker enters the metadata while the co-bot attached bar codes and shelves the books. As long as the human is not operating the robot, it is a colleague and not a tool.* **R10**
>
> *The basic perception of a colleague is an entity, human or otherwise, with whom/which one works closely and shares roles or their components. Therefore, co-bots are artificial colleagues.* **R01**
>
> *Co-bots being artificial workers must take instructions from their human colleagues who are far more intelligent than they are. Yes, co-bots can be junior colleagues performing routine, risky and tedious tasks under the guidance of a senior human colleague.* **R09**

Those who would treat co-bots as tools and not colleagues explained that being machines, their contribution to the job roles is limited. They merely extend the performance of human beings. Even where their performance is higher than the human being, they are merely offering support to the human being. They also emphasised that machines such as co-bots can only be colleagues with human workers if they are considered equal. In their opinion, there is no way co-bots will ever equal human beings. Therefore, regardless of their advancement, they will remain tools in the hands of the human workers. Some verbatim responses are hereunder:

> *Any entity which is artificial can be equalled to a living human being. By considering co-bots as colleagues, human beings would be imagining them as their equals. Yes, they may be stronger but can only serve as beasts of burden to their human operators.* **R04**
>
> *Machines make work easier for human beings. Machines have no interest in performing any work. Only human beings are able to attach value to work. Therefore, co-bots cannot appreciate the need for any work. They get no benefits from working. They do whatever they are assigned to do by human beings to make the work easier, convenient or safer for the human being.*

Therefore, they cannot be equals or colleagues with the humans. They are machines, tools, equipment. **R17**

Work is a God-given responsibility of human beings. Co-bots are tools helping humans to work better. They are not colleagues. **R16**

It can be concluded from the foregoing that information workers in Kenya are generally open to receiving co-bots in their workspaces. However, they would like to exercise absolute control over the co-bots either as a master or as a senior colleague.

Ethical Concerns About Considering Co-bots as Colleagues in the Information Workspaces

All the respondents had ethical concerns about considering co-bots in their workspaces as colleagues. One of the concerns, as has already been mentioned, revolved around the concept of work. The issue here is that work is a natural responsibility of humans. By being considered as colleagues, co-bots would be usurping God-given human responsibility which they lack the moral authority to do. In this regard, the respondents emphasised, as explained earlier, that co-bots can only support human beings in doing work. Their role is subordinate. The human being is the owner of work. Regardless of their contribution, co-bots can never be ethically considered as colleagues or equals with humans in matters of work. Besides, Lueg and Twidale (2018) argue that it is not possible to replicate human intellectual abilities in robots. They state that people are now more concerned about developing interfaces for robots than for "mammals who get tired, bored, excited, irritated, intrigued, or distracted, and who even change their minds about what they want to do" (p. 409). The views of the respondents tally with those of Pauliková et al. (2021) that co-bots should only supplement the work of humans and not replace them. Co-bots seek to change the concept of work. Indeed, García-Esteban et al. (2021) explain that the attribute which distinguishes co-bots from the other robots is the fact they can work independently without human intervention and, therefore, do not just complement human labour. They can own work and complete it without human support. Information professions in Kenya, who were the respondents in this study, hold the view that the concept of work and its ownership need to be clarified. This will help to answer the question

on whether co-bots, being machines, can work or not. They can only be considered colleagues if they can work.

Related to the issue above is the concept of the value of the human being. If co-bots can be considered colleagues, is the value attached to them equal to that attached to the human being? Will humans be treated as co-bots and vice versa? Is this not a recipe for chaos in society? With the debates about the rights and personality to ascribe to co-bots, this equality issue is serious. The scholars, such as Chomanski (2021), who advocate for co-bots to be given the same rights and personality humans have, are essentially equating the machines to humans. The rights include the ability to consent to instructions or not (Peeters and Haselager 2021). Bennett and Daly (2020) propose that co-bots should enjoy all the rights given to legal persons. The respondents in this study hold that humans are above co-bots and therefore can never be equated to them whether in the workplace or elsewhere in society. In Kenya, human life is held as sacred. The respondents could not see the conditions under which similar status can be given to machines which are created by other humans. They argued that as opposed to offspring, co-bots are not like the people who create them. They can never grow to become equals with their human creators. In their opinion, it is not possible to share human comradeship in workspaces with co-bots. They underscored the fact that co-bots can, indeed, collaborate with human beings in performing specific roles. However, this does not qualify the robots to be considered as colleagues to the human beings because that is tantamount to equating the artificial with the natural.

Another issue relates to the fear that co-bots will replace humans or at least reduce their role in the workspace. Besides denying humans their God-given right to work, this will deny them opportunity to make a living from working. The big question here is: Would it be ethical to give robots work when many people in Kenya are jobless? Available statistics indicate that about 20% of youth in Kenya are unemployed and the level is likely to increase (Gachari and Korir 2020; Njogu 2015). As explained by Demianova (2018), the fear that robots will replace human beings in some job functions is real. This is partly because robots are bound to be cheaper than human labour. The respondents argue that any initiatives which replaces humans in the workspace with robots or reduces their presence therein give the implication that robots are more important than humans. Besides, denying any human beings work is tantamount to denying them life. Without humans, can there be work for robots?

Therefore, the respondents emphasised, the needs of human beings for opportunities for work should be met before robots can be given work.

The entry of co-bots into information workspaces will result in unintended changes in values, work ethics and moral standards which will have far-reaching consequences on society. For instance, working closely with co-bots will make human workers as cold, emotionally, as the robots. Similarly, it will reduce humans to work mechanically as machines without exhaustively applying their intellectual abilities. Furthermore, humans will develop relationships and bonds with their co-bot colleagues which may isolate them from fellow human beings both in the workspaces and in the society at large. In Kenya, and also in the rest of sub-Saharan Africa, people value social connections and togetherness, possibly more than wealth. The respondents argued that the benefits of robots should not break the social ties which bind communities together. As long as there is potential for this, co-bots will always be viewed with suspicion in Africa. In Africa, "anthropomorphisation" is viewed as immoral and evil. Therefore, deep relationships with devices as the American soldiers did with Boomer are unlikely to develop or be encouraged. Such relations will be perceived as perversion and will attract stigma and disdain.

The other ethical issue relates to trust. Can machines be trusted fully? Can machines make and keep promises? Can machines be confidants? Can robots keep secrets? Can robots advise humans on social or emotional issues? Can robots be role models? Indeed, Pinto et al. (2021) argued that even engineers find it difficult to fully trust robots which they develop. They find them to be volatile and unreliable in some instances. The respondents argued that it is not possible for humans to fully trust co-bots which do not have a sense of loyalty or moral standards. Therefore, co-bots cannot be true colleagues until humans are able to trust them enough to confide in them, seek advice from them, and accept them as role models. The respondents were unable to predict whether these concerns can be addressed through technological advancements alone. There is a need to socialise information professionals differently if they are to accept co-bots as colleagues who are worthy of trust. This will undoubtedly take a lot of time and effort and may not be realised in this generation.

Robots, to a large extent, are culturally dumb. This implies that they lack essential cultural sensitives like etiquette, values and codes of morality. Besides being professional, workspaces are also cultural (Wallace 2021). Although there are advances towards social robots (Jones 2017), it is not

possible to find robots which can fit perfectly in authentic communities like those in Kenya and the rest of sub-Saharan Africa. The respondents explained that most robots have been developed in exotic cultures. So far, no commercial robot is indigenous to Kenya. This means that they are unlikely to accommodate the way of life in Kenyan workspaces. This is exacerbated by the fact that there are limitations on what a robot can learn after development. It will be akin to teaching an old dog new tricks.

Limitations

No co-bots are currently deployed in Kenya's information workspaces. Therefore, the views presented here are not based on the actual presence of co-bots. It is likely that the respondents would consider the issues differently if they were already working in the presence of co-bots. Nonetheless, the opinions are backed up with literature from environments in which co-bots already exist.

Conclusions

From the foregoing, this chapter concludes that Kenyan information workers are willing to welcome co-bots into their spaces but as tools. They are of the view that co-bots do not meet the threshold of brotherhood and, therefore, cannot be considered as colleagues. The respondents explained that many factors influence the acceptability of co-bots into Kenya's information workspaces. These include basic enablers such as skills and infrastructure. However, they argued that there are ethical issues which are more deep and paramount. These include the concept of work as a divine gift to humanity which cannot be shared with machines; the notion that treating co-bots as legal persons equates them to human beings which is viewed as demeaning to humanity; the fear that co-bots will dominate and eventually replace humans in ordinary workspaces thereby denying them not just an opportunity to work but to livelihood; fear of unintended social consequences of "anthropomorphisation" which drive society to oblivion; lack of trust for machines created by limited humans to offer unlimited services and companionship; and discomfort with exotic robots entering professional and indigenous spaces. These ethical concerns need to be addressed comprehensively to enhance the acceptability of co-bots in information workspaces in Kenya. While it is relatively easy to address basic concerns like the need for facilitative skills,

Table 1 *Characteristics of co-bots* Adapted from Cohen et al. (2021)

Characteristic	*Industrial robots*	*Co-bots*
Role	Replacing a worker	Assisting a worker
Human interaction	Commands, programming, assigning locations, movements and gripping	Intelligent interaction: gesture recognition, speech recognition and anticipating operator moves
Camera and computer vision	External camera and external system when they exist	Built-in standard (as part of the co-bot), coupled with artificial intelligence
Workspace	Separate safe workspace for robots and operators usually fenced	Sharing the same workspace. No fencing is necessary
Work envelope	Essential and rigid	Not relevant; flexible and spontaneous
Handling of disruptions and obstruction	Usually needs a full set-up after disruption	Built-in standard to handle disruptions and obstruction; no need to restart
Re-programming	Rare	Frequent
Physical disruptions	Mostly hazardous; set-up required for re-initiation	Safe response to disruptions with easy re-initiation protocol
System self-awareness	Basic failure detection	Real-time monitoring of load on each axis and segment, tactile pressure and axis locations
Agility	Rapid motions	Slow motions
Payload	May be heavy	Not heavy
Acquisition cost	High	Low
Ability to work in dynamic environment, possibly with moving entities	No	Yes

policies and infrastructure, the ethical concerns will take much longer to mediate. It is therefore improbable that information workers in Kenya will unreservedly welcome co-bots into the workspaces in the near future.

References

Abok, V. A. and Kwanya, T. 2016. Maximising the potential of social media to deliver academic library services to students: A case study of The Technical University of Kenya Library. *Inkanyiso*, 8(2), pp. 147–155.

Achieng, M. E., Awino, Z. B., K'Obonyo, P. and Kitiabi, R. 2020. Advanced manufacturing technology, organizational resources and performance of large manufacturing companies in Kenya. *DBA Africa Management Review*, 10(2), pp. 1–33.

Adriaensen, A., Costantino, F., Di Gravio, G. and Patriarca, R. 2021. Teaming with industrial cobots: A socio-technical perspective on safety analysis. *Human Factors and Ergonomics in Manufacturing & Service Industries*. Available at: https://onlinelibrary.wiley.com/doi/epdf/10.1002/hfm.20939 (Accessed: 01 January 2022).

Anderson, S. L. 2008. Asimov's "three laws of robotics" and machine metaethics. *Ai & Society*, 22(4), pp. 477–493.

Anderson, M. and Anderson, S. L. 2010. Robot be good. *Scientific American*, 303(4), pp. 72–77.

Anderson, M., Anderson, S. L. and Armen, C. 2005, November. MedEthEx: Toward a medical ethics advisor. In *AAAI Fall Symposium: Caring Machines* (pp. 9–16).

Anitah, J. N., Nyamwange, S. O., Magutu, P. O., Chirchir, M. and Mose, J. M. 2019. Industry 4.0 technologies and operational performance of unilever Kenya and L'Oreal East Africa. *Noble International Journal of Business and Management Research*, 3(10), pp. 125–134.

Arunda, B. 2020. The impacts of emerging technologies in the future of law and legal practice: A case of Kenya. *Statistica*. Available at: http://journalofcmsd.net/wp-content/uploads/2021/04/The-Impacts-of-Emerging-Technologies-in-the-Future-of-Law-and-Legal-Practice.pdf (Accessed: 01 January 2022).

Asimov, I. 1941. Three laws of robotics. *Asimov, I. Runaround*. Available at: http://wilearncap.asuscomm.com:81/wikipedia_en_all_novid_2017-08/A/Zeroth_Law_of_Robotics.html (Accessed: 01 January 2022).

Asimov, I. and Robertson, R. G. 1976. *The Bicentennial Man and Other Stories* (Vol. 53). Garden City, NY: Doubleday.

Banga, K. and te Velde, D. W. 2018. Digitalisation and the future of African manufacturing. *Supporting Economic Transformation*. Available at: https://set.odi.org/wp-content/uploads/2018/03/SET_Future-of-manufacturing_Brief_Final.pdf (Accessed: 01 January 2022).

Bennett, B. and Daly, A. 2020. Recognising rights for robots: Can we? Will we? Should we? *Law, Innovation and Technology*, 12(1), pp. 60–80.

Bertolini, A. and Aiello, G. 2018. Robot companions: A legal and ethical analysis. *The Information Society*, 34(3), pp. 130–140.

Boada, J. P., Maestre, B. R. and Genís, C. T. 2021. The ethical issues of social assistive robotics: A critical literature review. *Technology in Society*, 67, pp. 101726.

Bramann, J. U. 2017. Building ICT entrepreneurship ecosystems in resource-scarce contexts: Learnings from Kenya's "Silicon Savannah". *Digital Kenya*, *227*. Available at: https://library.oapen.org/bitstream/handle/20.500.12657/28113/1001881.pdf (Accessed: 01 January 2022).

Braun, V. and Clarke, V. 2021. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, 13(2), pp. 201–216.

Calitz, A. P., Poisat, P. and Cullen, M. 2017. The future African workplace: The use of collaborative robots in manufacturing. *SA Journal of Human Resource Management*, 15(1), pp. 1–11.

Calo, R., Froomkin, A. M. and Kerr, I. Eds. 2016. *Robot Law*. Cheltenham: Edward Elgar Publishing.

Calverley, D. J. 2006. Android science and animal rights, does an analogy exist? *Connection Science*, 18(4), pp. 403–417.

Campbell, J. 2021. 7 steps to cobot deployment: Select a collaborative robot: Consider risks and if a system integrator is needed. *Control Engineering*, *68*(8), pp. 17–19.

Carpenter, J. 2013. Just doesn't look right: Exploring the impact of humanoid robot integration into explosive ordnance disposal teams. In *Handbook of Research on Technoself: Identity in a Technological Society* (pp. 609–636). IGI Global.

Castillo, J. F., Ortiz, J. H., Velásquez, M. F. D. and Saavedra, D. F. 2021. COBOTS in industry 4.0: Safe and efficient interaction. Available at: https://www.intechopen.com/chapters/77896 (Accessed: 01 January 2022).

Chigbu, B. I. and Nekhwevha, F. H. 2021. The future of work and uncertain labour alternatives as we live through the industrial age of possible singularity: Evidence from South Africa. *Technology in Society*, 67, p. 101715.

Chinyamurindi, W. T. and Mey, M. R. 2017. Technology in the African workplace to the special edition. *SA Journal of Human Resource Management*, 15(1), pp. 1–2.

Chomanski, B. 2021. If robots are people, can they be made for profit? Commercial implications of robot personhood. *AI and Ethics*, 1(2), pp. 183–193.

Clabaugh, C. and Matarić, M. 2018. Robots for the people, by the people: Personalizing human–machine interaction. *Science robotics*, 3(21), p. eaat7451.

Cohen, Y., Shoval, S. and Faccio, M. 2019. Strategic view on cobot deployment in assembly 4.0 systems. *IFAC-PapersOnLine*, 52(13), pp. 1519–1524.

Cohen, Y., Shoval, S., Faccio, M. and Minto, R. 2021. Deploying cobots in collaborative systems: Major considerations and productivity analysis. *International Journal of Production Research*, pp. 1–17.

Colgate, J. E., Edward, J., Peshkin, M. A. and Wannasuphoprasit, W. 1996. Cobots: Robots for collaboration with human operators.

Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.37.7236&rep=rep1&type=pdf (Accessed: 01 January 2022).

Cookson, C. 2015. Scientists appeal for ethical use of robots. *The Financial Times, UK*. Available at: https://www.ft.com/content/fee8bacc-9f37-11e5-8613-08e211ea5317 (Accessed: 01 January 2022).

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M. L., Koay, K. L. and Werry, I. 2005, August. What is a robot companion-friend, assistant or butler? In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1192–1197). IEEE.

Dauth, W., Findeisen, S., Südekum, J. and Woessner, N. 2017. German robots—The impact of industrial robots on workers.

de Graaf, M. M. 2016. An ethical evaluation of human–robot relationships. *International Journal of Social Robotics*, 8(4), pp. 589–598.

Demianova, Y. 2018. The impact of robotics on the labor market. *Topical Issues of Humanities, Technical and Natural Sciences*, pp. 13–15.

Dongre, A. R., Deshmukh, P. R., Kalaiselvan, G. and Upadhyaya, S. 2010. Application of qualitative methods in health research: An overview. *Online Journal of Health and Allied Sciences*, 8(4), p. 3.

Dunn, H. S. 2021. Globalisation from within: Enhancing digital productivity and technology transformation in the south. In *Re-imagining Communication in Africa and the Caribbean* (pp. 33–50). Palgrave Macmillan, Cham.

Fast-Berglund, Å., Palmkvist, F., Nyqvist, P., Ekered, S. and Åkerman, M. 2016. Evaluating cobots for final assembly. *Procedia CIRP*, 44, pp. 175–180.

Fiske, A., Henningsen, P. and Buyx, A. 2019. Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research*, 21(5), p. e13216.

Forbes, E. S., Caylor, K. K., Hirsch, M. E., Schimel, J. P., Young, T. P. and Young, H. S. 2020. Wild robots: Developing DIY technology to investigate soil carbon flux in a long-term, landscape-scale, large herbivore exclosure experiment in a central Kenya savanna. In *2020 ESA Annual Meeting (August 3–6)*. ESA.

Fridin, M. 2014. Kindergarten social assistive robot: First meeting and ethical issues. *Computers in Human Behavior*, 30, pp. 262–272.

Fryman, J. and Matthias, B. 2012, May. Safety of industrial robots: From conventional to collaborative applications. In *ROBOTIK 2012; 7th German Conference on Robotics* (pp. 1–5). VDE.

Gachari, J. M. and Korir, J. K. 2020. Effect of fiscal policy on unemployment in Kenya. *Journal of Economics and Finance (IOSR-JEF)*, 11(1), pp. 19–31.

Galin, R. and Mamchenko, M. 2020, October. Human–Robot collaboration in the society of the future: A survey on the challenges and the barriers. In

International Conference on Futuristic Trends in Networks and Computing Technologies (pp. 111–122). Springer, Singapore.

Galin, R., Meshcheryakov, R., Kamesheva, S. and Samoshina, A. 2020, May. Cobots and the benefits of their implementation in intelligent manufacturing. In *IOP Conference Series: Materials Science and Engineering* 862(3), p. 032075.

García-Esteban, J. A., Piardi, L., Leitão, P., Curto, B., and Moreno, V. 2021, May. An interaction strategy for safe human co-working with industrial collaborative robots. In *2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)* (pp. 585–590). IEEE.

Gates, B. 2007. A robot in every home. *Scientific American*, 296(1), pp. 58–65.

Giacometti, A., and Larsson, J. 2017. The rise of the machines: A study of the developing acceptance of collaborative robots. Available at: https://gupea.ub.gu.se/bitstream/2077/53579/1/gupea_2077_53579_1.pdf (Accessed: 01 January 2022).

Gichora, F.G. and Kwanya, T. 2015. The impact of Web 2.0 tools on academic libraries in Kenya. *International Journal of Library and Information Science*, 7(2), pp. 21–26.

Gisginis, A. 2021. Production line optimization featuring cobots and visual inspection system. Available at: https://www.diva-portal.org/smash/get/diva2:1574281/FULLTEXT02.pdf (Accessed: 01 January 2022).

Gjeldum, N., Aljinovic, A., Crnjac Zizic, M. and Mladineo, M. 2021. Collaborative robot task allocation on an assembly line using the decision support system. *International Journal of Computer Integrated Manufacturing*, pp. 1–17.

Gobinath, V. M. 2021. An overview of industry 4.0 technologies and benefits and challenges that incurred while adopting It. *Advances in Industrial Automation and Smart Manufacturing*, pp. 1–12.

Gunkel, D. J. 2018. *Robot rights*. MIT Press.

Gupta, A., Singh, A., Bharadwaj, D. and Mondal, A. K. 2021. Humans and robots: A mutually inclusive relationship in a contagious world. *International Journal of Automation and Computing*, pp. 1–19.

Haring, K. S., Novitzky, M. M., Robinette, P., De Visser, E. J., Wagner, A. and Williams, T. 2019, March. The dark side of human–robot interaction: Ethical considerations and community guidelines for the field of HRI. In *2019 14th ACM/IEEE International Conference on Human–Robot Interaction (HRI)* (pp. 689–690). IEEE.

Hooker, J. and Kim, T. W. 2019. Humanizing business in the age of artificial intelligence.

International Federation of Robotics – IFR. 2015. *World Robotics Executive Summary*. Frankfurt am Main: IFR. Available at: http://www.worldrobotics.org/uploads/media/Executive_Summary_WR_2015.pdf (Accessed: 01 January 2022).

Iqbal, J., Khan, Z. H. and Khalid, A. 2017. Prospects of robotics in food industry. *Food Science and Technology*, 37(*2*), pp. 159–165.

Isa, M. 2018. Should robots be granted human rights?—Artificial intelligence. *finweek*, 2018(12), pp. 22–22.

Jones, R. A. 2017. What makes a robot 'social'? *Social Studies of Science*, 47(4), pp. 556–579.

Kadir, B. A., Broberg, O., and Souza da Conceição, C. 2018. Designing human–robot collaborations in industry 4.0: Explorative case studies. In *DS 92: Proceedings of the DESIGN 2018 15th International Design Conference* (pp. 601–610).

Kahlke, R.M. 2014. Generic qualitative approaches: Pitfalls and benefits of methodological mixology. *International Journal of Qualitative Methods*, *13*(1), pp. 37–52.

Kaur, M., Kim, T. H. and Kim, W. S. 2021. New frontiers in 3d structural sensing robots. *Advanced Materials*, 33(19), p. 2002534.

Keet, C. M. 2021. *Natural language generation requirements for social robots in Sub-Saharan Africa*. Available at: http://www.meteck.org/files/ISTAfrica21NLGrobotsCRC.pdf (Accessed: 01 January 2022).

Kelley, R., Schaerer, E., Gomez, M. and Nicolescu, M. 2010. Liability in robotics: An international perspective on robots as animals. *Advanced Robotics*, 24(13), pp. 1861–1871.

Kibor, S. J. and Obwoge, M. E. 2014. Robotics innovation in Kenya as an opportunity for industrial growth: Challenges and prospects for TVET institutions and universities. *The International Journal of Science and Technoledge*, 2(6), p. 336.

Kildal, J., Tellaeche, A., Fernández, I. and Maurtua, I. 2018. Potential users' key concerns and expectations for the adoption of cobots. *Procedia CIRP*, 72, pp. 21–26.

Konijn, E. A., Smakman, M. and van den Berghe, R. 2020. Use of robots in education. *The International Encyclopedia of Media Psychology*, pp. 1–8.

Kopp, T., Baumgartner, M. and Kinkel, S. 2021. Success factors for introducing industrial human–robot interaction in practice: An empirically driven framework. *The International Journal of Advanced Manufacturing Technology*, 112(3), pp. 685–704.

Kwanya, T. 2021. Perception of robots in Kenya's infosphere: Tools or colleagues? In Ocholla, D. N., Britz, J., Mutula, S. and Evans, N. (Eds). *Information and Knowledge for Development in Africa—Cross-Cutting Issues* (pp. 47–72). Durbanville: African Online Scientific Information Systems.

Kwanya, T., Kogos, A. C., Kibe, L. W., Ogolla, E. O. and Onsare, C. 2021. Cyber-bullying research in Kenya: A meta-analysis. *Global Knowledge, Memory and Communication*. Available at: https://doi.org/10.1108/GKMC-08-2020-0124 (Accessed: 01 January 2022).

Kwanya, T., Stilwell, C. and Underwood, P. 2014. *Library 3.0: Intelligent Libraries and Apomediation*. Elsevier.

Lambrechts, W., Klaver, J. S., Koudijzer, L. and Semeijn, J. 2021. Human factors influencing the implementation of cobots in high volume distribution centres. *Logistics*, 5(2), p. 32.

Liang, T. P., Robert, L., Sarker, S., Cheung, C. M., Matt, C., Trenz, M. and Turel, O. 2021. Artificial intelligence and robots in individuals' lives: How to align technological possibilities and ethical issues. *Internet Research*.

Lin, P., Abney, K. and Bekey, G. 2011. Robot ethics: Mapping the issues for a mechanized world. *Artificial Intelligence*, 175(5–6), pp. 942–949.

Lueg, C. and Twidale, M. 2018. Designing for humans, not robots (or vulcans). *Library Trends*, 66(4), pp. 409–421.

Macharia, J. 2019. Sustainable development in Kenya. *Horizons: Journal of International Relations and Sustainable Development*, (13), pp. 172–183.

Maddahi, Y., Kalvandi, M., Langman, S., Capicotto, N. and Zareinia, K. 2021. RoboEthics in COVID-19: A case study in dentistry. *Frontiers in Robotics and AI*, *8*. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/pmc8131836/ (Accessed: 01 January 2022).

Magachi, J. K., Gichunge, E. and Senaji, T. 2017. Relationship between industrial robots on competitiveness of listed manufacturing firms in Kenya. *African Journal of Co-operative Development and Technology*, 2(1), pp. 96–105.

Malik, A. A. and Bilberg, A. 2019. Complexity-based task allocation in human–robot collaborative assembly. *Industrial Robot: The International Journal of Robotics Research and Application*, 46(4), pp. 471–480.

Malterud, K., Siersma, V. D. and Guassora, A. D. 2016. Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), pp. 1753–1760.

Margherita, E. G. and Braccini, A. M. 2021. Managing industry 4.0 automation for fair ethical business development: A single case study. *Technological Forecasting and Social Change*, 172, p. 121048.

Marvel, J. A. and Norcross, R. 2017. Implementing speed and separation monitoring in collaborative robot workcells. *Robotics and Computer-Integrated Manufacturing*, 44 (April), pp. 144–155.

Mayer, J. 2018. Robots and industrialization: What policies for inclusive growth? *Working Paper*. Available at: https://www.g24.org/wp-content/uploads/2018/08/Mayer_-_Robots_and_industrialization.pdf (Accessed: 01 January 2022).

Mazzolai, B., Kraus, T., Pirrone, N., Kooistra, L., De Simone, A., Cottin, A. and Margheri, L. 2021, September. Towards new frontiers for distributed environmental monitoring based on an ecosystem of plant seed-like soft robots. In *Proceedings of the Conference on Information Technology for Social Good* (pp. 221–224).

McQuillen, J. 2021. *Comfort, Acceptance, and Preferences: The Designing of a Human–Robot Workstation that Puts the Human First* (Doctoral dissertation, Clemson University).

Ménissier, T. 2020. A "machiavellian moment" for artificial intelligence? The Montreal declaration for the responsible development of Ai. *Raisons Politiques*, (1), pp. 67–81.

Musundi, L. C., Wahome, A. K. and Odula, B. E. 2021. Driving youth participation in agriculture: A synopsis of the influence of existing agricultural policies in selected counties in Kenya. *International Journal of Academic Research in Business and Social Sciences*, 11(9), pp. 991–1015.

Mvurya, M. 2020. The extent and use of artificial intelligence to achieve the big four agenda in Kenya. *Multidisciplinary Journal of Technical University of Mombasa*, 1(1), pp. 1–7.

Nakitare, J., Sawe, E., Nyambala, J. and Kwanya, T. 2020. The emerging roles of academic librarians in Kenya: Apomediaries or infomediaries? *Library Management* 41(6/7), pp. 339–353.

Naudé, W. 2017. Entrepreneurship, education and the fourth industrial revolution in Africa. Available at: https://www.econstor.eu/bitstream/10419/170839/1/dp10855.pdf (Accessed: 01 January 2022).

Nganga, O. B. 2020. *Exploring the Applicability and Challenges of Implementing Industry 4.0 Technologies in the Small and Medium Sized Industries in Kenya* [paper presentation]. 27th IEK International Conference. Mombasa, Kenya. Available at: https://iekenya.org/27thIEKConference/27thIEKConferencePapers/Exploring%20the%20Applicability%20and%20Challenges%20of%20implementing%20Industry%204.0.pdf (Accessed: 01 January 2022).

Nijssen, S. R., Müller, B. C., Baaren, R. B. V. and Paulus, M. 2019. Saving the robot or the human? Robots who feel deserve moral care. *Social Cognition*, 37(1), pp. 41-S2.

Njogu, L. K. 2015. The effect of value added tax on economic growth in Kenya. *International Academic Journal of Economics and Finance*, 1(5), pp. 10–30.

Nomura, T. 2017, August. Cultural differences in social acceptance of robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 534–538). IEEE.

Nyholm, S. and Smids, J. 2020. Can a robot be a good colleague? *Science and Engineering Ethics*, 26(4), pp. 2169–2188.

Ocholla, D. N. and Ocholla, L. 2020. Readiness of academic libraries in South Africa to research, teaching and learning support in the Fourth Industrial Revolution. *Library Management*, 41(6/7), pp. 355–368.

Pagallo, U. 2013. What robots want: Autonomous machines, codes and new frontiers of legal responsibility. In *Human law and computer law: Comparative perspectives* (pp. 47–65). Springer, Dordrecht.

Pagallo, U. 2018. Vital, Sophia, and Co.—The quest for the legal personhood of robots. *Information*, 9(9), p. 230.

Parviainen, J. and Coeckelbergh, M. 2020. The political choreography of the Sophia robot: Beyond robot rights and citizenship to political performances for the social robotics market. *AI & SOCIETY*, pp. 1–10.

Pauliková, A., Gyurák Babeľová, Z. and Ubárová, M. 2021. Analysis of the impact of human–cobot collaborative manufacturing implementation on the occupational health and safety and the quality requirements. *International Journal of Environmental Research and Public Health*, 18(4), p. 1927.

Pazienza, A., Macchiarulo, N., Vitulano, F., Fiorentini, A., Cammisa, M., Rigutini, L., ... and Trevisi, A. 2019. A novel integrated industrial approach with cobots in the age of industry 4.0 through conversational interaction and computer vision. In *CLiC-it*.

Peeters, A. and Haselager, P. 2021. Designing virtuous sex robots. *International Journal of Social Robotics*, 13(1), pp. 55–66.

Peshkin, M. and Colgate, J. E. 1999. Cobots. *Industrial Robot: An International Journal,* 26(5), pp. 335–341.

Pinto, A., Ramos, P. N., Carvalho, C. and Silva, C. 2021, June. Do future engineers trust cobots? In *2021 4th International Conference of the Portuguese Society for Engineering Education (CISPEE)* (pp. 1–6). IEEE.

Pransky, J. 2001. AIBO-the No. 1 selling service robot. *Industrial Robot: An International Journal*, 28(1), pp. 24–26.

Rapanyane, M. B. and Sethole, F. R. 2020. The rise of artificial intelligence and robots in the 4th Industrial Revolution: Implications for future South African job creation. *Contemporary Social Science*, 15(4), pp. 489–501.

Renzullo, D. 2019. Anthropomorphized AI as Capitalist Agents: The Price We Pay for Familiarity.

Retto, J. 2017. Sophia, First Citizen Robot of the World. Available at: https://www.researchgate.net/profile/Jesus-Retto/publication/321319964_SOPHIA_FIRST_CITIZEN_ROBOT_OF_THE_WORLD/links/5a1c8aa2a6fdcc0af3265a44/SOPHIA-FIRST-CITIZEN-ROBOT-OF-THE-WORLD.pdf (Accessed: 01 January 2022).

Riaz, M., Nawaz, I. and Sohail, M. 2020. Novel concepts of soft multi rough sets with MCGDM for the selection of humanoid robot. *Punjab University Journal of Mathematics*, 52(2), pp. 111–137.

Rossato, C., Pluchino, P., Cellini, N., Jacucci, G., Spagnolli, A. and Gamberini, L. 2021. Facing with collaborative robots: The subjective experience in senior and younger workers. *Cyberpsychology, Behavior, and Social Networking*, 24(5), pp. 349–356.

Scopelliti, M., Giuliani, M. V., D'amico, A. M. and Fornara, F. 2004. If I had a robot at home... Peoples' representation of domestic robots. In *Designing a more inclusive world* (pp. 257–266). Springer, London.

Sharkey, A. and Sharkey, N. 2012. Granny and the robots: Ethical issues in robot care for the elderly. *Ethics and Information Technology*, 14(1), pp. 27–40.

Simoes, A. C., Soares, A. L. and Barros, A. C. 2019, May. Drivers impacting cobots adoption in manufacturing context: A qualitative study. In *International Scientific-Technical Conference Manufacturing* (pp. 203–212). Springer, Cham.

Simões, A. C., Soares, A. L. and Barros, A. C. 2020. Factors influencing the intention of managers to adopt collaborative robots (cobots) in manufacturing organizations. *Journal of Engineering and Technology Management*, 57, p. 101574.

Singh, B., Sellappan, N. and Kumaradhas, P. 2013. Evolution of industrial robots and their applications. *International Journal of Emerging Technology and Advanced Engineering*, 3(5), pp. 763–768.

Sladić, S., Lesjak, R., Luttenberger, L. R. and Musa, M. Š. 2021. Trends and progress in collaborative robot applications. *Polytechnica*, 5(1), pp. 32–37.

Solaiman, S. M. 2017. Legal personality of robots, corporations, idols and chimpanzees: A quest for legitimacy. *Artificial Intelligence and Law*, 25(2), pp. 155–179.

Sony, M., Antony, J., Mc Dermott, O. and Garza-Reyes, J. A. 2021. An empirical examination of benefits, challenges, and critical success factors of industry 4.0 in manufacturing and service sector. *Technology in Society*, 67, p. 101754.

Sowa, K., Przegalinska, A. and Ciechanowski, L. 2021. Cobots in knowledge work: Human–AI collaboration in managerial professions. *Journal of Business Research*, 125, pp. 135–142.

Steer, C. 2017. Global commons, cosmic commons: Implications of military and security uses of outer space. *Georgetown Journal of International Affairs*, 18(1), pp. 9–16.

Stein, M. K. and Kaivo-Oja, J. 2020. Collaborative robots: Frontiers of current literature. *Journal of Intelligent Systems: Theory and Applications*, 3(2), pp. 13–20.

Sun, Y. 2016. *Toward Never-Ending Object Learning for Robots* (Doctoral dissertation). Washington University. https://digital.lib.washington.edu/researchworks/handle/1773/36549.

Tan, S. Y., Taeihagh, A. and Tripathi, A. 2021. Tensions and antagonistic interactions of risks and ethics of using robotics and autonomous systems in long-term care. *Technological Forecasting and Social Change*, 167, p. 120686.

Tella, A. 2020. Robots are coming to the libraries: Are librarians ready to accommodate them? *Library Hi Tech News* 37(8), pp. 13–17.

Tolksdorf, N. F., Siebert, S., Zorn, I., Horwath, I. and Rohlfing, K. J. 2021. Ethical considerations of applying robots in kindergarten settings: Towards an approach from a macroperspective. *International Journal of Social Robotics*, 13(2), pp. 129–140.

Veloso, M., Biswas, J., Coltin, B. and Rosenthal, S. 2015, June. Cobots: Robust symbiotic autonomous mobile service robots. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Vicentini, F. 2020. Terminology in safety of collaborative robotics. *Robotics and Computer-Integrated Manufacturing*, 63, p. 101921.

Vojić, S. 2020. Applications of collaborative industrial robots. *Machines Technologies Materials*, 14(3), pp. 96–99.

Wallace, J. 2021. Getting collaborative robots to work: A study of ethics emerging during the implementation of cobots. *Paladyn, Journal of Behavioral Robotics*, 12(1), pp. 299–309.

Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X. V., Makris, S. and Chryssolouris, G. 2019. Symbiotic human–robot collaborative assembly. *CIRP Annals*, 68(2), pp. 701–726.

Weller, C. 2017. Meet the first-ever robot citizen—A humanoid named Sophia that once said it would 'destroy humans'. *Business Insider*, 7–10. Available at: https://jewishcamp.org/wp-content/uploads/2018/05/Sophia-Article-Golem.pdf (Accessed: 01 January 2022).

Xia, Y. and LeTendre, G. 2021. Robots for future classrooms: A cross-cultural validation study of "negative attitudes toward robots scale" in the US context. *International Journal of Social Robotics*, 13(4), pp. 703–714.

Yilma, B. A., Naudet, Y. and Panetto, H. 2020, July. A new paradigm and metamodel for cyber-physical-social systems. In *21st IFAC World Congress, IFAC 2020*. Elsevier.

Yu, C. E. 2020. Emotional contagion in human–robot interaction. *E-review of Tourism Research*, 17(5).

Yulianto, B. 2019. Philosophy of information technology: Sex robot and its ethical issues. In *Rapid Automation: Concepts, Methodologies, Tools, and Applications* (pp. 1458–1467). IGI Global.

Zhu, Q., Wei, P., Shi, Y. and Du, J. 2020, November. Cognitive benefits of human–robot collaboration in complex industrial operations: A virtual reality experiment. In *Construction Research Congress 2020: Infrastructure Systems and Sustainability* (pp. 129–138). Reston, VA: American Society of Civil Engineers.

Artificial Intelligence in Africa: Emerging Challenges

Abejide Ade-Ibijola and Chinedu Okonkwo

Introduction

The adoption and use of Artificial Intelligence (AI) are rapidly increasing (Borenstein and Howard 2021) around the world. In Africa, AI creation and implementation are transforming our lives and cultures in a variety of ways including economically, socially, and politically (Luan et al. 2020; Roos 2018). These developments are always difficult to comprehend and predict, and they are only becoming more so as the COVID-19 pandemic continues (Borenstein and Howard 2021). According to a research report, increasing the adoption of AI in Africa requires the development of vibrant ecosystems based on five stakeholders who form the foundation of building AI success including policymakers, universities, large companies, start-ups, and multi-stakeholder partnerships (Schoeman et al. 2021).

A. Ade-Ibijola (✉) · C. Okonkwo
Research Group on Data, Artificial Intelligence, and Innovations for Digital Transformation, Johannesburg Business School, University of Johannesburg, Johannesburg, South Africa
e-mail: abejide@jbs.ac.za

C. Okonkwo
e-mail: chineduo@uj.ac.za

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_5

AI refers to a set of technologies that allow machines to function intelligently and mimic human sensing, comprehension, and action (Li et al. 2018; Hamet and Tremblay 2017; Adamopoulou and Moussiades 2020). The technology will have a significant impact in almost all areas where human intelligence is involved. It can be used by businesses and institutions to personalise activities, optimise operations, promote innovations, and empower and supplement staff (Schoeman et al. 2021). Essentially, AI changes the way people function, allowing for more efficient resource allocation, which leads to increased productivity and allows better government service delivery to the public (Borenstein and Howard 2021; Roos 2018). Specifically, the application of AI in education, health care, agriculture, commerce, and governance is showing a significant impact on their various activities.

AI is a rapidly growing technical field that has the potential to change every aspect of human social interactions (Pedro et al. 2019). AI-powered services pervade many aspects of human life across the globe; however, adoption rates vary between developed and developing nations (Francesc et al. 2019). It is evident that AI is penetrating the African social system through various operations (Schoeman et al. 2021; Marino Garcia and Kelly 2020). Chatbots in Kenya, for example, now provide healthcare services to people without visiting doctors, and a data-driven platform called Zenvus in Nigeria delivers insights to farmers (Francesc et al. 2019). Finance and other sectors are not excluded; AI-powered technologies are changing the ways of operations in these sectors. In *South Africa, Mama Money* and *Mukuru* enable easy and quick transfer of money across different countries in Africa. Kudi is a Nigerian AI-powered Chatbot that aims to provide financial services to underprivileged people (Kudi 2018). Other online shopping facilities are: TakeAlot (South Africa) and Konga (Nigeria).

Because of the widespread adoption of mobile technologies in Africa, there is an optimism that AI technologies will be the next wave of technologies to receive wide acceptance. However, with the exception of a few nations (such as South Africa, Nigeria, Ethiopia, Kenya, and Ghana), widespread adoption of AI applications in Africa is not yet a reality (Gadzala 2018). The crucial factors needed for technology adoption are sadly lacking across most of Africa, and many countries in Africa are still lacking the necessary infrastructure, governance, data ecosystem, STEM education, and other factors necessary for AI.

The purpose of this paper is to explore the challenges facing a widespread adoption of AI across African regions. Therefore, the research question is, "what are the challenges facing the design, development, deployment and application of AI in Africa?". To achieve this aim, a desk research to review literature (including reports) on AI in Africa was conducted. AI technologies are increasingly and significantly currently being developed, deployed, and applied more widely in the developed countries of the world than in Africa. This is as a result of contextual factors that can be conceptualised as barriers that require attention so as to improve AI capacities. This chapter highlights these barriers/challenges and makes recommendations that can improve the development, deployment, and usage of AI technologies in Africa.

This chapter starts with an overview of the meaning and nature of AI followed by the examination of the emerging challenges to AI in Africa. Section "Emerging Challenges to AI Adoption in Africa" discusses the implications of the challenges, while section "Addressing AI Adoption Challenges" suggests possible ways to address these issues. Section "Discussion" makes some practical suggestions and recommendations for diverse stakeholders, and finally, section "Conclusions" concludes the chapter.

Overview of Artificial Intelligence

Artificial Intelligence (AI) is a branch of science that studies and develops intelligent machines (Luxton 2016). This field of science was established in the 1950s, and by that time, AI was described as a new science which would methodically examine the phenomenon of "intelligence". Computer simulations of intelligent processes were to be used to achieve this goal (Brey and Søraker 2009). The authors further explained that in AI, intelligence is viewed as a broad mental skill that incorporates a variety of more specialised abilities, including reasoning, planning, problem-solving, concepts understanding, use of language, and learning. To test computer intelligence, in 1950, Turing and Haugeland (1950) introduced the *Turing Test* which is still invoked though with some criticism (Hayes and Ford 1995).

AI refers to technology that can perform tasks that require a certain level of intelligence—that is, a machine or tool that has been trained to perform work like that of a human. AI is categorised into two basic types including weak and strong AI or narrow and general AI

(Chang 2020). Weak AI portrays AI as a problem-solving tool, whereas strong AI entails the creation of a "real" mind (Flowers 2019). AI portfolio involves Natural Language Processing (NLP), Robotics, Machine and Deep Learning, Cognitive Computing, and Reinforcement Learning (Chang 2020). The idea of AI as a science that examines the phenomena of intelligence has been somewhat overtaken in recent years by a perception of AI as an engineering discipline in which scholars focus on producing usable programmes and tools that operate in areas that ordinarily need intelligence (Brey and Søraker 2009). AI technologies are already in use all around us, in nearly every aspect of life. It can be used in a variety of corporate roles to help employees at work by reducing their workload and in areas such as commerce, education, agriculture, and finance. In addition, Artificial Intelligence is being used in transportation, automobiles, manufacturing, and weather forecasting.

As a technology, AI is still in its infancy. Today's AI systems have only a rudimentary understanding of human expression, tone, emotion, and the intricacies of human interaction (Smith and Shum 2018). Before computers can truly act like humans, there is still a long way to go. While AI research has a long history of enthusiasm followed by extended disappointment, we are currently in the midst of an extraordinary period of technological innovation across multiple industries, which is fuelling the rise of AI. AI will have a big influence by enabling quicker and deeper advancement in practically every sector where (human) intellect plays a role. It may be used by businesses or organisations to engage customers, revolutionise product creation, optimise operations, and empower staff (Akerkar 2019). But, more crucially, AI can assist society in overcoming some of its most overwhelming challenges (Nishant et al. 2020).

AI solutions are being successfully deployed at scale in some African countries and especially in Kenya, Nigeria, Ghana, Ethiopia, and South Africa (Gadzala 2018). Most solutions currently target the financial services, agriculture, and healthcare sectors (Okonkwo and Ade-Ibijola 2021). South Africa leads the continent in AI adoption with a robust ecosystem that includes numerous technology hubs, research groups, and various. Many companies in South Africa are either integrating AI solutions into their existing operations or developing new solutions using AI (Ferrein and Meyer 2012; Gwagwa et al. 2020).

Emerging Challenges to AI Adoption in Africa

AI is a game-changing innovation with the potential to improve all sectors of the African social system. However, the adoption and use of AI applications in African society raise some issues including skills acquisition, ethics, programming, data integration, user attitude, government policy, and insufficient infrastructure and network connectivity. These challenges emerged from the review of AI-related literature and are presented and discussed in the subsequent sections.

Skills Acquisition

This entails learning the theoretical and practical skills required for the development, implementation, and use of AI applications. Technological advancement necessitates the acquisition of technical skills. Business and IT leaders in Africa agree that to achieve the adoption and use of AI technologies, the stakeholders' knowledge base must be changed or improved (Mzmkandaba 2019). One of the key essential competences in the era of the Fourth Industrial Revolution (4IR) is programming skills. Furthermore, Bianco (2021) reported that one of the major barriers to the adoption of this modern technology—AI—is skill. Any project requires the right expertise to succeed, and AI is no exception. Of all, AI skills are more difficult skills to master, and there is undoubtedly a demand–supply imbalance in the market.

AI, as a new and growing innovation, will improve existing jobs while also creating new ones, necessitating the acquisition of new skills. According to a Gartner research circle survey, 56% of participants believe that learning new skills would be essential to execute both existing and newly developed jobs. (Bianco 2021). In the development and implementation of an AI system, there is a need to incorporate expert knowledge (Abu-Alsaad 2019). Despite the fact that IT professionals (software developers and engineers) design and develop AI applications, they are not the primary users of AI. In developing markets such as the African market, a scarcity of AI-ready workers is a major issue (Ajadi 2020).

The development of AI involves machine learning and NLP processes which consist of complex algorithms; thus, programming skills are needed. The issue is "How can AI be programmed to perform accurate operations?". For example, in an interaction with a Chatbot, a user's questions can come in various forms demanding the same answer. For

example, "What is the time" and "Could you check the time". The Chatbot system may correctly answer the first question but incorrectly answer the second. According to Grosz (Grosz 2018), computational linguistics and NLP systems also raise some of the most serious potential issues, such as dialogue system failure, the impact of social Chatbots on how people communicate with one another, and system performance issues. As a result, ICT efficacy and programming competences are among the required skills for effective adoption and use of AI applications (Komarova et al. 2019).

Uncertainty

Although the adoption and use of AI applications are becoming common, it is still difficult for some business leaders to quantify the benefits associated with the technology. There are some well-known benefits of AI, such as instant response, time saving, medical advances, and revenue generation; other benefits like automation of process, enhanced learning, and customer experience are still hard for users in Africa to comprehend (Mzmkandaba 2019, Okonkwo and Ade-Ibijola 2021). Because these technologies mimic human intelligence, that is doing human jobs in a different manner, the question is therefore: Is the adoption of AI applications beneficial or not? While some business leaders and stakeholders believe and trust in this technology, others are afraid that implementation of AI may disrupt their traditional ways of working (Coetzee 2018; Smith and Neupane 2018). As a developing continent, Africa, the knowledge of AI is still at an early stage and the population is still not certain of the advantages. The fear of the unknown poses a great challenge to the adoption and use of AI in Africa.

Lack of Structured Data Ecosystem

AI initiatives rely on the quality and quantity of data contents to provide accurate information or responses to users in each situation. In many cases, an AI will fail if the data that is used to train the AI system does not reflect the demographic variables in the targeted population. A Chatbot system, for example, requires comprehensive information about its operations to provide correct responses to users; if the information requested by the user is not in the data bank, the system will fail. Data shortages in Africa are well known in the context of development, where

high-quality data are essential indicators of growth in relation to the Sustainable Development Goals (SDGs) and a key input for the development of modern technologies. The UN Economic Commission for Africa (UNECA) stated that African data ecosystems are at "nascent stages of the African data revolution and the private sector is increasingly becoming a critical and dynamic player within African data ecosystems" (UNECA 2016). Machine learning methods are only as good as the data they are given. AI algorithms include prejudices found in data or even in the individual who created the process, spreading social disparities. This is especially important in Africa, where users are more likely to import machine learning algorithms built and trained abroad using data that may not recognise or be biased against substantial parts of the African population (Kathryn Hume 2017). To enable researchers, developers, and users to adopt AI solutions, a deeper, larger, and more accessible pool of data is needed. In developing markets, particularly in unstable or conflict-affected areas, high-quality data is not always available or accessible (Ajadi 2020).

Lack of Relevant Government Policies

As AI-powered technologies are beginning to sweep over business, governance, and educational activities, there is a need for a policy on AI implementation strategies in African countries as seen in the developed countries such as Australia, China, France, and the United States (Pedro et al. 2019). Although some African countries, such as Mauritius, Egypt, Zambia, Tunisia, and Botswana, have recognised the potential of AI to boost GDP and have developed national AI strategies, and South Africa, Nigeria, and Kenya have passed data protection laws, all are still in their infancy (Pedro et al. 2019; Effoduh 2020). The African Union (AU) proposed the Promulgation of AI laws and regulations, called structured regulation of AI to manage the benefits of the technology for Africans (Effoduh 2020). Most of the African population is a late majority and laggard adopter of innovation; they take a "wait-and-see" approach to technology adoption (Okonkwo et al. 2020, 2021). AU needs to speed up to establish a well-structured adoption and implementation of AI-powered technology to boost its adoption among the African population. Overall, there is a general lack of relevant policies that can prioritise the design and implementation of AI as well as address the potential impacts on society.

Ethics

In terms of technology development, ethics refers to a set of principles based on public acceptance, religious beliefs, and cultural norms on the best behaviour that can be observed and followed during the development and deployment of innovative and emerging technologies. In Africa, ethics form the basis of human activities which can promote African cultures and help to build confidence in the development and applications of technologies in Africa (Dugbazah et al. 2021).

While AI has enormous potential, it also poses major difficulties for businesses and governments, notably in terms of ethics. The moral, economic, and social repercussions of the Second and Third Industrial Revolutions are still being debated in many African countries (Oosthuizen 2020). AI has already been implicated in several examples of ethical issues. Studies have revealed some major areas of AI possible implications on the African social world including accountability, data bias, transparency, and socio-economic risks (Ruane et al. 2019). AI technologies are systems that mimic human intelligence. AI undermines established moral and legal paradigms that place human agency solely in the hands of humans (Tegmark 2018). Using biased data, AI has been noted to create socio-economic inequality (Larsson et al. 2019). In addition, the design of AI systems involves some complex algorithms which in turn compromise trust and transparency. Data is used to train these algorithms. It has been claimed that there is a data scarcity in Africa (Microsoft 2018), and that the majority of acquired data does not correctly reflect the African experience, implying that many algorithms may not be appropriately adapted to the features of local populations (Mahomed 2018). To provide an acceptable basis for AI adoption in Africa, stakeholders must have open discussions on the ethical implications of AI and take necessary steps.

User Attitudes

Another challenge facing the adoption of AI systems in Africa is the users' attitude. An adopter's attitude towards the adoption and rejection of an invention can be favourable, negative, and apprehensive (Okonkwo et al. 2019). According to (Wang et al. 2008), attitudes are a primary predictive factor impacting the adoption of a new product; hence, a better knowledge of attitudes in a well-defined manner is required. Africans are very sceptical in adopting and using new technology due to culture and social

influences. In a study on the adoption of AI in higher education, Chatterjee and Bhattacharjee (2020) revealed that individuals' behavioural intentions to use AI in higher education are influenced by their attitudes. Likewise, another research on the adoption of software engineering products proved that user attitude influences the adoption of software tools (Okonkwo et al. 2019). This leads to the conclusion that higher education authorities would find it useful and beneficial to mould stakeholders' attitudes to shape their intentions and behaviour (Chatterjee and Bhattacharjee 2020). As a result, if students have negative perceptions of Chatbot technology applications in education, they will be hesitant to adopt and use the technology. Positive perception of an innovation accelerates adoption.

Insufficient Infrastructure and Network Connectivity

Inadequate infrastructure and a dearth of network affordability are some of the major hurdles of AI adoption in Africa. The growth of infrastructural development as well as mobile technology network connectivity in Africa is slow (Marino Garcia and Kelly 2015). A good percentage of Africa's population are unconnected and not having access to the internet. Adoption of AI requires adequate availability of wireless network connectivity. In addition, African countries have the world's most expensive broadband. The Alliance for Affordable Internet (A4AI) reported that African countries inhabited nine of the ten least affordable spots in terms of internet access, with expenditures ranging from 12 to 44% of GDP (Marino Garcia and Kelly 2020).

Addressing AI Adoption Challenges

Adoption and use of any innovation including AI require the necessary competence. First and foremost, AI skill is much more difficult to perfect, and there is a greater demand for AI expertise. It is important to create a conducive environment in business, health, education, and public ecosystems and encourage the employees to have interests on the use of AI to perform their operations. This will make them learn the skills. There is a need to improve the educational curriculum to integrate the teaching of AI skills from the secondary level. Improving people's learning abilities in mathematics and computer programming will help them acquire the necessary AI skills. Because the use of AI systems is permeating all aspects

of life, the knowledge of AI is for all. Adding introductory programming and computer basics subjects to all fields of study may also aid in the acquisition of AI skills.

All stakeholders (citizens, policymakers, and technical experts) should be involved early in the architectural design process so that societal expectations, fears and concerns are taken into account and no patchwork is necessary later as an afterthought. This will improve users' knowledge of the system, which will reduce uncertainty or fear of the unknown. A responsible data management framework that takes into account data diversity is recommended for AI developers. The framework will make it easier to collect comprehensive data for system development and will improve the accuracy of AI system operations.

Ethical challenges are big concerns regarding adoption of AI systems in Africa. Many initiatives, including educational institutions, government agencies, non-governmental groups, and industry, have attempted to address the ethical and legal challenges that have arisen because of AI technology, but the impact of these efforts is still insignificant in Africa (Borenstein and Howard 2021). The selection, design, deployment, and usage of AI technology have ethical implications. We therefore agree with the recommendations made by Borenstein and Howard, (2021) that it is critical for developers to recognise that the technology they are creating is entwined with ethical dimensions, and that they have a critical role and obligation to engage with ethical considerations as developers. Making developers aware of their professional responsibilities and moral implications while developing will aid in reducing ethical issues. Also, government and other professional bodies should strengthen their ethics policies to guide the development process of AI.

Creating a conducive environment encourages innovative minds. African nations are still struggling on infrastructural development. Most government policies concentrate on urban development while rural areas suffer. Government should ensure that they extend their developments to rural locations, bringing stable and adequate network coverage across all areas. For example, most of the farming industries that need AI to boost their operations and increase productivity are in rural areas. In addition, the African Union and each respective African nation should establish AI strategies and policies that will be a foundation for the development and implementation of AI technologies.

Finally, Africa has developing economies that require technological advancements to accelerate growth. As a result, it is critical that the continent encourages the use of AI in a variety of ways to perform various tasks. Establishing various agencies that can assist in dealing with citizens' doubts and fears should be encouraged as part of an awareness campaign. These agencies should improve Africa's appreciation and adoption of AI and establish efficient ways of addressing societal impacts. Technology start-ups or tech hubs should be established to train the next generation of AI experts, and local technological innovations should be supported by the government and private sector.

Discussion

The adoption of AI technologies in Africa is facing some challenges including lack of technical skills, uncertainty, lack of structured data, lack of government policies, ethics, and user attitude. AI has the potential to improve productivity at a firm while also keeping competitive and gaining a deeper knowledge of their consumers. However, there are not enough individuals with the necessary expertise and abilities to operate these applications. A strong skill set is required to create an AI technology with good content, system, and service qualities, and adequate knowledge of how to use the product is required to use the technology effectively. These skills extend beyond basic technological knowledge and may address other difficulties such as a lack of adequate managerial knowledge or even development of business ideas. In terms of maintenance, AI tools require routine maintenance and upgrade, and this calls for expert knowledge of programming. In other words, a company that wants to employ the services of AI technologies needs an in-house engineer or trusted vendor for maintenance and services. As a result, business leaders, other stakeholders, and members of the public who may be engaged in the implementation and usage of any sort of AI must acquire the core skills required to adapt, learn, govern, and utilise the technology in their respective fields of work. This implies that all the stakeholders including the government should develop a means to incorporate the required skills into the people starting from the young age of primary education to professional level. Lack of technical skills hinders the adoption and use of technologies (Mtega et al. 2012).

The application of AI technologies raises some ethical concerns, particularly in the areas of role, privacy, transparency, trust, personality, and

culture. There are numerous ethical questions surrounding the adoption of AI: Are AIs job killers? How is data stored and utilised? Is AI smarter than humans? and so forth. These questions instil fear and scepticism in the adoption of AI technologies. For example, in the healthcare system, AI is being used to collect various forms of personal information from patients. Gathering user data generates several privacy concerns; some may have legal backing and are protected by data protection laws (Ruane et al. 2019). When incorporating AI into any aspect of life, it is critical to consider user privacy. As an intelligence machine that is not human, it is best to inform users of its AI status so that they can make informed decisions about how to interact with the system. The human–machine relationship differs from the human-to-human relationship. Knowing the nature of an AI system will increase the user's understanding and trust. To accommodate different users, user groups, interests, characteristics, and context should be considered during the design and development of AI tools. Culture plays an important role in ethics. It is concerned with the social behaviour of people in a specific area. African nations are multilingual and have diverse cultural backgrounds, which can have an impact on the adoption of technological innovation. Culture, according to research conducted in Nigeria and South Africa, is a determining factor in technological innovation (Bankole et al. 2011; Lekhanya 2013).

Individual beliefs and feelings should be considered when designing AI systems. The user's perceptions of a new technology influence its adoption. A positive perception encourages users to adopt and use AI systems, whereas a negative perception leads to rejection of the innovation. If users do not have sufficient information about an innovation, they will be sceptical of its usefulness, leading to uncertainty—fear of the unknown. Because most Africans are laggards and late majority adopters, they are uncertain of new technological innovations, and they employ a wait-and-see approach to adopting new innovations. The greater the level of uncertainty, the lower the adoption of AI in Africa. Developers and other stakeholders should always strive to make AI systems user-friendly to reduce the uncertainty associated with new technology. Governments should step up infrastructural development and expansion of network connectivity especially to rural locations in Africa. Furthermore, African governments and stakeholders, including the African mobile ecosystem, must establish well-structured guidelines and policies to ensure proper AI systems development, implementation, and adoption in Africa.

AI technologies are rapidly evolving and being used for a wide range of purposes, including education, business, politics, and social activities. In Africa, there is still a need for sustainable development, and the use of AI technologies may help to achieve this. Using AI technologies as supporting tools will be impossible if they are not properly adopted, and for AI to gain widespread adoption, these challenges must be addressed.

Conclusions

This chapter has highlighted some of the emerging challenges facing the effective design, development, adoption, and use of AI technologies in Africa. AI technologies have enormous potential to aid the growth of African economies and human flourishing. AI technology has the potential to alter business operations and performance, enhance productivity, and improve health care, education, and transportation in Africa. Without addressing these challenges or barriers, Africa will continue to lag behind countries in the global world. Stakeholders in Africa, especially policymakers, need to establish robust governance structures and infrastructures to improve not only the design and development of AI but the adoption and use of AI technologies. In addition to policymakers, industry stakeholders also have a duty to focus on capacity development to ensure that adequate AI skills are acquired for the responsible use of AI to address African-specific needs and problems. It is therefore hoped that a proper understanding and appreciation of these challenges and recommendations will be had to ensure accelerated design, development, and adoption of AI, which will have a big influence on Africa's economy.

References

Abu-Alsaad, H.A., 2019. Agent applications in e-learning systems and current development and challenges of adaptive e-learning systems, in: 2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), IEEE. pp. 1–6.

Adamopoulou, E., and Moussiades, L., 2020. An overview of chatbot technology, in: IFIP International Conference on Artificial Intelligence Applications and Innovations, Springer. pp. 373–383.

Ajadi, S., 2020. Can AI help tackle the most pressing challenges in developing countries? GSM Association. https://www.gsma.com/mobilefordevelopment/region/africa/can-ai-help-tackle-the-most-pressing-challenges-in-developing-countries/ [Accessed 8 July 2021].

Akerkar, R., 2019. Employing AI in business, in: Artificial intelligence for business. Springer, Cham, pp. 63–74.

Bankole, F.O., Bankole, O.O., and Brown, I., 2011. Mobile banking adoption in Nigeria. *The Electronic Journal of Information Systems in Developing Countries* 47, 1–23.

Bianco, M., 2021. Overcoming the social barriers of AI adoption.

Borenstein, J., and Howard, A., 2021. Emerging challenges in AI and the need for AI ethics education. *AI and Ethics* 1, 61–65.

Brey, P., and Søraker, J.H., 2009. Philosophy of computing and information technology, in: Philosophy of technology and engineering sciences. Elsevier, pp. 1341–1407.

Chang, A.C., 2020. Intelligence-based medicine: Artificial intelligence and human cognition in clinical medicine and healthcare. Academic Press.

Chatterjee, S., and Bhattacharjee, K.K., 2020. Adoption of artificial intelligence in higher education: A quantitative analysis using structural equation modelling. *Education and Information Technologies* 25, 3443–3463.

Coetzee, J., 2018. Strategic implications of Fintech on South African retail banks. *South African Journal of Economic and Management Sciences* 21(1), 1–11.

Dugbazah J., Glover B., Mbuli B., and Kungade C., 2021. Ethical considerations on emerging technologies in Africa: A consultative roundtable report. African Union Development Agency. Available at: https://www.nepad.org/blog/ethical-considerations-emerging-technologies-africa-consultative-roundtable-report [Accessed 8 October 2021].

Effoduh, J.O., 2020. 7 ways that African states are legitimizing artificial intelligence. Openair.Africa. https://openair.africa/7waysthatafricanstatesarelegitimizingartificialintelligence/ [Accessed 6 July 2021].

Ferrein, A., and Meyer, T., 2012. A Brief Overview of Artificial Intelligence in South Africa. *AI Magazine* 33(1), 99–103.

Flowers, J.C., 2019. Strong and weak AI: Deweyan considerations, in: AAAI spring symposium: Towards conscious AI systems.

Francesc, P., Miguel, S., Axel, R., and Paula, V., 2019. Artificial intelligence in education: Challenges and opportunities for sustainable development—UNESCO Biblioteca Digital. Technical Report. UNESCO Working Papers on Education Policy. UNESCO. https://unesdoc.unesco.org. UNESCO.

Gadzala, A., 2018. Coming to life: Artificial intelligence in Africa. Atlantic Council, November 14.

Grosz, B.J., 2018. Smart enough to talk with us? Foundations and challenges for dialogue capable AI systems. *Computational Linguistics* 44, 1–15.

Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., and De Beer, J., 2020. Artificial Intelligence (AI) deployments in Africa: Benefits, challenges and policy dimensions. *The African Journal of Information and Communication* 26, 1–28.

Hamet, P., and Tremblay, J., 2017. Artificial intelligence in medicine. *Metabolism* 69, S36–S40.

Hayes, P., and Ford, K., 1995. Turing test considered harmful, in: IJCAI (1), Citeseer. pp. 972–977.

Kathryn Hume, 2017. Artificial Intelligence is the future—But it's not immune to human bias, Maclean's, December 27.

Komarova, N., Zamkovoi, A., Novikov, S., 2019. The fourth industrial revolution and staff development strategy in manufacturing. *Russian Engineering Research* 39, 330–333.

Kudi, 2018. A chatbot for financial services. Kudi. https://kudi.com/faq [Accessed 5 July 2021].

Larsson, S., et al. (2019). Sustainable AI: An inventory of the state of knowledge of ethical, social, and legal challenges related to artificial intelligence. Lund University Online Publications. Available at: https://www.semanticscholar.org/paper/Sustainable-AI%3A-An-inventory-of-the-state-of-of-and-Larsson-Anneroth/087f015b87428301a61eaba3ba92918b3fd56284

Lekhanya, L.M., 2013. Cultural influence on the diffusion and adoption of social media technologies by entrepreneurs in rural South Africa. *The International Business & Economics Research Journal*.

Li, L., Lin, Y.L., Zheng, N.N., Wang, F.Y., Liu, Y., Cao, D., Wang, K., and Huang, W.L., 2018. Artificial intelligence test: A case study of intelligent vehicles. *Artificial Intelligence Review* 50, 441–465.

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S.J., Ogata, H., Baltes, J., Guerra, R., Li, P., and Tsai, C.C., 2020. Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology* 11.

Luxton, D.D., 2016. An introduction to artificial intelligence in behavioural and mental health care, in: Artificial intelligence in behavioural and mental health care. Elsevier, pp. 1–26.

Mahomed, S. (2018). Healthcare, artificial intelligence and the Fourth Industrial Revolution: Ethical, social, and legal considerations. *South African Journal of Bioethics and Law* 11(2), 93 [online]. https://doi.org/10.7196/sajbl.2018.v11i2.00664.

Marino Garcia, J., and Kelly, T., 2015. The economics and policy implications of infrastructure sharing and mutualisation in Africa.

Marino Garcia, J., and Kelly, T., 2020. Artificial intelligence for Africa: An opportunity for growth, development, and democratisation. Access Partnership. https://www.up.ac.za/media/shared/7/ZP_Files/ai-for-africa.zp165664.pdf.

Microsoft, (2018). Artificial intelligence for Africa: An opportunity for growth, development, and democratisation. The University of Pretoria [online]. Available at: https://www.up.ac.za/media/shared/7/ZP_Files/ai-forafrica.zp165664.pdf.

Mtega, W.P., Bernard, R., Msungu, A.C., and Sanare, R., 2012. Using mobile phones for teaching and learning purposes in higher learning institutions: The case of Sokoine University of agriculture in Tanzania.

Mzmkandaba, S., 2019. Gartner lifts the lid on barriers to AI adoption. iTweb. https://www.itweb.co.za/content/mQwkoM6KJZbq3r9A. [Accessed 6 July 2021].

Nishant, R., Kennedy, M., and Corbett, J., 2020. Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management* 53, 102104.

Okonkwo, C.W., and Ade-Ibijola, A., 2021. Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 100033.

Okonkwo, C., Huisman, M., and Taylor, E., 2020. A framework for adoption and diffusion of mobile applications in Africa. *Advances in Science, Technology and Engineering Systems Journal* 5, 1577–1592.

Okonkwo, C.W., Huisman, M., and Taylor, E., 2019. The adoption of m-commerce applications: Rural dwellers perspectives, in: 12th, IADIS, international conference. Information systems.

Okonkwo, C.W., Huisman, M., and Taylor, E., 2021. Factors that influence Africa's refusal and discontinuation of mobile applications use. *African Journal of Science, Technology, Innovation and Development*, 1–10.

Oosthuizen, M., 2020. Africa's 4th industrial revolution—Endless opportunities. Institute for Security Studies [online]. Retrieved, 18.

Pedro, F., Subosa, M., Rivas, A., and Valverde, P., 2019. Artificial intelligence in education: challenges and opportunities for sustainable development.

Roos, S., 2018. Chatbots in education: A passing trend or a valuable pedagogical tool?

Ruane, E., Birhane, A., and Ventresque, A., 2019. Conversational AI: Social and ethical considerations., in: AICS, pp. 104–115.

Schoeman, W., Moore, R., Seedat, Y., and Chen, J.Y.J., 2021. Artificial intelligence: Is South Africa ready?

Smith, B. and Shum, H., 2018. The future computed. Microsoft.

Smith, M. L., and Neupane, S. (2018). Artificial intelligence and human development: Toward a research agenda. White Paper. International Development Research Centre (IDRC). https://idl-bnc-idrc.dspacedirect.org/handle/10625/56949.

Tegmark, M. (2018). Benefits & risks of artificial intelligence. Future of Life Institute [online]. Available at: https://futureoflife.org/background/benefitsrisks-of-artificial-intelligence/?cn-reloaded=1 [Accessed 21 May 2021].

Turing, A.M., and Haugeland, J., 1950. Computing machinery and intelligence. MIT Press, Cambridge, MA.

UNECA, 2016. The Africa data revolution report. 2016. https://www.uneca.org/sites/default/files/uploaded-documents/ACS/africa-data-revolution-report-2016.pdf [Accessed 26 May 2021].

Wang, G., Dou, W., and Zhou, N., 2008. Consumption attitudes and adoption of new consumer products: A contingency approach. *European Journal of Marketing*.

The Use of Gendered Chatbots in Nigeria: Critical Perspectives

Favour Borokini, Kutoma Wakunuma, and Simisola Akintoye

Introduction

Chatbots are "artificial intelligence (AI) software that can simulate a conversation (or a chat) with a user in natural language through messaging applications, websites, mobile apps or through the telephone".[1] In Africa, there has been a significant increase in the development and use of chatbots and their adoption is becoming widespread amongst data and service providers who use their capabilities to meet customer demands and needs amid an expanding customer base. The situation reflects how more businesses are turning towards AI to power their

[1] https://www.expert.ai/blog/chatbot/.

F. Borokini (✉)
University of Nottingham, Nottingham, UK
e-mail: favour.borokini@nottingham.ac.uk

K. Wakunuma · S. Akintoye
De Montfort University, Leicester, UK
e-mail: kutoma@dmu.ac.uk

S. Akintoye
e-mail: simi.akintoye@dmu.ac.uk

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_6

products and services in Africa, a growth not unrelated to the explosion of data and data collection and processing capabilities on the continent.

The use of chatbots is not limited to specific institutions as their use has been noted in banking and other financial institutions as well as in insurance, transportation and health care. However, although chatbots have garnered glowing reviews from various angles for the benefits they bring towards improving productivity and profit-making, the frequent anthropomorphisation of them as female, merits close scrutiny because of the impact on perceptions of women, as well as existing socio-cultural expectations, stereotypes and demands regarding how women are expected to act in society. In a UNESCO report titled "I'd blush if I Could"(UNESCO, 2019), UNESCO detailed the potential negative impacts of chatbots or voice-based conversational agents on societal perceptions of gender. According to the report, the proliferation of female-gendered conversational agents was primarily driven by customer preference and a non-critical examination of the product development decisions by product teams, which could entrench and perpetuate biases about women today.

Based on these reflections, this chapter will look at the deployment and integration of gendered chatbots in Nigerian institutions and the potential impact of this deployment on Nigerian women. Drawing from statistics about the industries and positions typically occupied and dominated by women, as well as the presence of women in the finance and technology development space, our contribution will evaluate the origins of the preference for these bots and proffer recommendations on ways to curb the negative effects of their deployment.

As chatbots are increasingly being used in the financial sector this chapter will focus on the use of chatbots in Nigerian commercial banks followed by a review of their use in other sectors. In 2018, Nigeria's United Bank for Africa (UBA) launched the very first chatbot by a commercial bank in the country (Eleanya, 2018). This move was precipitated by the growing need to improve financial inclusion and improve customer experience by simplifying financial transactions (Nelson, 2019).

Since then, chatbots have been used to provide a variety of services by banks in Nigeria, ranging from airtime top-up, account enquiry and customer complaints. In late 2021, almost 50% of the 22 commercial banks listed on the Central Bank of Nigeria's website (CBN Nigeria, 2021) had deployed chatbots in some shape or form. Although the rationale justifying their deployment by banks does not say anything about the

gender selection process, it is clear that the logic behind their deployment has largely revolved around the need to promote customer satisfaction, convenience and safety (Adesanya, 2020; Zenith Bank, n.d.) on the part of the banks, in line with global trends.

The following section will cover the methodological approach used to collect the data for this study.

Methodology

Using secondary data collection methods, this study analysed chatbots deployed in commercial banks and other institutions in Nigeria. This is based on information available on websites such as those of the Central Bank of Nigeria (CBN), and national newspapers. This research is therefore based on the findings of the analysis conducted.

This study focuses primarily on the 22 commercial banks on the CBN website. The Supervisory Framework of the Central Bank of Nigeria is structured into two departments: "Banking Supervision" and "Other Financial Institutions".[2] Commercial banks, alternatively described as deposit money banks, as well as discount houses, are under the jurisdiction of the former, while financial institutions such as Microfinance Banks, Bureaux-de-Change, Development Finance Institutions, Primary Mortgage Institutions and Finance Companies, which are described as "other", are the purview of the latter.

Information regarding the presence and availability as well as the features and descriptions of chatbots on the websites of the 22 commercial banks was collected. For the purposes of the research, chatbots available on the websites of the banks, in addition to those deployed on social media platforms such as Facebook, WhatsApp, Instagram and Telegram were considered. It was possible to access chatbots hosted on the latter three through links provided on the social media of the banks, and through phone numbers registered to the chatbots.

The availability or existence of the bots was predicated on live links on the websites of each of the banks, social media posts alluding to them, and news reports. In cases where there were social media posts and news reports alluding to their existence, but no live links were found, for the purposes of this study, such banks were described as not having chatbots.

[2] https://www.cbn.gov.ng/Supervision/framework2.asp.

The analysis was based on three identifiable chatbot cues: name, avatar and a last category, tagged "other descriptor" for cues that did not fit the first two categories, to categorise and classify their gender presentation into three categories: male/masculine, female/feminine and gender-neutral. For the purpose of this study, gender-neutral means that "something is not associated with either women or men"[3] as the European Institute for Gender Equality opined. In this study, the term is used to refer to features that could be borne by both men and women as well as those which are neither stereotypically male nor female.

In addition to searches conducted on the websites of these banks, the research also sourced information about the chatbots from blog posts, news reports and journal publications, leveraging the information available on these platforms.

Findings—The Use of Chatbots in Nigerian Banks

The research found that 10 of Nigeria's 22 commercial banks have either currently or in the past integrated chatbots into their product and service delivery. Based on their gender presentation, these chatbots are categorised based on either of these three features (Table 1).

Of these ten, based on the name given to the bot, the assigned gender based on pronouns or descriptions on the banks' social media or official website and the graphic depiction of the avatars, 7 chatbots are gendered female. These are bots belonging to Zenith Bank, United Bank for Africa, Sterling Bank, First City Monument Bank, Fidelity Bank, Ecobank and Access Bank. On the other hand, Heritage Bank's chatbot, which was integrated into its Octopus app was not gendered in any way and merely appeared as a feature of the mobile app while Keystone Bank's Oxygen was presented with a robot arm or a stylised "O2". Leo, United Bank for Africa's chatbot, was the only male-gendered chatbot on the list.

The breakdown of the chatbots in Table 2, which shows the names, assigned gender and avatar of the bots in the 10 banks, reveals that the majority of them are female-gendered and designed with a preference for female characteristics revealed in stereotypically female names or avatars or both, except for three banks. These banks are Keystone Bank, which has a robot as its avatar and a neutral name, Heritage Bank, which has a

[3] https://eige.europa.eu/thesaurus/terms/1190.

Table 1 Table illustrating the three cues used in determining the gender of chatbots

S/N	*Chatbot Feature*	*Explanation*
1	Names	The names given to these chatbots are often either stereotypically masculine or feminine, with outliers having gender-neutral names. Through these names, it is often possible to determine if the bots are designed to project a male or female gender(ed appearance). Where the bots are not assigned extrinsically male or female names, they are either assigned neutral names, i.e. names which may be borne by either male or female persons, or acronyms or the name of an animal species
2	Avatars	In addition to names, bots may also be classified as male or female based on the avatars which represent their identity. Avatars may refer to graphic depictions of these bots, showing physical features such as hairstyle, or clothing typically associated with a certain gender. For instance, UBA's Leo, the only male-gendered chatbot on the list from a total of 10, has an avatar often dressed in a T-shirt and jeans with a low cut. Leo's avatar is also portrayed with a sharp, chiselled jaw. Some other portrayals have Leo wearing a suit and dressed as a Yoruba man wearing the traditional Yoruba *fila* Ambiguity in gender presentation, where the avatars are gendered neither male nor female, is achieved due to an absence of he/she pronouns, due to the avatar being an animal specie or being an object such as a robot
3	Other descriptor	Asides from names, certain descriptors such as pronouns, adjectives or certain patterns of behaviour often attributed to or expected of a particular gender may also be used to identify what gender the bot has been assigned

neutral name for its bot and no avatar and United Bank for Africa, whose chatbot has a male-gendered avatar and name.

Analysis of the Table 2 reveals the following about the chatbot names:

a. Only one of the names, Leo, is a stereotypical male name. It accounts for 10% of the chatbot names.
b. Only 30% of the chatbot names are stereotypically female names. These are Kiki, Ivy and Tamada (a portmanteau of two feminine names), Tamara (a Hebrew word for date) and Ada (an Igbo name given to the first female child of a family).

Table 2 List of Commercial Banks in Nigeria using chatbots as at December 2021

S/N	*Bank Name*	*Chatbot/Virtual Assistant Available*	*Name of Chatbot/Virtual Assistant Name*	*Assigned gender of Chatbot/Virtual Assistant*	*Avatar Image*
1	Zenith Bank	Yes	Zenith Intelligent Virtual Assistant (ZIVA)	Female-gendered	
2	United Bank for Africa (UBA)	Yes	Leo	Male-gendered	
3	Sterling Bank	Yes	Kiki	Female-gendered	

S/N	*Bank Name*	*Chatbot/Virtual Assistant Available*	*Name of Chatbot/Virtual Assistant Name*	*Assigned gender of Chatbot/Virtual Assistant*	*Avatar Image*
4	Stanbic IBTC	Yes	Sami	Female-gendered	
5	Keystone Bank	No (discontinued)	Oxygen	Gender Neutral	
6	Heritage Banking Company	No (discontinued)	Nil	Gender Neutral	Nil
7	First City Monument Bank (FCMB)	Yes	Temi	Female-gendered	

(continued)

Table 2 (continued)

S/N	*Bank Name*	*Chatbot/Virtual Assistant Available*	*Name of Chatbot/Virtual Assistant Name*	*Assigned gender of Chatbot/Virtual Assistant*	*Avatar Image*
8	Fidelity Bank	Yes	Ivy	Female-gendered	
9	Ecobank	Yes	Rafiki	Gender Neutral	
10	Access Bank	Yes	Tamada - a portmanteau of Ada and Tamara	Female-gendered	

c. 60% of the chatbot names are gender-neutral. These names are Sami (which could be short for Samantha, a female name or a Samuel, a common male name), Temi (a Yoruba word meaning "my own" which precedes many gender-neutral Yoruba names such as Temiloluwa and Temitope), ZIVA (an acronym for Zenith Intelligent Virtual Assistant) and Rafiki (a Swahili word for friend). Heritage Bank's chatbot had no name and was simply a feature provided on their Octobus bank app and Oxygen, Keystone Bank's chatbot is a colourless, odourless gas.

For the avatars:

a. 10% of the avatars had a male avatar who sported a low cut and wore a T-shirt and jeans, both commonly worn by Nigerian men.
b. 20% of the avatars presented as gender-neutral. These were Oxygen, which was often depicted with a stylised "O2" or a robot or robot arm and Heritage Bank's chatbot which had no avatar.
c. 70% of the avatars are presented wearing common Nigerian female makeup, hairstyles and clothes. These are ZIVA, Kiki, Ivy, Sami, Temi, Rafiki and Tamada.

Other descriptors

Other descriptors used to verify the assigned gender granted to the chatbots where gender-neutral names but feminine avatars were present include pronouns such as "she".

It is clear that based on the names, it is difficult to ascertain the assigned gender presentation of some of the chatbots such as Sami and ZIVA. However, from the gender presentation of their avatars and the pronouns used in their description, it is clear that they are unambiguously gendered females. From our analysis, therefore, 70% of the identified chatbots are gendered female.

Discussion on Gendered Chatbots

The criticism and support of the female-gendered chatbots and other conversational agents have been the subject of various conversations in AI ethics. Support for chatbots has been predicated on their efficiency with

the conversational agents being forecast to save businesses as much as $7.3 billion by 2023 (Vogeler, 2019). The female-gendering of chatbots has also been pursued to promote the wider acceptance of bots. According to Borau et al. (2021), in specific domains such as health care, and with self-driving cars, where consumers are loath to trust recommendations and replies proffered by chatbots and other AI systems, the female-gendering of AI systems plays on human perceptions and stereotypes of warmth and friendliness associated with women. This is interesting to note, particularly as the world's first chatbot was named Eliza.[4,5]

However, there are important considerations stemming from the development, use and impact of female-gendered chatbots. The attribution of the female gender to chatbots appears to be a predominant one as research analysing the design of 1,375 chatbots showed (Feine et al., 2019). The trend was especially visible in such sectors as customer service, sales and brand representation which are usually customer-facing roles. Women operating in these fields, or who work in fields requiring them to operate in the public eye, often face intense scrutiny while contending with stereotypes about their personality and appearance. For instance, when asked to describe two female presenters, one of whom spoke in a more stereotypically female voice than the other, respondents classified the former as less intelligent and trustworthy but more empathetic and warm than the latter (Voelker, 1994). Long-held social stereotypes and assumptions regarding women such as these may inform the creation and build of female-gendered conversational agents, which in turn may inform and influence social norms regarding women's capacity and nature. The descriptions used by banks and blog writers to describe their roles and personalities are also telling of certain societal biases about the roles persons in the service industry, where women are well represented (PWC, 2020), are expected to play. In addition, they are revealing of societal stereotypes about the expected behaviour of women which is assumed to be solicitous, polite, always available and good communicators, hence their customer-facing roles. For instance, in a bid to encourage its use, Sterling Bank's Kiki is said to be one to never "air"—a slang that means

[4] https://medium.com/the-official-integrate-ai-blog/so-should-chatbots-be-female-fb3f2854ad37.

[5] http://psych.fullerton.edu/mbirnbaum/psych101/Eliza.htm.

to ignore—messages[6] (Sterling Bank Plc, 2020). This suggests an existing belief that female-gendered chatbots are just as good as—or better than women—at communicating, a field considered well-suited to women. This belief underpins hiring and training practices which have seen women relegated and restricted to soft and feminine roles and their exclusion from those considered hard and serious such as programming or engineering, considered better-suited to men. In essence, such a gendered categorisation of women essentialises women's position as good communicators, or even more concerning, sexualises women in so far as female chatbots may be used to attract more customers or grow an institution's profit through their soft, feminine voices and looks. Thus ignoring serious political questions of AI for social good, ethically and responsibly and as a technology that can play a role in empowering women.

A 2021 report from the International Finance Corporation (IFC)[7] (International Finance Corporation, 2021) which assessed the workplace gender equality policies of the 30 most capitalised companies listed on the Nigerian exchange revealed that although Nigerian women made up one-third of the workforce, this number lagged behind the global average by 5%. In addition, according to the report, women's representation at the highest leadership levels ranged from 20% to 27%, in line with global trends which range from 17% to 25%. At the managerial level, in the Central Bank of Nigeria and amongst the five biggest Nigerian banks, namely: First Bank, United Bank for Africa, Guaranty Trust Bank, Access Bank, and Zenith Bank, which are collectively known as FUGAZ, 80% of the appointed Executive Directors are male (Ushedo, 2021). However, while it is difficult to establish the precise number of women in the banking sector due to the absence of gender-disaggregated data collected by the Nigerian Bureau of Statistics,[8] Nigerian women are well represented in the service industry, which the financial sector is a part of, although they are, in contrast, underrepresented in management positions in the same sector. On the other hand, there is also a corresponding dearth in the number of women working in ICT. According to UNESCO's report, which details the general preference

[6] https://www.facebook.com/search/posts?q=sterling%20bank%20kiki&_rdc=1&_rdr.

[7] https://www.ifc.org/wps/wcm/connect/8911c51f-1b19-4395-8de3-bb142d589118/Gender+Equality+in+Nigeria%E2%80%99s+Private+Sector+Report.pdf?MOD=AJPERES&CVID=nM7Dn7a.

[8] https://nigerianstat.gov.ng/elibrary/read/1241018.

for female-gendered chatbots and other conversational agents, "a related or concurrent explanation for the predominance of female voice assistants may lie in the fact that they are designed by workforces that are overwhelmingly male" (UNESCO, 2019). Another report from the IFC report[9] also revealed that just 18% of the total developer population in Nigeria were women. The homogenous male nature of technology talent and the overrepresentation of women in the service industry, when considered with the prevalence of social stereotypes about the role of women, therefore reveal a direct relationship with the ubiquity of female-gendered chatbots in financial institutions in Nigeria. The reinforcement of gender difference in technology is further seen through Green's (2002) work when she states that because more women work in net-based professional settings, men are creating niche areas of internet-focused activity such as security. Green (2002, p. 188) indicates that such areas are "the *'masculinised' technocultural domain: which command high fees*". As a result, women are more likely to work in less high-profile areas. This kind of occupational segregation can be observed in work carried out by (Kotamraju, 2003) in the US, where she discovered that even though web design employs both men and women, they work in different areas, noting the division between graphic designers who did design and layouts and programmers who develop software. The graphic designers who were mostly women were lowly paid while the software development team was on higher rates. This was despite both groups being considered as web designers and dependent on each other to achieve shared company goals. Hafkin (2006) confirms this finding when she reveals that there are far fewer women who are systems analysts and programmers and even fewer women working as software and hardware engineers.

While the link between the performance of chatbots and corresponding assigned genders remains unclear, there is some evidence that developers often imbue conversational agents with specific ideas and feelings flowing from their personal perceptions of how their creations should work. In the case of Cortana, the Microsoft UX Lead, Jonathan Foster, had this to say: "we continue to endow her with make-believe feelings, opinions, challenges, likes and dislikes, even sensitivities and hopes. Smoke and mirrors, sure, but we dig in knowing that this imaginary world is invoked by real people who want detail and specificity" (Foster, 2019). But who are the

[9] https://www.ifc.org/wps/wcm/connect/e358c23f-afe3-49c5-a509-034257688580/e-Conomy-Africa-2020.pdf?MOD=AJPERES&CVID=nmuGYF2.

real people who want detail and specificity? More likely than not, it is the creators and developers as Foster (2019) pointed out. Therefore as the UNESCO (2019) report indicates, designers and developers are overwhelmingly male who embed their own values and sense of what a female chatbot should sound or look like. As such, there is a serious argument to be made for a more serious examination of AI beyond what may be perceived as its neutrality by developers and designers by looking more closely at its negative impact on women.

The gendering of chatbots can have the unintended impact of entrenching and reflecting unfounded biases about the capabilities and abilities of both genders. Such a situation could occur when male-gendered bots handle requests more efficiently than female-gendered ones. This has been noted from the perception of the United Bank of Africa's *Leo* which is Nigeria's only male-gendered banking chatbot which has been severally tagged the smartest banking chatbot in Nigeria (Moses-Ashike, 2021; Nweze, 2021). This could also be the case where male-gendered chatbots are deployed for use in typically male-dominated settings. Complications posed by the former can be likened to and result in what feminist researchers have tagged the glass elevator, a term coined to describe "the advantages that men receive in the so-called women's professions (nursing, teaching, librarianship, and social work), including the assumption that they are better suited than women for leadership positions" (Williams, 2013). The implications of this rather niche variety of technology-facilitated gender-based injustice against women could mean female representation in digital spaces is worth considering.

Similarly, the deployment of male-gendered bots to perform roles typically carried out by men also embeds conventions about the types of persons most suited to perform these roles. This is the case with Translators without Borders' *Shehu*, a chatbot designed to facilitate understanding and answer questions about COVID-19 in North-Eastern Nigerian. Shehu is a multilingual bot that speaks English, Hausa and Kanuri, commonly-spoken languages in the region (TWB Communications, 2021b). According to the Translators without Borders website, the word "Shehu" "is an official title for a scholar, and refers to someone learned and knowledgeable". However, while by definition, Shehu appears to be gender-neutral and is described using the gender-neutral pronoun "it" (TWB Communications, 2021a), it is both a title and a name typically borne by male scholars and male children. Furthermore, Shehu is also visually depicted as male. *Shehu* is depicted as wearing

a traditional cap, typically worn by men as as part of native attire. The combination of these features is arguably sufficient to determine that in spite of the gender-neutral definition of the name and the use of gender-neutral pronouns, Shehu is in fact designed to be male. It projects the appearance of a strong, intelligent and authoritative male figure which aligns with existing socio-cultural perception of Islamic scholars who are male and rarely female.

Gendered chatbots in communities in Africa is also worth considering as their use could introduce and, in some cases, further complicate gender relations in a way and manner akin to the impact of colonialism on indigenous women's rights which were eroded in colonial and post-colonial societies. Women's rights scholars have noted that in an attempt to undermine existing social hierarchies and structures, the activities of women's groups were suppressed, resulting in the removal of institutions such as women's support groups and chieftaincy titles (Alapo, 2014). It has also been noted that even Western education, typically appreciated for its role in the emancipation of women from traditional oppression, did not always have this result, as colonial education emphasised the preparation and training women for domestic roles rather than leadership within society (Okome, 2002). Already, several studies have highlighted the many ways technology today is a tool of neocolonialism reminiscent of colonial extractivist activities (Iyer, 2018). AI-powered chatbots could therefore introduce and impose new forms of gendered expectations upon women. Many chatbots today already are marketed on the premise of an ever-ready, ever-available, polite assistant. For instance, First City Monument Bank's Temi is given the following description: "Hi! I'm Temi, your personal person. I'll always have time for you any time of the day. Ready to discuss your plans be it health, travel or even future goals. The good news is, I get things done and I'll never reply to you with a 'k'".[10]

Customer expectations regarding chatbots also highlight the need for a chatbot that is more responsive and always present, an expectation said to be defeated by having to deal with customer service representatives on their websites (Abdulquadri et al., 2021). Researchers have also noted the use of online platforms to spread novel and existing, i.e. merely technologically-assisted variants of technology-facilitated violence as the

[10] https://twitter.com/myfcmb/status/1070234891353231360.

divide between online and offline spaces decreases (Henry and Powell, 2015).

Recommendations

At the moment, there is a huge increase in the use of AI-enabled chatbots; it is crucial to question the need for chatbots at the moment in a country where the unemployment rate has risen to 33.30% in the fourth quarter of 2020 from 27.10% in the second quarter of 2020. Exemplified by the Silicon Valley-esque "move fast and break things" approach to innovation, which has defined the uncritical development and deployment of AI systems and technology generally, today, techno chauvinism has been implicated severally in the decision of Nigerian banks and other institutions to integrate chatbots into service delivery as opposed to simply hiring more staff. In her book, Broussard (2018) defines techno chauvinism to mean the belief that technology is always the solution. In Nigeria, customer services at financial services are notoriously poor and often the subject of numerous social media complaints (Benson, 2018). This dismal state of affairs may be explained by hiring practices in the industry. According to the National Bureau of Statistics,[11] approximately 42% of bank employees were contract staffers in 2020, and only 95,026 persons were employed in total throughout the industry, down from 103,610 from the previous year.[12] Judging from this, perhaps a more worthwhile response to the challenges of a growing customer base might be to channel resources towards the employment and training of more staff rather than the costly pursuit and development of chatbots to improve service delivery.

Corporations that have deployed gendered chatbots and other conversational agents have often justified this move citing the need to innovate in line with customer preferences and expectations for (personality traits such as humaneness and warmth typically associated with) female bots (Guo et al., 2020). A suitable response to this may be derived from similar arguments in Science and Technology Studies (STS) research where rich criticism of racialisation and whiteness in Artificial Intelligence development as well as the demeaning depictions of people of colour in

[11] https://nigerianstat.gov.ng/elibrary/read/1241018.

[12] https://nigerianstat.gov.ng/download/1081.

science fiction exists (Sparrow, 2020). There, some Science and Technology Studies scholars have suggested that one way to address such troubling depictions of people of colour is by deracialising the depiction of robots such that future robots or their avatars are designed to have blue or green skin. These suggestions have been countered by research showing that the presence of racial and ethnic minorities in media is important for representation in some cases. In their commentary, Cave and Dihal (2021) mention that racialised bots might help build trust and increase interactions, particularly for and between marginalised groups. Applying this to gendered chatbots, it might be an argument worth considering that gendered chatbots are useful for representation and in imbuing a sense of care in the attitudes of engineers and developers towards the final product. However, there are serious concerns worth noting about the underlying profit-making motivations of these corporations who are compelled to make a profit by centring customer needs and playing to societal stereotypes and expectations at the expense of the dehumanisation of women. Moving forward, product development teams need to reflect on the potential cultural harm gendered bots may pose societally. It is crucial for product teams to adequately consider if bots must be gendered during the development and iteration process. To achieve this, there is the need to develop guidelines for gender-equal design of chatbots that will help engineers in the diminishing possible gender stereotypes that could become embedded in the process.

Another option worth considering, as argued by Cave and Dihal (2021), is the use of AI to subvert stereotypes. Subverting stereotypes could be achieved by creating female bots or avatars for positions and roles in which women and even men are traditionally underrepresented. A great example of a male bot performing extremely well in a female-dominated field is United Bank for Africa's Leo. It is possible to change or subvert the narrative about female capacity by depicting female-gendered bots breaking stereotypes regarding their capabilities without further entrenching gender biases.

For women to be active participants in AI as well as to have a sense of the social and economic potential of AI, much lies in the inclusion of their experiences and needs in technology policies (UNICTTF, 2002; Jorge, 2006) "gender-specific projects and programs, regulations that facilitate affordable access to women and the poor, establishment of universal access programs targeting women, licensing regimes that favour companies with gender-equality policies, and programs that consider women's

needs and realities" (Jorge, (2006, p. 74). These are regarded as crucial considerations in technology policies in order for women to be a part of the technology development process. The problem though, as Jorge concedes, is that although policies may mention gender-equality concerns, in most cases, they are not followed through at the regulation and implementation stage and thus remain merely as desirable add-ons.

On the part of the government, there is the need to develop gender-inclusive policies which prioritise inclusiveness in product and service delivery. This would reflect inclusiveness and respect for female employees and customers on the part of banks. Banking institutions need to reaffirm their commitment to guidelines such as the Nigerian Sustainable Banking Principles 2012 (CBN Nigeria, 2012). This can be achieved through the inclusion of women in various departments of these institutions. Such inclusion will require a political change in attitude simply because it calls for a reflection of and consequent change in power relations in order for women's needs, aspirations and interests to be realised. For this to happen, women's situation needs to be understood as a linkage between women's human poverty, globalisation as well as gender inequality. Chatbots, therefore, need to be designed, developed, implemented and used in a gender-sensitive way that sees women as equal and not categorised as less than or other to men.

Lastly, there is the need for critical discussions on issues around gendered chatbots and societal perceptions. This requires ongoing interaction between researchers, practitioners, developers and users to address pertinent questions such as promoting diversity amongst chatbot developers, identifying gender bias in chatbot development, and avoiding "female-by-default" chatbot designs and ethical considerations of organisations.

Conclusion

Resolving technology-facilitated violence perpetuated and engendered through the use of gendered chatbots requires a combination of efforts from various stakeholders, including the government and private industry, requiring a multi-pronged approach in the form of awareness campaigns, training, research and policy development, and paying close attention to the implications of their development.

Furthermore, solutions cannot be implemented without the ignition of conversations regarding technology and Artificial Intelligence ethics in the

technology space. To change the narrative, a greater discussion is about how the dearth of critical Nigerian technology studies and research about societal biases and stereotypes and the huge gap in the number of women working in this space have occasioned the misuse of technology. These conversations will facilitate discussions and create environments where product and engineering teams can work with researchers and individuals in the third sector to develop more humane, gender-responsive and respectful technology.

References

Abdulquadri, A. et al. (2021) Digital Transformation in Financial Services Provision: A Nigerian Perspective to the Adoption of Chatbot. *Journal of Enterprising Communities: People and Places in the Global Economy*, 15(2), pp. 258–281.

Adesanya, A. (2020) *Sterling Bank Takes Banking Services to WhatsApp | Business Post Nigeria*. Available from: https://businesspost.ng/banking/sterling-bank-takes-banking-services-to-whatsapp/ [Accessed 24 February 2022].

Alapo, R. (2014) Gender and Ethnicity in Nigeria: Post-Colonial Societal Constructs on Culture and Class.

Benson, E.A. (2018) *Industries with the Worst Customer Service in Nigeria*. Available from: https://nairametrics.com/2018/10/05/poor-customer-service-in-nigerian-companies/ [Accessed 22 February 2022].

Borau, S. et al. (2021) The Most Human Bot: Female Gendering Increases Humanness Perceptions of Bots and Acceptance of AI. *Psychology & Marketing*, 38(7), pp. 1052–1068.

Broussard, M. (2018) *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press.

Cave, S. and Dihal, K. (2021) Race and AI: The Diversity Dilemma. *Philosophy & Technology*, 34(4), pp. 1775–1779.

CBN Nigeria (2021) *Central Bank of Nigeria | List of Commercial Banks*. Available from: https://www.cbn.gov.ng/supervision/inst-dm.asp [Accessed 22 February 2022].

CBN Nigeria (2012) *Nigerian Sustainable Banking Principles Guidance Note*. Lagos, Nigeria.

Eleanya, F. (2018) *From LEO to Ada: How Nigerian Banks Get Creative with Chatbots* [Online]. Businessday NG. Available from: https://businessday.ng/technology/article/leo-ada-nigerian-banks-get-creative-chatbots/ [Accessed 24 February 2022].

Feine, J. et al. (2019) Gender Bias in Chatbot Design. In: *International Workshop on Chatbot Research and Design*. Springer, pp. 79–93.

Foster, J. (2019) *What Did We Get Ourselves Into?* [Online]. Microsoft Design. Available from: https://medium.com/microsoft-design/what-did-we-get-ourselves-into-36ddae39e69b [Accessed 25 February 2022].

Green, L. (2002) *Communication, Technology and Society*. London: SAGE Publications.

Guo, Y. et al. (2020) "She Is Not Just a Computer": Gender Role of AI Chatbots in Debt Collection. *ICIS 2020 Proceedings* [Online]. Available from: https://aisel.aisnet.org/icis2020/hci_artintel/hci_artintel/20.

Hafkin, N.J. (2006) Women, Gender and ICT Statistics and Indicators. *Cinderella or Cyberella*, pp. 49–69.

Henry, N. and Powell, A. (2015) Embodied Harms: Gender, Shame, and Technology-Facilitated Sexual Violence. *Violence Against Women* [Online]. Available from: https://doi.org/10.1177/1077801215576581 [Accessed 24 February 2022].

Iyer, V.G. (2018) Strategic Environmental Assessment (SEA) Process for Green Materials and Environmental Engineering Systems towards Sustainable Development-Business Excellence Achievements. In: *2018 4th International Conference on Green Materials and Environmental Engineering (GMEE 2018)*. DEStech Transactions on Environment Energy and Earth Sciences. Adv Sci & Ind Res Ctr.

Jorge, N.S. (2006) *Engendering ICT Policy And Regulation: Prioritising Universal Access For Women's Empowerment*. Kumarian Press, Bloomfield, CT, USA.

Kotamraju, N.P. (2003) Art Versus Code: The Gendered Evolution of Web Design Skills. *Society Online: The Internet in Context*, pp. 189–200.

Moses-Ashike, H. (2021) *UBA's Artificial Intelligence Chatbot Hits 3m USER Mark* [Online]. Businessday NG. Available from: https://businessday.ng/news/article/ubas-artificial-intelligence-chatbot-hits-3m-user-mark/ [Accessed 24 February 2022].

Nelson, G.S. (2019) Bias in Artificial Intelligence. *North Carolina Medical Journal*, 80(4), pp. 220–222.

Nweze, C. (2021) Expanding Banking Services with Intelligence Chatbot LEO [Online]. *The Nation Newspaper*. Available from : https://thenationonlineng.net/expanding-banking-services-with-intelligence-chatbot-leo/ [Accessed 24 February 2022].

Okome, M.O. (2002) Domestic, Regional, and International Protection of Nigerian Women Against Discrimination: Constraints and possibilities. *African Studies Quarterly*, 6(3), pp. 33–63.

PWC (2020) Impact of Women on Nigeria's Economy, p. 35.

Sparrow, R. (2020) Robotics Has a Race Problem. *Science, Technology, & Human Values*, 45(3), pp. 538–560.

TWB Communications (2021a) *TWB Chatbots* [Online]. Translators Without Borders. Available from: https://translatorswithoutborders.org/chatbots/ [Accessed 24 February 2022].

TWB Communications (2021b) *TWB Develops Chatbot to Improve COVID-19 Communication in Nigeria* [Online]. Translators without Borders. Available from: https://translatorswithoutborders.org/chatbot-release-northeast-nigeria/ [Accessed 24 February 2022].

UNESCO (2019) *I'd Blush If I Could: Closing Gender Divides In Digital Skills Through Education*. Available from: https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1 [Accessed 24 February 2022].

UNICTTF (2002) Information and Communication Technologies and Their Impact on and Use as an Instrument for the Advancement and Empowerment of Women.

Ushedo, K. (2021) *Should CBN Enforce Gender Parity on the Board of Nigerian Banks?—Nairametrics*. Available from: https://nairametrics.com/2021/07/21/should-cbn-enforce-gender-parity-on-the-board-of-nigerian-banks/ [Accessed 24 February 2022].

Voelker, D.H. (1994) *The Effects of Image Size and Voice Volume on the Evaluation of Represented Faces*. PhD Thesis. Stanford University.

Vogeler, W. (2019) *Why Are so Many Company Chatbots Female?* [Online]. FindLaw. Available from : https://www.findlaw.com/legalblogs/in-house/why-are-so-many-company-chatbots-female/ [Accessed 24 February 2022].

Williams, C.L. (2013) The Glass Escalator, Revisited: Gender Inequality in Neoliberal Times, SWS Feminist Lecturer. *Gender & Society* [Online]. Available from: https://doi.org/10.1177/0891243213490232 [Accessed 24 February 2022].

Zenith Bank. *Zenith Bank Launches Intelligent Chatbot, ZiVA* [Online]. Zenith Bank Plc. Available from: https://www.zenithbank.com/media/news/zenith-bank-launches-intelligent-chatbot-ziva/ [Accessed 24 February 2022].

AI Policy as a Response to AI Ethics? Addressing Ethical Issues in the Development of AI Policies in North Africa

Bernd Carsten Stahl, Tonii Leach, Oluyinka Oyeniji, and George Ogoh

Introduction

The concept of artificial intelligence (AI), despite its 70-year history (McCarthy et al. 2006), has become highly prominent since the mid-2010s. It is commonly accepted that this is the result of the availability of large datasets and increasing computing power which allowed established and novel approaches to machine learning, such as deep learning and artificial neural networks to demonstrate their capabilities (Hall and Pesenti

B. C. Stahl (✉) · T. Leach · O. Oyeniji · G. Ogoh
Centre for Computing and Social Responsibility, De Montfort University, Leicester, UK
e-mail: bernd.stahl@dmu.ac.uk

T. Leach
e-mail: antonia.leach@dmu.ac.uk

O. Oyeniji
e-mail: p2623648@my365.dmu.ac.uk

G. Ogoh
e-mail: george.ogoh@dmu.ac.uk

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_7

2017). There is broad recognition that the potential of this development is far from achieved and that AI applications are likely to radically change many processes in business and administration, but also in the personal lives of citizens and consumers.

These developments are welcomed by many, as they raise the vista of increased economic well-being but also the resolution of many challenges that can benefit from new computational approaches, for example, in diagnosis and delivery of health care, data analysis in big science or the reduction of climate-relevant emissions. At the same time, there are steadily increasing concerns about the potentially negative impacts on ethics and human rights that these technologies may have.

Policymakers and decision-makers in industry, civil society and elsewhere are thus faced with the challenge of harnessing the benefits of AI while managing the ethical and other risks these technologies may pose. This is exacerbated by the apparent perception of competition between countries, all of which want to be the first to benefit from AI and want to strengthen their AI industry which is seen as a source of economic but also political competitiveness.

This situation which is characterised by a strategic dominance of the big economic and technical blocs that hold much of the knowledge, technology and data to make AI succeed (USA, China, EU) raises the question of how countries outside of these power blocs can approach AI. One crucial question that forms part of this broader issue is whether and how current AI policies can address the ethical and human rights issues linked to AI.

This chapter contributes to knowledge by offering an analysis of AI policies in North Africa and offers the first analysis of the way in which ethical issues have been considered in these policies. This analysis highlights strengths and limitations in existing policies. These insights are relevant to policymakers who aim to develop AI strategies, in Africa and elsewhere.

The chapter starts with a brief overview of the AI ethics discourse, suggesting that it would benefit from a systems perspective. It then provides an overview of current AI policies. This is followed by a discussion of AI strategies in North Africa. The discussion section explores how ethical concerns are covered in the North African AI strategies and which gaps and opportunities in terms of coverage of ethical issues arise from the current state of these policies.

Ethics of AI

The chapter rests on the assumption that policy can be used to address ethical concerns. This implies a broad understanding of ethics as anything having to do with the distinctions between right and wrong, good and bad, appropriate and inappropriate. Using such a broad starting point, it is probably uncontentious to say that at least some policy initiatives are driven by ethical concerns. The question of this chapter is whether in the field of AI such concerns can and should motivate or drive policy agendas.

An Overview of AI Ethics Issues

Ethics as a discipline of philosophy has produced numerous theoretical positions using different aspects to distinguish right from wrong. These include duty-based (deontological) positions which use the intention of an agent to provide an ethical evaluation of actions (Kant 1797, 1788). Further well-established ones look at the consequences of actions for this evaluation (Bentham 1789; Mill 1861) or the character of the agent (Aristotle 2007). There are furthermore numerous more recent theoretical positions specifically aimed at computing and information technologies (Bynum 2001; Floridi 1999; Introna 2005), including recent attempts to formulate ethical positions specifically for AI (AI HLEG 2019; Borenstein et al. 2021; Dignum 2019). The dominant conceptual approach to the ethics of AI, however, is not to use philosophical theories but to refer to mid-level principles and guidelines (Jobin et al. 2019).

The aspect of AI ethics that may motivate policy developments, however, is usually not the theoretical reflection and evaluation of ethical concerns, but the concerns themselves. We use the term 'ethical issue' to denote phenomena that are perceived as ethically problematic. These ethical issues are subject to empirical investigation and often to intervention. There are numerous ethical issues of AI that are prominently discussed, and we now highlight some of the most significant ones.

A discussion of the ethics of AI needs to be informed by the capabilities of AI which are the reason for the rapidly increasing use of AI technologies. This, in turn, requires a definition of the term AI. While AI has been a part of computer science since the 1950s (McCarthy et al. 2006), It has only been during the last 10 years that some of the approaches that constitute AI have made significant progress in solving relevant problems (Boden 2018; Stone et al. 2016). This refers in particular to machine learning that has benefited from the availability of computing power, large data sets and improved algorithms (Hall and Pesenti 2017; UKRI 2021).

While the field of AI is broader than machine learning (Elsevier 2018), it has been the success of machine learning that has led to high expectations for AI.

The benefits of machine learning arise from its ability to analyse large datasets and identify patterns in data. The consequences and benefits of the capability depend on the specific application. Very generally, one can say that AI can improve the use of large data sets and thereby help address a broad range of questions. This can translate into increased economic efficiency, but also better understanding of environmental challenges (AI Council 2021) or ways of promoting the UN's Sustainable Development Goals (AI HLEG 2019) and be used for 'good' (Berendt 2019), i.e. for morally desirable purposes.

The flip side of these benefits of AI, however, is similarly visible and important. Current machine learning techniques require access to large data sets in order to extract information and construct models. This raises concerns about privacy and data protection (Buttarelli 2018; EDPS 2020). Ongoing data use can constitute problematic surveillance, in particular where novel biometric data such as emotional data are used (Dignum 2019). AI systems that are used for decision support can perpetuate and even exacerbate existing biases (Access Now 2018) which can lead to unfair discrimination (Latonero 2018) on the basis of sensitive characteristics such as gender, age or race. With the growing importance of AI systems, their reliability and security (Brundage et al. 2018) become pressing questions, including on the national level (Babuta et al. 2020).

The growing use of AI across all areas of society, private organisations and public administration raises concerns about ethical consequences of their large-scale deployment. AI has already reshaped large parts of the economic system and contributed to new ways of wealth creation. While the potential of AI to create wealth is undisputed, the justice of the subsequent distribution of this wealth is a major cause for concern (Zuboff 2019). At present, the main beneficiaries are big organisations that hold data and computing resources, so that AI perpetuates and exacerbates inequality on a local and global scale (European Parliament 2020). Furthermore, AI is likely to have consequences for employment (Kaplan and Haenlein 2019; Rai et al. 2019), even though the full impacts of AI on employment remain contested (Willcocks 2020).

Economic consequences of AI spill over into the political field where there are worries that AI technologies can damage democratic processes, for example, by disrupting elections or contributing to political decisions in opaque ways (Yeung 2018). The concentration of wealth in the economic field can support a similar concentration in political power

which has led to calls for political and legal interventions (Coeckelbergh 2020; Nemitz 2018). In addition, AI can lead to changes in ethically relevant areas. It has the potential to change the nature of warfare by introducing autonomous weapons (Guterres 2020; Richards et al. 2020). The growing use of AI requires significant energy which is an important part of AI's environmental impact (Nishant et al. 2020; UNESCO 2020).

A further set of ethical concerns about AI is how it affects human freedom and possible actions. Autonomous decision-making is one of the capabilities of AI, but it is highly contested whether and to what degree this should be encouraged or permitted. There are strong calls for ensuring human control of AI (Council of Europe 2019). Such calls do little to change the fact that technology, increasingly including AI, structures human spaces for action. What we perceive to be possible or not is affected by the technical capabilities that surround us.

A final set of ethical concerns that figure prominently in the AI ethics debate has to do with the potential of truly autonomous machines. This is sometimes referred to as Artificial General Intelligence (AGI) or general AI. These are machines that would have truly human (or super-human) reasoning abilities and emulate humans' cognitive processes (Shneiderman 2020). Such machines do not currently exist and it is unclear whether they will ever be possible using existing technological principles (Smith 2019). They nevertheless figure prominently in popular culture and inspire ethical questions such as whether machines could develop consciousness, deserve to be assigned rights, and have personhood.

An Ecosystems Perspective on AI Ethics

The above quick overview of some of the key discussion points of AI ethics has shown the breadth of the debate, based on different concepts of AI and the almost infinite set of possible applications. This calls into question what might be called the common-sense approach to the ethics of AI. By this we mean, the view that there is a clearly identifiable technology called AI which, when used has clearly identifiable consequences that translate into ethical concerns. Once recognised, these can then be addressed and rectified.

This common-sense view which is not usually spelt out in detail, arguably underlies much of the discussion of how to deal with ethics of

AI. When spelt out in the simple form just suggested it is clear, however, that this view is not tenable. AI is not a clearly identifiable technology but a set of families of technologies and techniques that have very little in common. The current focus on machine learning has the advantage of narrowing down the candidate technologies but it misses the fact that there are other fields of AI such as expert systems or fuzzy logic which justifiably use the same label but do not share the same technical characteristics. Furthermore, the ethical issues raised by AI are context-dependent. AI does not 'have' ethical issues, but these emerge in contexts of use depending on the stakeholders that are involved. Whether the ability of AI to detect patterns in the data and propose actions on this is ethically problematic has little to do with the technical implementation and more with the moral sensitivities of the people who are involved.

Elsewhere we have therefore proposed to look at AI through the lens of a system, more specifically to see AI as a (set of interlinking) innovation ecosystems (Stahl 2021; Stahl et al. 2021). AI as instantiations of computer systems can easily be described from a systems perspective. There is a long-standing tradition, in particular in the field of information systems to highlight the social side of systems by focusing on the concept of socio-technical systems (Avgerou and McGrath 2007; Leonardi 2012; Mumford 2006). We are suggesting that this socio-technical approach is important to be included in the ethics of AI debate, as it can explain the interdependence of actors and technologies, the difficulty of predicting outcomes and the challenge of delineating the system and drawing clear boundaries.

We are furthermore pointing to the literature on innovation ecosystems (Adner 2006; Carayannis et al. 2021; Moore 1993) as a further source of inspiration that would benefit the AI ethics debate. This literature has a focus on the process of innovation and the creation of new socio-technical systems, which is the stage where AI remains. It has developed a number of concepts and methods to understand the state of an innovation ecosystem as well as means of intervening and governing such systems (Lis and Otto 2021; Wareham et al. 2014). Combining systems theory and ethics is not trivial and it goes beyond this chapter to explore their relationship. For our purposes, it is sufficient to point to the socio-technical innovation ecosystems nature of AI to highlight that addressing their ethical consequences is not straightforward. This allows us to return to the main point of this chapter, namely to explore whether and how AI

policies and strategies take into account ethical concerns and may even offer avenues for addressing them.

AI Policies and Strategies

Since 2016 a broad range of soft-law (non-regulatory) approaches to AI have been released by governments, international organisations, multinational corporations, civil society organisations and non-governmental organisations. With the overarching purpose of promoting the benefits and addressing the risks of AI (Roberts et al. 2021), this body of documents includes policies, strategies, reports, white papers, guidelines and principles. While not legally binding, these documents can frame the thinking and can influence decision-making within stakeholder groups about AI (Jobin et al. 2019).

Concepts

Within the scholarly community, there is some debate around the relative merits of creating strategies versus policies in order to address the risks and promote the benefits of AI. Fatima et al. (2020) argue that strategies and policies can be distinguished in the following way; strategies set out the vision and ambitions along with key priority areas and rationale behind these choices, whereas policy involves the operationalisation of the strategy into tangible objectives and sets of actions. However, much of the research around AI, ethics and governance, uses the terms 'policy' and 'strategy' interchangeably (Roberts et al. 2021; Robinson 2020). As such, and in order to consider a suitably wide range of approaches to ethical issues in AI policy/strategy, for the purposes of this chapter, we will consider policies and strategies jointly, adopting Calo's (2017) broad definition of 'policy' as a concept used to describe "societal efforts to channel AI in the public interest", which clearly coincides with the intention to promote the benefits and address the risks of AI inherent to both policies and strategies that consider AI. In this chapter, we draw from a selection of AI policies described in more detail in Ulnicane et al. (2020).

The AI Policy Landscape

By their very nature national policies imply a fundamental level of competitiveness over collaboration, through the prioritisation of the impacts of

AI likely to be felt within that nation, and the promotion of the benefits and mitigation of the risks of AI for this prime stakeholder group—in fact, the very production of a national AI policy seems to itself have morphed into a statement on the seriousness with which the consequences of AI (to wit, the benefits of being first and a world-leader versus the detriment of being left behind) are being considered.

The consideration of AI policies and the extent to which they substantively address ethical issues requires an understanding of the vision for AI being promoted by policymakers. These visions vary greatly globally and can contribute to the prioritisation of certain ethical issues over others. China's AI policies (China State Council 2015, 2017) focus heavily on increasing economic and military competitiveness on a global scale and view AI as a strategic tool in promoting its global position (Allen 2019). The United States also derives its policy from the perspective of the protection of American interests globally through strategic military dominance ("DOD Adopts Ethical Principles for Artificial Intelligence" 2020; The Executive Office of the President of the United States 2020). Russia's AI policy follows clearly Vladimir Putin's assertion that "whoever becomes the leader in the field [of AI] would rule the world", espousing a vision of technological sovereignty and concomitant global competitiveness (Office of the President of the Russian Federation 2019). European national policies, however, have a much clearer focus on the societal impact of AI and seek to address the impact of AI on people's daily lives, ensuring that the benefits are shared and risk of harms mitigated for citizens (see, e.g. AI4Belgium 2019; Government Offices of Sweden 2018; House of Lords 2018). They broadly consider the need to promote and protect the fundamental rights of the individual, and that these rights should not be elided in the interest of promoting the economic benefits of AI (Cedric Villani 2018; European Union Agency for Fundamental Rights 2020). This is in no way to suggest that military applications of AI are not considered within the EU—simply that military applications are specifically eliminated from the scope of current policy and legislation (European Commission 2021). Furthermore, the very nature of the EU as a supranational organisation means that EU AI policy focuses strongly on value alignment across the member states and takes a highly collaborative approach to envisioning an AI future (The European Commission 2020).

There is a consensus amongst policymakers that AI, as a technology, creates ethical dilemmas (see, for example Campolo et al. 2017; Cedric

Villani 2018; European Commission 2018; World Economic Forum 2018). However, there is a breadth of consideration for what those ethical issues actually are: the primary issues identified by Ouchchy et al. (2020) differ from those identified by Stahl et al. (2016), which differ again from those enumerated by Stahl (2021). In order to maintain a focus on a clear vision for an AI future, many policies focus on a small number of specific areas of ethical concern. However, the consideration of these issues in policy, and substantively, the selection of issues for consideration, appears to vary greatly across policy. What follows is a brief overview of some of the more commonly addressed ethical concerns, and the approach taken to addressing these in policy.

One of the most robustly considered issues within AI policy is automation, and in particular the risk of jobs being displaced and replaced by AI. In general, policy focuses much more heavily on attempting to mitigate the risk of mass unemployment than trying to drive a benefit of AI—in the form of productivity gains—which is unsurprising given the strength of both feeling and rhetoric regarding unemployment, income preservation and state financial support for individuals (Thierer et al. 2017). While automation poses clear ethical questions about which jobs are likely to be automated (Big Innovation Centre 2017a), who is likely to be most affected (Big Innovation Centre 2017b), and where the responsibility for ensuring people's livelihoods lies (IPPR 2017), in practice policy takes a fiscal approach to mitigating the risks of job loss through automation; where job losses are expected to be manageable, policies tend to consider investment in retraining and upskilling programmes (House of Lords 2018). Where the expectation of job loss is much more severe, state financial support for the individual and universal basic income ideas tend to be promoted (Thierer et al. 2017).

Many AI systems are dependent on being trained on large volumes of data in order to perform effectively, raising concerns around the privacy of the individual and data protection (Campolo et al. 2017). Many of the concerns raised around privacy and data protection relate to the need to ensure that the use of data infringes on the rights of the individual to privacy as little as possible (House of Lords 2018), including considering issues of surveillance or monitoring which may become concomitant with the sharing of personal data (Campolo et al. 2017; The 2015 Study Panel 2016), while others focus on issues of re-identification of individuals through the process of data mining (World Economic Forum 2018). In

attempting to address some of these issues, some policies adopt the position of supporting a legal perspective (the human right to privacy [Ponce Del Castillo 2017]) or look to promote positive privacy steps through consideration of parties responsible for such breaches (UNI Global Union 2017).

In the light of a range of recent, high-profile scandals (see, e.g. Akter et al. 2021) many policies seek to robustly address the impact of biased and discriminatory AI. However, there are difficulties in trying to address these very different, yet interrelated issues through policy. Bias, the differentiation of outcome based on preference or likelihood, can occur when training an AI system on data which may reflect past decisions and therefore historic injustices and inequalities and unconscious discriminatory attitudes (Ferrer et al. 2021). A strong debate around whether bias should, or even can, be removed from datasets (House of Lords 2018) has created a lack of clarity about concrete steps that could be taken to address this issue (The 2015 Study Panel 2016), and in some cases has elicited allegations of the intention to render a social problem as a technical one (Jobin et al. 2019). Discrimination is considered an outcome that unfairly disadvantages or detriments one group of people in favour of another. Many national policies in Europe address issues of AI discrimination through the lens of the existing GDPR (European Parliament and European Council 2016) (under the extension of rights regarding automated decision-making (Centre for Data Ethics 2020; ICO 2017), and equality legislation (in relation to discrimination relating to protected characteristics) (European Union Agency for Fundamental Rights 2020).

Certain groups within society can be disadvantaged by AI across a number of vectors. Those in jobs likely to be displaced by AI is one group, and those in groups likely to be un- or underrepresented in data (resulting in bias and discriminatory outcomes) is another. However, the impact of AI may not simply amount to those that are negatively impacted by AI—consideration is also given for the fact that certain groups may not be in a position to benefit from AI (House of Lords 2018; IPPR 2017). The causes of this may be broad, and can include; a prohibitive cost of technology, lack of infrastructure and digital illiteracy, amongst other factors. Current policies identify this risk, but further exploration of a range of measures to mitigate this risk could prove fruitful.

Many policies address what is seen as a fundamental issue around the development, deployment and adoption of AI; namely, the concept of trust. This issue focuses around two key areas; trust in the AI systems

themselves (Campolo et al. 2017), and trust in the companies that develop and deploy AI systems (AI Now Institute 2018), and the argument is made that public trust in AI is vital in ensuring that it can be developed and deployed to the benefit of society (House of Lords 2018). There is a plurality of approaches to addressing this issue. Some policies advocate for regulation to boost public confidence in AI (European Commission 2018), others suggest that improving public understanding of AI (and its limitations) may help to address this issue (World Economic Forum 2018), some focus on the role of government in assuring the public of the safety of AI (IEEE-USA 2017), and yet more suggest co-creation strategies as a method of improving public confidence (The Federal Government 2018). Whilst there are such a wide range of strategies considered in these policies, few of them have been tested to determine their success in building public confidence—a determination made more difficult by the long-term nature of some of the impacts and effects of AI.

As well as addressing relevant and practical ethical issues, many policies also include some consideration for the future development of an AGI, for example, by asking how humanity can be protected from superintelligent machines and ensuring that such an AI would act safely (Big Innovation Centre 2017c; World Economic Forum 2018). While, as a point of philosophical discussion, these questions are clearly of interest to many scientists, policymakers and members of the public, it is worth raising the question as to the value added in the inclusion of this debate to policy, given the short- to mid-term nature of specific policy impacts and the likelihood that, if AGI were ever to be developed, it would be an emergence in the long term (Crawford and Whittaker 2016).

AI Strategies in Africa

Based on the understanding of AI ethics and AI policy developed in the preceding sections, we now evaluate how these topics are considered in Africa from the regional perspective by the African Union and as concentrated in each individual Northern African country. The North Africa region was selected as a case study for this purpose due to the relative proliferation of AI specific strategies in this area, as opposed to the more generalised strategies which include an AI element common in other regions of Africa (such as those relating to the 4th Industrial Revolution or the Digital Economy).

Processes for the Creation of AI Strategies in North Africa

Africa as a continent is not left out of the discussion regarding deployment of AI in different sectors of economic and social development. The diffusion of AI across different sectors of its application in Africa includes different stakeholders. The concerns highlighted in other parts of the world form the basis of a few national strategies such as Egypt as will be discussed below. Its deployment has seen active participation by the private and public sectors. AI Technology has been used in health, agriculture, fintech, public transportation as well as language translation. Academia has also developed different initiatives to ensure its development while the continent boasts of hundreds of AI hubs (Gwagwa et al. 2020). However, the absence of national policies regulating AI technologies has been of major concern (Candelon et al. 2021).

The negative impact of this challenge is not lost on the African Union which instituted a Task Force on the 26th October, 2019, mandating member states to "establish a working group on Artificial intelligence to study the creation of a common African stance on AI, the development of an Africa wide capacity to building framework and establishment of an AI think thank to assess and recommend projects to collaborate on in line with Agenda 2063 and the UNSDGs" (African Union 2020). African countries have established groups and initiatives to govern the technology. In 2019, the African Working Group on AI was established amongst other things, "to establish a common AI strategy for Africa" (OECD 2019).

In North Africa, comprising Morocco, Algeria, Libya, Egypt, Tunisia and Sudan, almost all the countries in the region have already instituted processes of framing policies to ensure safe deployment, harness economic potentials and ensure ethical use of AI. For example, Algeria presented its National Artificial Intelligence Strategy 2020–2030 on the 18th of January, 2021, to "improve Algerian skills in the field of AI through education, training and research, on the one hand, and strengthen these capacities as a development tool allowing socio-economic sectors to iron out the obstacles hindering the digital transition underway, on the other hand" ("Strategy for research in artificial intelligence launched" 2021).

Tunisia also joined the race by creating a Task Force to "devise a methodology and an action plan to produce the country's National AI Strategy (Ministère de l'Enseignement Supérieur et de la Recherche Scientfique 2018).

In crafting its national strategy, the Egyptian government formed the National Council for AI "as a partnership between the governmental institutions, prominent academics and practitioners from leading businesses in the field of AI" with a responsibility amongst others, to "identify AI applications that provide smart, safe and sustainable solutions and services". The country expects to "track and monitor the implementation of the strategy, laws and regulations, ethical principles and guidelines" (MCIT 2020).

These developments identify ongoing discourse on how national and regional policies on AI in Africa already anticipate ethical concerns in their quest to harness the full potentials which the technology is capable of on the continent. Government has been described as "both an enabler of innovation" and as "a driver of demand for AI". Thus, governments of African countries have the onerous responsibility of framing policies, establishing research centres while effective regulation must address issues relating to transparency, accountability, safety, etc. ("Developing an artificial intelligence strategy," n.d.; Kiemde and Kora 2021).

Future national policies, strategies and initiatives will further drive ethical considerations along with economic, social, impact on work, etc., as well as attainment of the United Nations Sustainable Development Goals (Vinuesa et al. 2020).

Specific Aspects of AI Strategies in North Africa

As already pointed out, only three of the 6 countries in North of Africa have made positive steps in the development of National AI Strategies. Of these countries, Egypt is the sole country that has completed the process of creating a policy document and has made this document publicly available online (MCIT 2020). It is interesting to see that Egypt's national AI strategy document titled "National AI Strategy" has in its mission the creation of governance mechanisms to ensure the sustainability and competitiveness of the AI industry in Egypt. This is important considering the dominance of the developed countries and their corporations in the AI industry and how adept they are at commercialising AI technologies. Although no justification has been provided, Egypt has identified 5 priority sectors for the development of AI namely Agriculture/environment and water management, Health care, National Language Processing, Economic planning and Manufacturing and infrastructure management. Also, the document says very little about ethical

governance of AI and related technologies except that there is an ambition to track and monitor implementation of strategy, laws and regulations, ethical principles and guidelines.

Although Algeria is said to have developed a national AI strategy, a policy document that fully spells out policy considerations of the strategy was not found during the development of this chapter, perhaps due to language barriers as it may be available in other languages other than English. What is publicly available, is a summary of an event to launch the Algerian AI strategy titled "the national strategy of research and innovation on Artificial Intelligence (2020–2030)" which can be found on the websites of the Algerian Embassy (2021). The strategy which appears to be directed at higher education aims to develop expertise in AI through teaching, training and research. According to the Algérie Press Service (APS 2021) while higher education is a priority sector for the strategy, the other priority sectors are health, energy and technologies. It appears there are little or no considerations for issues relating to ethics, regulation or governance of AI in the Algerian AI strategy as ethical considerations do not feature in the available documents.

While Tunisia has also taken important steps towards the development of a national AI strategy, very little information is publicly available about the ongoing development efforts. In 2018, the Tunisian Secretary of State for Research set up a Task Force and Steering Committee to develop the national AI strategy of Tunisia. The Tunisian National Agency for the Advancement of Scientific Research suggests that this strategy which aims at securing a respectable place and proactive role for Tunisia in the global AI sector also considers the ethical challenges of AI (ANPR 2018). A summary of the developments in the creation of AI strategies in North Africa can be seen in Table 1.

The above analysis gives some indication of how ethics have been addressed in the strategy documents of the North African countries. It has shown that very little has been said about AI ethics in the countries that have started developing AI strategies. While it would be interesting to understand why these policy documents barely mentioned AI ethics, there is very little information in the available documents to indicate why this is the case. Therefore, any explanation provided here would be mere speculation, rather than fact and that is not very helpful. What is clear from the document is that AI ethics is not a priority area in the countries that have begun devising AI strategies in North Africa.

Table 1 Overview of existing North-African AI strategies

Country	*Title of AI strategy document*	*Status*	*Policy areas/priority segments or sectors*	*Regulatory/ethical considerations*
Algeria	The national strategy of research and innovation on Artificial Intelligence (2020–2030)	Completed	– Higher education – Health – Energy – Technologies	
Egypt	National AI Strategy	Completed	– Agriculture/environment and water management – Healthcare – Natural Language Processing – Economic planning – Manufacturing and infrastructure management	Track and monitor implementation of strategy, laws and regulations, ethical principles and guidelines
Tunisia	National Artificial Intelligence Strategy	In progress		Sustainable, equitable development, and ethical challenges

Discussion

The above indicates that parallels can be drawn from the trajectories of national AI policies in North Africa with those on the global scene. For example, in recognition of the power of AI to transform lives and improve the economic landscape of nations, it can be seen that generally, AI strategies are geared towards national economic and technological advancement. Like other countries, the strategies of the North African countries have been designed to target specific sectors that can enable the achievement of such ambitions. By targeting sectors that have been historically neglected by the government such as health care and infrastructure development (e.g. energy), the AI strategies of North Africa are seeking to channel the productivity that AI enables for economic development. Also, in recognition of the fact that AI will play a major role in transformative technologies that will emerge in the next decade, North African

countries like Algeria have included education in the list of priority areas for AI.

Interestingly, unlike the United States and China, the North African countries appear uninterested in increasing their military competitiveness with the aid of AI as their national AI strategies have remained quiet on this subject. If it turns out that it is the case these countries are uninterested in the military uses of AI, then they cleverly avoid the ethical dilemmas associated with the use of AI for military purposes. As Pfaff (2020) points out, although the deployment and employment of AI in the military context might have its advantages, they could also be problematic, for example, by resulting in "atrocities for which no one is accountable". It is such concerns that have prompted the European Parliament (Legislative Observatory 2021) to stress that "autonomous weapons systems raise fundamental ethical and legal questions about the ability of humans to control these systems" and may therefore only be used as a last resort and must be subject to human control.

Unlike many of the big economic and technical blocks that dominate the AI landscape including the development of national AI strategies, the North African AI strategies contain little or nothing on ethical or human rights considerations for the deployment and employment of the technology nationally. For example, the Tunisian AI strategy only mentions a consideration of the ethical challenges of AI in its quest to attain a secure and respectable place in the global AI race. The strategy does not go into any detail about what type of ethical challenges it would focus on, or how this would be achieved. Similarly, Egypt, the only other country that has some sort of ethical consideration in its national AI strategy, appears to only be interested in the tracking and monitoring the implementation of ethical principles, guidelines, laws and regulations of AI. What it means by tracking and monitoring the implementation of ethical principles is ambiguous because there is no indication of the ethical principles that it considers important, nor is there a plan for achieving that in the strategy. There is, therefore, no substance to the statement on ethics and the interest in ethical AI can be said to be superficial.

As already pointed out in previous sections, the growing use of AI raises concerns about the ethical consequences of their deployment and employment across all areas of society. Considering the ethical concerns raised by the application and deployment of AI, the promotion of its benefits should never be undertaken without appropriate safeguards for the mitigation of its risks. In this respect, lessons can be learned from

other countries that have embedded clear mechanisms for the consideration of ethical governance of AI in their national AI strategies. For example, the UK which published its first dedicated national AI strategy on 22 September 2021 stresses that the UK must get the national and international governance of AI right to encourage innovation, investment and the protection of the public and fundamental values (UK Government 2021).

The UK AI strategy has, therefore, included themes on the ethical governance of AI throughout the document and indicates a clear mechanism for the attainment of this ambition giving an indication of the seriousness with which it considers this matter. Measures identified in its short, medium and long-term plan include governance and regulatory regimes that keep pace with the fast-changing demands of AI, a plan for horizon scanning to increase governments awareness of AI safety, publishing of a white paper on the national position on governing and regulation of AI, the development of cross-government standards, updating guidance on AI ethics and safety in the public sector, actions that can safely advance AI and the mitigation of risks, backing of diversity in AI, Measures identified in its short, medium and long-term plan include governance and regulatory regimes that keep pace with the fast-changing demands of AI, a plan for horizon scanning to increase governments awareness of AI safety, publishing of a white paper on the national position on governing and regulation of AI, the development of cross-government standards, updating guidance on AI ethics and safety in the public sector, actions that can safely advance AI and the mitigation of risks, and the backing of diversity in AI,

To avoid being forced to adopt or accept the standards set by other countries, the North African countries and by extension, countries in other parts of Africa would be well-advised to start now to seriously include ethical considerations with clear mechanisms for the attainment of ethical AI in their national AI strategies. As much as they are encouraged to learn from other countries in developing strategies for ethical AI, they must realise that the different cultures, values and norms in their home countries would likely require different ethical approaches. While they cannot get away with simply mentioning ethics in their strategies, they cannot simply copy the ethical standards set by others. Any ethical principles, standards, and governance mechanisms that are developed must also be done from the perspective of the protection of their people, their culture and their values.

Conclusion

We considered the different traditional definitions given to AI as a technology in writing this chapter. We then suggested a relatively novel perspective of defining AI as a family of technologies and techniques as a background for determining how ethics of the technology may be properly conceptualised. As a contribution to new knowledge, this chapter is framed to consider the subject of ethics as a proposed catalyst to activate, motivate and ensure that national policies on AI are prioritised on the continent, referencing developments in the northern region. We also introduced the socio-technical approach to be included in debates on ethics of AI to explain interdependence of actors and technologies. We suggest that this will help in properly shaping the question and resolution of what may be ethical about the deployment of technologies making up AI within different sectors. These are new perspectives which contribute to knowledge and suggest future discourse as may be directly related to the nuances of an heterogeneous continent, using the northern region for reference.

In this chapter, we have looked at key ethical concerns that AI raises and explored whether these are reflected in AI strategies and policies. While such strategies and policies have multiple goals, we found that addressing ethical concerns is often part of these. We then looked at the landscape of AI policies in Africa. We could only identify a limited number of these in North Africa but found limited engagement with ethical concerns in these North African documents.

The efforts at developing national strategies for deployment of AI are gaining momentum. We have been able to establish that regional bodies including the African Union recognise AI has a priority technology in the successful digitisation of processes for use in the private and public sectors. The reports of the United Nations (Hu et al. 2019) specifically tailored to examine AI readiness by African nations point to different challenges which militate against nations being able to develop AI strategies. Issues such as lack of technical knowledge, paucity of funds and critical infrastructure are just some of the pressing needs which the African States require to overcome. Thus, there is a pressing need to protect African citizens by ensuring deployed AI is safe for use, assurance that it will not lead to job losses or raise problems inclusion, recognition of diversities and gender equity are ethical concerns which should be incorporated into future national AI Policies.

In this instance, Africa will not need to play catch-up even though there are worries that it will be left behind in the adoption of technology. Proactive national policies which embody suitable ethical frameworks can be expected to emerge. Governments of African states could play more active roles with regards to the importance of safe and ethical AI to frame national policies which govern AI adoption by both the private and public sectors. While adoption of the technology has been accelerated and largely driven by the private sector, harnessing the full economic potentials requires a governance framework which is the responsibility of states. The establishment of such policy frameworks will assure investors and citizens alike and help countries benefit from AI while mitigating its undesirable side effects.

To effectively harness the benefits of AI while managing the ethical and other risks that AI technologies pose, we recommend that governments that are developing AI strategies or policies should not only set out their vision and ambitions along key economic segments or sectors but also seriously address the ethical issues of AI. This implies clearly spelling out mechanisms for the mitigation of such issues. Based on the insights developed in this chapter, we believe that it is possible to learn from the examples of other countries that have successfully developed an AI strategy that clearly balances their ambitions with ethical considerations and mechanisms. However, we suggest that in developing new strategies and policies governments should try to avoid copying strategies from other countries and instead base theirs on local considerations of values and cultures.

References

Access Now, 2018. Human Rights in the Age of Artificial Intelligence. Access Now.

Adner, R., 2006. Match Your Innovation Strategy to Your Innovation Ecosystem. Harv. Bus. Rev. 84, 98–107.

African Union, 2020. The Digital Transformation Strategy for Africa (2020–2030). African Union, Addis Ababa, Ethiopia.

AI4Belgium, 2019. AI4Belgium.

AI Council, 2021. AI Roadmap. Office for Artificial Intelligence. Department for Business, Energy & Industrial Strategy, and Department for Digital, Culture, Media & Sport, London.

AI HLEG, 2019. Ethics Guidelines for Trustworthy AI. European Commission—Directorate-General for Communication, Brussels.

AI Now Institute, 2018. AI Now Report 2018. New York.
Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y.K., D'Ambra, J., Shen, K.N., 2021. Algorithmic Bias in Data-Driven Innovation in the Age of AI. Int. J. Inf. Manag. 102387. https://doi.org/10.1016/j.ijinfomgt.2021.102387
Algerian Embassy, 2021. Higher Education: National Artificial Intelligence Strategy 2020–2030 presente [WWW Document]. http://algerianembassy.org/science-technology-environment/higher-education-national-artificial-intelligence-strategy-2020-2030-presented.html (accessed 6.2.21).
Allen, G.C., 2019. Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security. Center for New American Security.
ANPR, 2018. National AI Strategy: Unlocking Tunisia's Capabilities Potential. Agence Nationale de la Promotion de la Recherche scientifique. http://www.anpr.tn/national-ai-strategy-unlocking-tunisias-capabilities-potential/ (accessed 10.14.21).
APS, 2021. Presentation of the National Artificial Intelligence Strategy 2020–2030 [WWW Document]. https://www.aps.dz/sante-science-technologie/116102-enseignement-superieur-presentation-de-la-strategie-nationale-de-l-intelligence-artificielle-2020-2030 (accessed 6.2.21).
Aristotle, 2007. The Nicomachean Ethics. Filiquarian Publishing, LLC.
Avgerou, C., McGrath, K., 2007. Power, Rationality, and the Art of Living Through Socio-Technical Change. MIS Q. 31, 295–315.
Babuta, A., Oswald, M., Janjeva, A., 2020. Artificial Intelligence and UK National Security—Policy Considerations (Occasional Paper). Royal United Services Institute for Defence and Security Studies.
Bentham, J., 1789. An Introduction to the Principles of Morals and Legislation. Dover Publications Inc.
Berendt, B., 2019. AI for the Common Good?! Pitfalls, challenges, and ethics pen-testing. Paladyn J. Behav. Robot. 10, 44–65. https://doi.org/10.1515/pjbr-2019-0004
Big Innovation Centre, 2017a. Governance, Social and Organisational Perspective for AI. A theme report based on the 5th meeting of the All-Party Parliamentary Group on Artificial Intelligence [APPG AI]. Big Innovation Centre, London.
Big Innovation Centre, 2017b. Inequality, Education, Skills, and Jobs. A theme report based on the 6th meeting of the All-Party Parliamentary Group on Artificial Intelligence [APPG AI]. Big Innovation Centre, London.
Big Innovation Centre, 2017c. International Perspective and Exemplars, All-Party Parliamentary Group on Artificial Intelligence. Big Innovation Centre, London.

Boden, M.A., 2018. Artificial Intelligence: A Very Short Introduction, Reprint edition. OUP, Oxford, United Kingdom.

Borenstein, J., Grodzinsky, F.S., Howard, A., Miller, K.W., Wolf, M.J., 2021. AI Ethics: A Long History and a Recent Burst of Attention. Computer 54, 96–102. https://doi.org/10.1109/MC.2020.3034950

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G.C., Steinhardt, J., Flynn, C., hÉigeartaigh, S.Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., Amodei, D., 2018. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. ArXiv180207228 Cs.

Buttarelli, G., 2018. Choose Humanity: Putting Dignity Back into Digital.

Bynum, T.W., 2001. Computer Ethics: Its Birth and Its Future. Ethics Inf. Technol. 3, 109–112. https://doi.org/10.1023/A:1011893925319

Calo, R., 2017. Artificial Intelligence Policy: A Primer and Roadmap. UC Davis Law Rev. 51, 399.

Campolo, A., Sanfilippo, M., Whittaker, M., Crawford, K., 2017. AI Now 2017 Report. AI Now Institute, New York.

Candelon, F., El Bedraoui, H., Maher, H., 2021. Developing an Artificial Intelligence for Africa Strategy [WWW Document]. OECD Dev. Matters. https://oecd-development-matters.org/2021/02/09/developing-an-artificial-intelligence-for-africa-strategy/

Carayannis, E.G., Grigoroudis, E., Stamati, D., Valvi, T., 2021. Social Business Model Innovation: A Quadruple/Quintuple Helix-Based Social Innovation Ecosystem. IEEE Trans. Eng. Manag. 68, 235–248. https://doi.org/10.1109/TEM.2019.2914408

Cedric Villani, 2018. For a Meaningful Artificial Intelligence. Towards a French and European Strategy. French Parliament, Paris.

Centre for Data Ethics, 2020. AI Barometer Report.

China State Council, 2015. Made in China 2025.

China State Council, 2017. A Next Generation Artificial Intelligence Development Plan.

Coeckelbergh, M., 2020. AI Ethics. The MIT Press.

Council of Europe, 2019. Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights.

Crawford, K., Whittaker, M., 2016. The AI Now Report. The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term. AI Now Institute, New York.

Developing an Artificial Intelligence Strategy: National Guide [WWW Document], n.d. ESCWA. http://www.unescwa.org/publications/developing-artificial-intelligence-strategy-national-guide (accessed 10.4.21).

Dignum, V., 2019. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way, 1st ed. Springer.
DOD Adopts Ethical Principles for Artificial Intelligence [WWW Document], 2020. US Dep. Def. https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/ (accessed 10.2.21).
EDPS, 2020. EDPS Opinion on the European Commission's White Paper on Artificial Intelligence—A European Approach to Excellence and Trust (Opinion 4/2020) (Opinion No. 4/2020). EDPS.
Elsevier, 2018. ArtificiaI Intelligence: How Knowledge Is Created, Transferred, and Used—Trends in China, Europe, and the United States. Elsevier, Amsterdam.
European Commission, 2018. Artificial Intelligence—A European Perspective. European Commission, Luxembourg.
European Commission, 2021. Proposal for a Regulation on a European Approach for Artificial Intelligence (No. COM(2021) 206 final). European Commission, Brussels.
European Parliament, 2020. The Ethics of Artificial Intelligence: Issues and initiatives (No. PE 634.452). EPRS | European Parliamentary Research Service.
European Parliament, European Council, 2016. General Data Protection Regulation.
European Union Agency for Fundamental Rights, 2020. Getting the Future Right: Artificial Intelligence and Fundamental Rights.
Fatima, S., Desouza, K.C., Dawson, G.S., 2020. National Strategic Artificial Intelligence Plans: A Multi-dimensional Analysis. Econ. Anal. Policy 67, 178–194. https://doi.org/10.1016/j.eap.2020.07.008
Ferrer, X., van Nuenen, T., Such, J.M., Coté, M., Criado, N., 2021. Bias and Discrimination in AI: A Cross-Disciplinary Perspective. IEEE Technol. Soc. Mag. 40, 72–80.
Floridi, L., 1999. Information Ethics: On the Philosophical Foundation of Computer Ethics. Ethics Inf. Technol. 1, 33–52.
Government Offices of Sweden, 2018. National Approach to Artificial Intelligence. Sweden.
Guterres, A., 2020. The Highest Aspiration—A Call to Action for Human Rights. United Nations.
Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., de Beer, J., 2020. Artificial Intelligence (AI) Deployments in Africa: Benefits, Challenges and Policy Dimensions. Afr. J. Inf. Commun. 26, 1–28. https://doi.org/10.23962/10539/30361

Hall, W., Pesenti, J., 2017. Growing the Artificial Intelligence Industry in the UK. Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy, London.

House of Lords (H. of L.), 2018. AI in the UK: Ready, Willing and Able? Report of Session 2017–19. Select Committee on Artificial Intelligence, London.

Hu, X., Bhanu, N., Flores, E., Lucia, Prateek, S., Macarena, R.L., 2019. Steering AI and Advanced ICTs for Knowledge Societies: A Rights, Openness, Access, and Multi-stakeholder Perspective. UNESCO Publishing.

ICO, 2017. Big Data, Artificial Intelligence, Machine Learning and Data Protection. Information Commissioner's Office.

IEEE-USA, 2017. Artificial Intelligence Research, Development & Regulation (Position Statement). IEEE-USA, Washington, DC.

Introna, L.D., 2005. Disclosive Ethics and Information Technology: Disclosing Facial Recognition Systems. Ethics Inf. Technol. 7, 75–86.

IPPR, 2017. Managing Automation: Employment, Inequality and Ethics in the Digital Age. IPPR, London.

Jobin, A., Ienca, M., Vayena, E., 2019. The Global Landscape of AI Ethics Guidelines. Nat. Mach. Intell. 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kant, I., 1788. Kritik der praktischen Vernunft. Reclam, Ditzingen.

Kant, I., 1797. Grundlegung zur Metaphysik der Sitten. Reclam, Ditzingen.

Kaplan, A., Haenlein, M., 2019. Siri, Siri, in My Hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence. Bus. Horiz. 62, 15–25.

Kiemde, S.M.A., Kora, A.D., 2021. Towards an Ethics of AI in Africa: Rule of Education. AI Ethics. https://doi.org/10.1007/s43681-021-00106-8

Latonero, M., 2018. Governing Artificial Intelligence: Upholding Human Rights & Dignity. Data&Society.

Legislative Observatory, 2021. Artificial Intelligence: Questions of Interpretation and Application of International Law in so far as the EU Is Affected in the Areas of Civil and Military Uses and of State Authority Outside the Scope of Criminal Justice. European Parliament, Brussels.

Leonardi, P., 2012. Materiality, Sociomateriality, and Socio-Technical Systems: What Do These Terms Mean? How Are They Related? Do We Need Them?, in: Leonardi, P.M., Nardi, B.A., Kallinikos, J. (Eds.), Materiality and Organizing: Social Interaction in a Technological World. Oxford University Press, Oxford, pp. 25–48.

Lis, D., Otto, B., 2021. Towards a Taxonomy of Ecosystem Data Governance, in: Proceedings of the 54th Hawaii International Conference on System Sciences. Presented at the Hawaii International Conference on System Sciences, pp. 6067–6076. https://doi.org/10.24251/HICSS.2021.733

McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E., 2006. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, august 31, 1955. AI Mag. 27, 12–12.

MCIT, 2020. Artificial Intelligence: Egypt's AI Strategy [WWW Document]. Online. https://mcit.gov.eg/en/Artificial_Intelligence (accessed 10.14.21).

Mill, J.S., 1861. Utilitarianism, 2nd Revised edition. Hackett Publishing Co, Inc.

Ministère de l'Enseignement Supérieur et de la Recherche Scientfique, 2018. National AI Strategy: Unlocking Tunisia's capabilitiespotentia. http://www.jaist.ac.jp/~bao/AI/OtherAIstrategies/National%20AI%20Strategy:%20Unlocking%20Tunisia%E2%80%99s%20capabilities%20potential.%20%E2%80%93%20Agence%20Nationale%20de%20la%20Promotion%20de%20la%20Recherche%20scientifique.pdf (accessed 6.2.21).

Moore, J.F., 1993. Predators and Prey: A New Ecology of Competition. Harv. Bus. Rev. 71, 75–86.

Mumford, E., 2006. The Story of Socio-Technical Design: Reflections on Its Successes, Failures and Potential. Inf. Syst. J. 16, 317–342.

Nemitz, P., 2018. Constitutional Democracy and Technology in the Age of Artificial Intelligence. Philos. Trans. R. Soc. A 376, 20180089. https://doi.org/10.1098/rsta.2018.0089

Nishant, R., Kennedy, M., Corbett, J., 2020. Artificial Intelligence for Sustainability: Challenges, Opportunities, and a Research Agenda. Int. J. Inf. Manag. 53, 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104

OECD, 2019. Recommendation of the Council on Artificial Intelligence (OECD Legal Instruments). OECD.

Office of the President of the Russian Federation, 2019. Decree of the President of the Russian Federation on the Development of Artificial Intelligence in the Russian Federation.

Ouchchy, L., Coin, A., Dubljević, V., 2020. AI in the Headlines: The Portrayal of the Ethical Issues of Artificial Intelligence in the Media. AI Soc. https://doi.org/10.1007/s00146-020-00965-5

Pfaff, C.A., 2020. The Ethics of Acquiring Disruptive Technologies: Artificial Intelligence, Autonomous Weapons, and Decision Support Systems. Tex. Natl. Secur. Rev. 3, 17. http://dx.doi.org/https://doi.org/10.26153/tsw/8107

Ponce Del Castillo, A., 2017. A Law on Robotics and Artificial Intelligence in the EU? (Foresight Brief No. #02). European Trade Union Institute, Brussels.

Rai, A., Constantinides, P., Sarker, S., 2019. Next-Generation Digital Platforms: Toward Human–AI Hybrids. Mis Q. 43, iii–x.

Richards, L., Brockmann, K., Boulanini, V., 2020. Responsible Artificial Intelligence Research and Innovation for International Peace and Security. Stockholm International Peace Research Institute, Stockholm.

Roberts, H., Cowls, J., Hine, E., Mazzi, F., Tsamados, A., Taddeo, M., Floridi, L., 2021. Achieving a 'Good AI Society': Comparing the Aims and Progress

of the EU and the US (SSRN Scholarly Paper No. ID 3851523). Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3851523

Robinson, S.C., 2020. Trust, Transparency, and Openness: How Inclusion of Cultural Values Shapes Nordic National Public Policy Strategies for Artificial Intelligence (AI)—ScienceDirect. Technol. Soc. 63.

Shneiderman, B., 2020. Design Lessons from AI's Two Grand Goals: Human Emulation and Useful Applications. IEEE Trans. Technol. Soc. 1, 73–82. https://doi.org/10.1109/TTS.2020.2992669

Smith, B.C., 2019. The Promise of Artificial Intelligence: Reckoning and Judgment. The MIT Press.

Stahl, B.C., 2021. Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies. SpringerBriefs in Research and Innovation Governance. Springer International Publishing. https://doi.org/10.1007/978-3-030-69978-9

Stahl, B.C., Andreou, A., Brey, P., Hatzakis, T., Kirichenko, A., Macnish, K., Laulhé Shaelou, S., Patel, A., Ryan, M., Wright, D., 2021. Artificial Intelligence for Human Flourishing—Beyond Principles for Machine Learning. J. Bus. Res. 124, 374–388. https://doi.org/10.1016/j.jbusres.2020.11.030

Stahl, B.C., Timmermans, J., Mittelstadt, B.D., 2016. The Ethics of Computing: A Survey of the Computing-Oriented Literature. ACM Comput Surv 48, 55:1–55:38. https://doi.org/10.1145/2871196

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., 2016. Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel. Stanf. Univ. Stanf. CA Httpai100 Stanf. Edu2016-Rep. Accessed Sept. 6, 2016.

Strategy for Research in Artificial Intelligence Launched, 2021. AAH.JZR. https://jazairhope.org/en/strategy-for-research-in-artificial-intelligence-launched/ (accessed 10.11.21).

The 2015 Study Panel, 2016. Artificial Intelligence and Life in 2030 (One Hundred Year Study on Artificial Intelligence). Stanford University, Stanford.

The European Commission, 2020. On Artificial Intelligence—A European Approach to Excellence and Trust.

The Executive Office of the President of the United States, 2020. Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government.

The Federal Government, 2018. Artificial Intelligence Strategy. The Federal Government.

Thierer, A., Castillo, O., Russell, R., 2017. Artificial Intelligence and Public Policy. George Mason University, Arligton.

UK Government, 2021. National AI Strategy.

UKRI, 2021. Transforming Our World with AI—UKRI's Role in Embracing the Opportunity. UK Research and Innovation, Swindon.

Ulnicane, I., Knight, W., Leach, T., Stahl, B.C., Wanjiku, W.-G., 2020. Framing Governance for a Contested Emerging Technology: Insights from AI Policy. Policy Soc. 40, 158–177. https://doi.org/10.1080/14494035.2020.1855800

UNESCO, 2020. First Draft of the Recommendation on the Ethics of Artificial Intelligence (No. SHS/BIO/AHEG-AI/2020/4 REV.2). UNESCO, Paris.

UNI Global Union, 2017. Top Ten Principles for Ethical AI. UNI Global Union, Nyon Switzerland.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M., Fuso Nerini, F., 2020. The Role of Artificial Intelligence in Achieving the Sustainable Development Goals. Nat. Commun. 11, 233. https://doi.org/10.1038/s41467-019-14108-y

Wareham, J., Fox, P.B., Cano Giner, J.L., 2014. Technology Ecosystem Governance. Organ. Sci. 25, 1195–1215.

Willcocks, L., 2020. Robo-Apocalypse Cancelled? Reframing the Automation and Future of Work Debate. J. Inf. Technol. 35, 286–302. https://doi.org/10.1177/0268396220925830

World Economic Forum, 2018. Artificial Intelligence for the Common Good. Sustainable, Inclusive and Trustworthy. World Economic Forum.

Yeung, K., 2018. Algorithmic Regulation: A Critical Interrogation. Regul. Gov. 12, 505–523.

Zuboff, P.S., 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, 1st edition. Profile Books.

Towards Shaping the Future of Responsible AI in Africa

Damian Okaibedi Eke, Schmidt Shilukobo Chintu, and Kutoma Wakunuma

Introduction

The accelerated acceptance and deployment of digital technologies (especially Artificial Intelligence-AI) occasioned by the COVID-19 pandemic (Chang 2020; Kumar et al. 2021) is a pointer to the role technology has to play in our societies today. Public and private entities significantly increased the use of or employed a number of AI tools, digital platforms, big data and robotics as public service delivery tools, education platforms or work-based solutions during the global crisis. This clearly demonstrates that AI and other disruptive technologies are fast becoming critical foundations that enable human flourishing. AI is a major element

D. O. Eke (✉) · K. Wakunuma
De Montfort University, Leicester, UK
e-mail: damian.eke@dmu.ac.uk

K. Wakunuma
e-mail: kutoma@dmu.ac.uk

S. S. Chintu
University of Greenwich, London, UK
e-mail: schmidt.chintu@cyberacademy.ac.zm

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_8

and driver of what many have described as the fourth industrial revolution which is the functional convergence of AI, robotics, the internet of Things (IoT), 3D printing, genomics, quantum computing, blockchain and other disruptive technologies. Conceptualised by Schwab (2016), the idea of Fourth Industrial Revolution builds on the preceding industrial revolutions characterised by water and steam power, electric power and electronics and information technology. Schwab's argument was rooted in the fact that the convergence of disruptive technologies is blurring lines between the physical, digital and biological spheres as well as transforming every industry through unprecedented velocity, scope and systems impact. And at the heart of this epochal revolution is AI—a technology increasingly misconceived, misunderstood and mischaracterised. The focus on AI in this chapter is informed by the desire not to confuse its impacts with other emerging technologies.

AI offers exciting possibilities for African societies, promising among other things to alleviate poverty, reduce economic inequalities and improve access to public and private services in health, transportation and education. Whereas there is still a huge gap between Africa and other developed parts of the world in terms of capacity to develop datasets, AI systems development and deployment, AI is gaining traction in many aspects of life in Africa. One can say that, as many parts of Africa are yet to benefit fully from the first three industrial revolutions (electricity, mechanisation of production and automation of industries), a critical element of the Fourth Industrial Revolution (4IR), AI has made inroads into African socio-economic fabric. This raises the question of whether Africa should focus more on AI capacity building rather than embracing transferred technologies from the Global North in order to ensure that the full benefits of the 4IR are obtained for all sections of the society. On the other hand, AI not only creates unprecedented opportunities with a real bearing on people's lives, it raises fundamental questions on fairness, human rights, privacy, bias, security and the future of work among others. This chapter provides insights and perspectives on the AI landscape in Africa and a snapshot of how the future of AI development and deployment should look like. It explores the question; how can stakeholders in Africa ensure that Africa has sufficient capacity for Responsible AI? In this chapter, we present views on how future discussions of Responsible AI in Africa can be shaped. It starts with the presentation of the current landscape of AI deployment in Africa, highlighting potential socio-cultural impacts, ethical and legal impact of AI considering the unique cultural

dimensions in Africa and providing recommendations of how Africa can achieve Responsible AI. We conclude with mapping the roles Africa can play in the global discourse on Responsible AI. This contributes to the emerging debate on AI ethics, regulation, policy and governance in Africa.

The Meaning of AI

The meaning of AI is often contested across different disciplines. Whereas it was established in the 1950s as a field of study, narratives of intelligent machines have had a very long history (Cave et al. 2018). However, since Turing (1950) posed the question, "*Can machines think?*" AI has grown significantly as a major branch of computer science concerned with the design and application of machines capable of performing tasks that normally require human intelligence. Alan Turing's paper on *Computing Machinery and Intelligence* can be said to have laid the fundamental goals of AI and in 1956 John McCarthy organised a conference in Dartmouth where the term artificial intelligence was first adopted. McCarthy and Marvin Minsky, acknowledged as two of the pioneers of the field, subsequently co-founded the Artificial Intelligence Project (now the MIT AI Lab) in 1959 to explore the potential of AI. The term has since been defined differently in the research and innovation ecosystem. Minsky (1968, p. v) defined AI as "the science of making machines do things that would require intelligence if done by men". This definition reflected an earlier definition of AI proffered by McCarthy et al. (1955) as the problem of "making a machine behave in ways that would be called intelligent if a human were so behaving". In their work, *Artificial Intelligence: A Modern Approach*, Russell and Norvig (2002, p. viii) in answering Turing's question described AI as "the study of agents that receive precepts from the environment and perform actions". They went further to expose four different approaches that have characterised the history of AI including agents that think humanly, think rationally, act humanly and act rationally.

However, following an understanding of human intelligence as a "product of many factors and subject to innumerable influences" (Wechsler 1975) or as Gardner (2000, pp. 33–34) defined it, the "biopsychological potential to process information…to solve problems or create products that are of value in a culture", there is a question of whether a machine can achieve full human intelligence. Can a computer/machine acquire full socio-cultural, psychological intelligence like humans do?

These questions are at the heart of the categorisation of AI. AI applications that can only perform specific tasks are generally known as having artificial narrow intelligence (ANI) or weak AI (Shane 2019). On the other hand, some have described AI systems—artificial general intelligence (AGI) that can be able to "reason, plan, and solve problems autonomously for tasks they were never even designed for" (Kaplan and Haenlein 2019). Or as Searle (1980) described; designed in a way that the "computer is a mind, in the sense that computers can be literally said to understand and have other cognitive states". Kaplan and Haenlein (2019) also raise the possibility of a third category of AI called artificial super intelligence (ASI) "which are truly self-aware and conscious systems that, in a certain way, will make humans redundant".

For the purposes of this paper, the use of the concept of AI remains at the level of ANI which we define here as systems or applications that have the ability to interpret and learn from data for the performance of identified tasks in an agile way. Big data, machine and deep learning are critical drivers of AI which are increasingly applied in many aspects of our lives. Whereas AGI and ASI remain only possibilities and not currently available, access to ANI applications is becoming ubiquitous and pervasive even in many parts of the world including the developing economies of Africa. This conceptual clarification of AI is critically important to the present discourse because a mischaracterisation of what current AI applications can do, can affect its design, implementation, regulation and overall governance. Responsible AI governance should be rooted in a clear understanding of the nature, scope and potentials of AI. There is no attempt to engage with the conceptual tensions surrounding AI in this chapter but it was necessary to provide a clear view of what we mean by AI.

The Current Landscape of AI Deployment in Africa

There is an increasing level of AI (as defined in this chapter) deployment in Africa. Sectors where AI has been employed include; healthcare, education, transportation, financial services, agriculture, public services, security, business management and telecommunications. This chapter highlights some of the AI systems deployed in these fields. It is important to note that many of these systems are developed by local experts but most influenced, sponsored or controlled by big tech companies from the Global North.

Healthcare

AI is being deployed for a number of services in healthcare in Africa. For instance, MinoHealth AI Labs[1] in Ghana is using AI for automated diagnostics, forecasts and prognostics. BareApp[2] (also developed in Ghana) combines AI and skin expertise to diagnose black skin diseases and make recommendations for best treatment. In Nigeria, RxAll[3] is enabling pharmacies and patients to avoid buying counterfeit medicines online through a Deep Learning-Hyperspectral IoT platform for authenticating drugs in real time. Vinsighte[4] has developed an AI-powered system that detects eye diseases at an early stage and aids the visually impaired to read books and navigate their environments independently. InStrat[5] (also developed in Nigeria) uses AI to detect and predict possible disease outbreaks by electronically collecting and analysing clinical and non-clinical data. In Uganda, Chil AI Lab Group[6] has developed a system that combines AI and other emerging technologies to enhance management of female chronic diseases. The overall aim, however, is to provide accessible and affordable chronic disease prevention and management to women. AI is also being deployed for non-clinical healthcare purposes such as insurance. Deployed in countries such as Nigeria, Ghana and South Africa, Curacel[7] uses AI to optimise health insurance claims.

Financial Services

Financial services sector is another field where AI is being deployed in Africa. In Kenya, M-Shwari[8] has developed a system that relies on AI to review online loan applications, helping it to consider applications from

[1] https://www.minohealth.org.

[2] https://bareapp.ai/.

[3] https://www.rxall.net/.

[4] https://www.vinsighte.com.ng/index.php.

[5] http://instratghs.com/our-services/.

[6] https://chilailabgroup.com/.

[7] https://curacel.co/health-insurance/.

[8] https://www.safaricom.co.ke/personal/m-pesa/credit-and-savings/m-shwari.

customers who live far from bank branches. Similarly, In Egypt, Cassbana[9] relies on AI to create digital identities for underserved communities so as to integrate them into the banking system. It also manages their financial requests and builds a behaviour-based scoring system for them. In Nigeria, Debtors Africa[10] and Kudi[11] are two of the AI applications deployed in the financial sector. Debtors Africa uses AI to automatically update its independent, searchable database for recalcitrant and delinquent debtors, providing status of debtors in real time. Kudi on the other hand is a chatbot that responds to financial requests and allows users to send money and pay their bills.

Security Services

Private and public security services are also deploying AI systems. To provide a solution to the well documented risks of misidentification, unfair discrimination and bias against black people in the use of biometric surveillance technology, a Ghanaian company BACE Group[12] has developed BACE APIs that enables accurate identification of black people through facial recognition technology. This company provides secure identity verification as a security service to both the public and the private sectors. Similarly, in Kenya, the government has deployed AI-powered facial recognition technologies to complement policing efforts as part of the Safe city Project.[13] This technology was developed by the Chinese tech giant Huawei and concerns have been expressed regarding the dearth of regulatory provisions that can ensure the responsible use of this technology considering the well documented legal, ethical and socio-cultural concerns related to facial recognition (Feldstein 2019). This is because there have been reports of a similar technology being purchased by the Uganda government from Huawei to spy on political opponents (ibid.). Tabiri Analytics[14] (which is deployed in Rwanda, Kenya and Uganda) has

[9] https://www.cassbana.com/#how-it-works.

[10] https://www.debtorsafrica.com/Home/AboutUs.

[11] https://kudi.com/about-us/.

[12] https://www.bacegroup.com/.

[13] https://africatimes.com/2019/12/18/huaweis-surveillance-tech-in-kenya-a-safe-bet/.

[14] https://tabirianalytics.com/about/.

also been used to provide continuous monitoring service to prevent cyber threats. This system uses cloud computing, machine learning and AI to automate human analysis of IT system log data to achieve cybersecurity. Global Auto Systems[15] is also using data and system analytics to provide security systems for schools, colleges and universities in Uganda.

Education

In Uganda, M-Shule[16] has developed and deployed a Toolkit that helps learners build academic and life skills with interactive, self-paced and personalised learning over SMS, measure progress and performance and also keep stakeholders up-to-date with awareness campaigns and situational response information. It also helps to collect data and insights from stakeholders via SMS to make real-time decisions. In Kenya, Eneza Education[17] is providing primary and secondary school students with virtual tutorials on curriculum-aligned content in all subjects while in Nigeria Tuteria[18] uses AI to link qualified tutors to students within a particular area and budget. It also verifies tutors IDs, conducts background checks and evaluates tutors' performance. In South Africa, Botlhale AI[19] solutions specialises in conversational AI. With a suite of Natural Language processing tools, this company ensures that those who speak African languages do not miss out on the benefits of technologies.

Transportation and Logistics

AI is also changing the transport sector in some African countries. In Egypt, SWVL[20] and Softech Technologies[21] have deployed AI applications making huge impacts. Swvl uses AI to coordinate a fleet of private buses, allowing commuters to bypass often congested public transit networks while Softech uses AI to help commuters to plan their itinerary

[15] https://globalautosystems.co.ug.

[16] https://m-shule.com/index.html#toolkit.

[17] https://enezaeducation.com.

[18] https://www.tuteria.com/why-use-tuteria/.

[19] https://botlhale.ai/about-us/.

[20] https://www.swvl.com/.

[21] https://www.softec.ai/about.

by collecting and analysing data on transportation conditions. Softech provides both an AI-based digital planner and solutions for commercial fleets providing end-to-end visibility and command and control over B2B logistics, transportation and mobility operations. In Kenya, Amitruck[22] is being used to create a digital marketplace for trucking; connecting transporters and clients on a digital platform while Kamtar is used in Ivory Coast to connect shippers and carriers. Kobo360[23] is also using big data analytics and technology to reduce logistics frictions in Nigeria. The overall goals are to ensure efficiency and cost reduction in the supply chain.

Telecommunication

In Telecommunication, MTN[24] (a mobile Telecom operator which is in operation in Benin, Cameroon, Ghana, Guinea Bissau, Ivory Coast, Nigeria, Rwanda, South Africa Uganda and Zambia) has launched a chatbot that simplifies and enhances the quality of customers' experience. Similarly, Ooredoo[25] partners with PI works (an AI company) to enhance customer experience, network coverage and connectivity in the MENA region particularly Tunisia and Algeria. Safaricom[26] has also introduced its AI Chatbot assistant to popular messaging service WhatsApp to perform telecom-related tasks as well as answer queries regarding M-PESA (mobile money service) in Kenya.

Public Service Delivery

The most prominent use of AI for public service delivery can be found in Rwanda where robots[27] donated by the United Nations Development

[22] https://www.amitruck.com/our-solution.

[23] https://kobo360.com/NG/en/.

[24] https://www.mtnonline.com/about-us/how-big-data-artificial-intelligencell-rule-your-future/.

[25] https://www.ooredoo.qa/web/en/.

[26] https://www.safaricom.co.ke/about/media-center/publications/press-releases/release/893.

[27] https://edition.cnn.com/2020/05/25/africa/rwanda-coronavirus-robots/index.html.

Program (UNDP) were introduced to help in the fight against COVID-19 pandemic as part of public health intervention. These robots were used for a number of tasks including temperature screening, detecting people not wearing masks in healthcare settings and delivering medicine, food and other essentials in place of frontline health workers. In South Africa AI is being used to detect gunshots. Shotspotter[28] is used to fight wildlife poaching in Kruger National Park.

Politics

There is also documented evidence to indicate that AI tools have been employed in the African political landscape (e.g. in Kenya and Nigeria) as a tool for mis/disinformation. However, AI tools like the one developed by African check,[29] a South African organisation using AI technology to fact check political claims, are helping to fight misinformation. In the same vein, the protests that followed the #ENDSARS hashtag[30] against the Special Anti-Robbery Squad (SARS) in Nigeria in 2020 demonstrated the potential impact of AI systems in changing political narratives for the citizens. AI-powered digital tools are increasingly changing the way politics and public sector decisions are and while providing citizens with information on their rights.

Agriculture

AI Mozambique, Hello Tractor[31] is helping farmers share equipment. It also leverages machine learning to predict crop yields and facilitates access to financing for farmers. Other AI tools used in Agriculture include Agrix Tech[32] (in Cameroon) that helps farmers detect crop diseases and propose sustainable and environmentally friendly solutions to small-scale farmers

[28] https://phys.org/news/2018-09-south-africa-tech-vicious-gun.html.

[29] https://africacheck.org/.

[30] https://globalvoices.org/2020/10/14/lazy-nigerian-youth-mobilize-endsars-protest-from-social-media-to-the-streets/.

[31] https://hellotractor.com/.

[32] https://digestafrica.com/companies/agrix-tech.

and PlantVillage Nuru[33] in Kenya that serves as a crop disease diagnosis tool. In Egypt, AbuErdan[34] uses deep learning neural networks and predictive analytics algorithms to forecast chicken future performance. Apollo Agriculture[35] in Kenya uses satellite data and machine learning to advise farmers on credit decisions and automated operations to keep costs low and processes scalable. Similarly, Aerobotics[36] is being deployed in South Africa to assist farmers/growers to make informed decisions. This application uses AI to collect and analyse data on crop yields and subsequently predict future performances.

The Future of AI in Africa

The current landscape of AI in Africa is dotted with the presence of big tech companies from the Global North including Google, Facebook, Alibaba Group, Amazon, Microsoft and IBM Research. These companies bring improved capacity and enabling infrastructure to make data more pervasive and valuable which can be leveraged by AI and other emerging technologies to drive large-scale transformation in Africa and make the continent more competitive. This will require increased sharing, interoperability of data processing systems and significant convergence of emerging technologies. AI presents good opportunities for many sectors to optimise solutions to Africa's problems. The future of many public and private sectors in Africa will be intricately linked with the future of AI. In healthcare, AI can provide solutions to many challenges including in medical diagnostics, drug research and discovery, clinical trials, disease management, pharmacogenomics, improvement of patient outcomes, data management and clinical decision support tools. In banking and other financial services, AI can improve risk detection and management in addition to data management. There are also prospects of using AI as a new tool for counterterrorism in Africa. McKendrick (2019) and Ramanouski (2019) have described how AI can theoretically contribute to counterterrorism operations. Therefore, with the continued activities of terrorist groups such as Boko Haram and al-Shabaab in Africa,

[33] https://www.biorxiv.org/content/10.1101/2020.01.26.919449v1.ful.

[34] https://abuerdan.com/contact-us/.

[35] https://www.apolloagriculture.com.

[36] https://www.aerobotics.com.

AI can provide a potent tool for counterterrorism. Although a joint report by UNICRI and UNCCT have also detailed the possible malicious use of AI for terrorist purposes (United Nations 2021).

Furthermore, AI holds the promise of enabling a revolution in how agriculture is done in Africa. From crop and soil monitoring, improved plant and crop disease diagnosis, crop yield prediction and price forecasts, intelligent spraying, pest control, drought prediction, to agriculture robots and genomic precision, AI has the potential to transform the Agriculture industry in Africa. In addition, Africa's educational system has many challenges in personnel and facilities that AI can address in future. This includes among other things AI-powered virtual teaching assistants that can help both teachers and students in assessments and providing feedback. Considering the lack of physical infrastructure and insufficiency of training for teachers in Africa, the potential impact of AI on learning may be extensive. There is sufficient evidence to show that there will also be possible increases in the use of AI in politics, public service delivery, transportation (e.g. driverless cars) and the military in Africa in the next two decades as they are currently used in the Global North.

However, despite the potential benefits of this technology, the design and implementation of AI systems raise significant ethical, legal and socio-cultural challenges. There is a growing body of the literature to highlight that AI design and implementation are not only changing socio-cultural dynamics but also exacerbating existing societal inequalities, biases and stereotypes (Nelson 2019; Weber 2019). Many AI systems ranging from applications for predictive policing (McDaniel and Pease 2021), to facial recognition technologies (Raji et al. 2020) have shown to be inherently biased and discriminate against certain sections of the society. These biases can creep into AI through the underlying datasets or algorithms. Both the data and algorithms can include biased human decisions or reflect historical inequalities bordering on gender, race, social status and geographical location. Most importantly, black people have been shown to be disproportionately affected by unfair bias in current AI systems which puts Africa at a disadvantage. The impact of such biases inflicts hurt on those who are discriminated against. It brings mistrust and possible unacceptance which reduces its potential benefits to businesses and the society at large.

AI also raises challenges for human rights in a number of ways. From the creation of autonomous and intelligent agents like driverless cars, neurotechnologies that could clearly disrupt people's sense of identity

and agency (Yuste et al. 2017), to the possibility of digital authoritarianism, AI challenges established perceptions of human rights which need attention. The possibility of deploying AI-based surveillance technologies by governments is a challenge that requires attention since human rights reports in Africa do not look good. However, the above concerns related to unfair biases and discrimination and human rights together with the high energy requirements for AI systems contribute to conclusions that AI can inhibit some of the UN sustainable development goals (SDG) targets (Gupta et al. 2021; Vinuesa et al. 2020). These and many more negative impacts and unintended consequences of AI applications call for a design and implementation of this technology in a way that is ethically responsible, legally compliant and socio-culturally acceptable (Wakunuma et al., 2022).This is the concept of Responsible AI. Realising the full capacity of AI for human flourishing depends on Responsible AI and therefore should be a major agenda for AI discourse in Africa.

What Can Africa Do to Achieve Responsible AI?

AI and other emerging technologies are characterised by features such as logical malleability, ubiquity, pervasiveness, interactivity, possibility of augmentation and potentially autonomy. 'Logical malleability' (Moor 1985) makes it difficult to predict how AI systems can be used or others discussed under the concept of interpretive flexibility (Doherty et al. 2006). Its ubiquitous, pervasive and interactive nature continues to be more pronounced as new use cases emerge. Together with the potential to achieve autonomy, these features mean that both the design and deployment require the consideration of ethical, legal and socio-cultural values and principles because of possible intended and unintended consequences. These features inform the many uses of AI including for improving processes and efficiency, social control and to promote human flourishing (Stahl, 2021). However, for AI to promote human flourishing especially in Africa, the principle of responsibility needs to be integrated into its design.

Responsible AI is about how AI can be sensitive to human values (which is shaped by cultural beliefs and systems) and to societal needs, expectations, hopes and fears. As Dignum (2017) opined, Responsible AI rests on three pillars; the willingness of stakeholders to accept responsibility for the impact of AI, the development of mechanisms that can

enable AI systems to be sensitive to ethics and human values and appreciation of different impacts of AI in different cultures. These pillars shape initiatives in AI education, governance, regulation, risk assessment approaches and quality assurance. For instance, the promotion of AI governance comes from the willingness by public and private entities to address the impact of AI. However, the question for Africa is; do we have the necessary technical and socio-economic infrastructure to facilitate these? All available indexes point to the fact that Africa lags behind in the comparative global AI readiness. Recent reports on the current state of the art on AI in Sub-Saharan Africa confirms the evident lack of AI capacity in the AI ecosystem and suggest the need for greater capacity (Butcher et al. 2021; Gwagwa et al. 2021). One thing that is clear is that the role of ensuring that Responsible AI is achieved in Africa is firstly ours before any other person else. We are mainly responsible for ensuring that AI designed and deployed in Africa is sensitive to our socio-cultural contexts but not only ours.

The previous section has mapped the future of AI applications in Africa and their potentially historical impacts, we will now provide perspectives on how Africa can not only increase AI capacity but how to achieve Responsible AI considering the continued global discourse on AI ethics.

Framing the Role of AI for Africa

Africa is historically credited with contributing to the industrial revolutions of the eighteenth century. It was Karl Marx who wrote that the "turning of Africa into a warren for the commercial hunting of blackskins" contributed to the "rosy dawn of the era of capitalist production". Eric Williams (1944) also echoed the role slavery had on capitalism in his work titled *Capitalism and Slavery*. According to Parvanova (2017) this industrial revolution became the force behind colonialism because it created the need for Europe to expand; increased production capacity required more raw materials to satisfy demands. However, a number of factors including colonialism impeded the spread of industrial revolutions in the dependent countries of Africa. Alam (2012) has provided empirical evidence to demonstrate the impact of colonialism on industrial Revolutions in Africa. Unfortunately, since the end of colonialism, Africa has not caught up with subsequent industrial revolutions characterised by digital technology and the results are in the social, economic and digital inequalities between the Global North and the Global South. The first three

industrial revolutions have shaped the societies we live in today and it is safe to say that Africa has not benefited fully from them. About 40% of Africans still do not have access to electricity and fully automated productions are yet to be achieved. And with the 3rd industrial revolution came the challenge of digital divide. According to a 2021 report by the Ibrahim Forum, "89% of learners in sub-Saharan Africa do not have access to household computers. 82% lack internet access and at least 20 million live in areas not covered by a mobile network" (Mo Ibrahim Foundation 2021).

With AI driving the 4IR, there is a potential to further divide humans on a class level and Africa may be disconnected or will not receive the same level of benefit from AI systems. Some have also pointed out that AI raises the risk of neo-colonialism with regard to data and the algorithms that shape AI. AI is as good as the datasets and the algorithm that shaped it. That means that there is inherent power and control in datasets. Allowing the Global North to own and control the datasets that shape AI systems developed for Africa amounts to what many have described as 'data colonialism' (Couldry and Mejias 2019; Viera Magalhães and Couldry 2021).

Beyond data there are also the algorithms that are not neutral (Mittelstadt et al. 2016; Warfield 2020; Stinson 2021). According to Mittelstadt et al. (2016), ethical issues related to algorithms include both epistemic and normative concerns; possibilities of unjustified actions, opacity, bias, discrimination, challenges to autonomy and informational privacy. Birhane (2020) has also written about the possibility of 'algorithmic colonization' to show that algorithms contain the biased interests and values of those who develop them which are often overshadowed by the hype around AI. The argument here is that Africa needs to move beyond the hyperbolic language surrounding AI to understand the true facts about datasets and algorithms and their inherent power in AI. Those who have the data and the algorithms will hold great power and influence; great power to improve processes and procedures, for social control and to acquire better human flourishing but in a way to favour them. An example is the development of COVID-19 vaccines. Despite the potential benefits of equitable access to vaccines, the greatest barriers to adequate vaccine supply remains intellectual property (IP) protection governing the production and access to vaccines (Erfani et al. 2021). Those who produced the vaccines and own the IP remain in the position of power and lack the political and moral will to waive their IP rights to facilitate

equitable access to the vaccines. With these in mind, Africa needs to be deliberate in framing the role AI can play for the continent and in the development and definition of our data and our algorithm. Responsible AI in Africa means AI developed with African data and culturally sensitive algorithms. Policymakers in public and private sectors, researchers, industry players and all stakeholders need to decide on what kind of AI-driven society we want. The role of AI in Africa should be to provide fundamental solutions that can level up evident inequalities in production, in healthcare, education, gender and other spheres of life. AI in Africa needs to amplify our positive cultural contexts. Over reliance on importing AI systems driven by foreign values and principles can only exacerbate the risk of neo-colonialism. Africa's attitude towards the 4IR, especially AI does not need to mirror how the other industrial revolutions were handled.

Another thing to say here is that due to the pervasiveness of AI technology and the interconnectedness of the African societies, it will take the whole continent to frame the goal of AI around improving and promoting the unique cultural contexts Africa possesses. Africa's framing of the goals of AI needs to therefore rely on the fundamental cultural narratives of Africans, consisting of our stories, our beliefs, values, needs, expectations, fears and concerns. Values and power are central in AI design and implementation, however, inherently-power-driven AI narratives from the Global North should not determine the paths we chart for our AI journey. Africa should proactively frame how AI should be developed and deployed in our communities (Eke and Ogoh, 2022). Lessons from the Global North should shape such a framing to centre on creating an ecosystem where AI can thrive and on embedding critical African values into AI systems to ensure that AI in Africa can truly conform to the principles of science for and with the society.

Identification of Relevant African Values and Principles to Be Embedded into AI Systems

The centrality of data in AI means that AI aligns with human preferences, interests and values. But the critical question is what and whose values should AI applications align with. At the foundation of global discussion on AI alignment are established ethical traditions (Yu et al. 2018) such as utilitarianism (Roff 2020), deontology (Hooker and Kim 2018) or

virtue ethics (Neubert and Montañez 2020). Each of these ethical traditions are without roots in Africa and emphasise the importance of AI to respect the objective interests of humanity or a particular group of people. Impliedly, ethically-aligned AI discourse has focused on the principles and values from cultural contexts from which these frameworks emerged—the Global North.

In his book, *Human Compatible: Artificial Intelligence and the Problem of Control*, Stuart Russell (2019), observed that aligning AI to human values is the crucial goal of AI value alignment. This is particularly important because of the critical potential characteristic of autonomy of AI systems. The values embedded into AI, therefore, are critical to its impact on society. As AI systems are developed for African societies, it is critical to ensure that the values embedded in these systems represent objective interests and beliefs in Africa. In essence, to design AI for Africans requires the positive action of integrating African values and principles in the design and implementation. The first part of this is to normatively understand what values and principles ought to be embedded in AI systems. Africa has rich moral traditions built around core values of interconnectedness, solidarity, communality and respect conceptualised in ethical frameworks such as *Ubuntu* (translated broadly as "I am because we are") and or *ujamma* (the spirit of brotherhood). These and other relevant African value-systems should form the central focus of AI value alignment in Africa. AI discourse in Africa ought to focus not only on what AI *might* do for Africa but also how AI *should* be done for the benefit of Africans.

For AI to be truly for society, there must be an understanding that the technical design needs to reflect societal values, needs and expectations. Unlike the usual tick-box exercise evident in most research and innovation processes, this is about a proactive and continuous consideration of the social and ethical consequences and a conscious integration of values in the design as well as the deployment of AI systems. As Stahl (2021) suggested, this demands a fundamental rethink of the relationship of AI research and innovation and ethics. African ethical principles should be an integral part of AI's scientific excellence in a way that promotes solutions tailored for African societies. Our values should be central to the intended consequences of AI in Africa; included in the risk assessments and form part of the evaluation of trustworthiness and responsibility of the AI systems. Interventions needed to address identified risks should reflect contextual African values. However, the identification and applications of

these values and principles require the collaboration and contributions of diverse stakeholders from all AI ecosystems (Stahl 2021).

Increased Involvement of Relevant Stakeholders

In his book, Stahl (2021) highlighted the idea of AI ecosystems which is already established in the European Commission's AI White paper (European Commission 2020), global recommendations from OECD (2019) and UNESCO (2020) and also in the UK's Digital Catapult (2020). Using ecosystems as a metaphor, Stahl conceptualised AI ecosystems (consisting of individuals, organisations, innovation systems and landscapes) as examples of innovation ecosystems characterised by complex relationships between different and interdependent actors, willing to co-evolve and mutually learn as they drive change with openness. As he pointed out, there are elements of AI that are global but identifiable regional differences in the USA, China and Europe suggest separate ecosystems distinguished by geography, jurisdictions and other elements or environments within which the AI system is embedded. These environments range from "technical, policy, economic, legal, social, ethical and other aspects that closely interact with AI" (Stahl 2021, p. 93) and produce diverse stakeholders that can influence how societal impacts of AI are perceived, identified and can be addressed. It is important for Africa to identify these AI ecosystems (at the broader continental and national levels; relevant disciplines and sectors) and the stakeholders therein for a process of co-creation of frameworks for AI. There are diverse cultural contexts (languages, values, belief systems, etc.) and interests (educational, political, economic, legal, etc.) in Africa which need to be represented in the discourse on both *what* AI *might* do and *how* it *should* be done. The recognition of these contexts is the foundation of Responsible AI (Table 1).

The Role of Africa in the Global AI Discourse

As we have highlighted in earlier sections, every region, nation and cultural community has a role to play in shaping the discussions around Responsible AI. The approach to Responsible AI should be bottom-up rather than top-down. A clear conceptualisation of contextual values, needs and interests should precede acceptable international frameworks

Table 1 Roadmap to achieving Responsible AI in Africa

Framing the role of AI in Africa	Considering African • Expectations • Interests • Fears/concerns • Hopes • Needs • Innovation • Acceptance • Ownership • Data • algorithm
Identification of relevant African values and principles to be embedded into AI systems	Considering African • Unique cultural belief-systems and traditions • Ethical frameworks (characterised by communitarian principles e.g. Ubuntu, ujamaa, humanism etc.) • Cultural practices and values
Involvement of relevant stakeholders	*From different*: Nations, regions, jurisdictions, fields of study, gender, socio-economic backgrounds, public and private sector, languages (Policy makers, industry, academia and citizens) For: • Co-creation • Engagement • Inclusion • Participation/buy-in • Building trust • Consultation • Education

for AI governance and principles. Therefore, Africa nations in particular and the region in general, like other countries and regions in the Global North, have key roles to play due to the level of distribution of power that comes from global AI governance. We have interests, needs, hopes, fears, principles and values that need to be factored in the global consideration Responsible AI. This starts with having clear normative and epistemic understandings of unique African perspectives that AI design and implementation should align with. Such understanding should shape Africa's roles in global discussions on AI ethics. Africa should not only be included in the global discourse on Responsible AI but should approach

the proverbial table with their interests, expectations as well as their values and moral principles such as the communitarian principles exemplified in concepts such as *ubuntu* and *ujamma*.

In addition to this, Africa offers the global AI design and implementation landscape opportunities of diversification and generalisability in terms of datasets, skillsets and personnel to achieve Responsible AI. Disproportionate amount of data collected from the Global North contribute to the persistent challenges of unfair bias and discrimination in AI. The data, values and humans who build and deploy AI should include representation from backgrounds in Africa. This focus on diversification can ensure better AI outcomes informed by reliable insights from sufficiently representative datasets. A recent report on AI in Sub-Saharan Africa observed that such a representation should be more than a box-ticking exercise and must be seen as a moral imperative for all stakeholders (Ndung'u and Signe 2020) Stakeholders ought to appreciate the intrinsic values of diverse interests, expectations and perspectives inherent in the data that inform AI. The inclusion of new voices, perspectives and datasets provide new opportunities for designing solutions for more people. For instance, facial recognition systems developed by big tech companies in North America, Asia and Europe predominantly misidentify people based on race and gender owing to both the data and algorithms that shape these systems (Buolamwini and Gebru 2018). In contrast to these systems, a Ghanaian tech start-up (led by Ivorian researcher Charlette N'Guessan) has developed a facial recognition system for the local market trained with more diverse and representative datasets that can accurately identify black faces. Since AI is still at its introductory stage in Africa, there are opportunities of creating systematic diversity mechanisms to reduce the discriminatory effects of AI in the society. Building on from the growing body of AI ethics literature and practice, Africa can build a template for a sustainable consideration of diversity and inclusion in AI design and implementation.

Conclusion

Shaping the future of Responsible AI in Africa is a pertinent concept particularly when we consider the importance of developing Africa's own contribution to the discourse of AI. For a long time, there have been discussions around AI including those around benefits but also around the ethical challenges associated with the technology. These have mainly

been led by the Global North with little contribution from the Global South. This chapter has sought to make a contribution to the discourse of AI from an African perspective by focussing on what AI means and how it may look like for the future when we consider both the benefits and challenges of AI. The chapter has shown that there are clear benefits of AI by looking at a plethora of AI application areas in Africa. In particular, AI is being applied in healthcare, in finance, security, education, transport and logistics, telecommunication, public service delivery, agriculture as well as being used in politics. This is an interesting array of application areas and showcases the fact that although there is a limited discourse of AI on a global level, Africa has taken to AI and continues to do so in abundance. This indicates that there is a need to understand how AI is being applied in an African context and subsequently a need to understand what value-systems are being applied or can be applied as AI becomes mainstream in the African context. By cultivating this understanding, we can then begin to explore the possibilities that lie ahead in future of AI in Africa and as a consequence what can be learnt and shared by Africa in the global discourse of AI. Currently, this chapter notes that AI applicability is synonymous with big tech companies from the Global North, therefore raising concerns around dependency in terms of technology know-how, capability, capacity as well as the inculcation of value-systems from the Global North to the Global South. This raises further questions around the potential and possibility of digital/neo-colonialism which can leave Africa grappling with the technology and not being able to understand fully or find solutions for challenges that result from AI as it is applied in different domains. Simply put, the needs of the Global North are different from those of the Global South, as such, it goes without saying that the application of AI on the African continent may be different in terms of the problems it intends to solve and subsequent benefits the technology will have. Similarly, the ethical and social challenges that may result will differ in a number of ways when compared to the Global North. As such, this calls for Responsible AI particularly in as far as understanding the value-systems and human values that may be applicable when it comes to AI on the African continent. Africa's challenges are vast and include but are not limited to hunger, poverty, education, health, climate, gender disparities and various inequalities, climate among others. The application and use of AI will go a long way in mitigating some of these challenges, however, this can only be done by incorporating the continent's values in the technology and not being overly reliant and dependent on those from

the Global North which have been embedded in AI due to technology's origins. Despite AI's origins, as its use spreads across the globe, there is room to tailor it to the locale of its adoption and use for it to be effective and truly meaningful. It is for this reason that in this chapter we allude to and recognise the importance of embedding African value-systems and principles through philosophies like Ubuntu, Ujamma and others in our quest of framing a truly Responsible AI for Africa. In this case, Responsible AI means to think about, anticipate, design, implement, adopt, adapt and use AI that connects, is communal, respects and works in solidarity with different stakeholders for the common good and meeting head-on the challenges that Africa faces.

Acknowledgements The Work of DE is supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreements No. 945539 (HBP SGA3).

References

Alam, M.S., 2012. Colonialism and industrialization: Empirical results. Available SSRN 2031131.

Birhane, A., 2020. Algorithmic colonization of Africa. SCRIPTed 17, 389.

Buolamwini, J., Gebru, T., 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification, in: Proceedings of Machine Learning Research. Presented at the Conference on Fairness, Accountability, and Transparency, pp. 1–15.

Butcher, N., Wilson-Strydom, M., Baijnath, M., 2021. Artificial intelligence capacity in sub-Saharan Africa: Compendium report.

Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., Taylor, L., 2018. Portrayals and perceptions of AI and why they matter.

Chang, A.C., 2020. Artificial intelligence and COVID-19: Present state and future vision. Intell.-Based Med. 3–4, 100012. https://doi.org/10.1016/j.ibmed.2020.100012

Couldry, N., Mejias, U.A., 2019. Data colonialism: Rethinking big data's relation to the contemporary subject. Telev. New Media 20, 336–349.

Digital Catapault, 2020. Lessons in practical AI ethics: Taking the UK's AI ecosystem from "what" to "how." Digital Catapault, London.

Dignum, V., 2017. Responsible artificial intelligence: Designing AI for human values.

Doherty, N.F., Coombs, C.R., Loan-Clarke, J., 2006. A re-conceptualization of the interpretive flexibility of information technologies: Redressing the balance between the social and the technical. Eur. J. Inf. Syst. 15, 569–582.

Eke, D., Ogoh, G., 2022. Forgotten African AI narratives and the future of AI in Africa. Int. Rev. Inf. Ethics 31(1), [Online] Available from https://doi.org/10.29173/irie482 [Accessed 14/09/2022].

Erfani, P., Binagwaho, A., Jalloh, M.J., Yunus, M., Farmer, P., Kerry, V., 2021. Intellectual property waiver for covid-19 vaccines will advance global health equity. BMJ 374. https://doi.org/10.1136/bmj.n1837

European Commission, 2020. Report on the safety and liability implications of artificial intelligence, the internet of things and robotics. European Commission, Brussels.

Feldstein, S., 2019. The global expansion of AI surveillance. Carnegie Endowment for International Peace, Washington, DC.

Gardner, H.E., 2000. Intelligence reframed: Multiple intelligences for the 21st century. Hachette UK.

Gupta, S., Langhans, S.D., Domisch, S., Fuso-Nerini, F., Felländer, A., Battaglini, M., Tegmark, M., Vinuesa, R., 2021. Assessing whether artificial intelligence is an enabler or an inhibitor of sustainability at indicator level. Transp. Eng. 4, 100064. https://doi.org/10.1016/j.treng.2021.100064

Gwagwa, A., Kachidza, P., Siminyu, K., Smith, M., 2021. Responsible artificial intelligence in Sub-Saharan Africa: Landscape and general state of play.

Hooker, J.N., Kim, T.W.N., 2018. Toward non-intuition-based machine and artificial intelligence ethics: A deontological approach based on modal logic, in: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, pp. 130–136.

Kaplan, A., Haenlein, M., 2019. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus. Horiz. 62, 15–25.

Kumar, A., Pujari, P., Gupta, N., 2021. Artificial intelligence: Technology 4.0 as a solution for healthcare workers during COVID-19 pandemic. Acta Univ. Bohem. Merid. 24, 19–35.

McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E., 1955. A proposal for the Dartmouth summer research project on artificial intelligence.

McDaniel, J.L., Pease, K., 2021. Predictive policing and artificial intelligence. Routledge, Abingdon.

McKendrick, K., 2019. Artificial intelligence prediction and counterterrorism. R. Inst. Int. Aff.

Minsky, M., 1968. Semantic information processing. The MIT Press, Cambridge, MA.

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L., 2016. The ethics of algorithms: Mapping the debate. Big Data Soc. 3, 2053951716679679. https://doi.org/10.1177/2053951716679679

Mo Ibrahim Foundation, 2021. COVID-19 in Africa—One year on: Impact and prospects. Mo Ibrahim Foundation Forum.

Moor, J.H., 1985. What is computer ethics? Metaphilosophy 16, 266–275.

Ndung'u, N., Signe, L., 2020. The Fourth Industrial Revolution and digitization will transform Africa into a global powerhouse. Foresight Afr. Rep.

Nelson, G.S., 2019. Bias in artificial intelligence. N. C. Med. J. 80, 220–222. https://doi.org/10.18043/ncm.80.4.220

Neubert, M.J., Montañez, G.D., 2020. Virtue as a framework for the design and use of artificial intelligence. Bus. Horiz. 63, 195–204.

OECD, 2019. Recommendation of the council on artificial intelligence. OECD/LEGAL/0449.

Parvanova, D., 2017. The industrial revolution was the force behind the new imperialism. ESSAI 15, 30.

Raji, I.D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., Denton, E., 2020. Saving face: Investigating the ethical concerns of facial recognition auditing, in: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, pp. 145–151.

Ramanouski, V., 2019. Possible use of AI technologies in counterterrorism responses by Iraqi security establishment, in: Proceedings of the European Conference on the Impact of AI and Robotics, EM-Normandie Business School, Oxford, UK, pp. 261–265.

Roff, H.M., 2020. Expected utilitarianism. ArXiv200807321 Cs.

Russell, S., 2019. Human compatible: Artificial intelligence and the problem of control. Penguin.

Russell, S., Norvig, P., 2002. Artificial intelligence: A modern approach.

Schwab, K., 2016. The Fourth Industrial Revolution: What it means, how to respond. World Econ. Forum. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Searle, J.R., 1980. Minds, brains, and programs. Behav. Brain Sci. 3, 417–424.

Shane, J., 2019. You look like a thing and I love you. Hachette UK.

Stahl, B.C., 2021. Artificial intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies. Springer Nature.

Stinson, C., 2021. Algorithms are not neutral: Bias in collaborative filtering. ArXiv Prepr. ArXiv210501031.

Turing, A.M., 1950. Computing machinery and intelligence. Mind 49, 433–460.

UNESCO, 2020. First draft text of the recommendation on the ethics of artificial intelligence. Ad hoc expert group (AHEG) for the preparation of a draft text. UNESCO, Paris.

United Nations, 2021. Algorithms and terrorism: The malicious use of artificial intelligence for terrorist purposes. United Nations Office of Counter-Terrorism and United Nations Interregional Crime and Justice Research Institute, New York and Torino.

Viera Magalhães, J., Couldry, N., 2021. Giving by taking away: Big tech, data colonialism and the reconfiguration of social good. Int. J. Commun. 15, 343–362.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M., Fuso Nerini, F., 2020. The role of artificial intelligence in achieving the Sustainable Development Goals. Nat. Commun. 11, 233. https://doi.org/10.1038/s41467-019-14108-y

Wakunuma, K., et al. 2022. Responsible AI, SDGs, and AI Governance in Africa, in: 2022 IST-Africa Conference (IST-Africa). IEEE, pp. 1–13. https://doi.org/10.23919/IST-Africa56635.2022.9845598

Warfield, B.W.A., 2020. Not neutral: A look at bias in artificial intelligence algorithms that are assumed to be objective (PhD Thesis).

Weber, C., 2019. Engineering bias in AI. IEEE Pulse 10, 15–17.

Wechsler, D., 1975. Intelligence defined and undefined: A relativistic appraisal. Am. Psychol. 30, 135.

Williams, E., 1944. Capitalism and slavery. UNC Press Books.

Yu, H., Shen, Z., Miao, C., Leung, C., Lesser, V.R., Yang, Q., 2018. Building ethics into artificial intelligence. ArXiv Prepr. ArXiv181202953.

Yuste, R., Goering, S., Arcas, B.A. y, Bi, G., Carmena, J.M., Carter, A., Fins, J.J., Friesen, P., Gallant, J., Huggins, J.E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., Sullivan, L.S., Teicher, M., Wasserman, D., Wexler, A., Whittaker, M., Wolpaw, J., 2017. Four ethical priorities for neurotechnologies and AI. Nature 551, 159–163. https://doi.org/10.1038/551159a

Responsible Artificial Intelligence: Recommendations and Lessons Learned

Virginia Dignum

Introduction

Nowadays, Artificial Intelligence (AI) is almost ubiquitous. We can hardly open a newspaper or tune in to a news show without getting some story about AI. But AI means different things to different people.

As AI is increasingly impacting many aspects of life, the awareness that it has the potential to impact our lives and our world as no other technology has done before is rightfully raising many questions concerning its ethical, legal, societal and economical effects. However, whereas the dangers and risks of application of AI without due consideration of its societal, ethical or legal impact, are increasingly acknowledged, the potential of AI to contribute to human and societal well-being cannot be dismissed. A comprehensive analysis of the role of AI in achieving the Sustainable Development Goals (Vinuesa et al. 2020) in which I participated, concluded that it has the potential to shape the delivery of all 17 goals, contributing positively to 134 targets across all the goals, but it may also inhibit 59 targets (Vinuesa et al. 2020).

V. Dignum (✉)
Umeå University, Umeå, Sweden
e-mail: virginia@cs.umu.se

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3_9

Ensuring the responsible development and use of AI is becoming a main direction in AI research and practice. Governments, corporations and international organisations alike are coming forward with proposals and declarations of their commitment to an accountable, responsible, transparent approach to AI, where human values and ethical principles are leading. This is a much-needed development, one to which I have dedicated my efforts and research in the last few years.

Currently, there are over 600 AI-related policy recommendations, guidelines or strategy reports, which have been released by prominent intergovernmental organisations, professional bodies, national-level committees and other public organisations, non-governmental and private for-profit companies.[1] A recent study of the global landscape of AI ethics guidelines shows that there is a global convergence around five ethical principles: Transparency, Justice and Fairness, Non-Maleficence, Responsibility and Privacy (Jobin et al. 2019). Nevertheless, even though organisations agree on the need to consider these principles, how they are interpreted and applied in practice, varies significantly across the different recommendation documents.

At the same time, the growing hype around 'AI' is blurring its definition and shoving into the same heap concepts and applications of many different sorts. A hard needed first step in the responsible development and use of AI is to ensure a proper AI narrative, one that demystifies the possibilities and the processes of AI technologies, and that enables all to participate in the discussion on the role of AI in society. Understanding the capabilities and addressing the risks of AI requires that we have a clear understanding of what it is, how it is applied and what are the opportunities and risks involved.

The paper is organised as follows. After a brief discussion on the different perspectives on what consists of an AI system, I present in section "Ensuring the Responsible Development and Use of AI", current efforts towards the responsible and trustworthy development and use of AI. In section "From an Individualistic to a Social Conception of AI", I describe the need to extend current AI research, from an individualistic to a social conception of AI.

[1] See OECD's AI Observatory https://oecd.ai/.

What Is AI and Why Should We Care?

Technological developments have brought forward many potential benefits but at the same time, the risks and problems posed by AI-driven applications are increasingly being reported. All those many guidelines for AI governance and regulation have the risk to become void without an understanding of what AI is and what it can, and cannot, do. Current AI narratives bring forward benefits and risks and describe AI in many different ways, from the obvious next step in digitisation to some kind of magic. Magic in the sense that it can know all about us, and use that knowledge to decide about us or for us in possibly unexpected ways, either solving all our problems, or destroying the world in the process. The reality is, as usual, somewhere in the middle. In the following, I briefly describe some of the ways AI is often misunderstood and conclude with a reflection of the significance to the current efforts towards AI governance.

Currently, AI is mostly associated with Machine Learning (ML). Machine Learning, and in particular, Neural Networks or Deep Learning, is a subset of AI techniques that uses statistical methods to enable computers to perceive some characteristics of their environment. Current techniques are particularly efficient in perceiving images, written or spoken text, as well as the many applications of structured data. By analysing many thousands of examples (typically a few million), the system is able to identify commonalities in these examples, which then enable it to interpret data that it has never seen before, which is often referred to as prediction.

AI Is Not Intelligent

John McCarthy, who originally coined the term Artificial Intelligence, defined it as "the study and design of intelligent agents". In this definition, which is still one of the most common definitions of AI, the concept of intelligence refers to the ability of computers to perform tasks commonly associated with intelligent beings, i.e. humans or other non-human animals. The question remains of what is human (or animal) intelligence. Commonly associated with the ability of the mind to reach correct conclusions about what is true and what is false, and about how to solve problems (Colman 2015), there is no single accepted definition

of intelligence. Moreover, intelligence is a multifaceted concept. Psychologists debate on issues such as types of intelligence, the role of nature versus nurture in intelligence, how intelligence is represented in the brain, and the meaning of group differences in intelligence. Major theories include Sternberg's triarchic theory (Sternberg 1984), Gardner's theory of multiple intelligences (Gardner 2011) and Piaget's theory of development (Piaget 1964). Many characterise human intelligence as more than an analytical process and to include creative, practical and other abilities. These abilities, for a large part associated with socio-cultural background and context, are far from being possible to be replicated by AI systems, even if these may approach analytical intelligence for some (simple) tasks.

AI Is Not Artificial

AI is not magic. It will not solve all our problems, nor can it exist without the use of natural resources and the work of legions of people. In a recent book, 'The Atlas of AI' (Crawford 2021), Kate Crawford describes the field as a collection of maps that enable the reader to traverse places, their relations and their impact on AI as an infrastructure. From the mines where the core components of hardware originate, to the warehouses where human labourers are mere servants to the automated structure, in an uneasy reminder of Chaplin's *Modern Times*, to the hardship of data classification by low paid workers in data labelling farms, Crawford exposes the hard reality of the hidden side of AI success. Concluding with the powerful reminder that AI is not an objective, neutral and universal computational technique, but is deeply embedded in the social, political, cultural and economic reality of those that build, use and mostly control it (Dignum 2021).

AI Is Not the Algorithm

AI is based on algorithms. The concept of 'algorithm' is achieving magical proportions, used right and left to signify many things, *de facto* seen as a synonym to AI. But, even though AI uses algorithms, as does any other computer program or engineering process, AI *is not* the algorithm.

The easiest way to understand an algorithm is as a recipe, a set of precise rules to achieve a certain result. Every time you add two numbers, you are using an algorithm, as well as when you are baking an apple pie. However, by itself, the recipe has never turned into an apple pie; and, the

end result of your pie has as much to do with your baking skills and your choice of ingredients, as with the choice for a specific recipe. The same applies to AI algorithms: for a large part the behaviour and results of the system depends on its input data, and on the choices made by those that developed, trained and selected the algorithm. In the same way as we have the choice to use organic apples to make our pie, in AI we also have the choice to use data that respects and ensures fairness, privacy, transparency and all other values we hold dear. This is what Responsible AI is about, and includes demanding the same requirements from the ones that develop the systems that affect us.

Responsible AI

AI is first and foremost technology that can automatise simple, lesser, tasks. At the present, AI systems are largely incapable of understanding meaning. An AI system can correctly identify cats in pictures or cancer cells in scan images, but it has no idea of what a cat or a cancer cell is. Moreover, AI system can only do this if there are enough people performing the tasks (classification, collection, maintenance…) that are needed to make the system function, misleadingly, in an autonomous manner.

But it is much more, both in terms of techniques used, as in terms of societal impact and human participation. As such, AI can be best understood as a socio-technical ecosystem. In order to understand AI, it is necessary to recognise the interaction between people and technology, and how complex infrastructures affect and are affected by society and by human behaviour.

As such, AI is not just about the automation of decisions and actions, the adaptability to learn from the changes affected in the environment, and the interactivity required to be sensitive to the actions and aims of other agents in that environment, and decide when to cooperate or to compete. It is mostly about the structures of power, participation and access to technology that determine who can influence which decisions or actions are being automated, which data, knowledge and resources are used to learn from, and how interactions between those that decide and those that are impacted are defined and maintained.

A responsible, ethical, approach to AI will ensure transparency about how adaptation is done, responsibility for the level of automation on which the system is able to reason, and accountability for the results

and the principles that guide its interactions with others, most importantly with people. In addition, and above all, a responsible approach to AI makes clear that AI systems are artefacts manufactured by people for some purpose, and that those which make these have the power to decide on the use of AI. It is time to discuss how power structures determine AI and how AI establishes and maintains power structures, and on the balance between, those who benefit from, and those who are harmed by the use of AI (Crawford 2021).

Ensuring the Responsible Development and Use of AI

Ethical AI is not, as some may claim, a way to give machines some kind of 'responsibility' for their actions and decisions, and in the process, discharge people and organisations of their responsibility. On the contrary, ethical AI gives the people and organisations involved more responsibility and more accountability: for the decisions and actions of the AI applications, and for their own decision of using AI in a given application context. When considering effects and the governance thereof, the technology, or the artefact that embeds that technology, cannot be separated from the socio-technical ecosystem of which it is a component. Guidelines, principles and strategies to ensure trust and responsibility in AI, must be directed towards the socio-technical ecosystem in which AI is developed and used. It is not the AI artefact or application that needs to be ethical, trustworthy or responsible. Rather, it is the social component of this ecosystem that can and should take responsibility and act in consideration of an ethical framework such that the overall system can be trusted by the society. Having said this, governance can be achieved by several means, softer or harder. Currently several directions are being explored, the main ones are highlighted in the remainder of this section. Future research and experience will identify which approaches are the most suitable, but given the complexity of the problem, it is very likely that a combination of approaches will be needed.

Regulation

AI regulation is a hot topic, with many proposers and opponents. The recent proposal by the European Commission envisions a risk-based

approach to regulation that ensures that people can trust that AI technology is used in a way that is safe and compliant with the law, including the respect of fundamental human rights.

The proposal implements most of the 7 requirements of the Ethics Guidelines for Trustworthy AI into specific requirements for 'high-risk' AI. However, it does not deal explicitly with issues of inclusion, non-discrimination and fairness. Minimising or eliminating discriminatory bias or unfair outcomes is more than excluding the use of low-quality data. The design of any artefact, such as an AI system, is in itself an accumulation of choices and choices are biased by nature as they involve selecting an option over another. Technical solutions at dataset level must be complemented by socio-technical processes that help avoid any discriminatory or unfair outcomes of AI.

Moreover, successful regulation demands clear choices about what is being regulated: is it the technology itself, or the impact, or results of its application? By focusing on technologies, or methods, i.e. by regulating systems that are based on "machine learning, logic, or statistical approaches", such as described in the AI definition used in the European Commission's proposal, we run the risk of seeing organisations evading the regulation, simply by classifying their applications differently. Conversely, there are a plethora of applications based on, e.g. statistics that are not AI.

A future-proof regulation should focus on the outcomes of systems, whether or not these systems fall in the current understanding of what is 'AI'. If someone is wrongly identified, is denied human rights or access to resources, or is conditioned to believe or act in a certain way, it does not matter whether the system is 'AI' or not. It is simply wrong. Moreover, regulation must also address the inputs, processes and conditions under which AI is developed and used are at least as important. Much has been said about the dangers of biased data and discriminating applications. Attention for the societal, environmental and climate costs of AI systems is increasing. All these must be included in any effort to ensure the responsible development and use of AI.

At the same time, AI systems are computer applications, i.e. are artefacts, and as such subject to existing constraints, legislation, for which due diligence obligations and liabilities apply. That is, already now, AI does not operate in a lawless space. Before defining extra regulations, we need to start by understanding what is already covered by existing legislation.

A risk-based approach to regulation, as proposed by the European Commission, is the right direction to take, but needs to be informed by a clear understanding of what is the source of those risks. Moreover, it requires to not merely focus on technical solutions at the level of the algorithms or the datasets, but rather on developing socio-technical processes, and the corporate responsibility, to ensure that any discriminatory or unfair outcomes are avoided and mitigated. Independently of whether we call the system 'AI' or not.

Standardisation

Standards are consensus-based agreed-upon ways of doing things by providing what they consider to be the minimum universally-acknowledged specifications. Industry standards are proven to be beneficial to organisations and individuals. Standards can help reduce costs and improve efficiency of organisations by providing consistency and quality metrics, the establishment of a common vocabulary, good-design methodologies and architectural frameworks. At the same time, standards provide consumers with confidence in the quality and safety of products and services.

Most standards are considered soft governance; i.e. non-mandatory to follow. Yet, it is often in the best interest of companies to follow them to demonstrate due diligence and, therefore, limit their legal liability in case of an incident. Moreover, standards can ensure user-friendly integration between products (Theodorou and Dignum 2020).

AI standards work to support the governance of AI development and use is ongoing at ISO and IEEE, the two leading standards bodies. Such standards can support AI policy goals in particular where it concerns safety, security and robustness of AI, guarantees of explainability, and means to reduce bias in algorithmic decisions (Cihon 2019).

Jointly with IEC, ISO has established a Standards Committee on Artificial Intelligence (SC 42). Ongoing SC 42 efforts are, so far, limited and preliminary (Cihon 2019). On the other hand, IEEE's Standards Association global initiative on Ethically Aligned Design is actively working on vision and recommendations to address the values and intentions as well as legal and technical implementations of autonomous and intelligent systems to prioritise human well-being (IEEE 2016). This is the joint

work of over 700 international researchers and practitioners. In particular, the P7000[2] series aims to develop standards that will eventually serve to underpin and scaffold future norms and standards within a new framework of ethical governance for AI/AS design. Currently, the P7000 working groups are working on candidate standard recommendations to address issues as diverse as system design, transparency in autonomous systems, algorithmic bias, personal, children, student and employer data governance, nudging, or, the identification and rating the trustworthiness of news sources. Notably, the efforts on assessment of impact of autonomous and intelligent systems on human well-being is now available as an IEEE standard.[3]

Assessment

Responsible AI is more than the ticking of some ethical 'boxes' or the development of some add-on features in AI systems. Nevertheless, developers and users can benefit from support and concrete steps to understand the relevant legal and ethical standards and considerations when making decisions on the use of AI applications. Impact assessment tools provide a step-by-step evaluation of the impact of systems, methods or tools on aspects such as privacy, transparency, explanation, bias or liability (Taddeo and Floridi 2018).

It is important to realise, as described in Taddeo and Floridi (2018) that even though these approaches "can never map the entire spectrum of opportunities, risks, and unintended consequences of AI systems, they may identify preferable alternatives, valuable courses of action, likely risks, and mitigating strategies. This has a dual advantage. As an opportunity strategy, foresight methodologies can help leverage ethical solutions. As a form of risk management, they can help prevent or mitigate costly mistakes, by avoiding decisions or actions that are ethically unacceptable".

Currently, much effort is being put on the development of assessment tools.[4] The EU Guidelines for trustworthy AI are accompanied by

[2] See https://ethicsinaction.ieee.org/p7000/.

[3] See https://ieeexplore.ieee.org/browse/standards/reading-room/page?ranges=2020_2020_Year.

[4] A comprehensive list of existing frameworks is available at https://www.aiethicist.org/frameworks-guidelines-toolkits.

a comprehensive assessment framework which was developed based on a public consultation process.

Finally, it is important to realise that any requirements for trustworthy AI are necessary but not sufficient to develop human-centred AI. That is, such requirements need to be understood and implemented from a contextual perspective, i.e. it should be possible to adjust the implementation of the requirement such as transparency based on the context in which the system is used. That is requirements such as transparency should not have one fixed definition for all AI systems, but rather be defined based on how the AI system is used. At the same time, any AI technique used in the design and implementation should be amenable to explicitly consider all ethical requirements. For example, it should be possible to explain (or to show) how the system got to a certain decision or behaviour.

Assessment tools need to be able to account for this contextualisation, as well as ensuring alignment with existing frameworks and requirements in terms of other types of assessment, such that the evaluation of trust and responsibility of AI systems provides added value to those developing and using it, rather than adding yet another bureaucratic burden.

Codes of Conduct and Advisory Boards

A professional code of conduct is a public statement developed for and by a professional group to reflect shared principles about practice, conduct and ethics of those exercising the profession; describe the quality of behaviour that reflects the expectations of the profession and the community; provide a clear statement to the society about these expectations, and enable professionals to reflect on their own ethical decisions.

A code of conduct supports professionals to assess and resolve difficult professional and ethical dilemmas. While there in the case of ethical dilemmas there is not a correct solution, the professionals can give account of their actions by referring to the code. In line with other socially sensitive professions, such as medical doctors or lawyers, i.e. with the attendant certification of 'ethical AI' can support trust. Several organisations are working on the development of codes of conduct for data and AI-related professions, with specific ethical duties. Just recently ACM, the

Association for Computing Machinery, the largest international association of computing professionals, updated their code of conduct.[5] This voluntary code is "a collection of principles and guidelines designed to help computing professionals make ethically responsible decisions in professional practice. It translates broad ethical principles into concrete statements about professional conduct". This code explicitly addresses issues associated with the development of AI systems, namely issues of emergent properties, discrimination and privacy. Specifically, it calls out the responsibility of technologists to ensure that systems are inclusive and accessible to all and requires that they are knowledgeable about privacy issues.

At the same time, the role of an AI Ethicist is becoming a hot topic as large businesses are increasingly dependent on AI and as the impact of these systems on people and society becomes increasingly more evident, and not always for the best. Recent scandals both about the impact of AI in bias and discrimination, as on the way businesses are dealing with their own responsibility, specifically on the role and treatment of whistleblowers, have increased the demand for clear and explicit organisational structures to deal with the impact of AI.

Many organisations have since established the role of chief AI ethics officer, or similar. Others, recognising that the societal and ethical issues that arise from AI are complex and multi-dimensional, and therefore require insights and expertise from many different disciplines and an open participation of different stakeholders, have established AI ethics boards or advisory panels.

Awareness and Participation

Inclusion and diversity are a broader societal challenge and central to AI development. It is therefore important that as broad a group of people as possible have a basic knowledge of AI, what can (and can't) be done with AI, and how AI impacts individual decisions and shapes society. A well-known initiative in this area is Elements of AI,[6] initiated in Finland with the objective to train one per cent of EU citizens in the basics of

[5] See https://www.acm.org/code-of-ethics.

[6] See https://www.elementsofai.com/.

artificial intelligence, thereby strengthening digital leadership within the EU.

In parallel, research and development of AI systems must be informed by diversity, in all the meanings of diversity, and obviously including gender, cultural background and ethnicity. Moreover, AI is not any longer an engineering discipline and at the same time there is growing evidence that cognitive diversity contributes to better decision-making. Therefore, developing teams should include social scientists, philosophers and others, as well as ensuring gender, ethnicity and cultural differences. It is equally important to diversify the discipline background and expertise of those working on AI to include AI professionals with knowledge of, amongst others, philosophy, social science, law and economy. Regulation and codes of conduct can specify targets and goals, along with incentives, as a way to foster diversity in AI teams (Dignum 2020).

From an Individualistic to a Social Conception of AI

The dominant approach to AI has so far been an individualistic, rational one. Russell and Norvig's classic AI text book defines AI along two dimensions (Russell and Norvig 2010): how it reasons (human-like or rationally[7]) and what it 'does' (think or act). Human-like approaches aim to understand and model how the human mind works, and rational approaches aim at developing systems that result in the optimal level of benefit or utility for an individual. Both approaches are well aligned with the Western philosophy statement "I think therefore I am", fundamentally conceptualising an AI system as an individual entity.

Intelligent agents are typically characterised as bounded rational, acting towards their own perceived interests. For instance, by identifying and applying patterns in (human-generated) data, machine learning systems mimic and extend the human reasoning and actions embedded in that data, whereas symbolic logic approaches (the so called 'good old-fashioned AI', or GOFAI) aim to capture the laws of rational thought and action, resulting in an idealised model of human reasoning.

[7] Note that this terminology does not imply that human-like behaviour is not rational, but uses 'rational' to refer to utility optimising behaviour.

Human-Like or Rational?

That is, rationality is often a central assumption for agent deliberation (Dignum 2017). Moreover, intelligent systems are expected to hold consistent world views (beliefs), and to optimise action and decision based on a set of given preferences (often accuracy has highest priority). This view on rationality entails that agents are expected, and designed, to act rationally in the sense that they choose the best means available to achieve a given end, and maintain consistency between what is wanted and what is chosen (Lindenberg 2001).

The main advantages of a rationality assumption are their parsimony and applicability to a very broad range of situations and environments, and their ability to generate falsifiable, and sometimes empirically confirmed, hypotheses about actions in these environments. This gives conventional rational choice approaches a combination of generality and predictive power not found in other approaches.

Unfortunately, this type of rational behaviour fits mostly with strategic choices, where information is all available or can be gathered at will. It does not really suit most human behaviour which is based on split second decisions, on habits, on social conventions and power structures. When the aim of AI systems is to develop models of societal behaviour or to develop systems that are able to interact with people in social settings, rationality is not enough to model human behaviour. This was exemplified before by all the application areas, where the rational behaviour needs to be combined with different types of behaviour in order to be effective. In reality, human behaviour is neither simple nor rational, but derives from a complex mix of mental, physical, emotional and social aspects. Realistic applications must moreover consider situations in which not all alternatives, consequences and event probabilities can be foreseen. Thus, it is impossible to 'rationally' optimise utility, as the utility function is not completely known, neither are the optimisation criteria known. This renders rational choice approaches unable to accurately model and predict a wide range of human behaviours. Already in 2010, Dignum and Dignum (2010) show how different types of variations and models cater for different applications, while no generic model exists that serves as a foundation for all models.

Both the human-like and the rational perspectives on AI are suited for a task-oriented view on the purpose of AI systems. That is, the system is expected to optimise the result of its actions for a specific purpose. Even

though it is able to perceive its environment and adapt accordingly, it is mostly unaware of its own role in that environment, and of the fact that its actions contribute to change. Given the large impact of AI on society, a new modelling paradigm is needed that is able to account for this feedback loop of decision—action—context. That is, AI modelling needs to follow a social paradigm.

Social AI

Non-Western philosophies, and in particular Ubuntu, take a societal rather than an individual stance, which begs the question of how AI would be defined from the perspective of Ubuntu thought. Without trying to describe or fully understand Ubuntu philosophy, I will in the following apply some of its main tenets to show how these can be applied to AI concepts and development approaches.

Ubuntu expresses the deeply-held African ideals of one's personhood being rooted in one's interconnectedness with others, and emphasises norms for inter-personal relationships that contribute to social justice, such as reciprocity, selflessness and symbiosis. Community is at the core of Ubuntu, focusing on interconnectedness and caring for communal living, underpinned by values of cooperation and collaboration (Mugumbate and Nyanguru 2013). Solidarity, which requires people to be aware of and attentive to the needs of those around them, rather than focusing only on their own needs is therefore central in Ubuntu, with an emphasis on caring, caretaking and context (van Breda 2019).

As such, Ubuntu philosophy is essentially relational and defines morally right actions as those that that connect, rather than separate (i.e. honours communal relationships, reduces discord or promotes friendly relationships. The concept of community can best be understood as an (objective) standard that should guide what the majority wants, or what moral norms become central (Ewuoso and Hall 2019). This does not imply that individual rights are subordinated but that individuals pursue their own good through pursuing the common good (Lutz 2009).

Human rights set the foundational value of human dignity in terms of autonomy. This view, for a large part originating from Kantian philosophy, sets human rights as the ultimate ways of treating our intrinsically valuable capacity for self-governance with respect. It has therefore been argued that the collectivistic grounds of Ubuntu thought are at odds with this individual autonomy view. According to Metz, "*While the Kantian*

theory is the view that persons have a superlative worth because they have the capacity for autonomy, the present, Ubuntu-inspired account is that they do because they have the capacity to relate to others in a communal way" (Metz 2011, p. 544). Or, as Metz also describes "*Human rights violations are ways of gravely disrespecting people's capacity for communal relationship, conceived as identity and solidarity […]*" (ibid., p. 545). In Ubuntu, human nature is special and inviolable due to its capacity for harmonious relationships. At the same time, no individual's rights are greater than another, thus, every individual in a community, including both children and adults is important and should be heard and respected (Osei-Hwedie 2007).

With respect to the ethics of AI development and use, the above formulation of human dignity as the human capability to relate to others in a communal way, can account for, or justify, the resolution of moral dilemmas, where autonomy conflicts with beneficence or any of the other principles, as also proposed with respect to bioethics and medicine, or to ground the UN's Sustainable Development Goals (Ewuoso and Hall 2019).

Specifically, the formulation of Ubuntu described above may be used to justify decisions in face of ethical dilemma(s), for example, where such a decision favours the action that enhances communal relationships, or the capacity for the same (ibid.). As such, this framework could usefully supplement utilitarian, individualistic and deontological approaches that are often embedded in AI ethics decision-making. As proposed for the case of clinical contexts, ethical decision-making in the context of AI systems, can also be extended with rules that state "A breach of an ethical principle is justifiable if, on the balance of probabilities, such a breach is more likely to enhance communal relationships (…)" (Ewuoso and Hall 2019).

Towards a Social Paradigm for AI

As discussed in section "What Is AI and Why Should We Care?", AI is neither 'artificial' nor 'intelligent' but the product of choices involving theory and values. Current AI paradigm, as seen in section "Human–Like or Rational?", rely, and are bound, but individualistic, theories of intelligence, thinking, rationality and human nature. As such, this paradigm supports the implementation of different reasoning and action approaches, corresponding to individual understandings of contexts and

different reactions of agents to contexts. However, AI, like all technology, affects and changes our world, which in turn changes us. New paradigms are needed that address collective understanding and the effect of change in context and the feedback loop from change back to the individual and collective reasoning and behaviour. Modelling this feedback loop recognises that it is not just about which action is performed but, what kind of reasoning leads to that action, and which values and perceptions lead the observation of the context.

Given the transnational character of AI, it is also imperative to address the ways in which AI may impact or be accepted by society in various regions around the world. In particular, it is needed to position the African continent in global debates and policymaking in Responsible AI. For instance, initiatives such as Responsible AI Network—Africa[8] and the African Observatory on Responsible AI[9] are aiming to understand how AI may impact or be accepted by society in various regions around the world, deepen the understanding of AI and its effects in (Sub-Saharan) Africa, and promote the development and implementation of locally appropriate evidence-led AI policies and enabling legislation.

However, at the same time, it is as important to extend current conceptualisations of AI, with the relational worldview that characterise African thought, such as embedded in Ubuntu philosophy. Current AI paradigms strengthen existing power structures and prevent a truly societal understanding of the impact and challenges of AI for humanity and society. Given the impact AI systems can have on people, inter-personal interactions, and society as a whole, it seems to be relevant to consider a relational stance to approach the specification, development and analysis of AI systems. The deeply-held African vision of one's personhood being rooted in one's interconnectedness as expressed in Ubuntu philosophy can support integrating such a social perspective to AI—in terms of how it reasons (human-like or rationally) and what it 'does' (think or act), which would result in a new paradigm that considers social/collective reasoning and includes change, or reaction, as a third possible result of AI, next to thinking and acting. Figure 1 depicts this perspective on AI, extending

[8] See https://rainafrica.org/.

[9] See https://www.idrc.ca/en/project/african-observatory-responsible-artificial-intelligence.

	Human-like	Rational	**Social**
Think	Think humanly	Think rationally	**Think socially**
Act	Act humanly	Act rationally	**Act jointly**
Change/react	**Enhance human performance**	**Rational institutions**	**Social engagement**

Fig. 1 Social perspectives in AI

the well-known dimensions defined in (Russell and Norvig 2010), here depicted in grey shading.

Collective or social reasoning is about modelling societal values and norms, and how these ground and influence human and rational thinking. In social action, outcomes are relative to the actions of others, where others are not just seen as opponents or obstacles on the decision-making, but as a positive force for achieving a joint endeavour.

Addressing change from the human-like perspective leads to approaches to AI that aim at enhance, rather than replace, human performance, and from a rational perspective, it concerns the development and optimisation of institutional infrastructures that maximise the effects of rational behaviour.

All these perspectives need to be brought together to address the impact of AI for a socially grounded and engaged perspective. This is no easy feat, but one for which there are no single models, nor simple approaches. It will require multidisciplinary and multi-stakeholder participation and to accept that any solution is always contingent and contextual.

Conclusions

Increasingly, AI systems will be taking decisions that affect our lives, in smaller or larger ways. In all areas of application, AI must be able to take into account societal values, moral and ethical considerations, weigh the respective priorities of values held by different stakeholders and in multi-cultural contexts, explain its reasoning and guarantee transparency. As the capabilities for autonomous decision-making grow, perhaps the most

important issue to consider is the need to rethink responsibility. Being fundamentally tools, AI systems are fully under the control and responsibility of their owners or users. However, their potential autonomy and capability to learn, require that design considers accountability, responsibility and transparency principles in an explicit and systematic manner. The development of AI algorithms has so far been led by the goal of improving performance, leading to opaque black boxes. Putting human values at the core of AI systems calls for a mind-shift of researchers and developers towards the goal of improving transparency rather than performance, which will lead to novel and exciting techniques and applications. In particular, this requires to complement the currently predominant individualistic view of AI systems, to one that acknowledges and incorporates collective, societal, and ethical values at the core of the design, development and use of AI systems.

Biological evolution has long been revised from a 'ladder' view: a unilinear progression from 'primitive' to 'advanced'. The same revision is also seen in anthropology: the idea that cultural evolution follows a ladder model, with small-scale decentralised societies at the bottom and hierarchical, state, societies at the top, where the top would be technologically more advanced, has been shown to be not only demeaning but also inaccurate (Eglash 1999). These fields have long since moved to a more dynamic, branching type model.[10] It is high time that AI models embrace such a branching view. Only then, can AI align with the diversity that truly reflects worldwide differences in cultural and philosophical thought. In the same way as biology or culture, intelligence is not linear, it is branching.

Acknowledgements This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP), funded by the Knut and Alice Wallenberg Foundation and by the European Commission's Horizon 2020 Research and Innovation Programme project HumaneAI-Net (grant 825619).

[10] See Gould (1997) for a branching model of biological evolution, and Durham (1982) for cultural evolution.

References

Cihon, P., 2019. Standards for AI governance: International standards to enable global coordination in AI research & development. Future Humanity Inst. Univ. Oxf.

Colman, A.M., 2015. A dictionary of psychology. Oxford quick reference.

Crawford, K., 2021. The Atlas of AI. Yale University Press.

Dignum, V., 2017. Social agents: Bridging simulation and engineering. Commun. ACM 60, 32–34.

Dignum, V., 2020. Responsibility and artificial intelligence. Oxf. Handb. Ethics AI 4698, 215.

Dignum, V., 2021. AI-the people and places that make, use and manage it. Nature 593, 499–500.

Dignum, V., Dignum, F., 2010. Designing agent systems: State of the practice. Int. J. Agent-Oriented Softw. Eng. 4, 224–243.

Durham, William H, 1982. Interactions of genetic and cultural evolution: Models and examples. Hum. Ecol. 10 (2), 289–323.

Eglash, R., 1999. African fractals: Modern computing and indigenous design.

Ewuoso, C., Hall, S., 2019. Core aspects of Ubuntu: A systematic review. South Afr. J. Bioeth. Law 12, 93–103.

Gardner, H.E., 2011. Frames of mind: The theory of multiple intelligences. Hachette UK.

Gould, Stephen Jay, 1997. Redrafting the tree of life. Proc. Am. Philos. Soc. 141 (1), 30–54.

IEEE-EA, 2016. A vision for prioritizing human well-being with artificial intelligence and autonomous systems. IEEE Glob Initiat Ethical Consid. Artif Intell Auton Syst 13.

Jobin, A., Ienca, M., Vayena, E., 2019. The global landscape of AI ethics guidelines. Nat. Mach. Intell. 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Lindenberg, S., 2001. Social rationality versus rational egoism, in: Handbook of Sociological Theory. Springer, pp. 635–668.

Lutz, D.W., 2009. African Ubuntu philosophy and global management. J. Bus. Ethics 84, 313–328.

Metz, T., 2011. Ubuntu as a moral theory and human rights in South Africa. Afr. Hum. Rights Law J. 11, 532–559.

Mugumbate, J., Nyanguru, A., 2013. Exploring African philosophy: The value of Ubuntu in social work. Afr. J. Soc. Work 3, 82–100.

Osei-Hwedie, K., 2007. Afro-centrism: The challenge of social development. Soc. Work. Werk 43.

Piaget, J., 1964. Part I: Cognitive development in children: Piaget development and learning. J. Res. Sci. Teach. 2, 176–186.

Russell, S., Norvig, P., 2010. Artificial intelligence: A modern approach. Third Edit. Prentice Hall 10, B978–012161964.
Sternberg, R.J., 1984. Toward a triarchic theory of human intelligence. Behav. Brain Sci. 7, 269–287.
Taddeo, M., Floridi, L., 2018. How AI can be a force for good. Science 361, 751–752.
Theodorou, A., Dignum, V., 2020. Towards ethical and socio-legal governance in AI. Nat. Mach. Intell. 2, 10–12.
van Breda, A.D., 2019. Developing the notion of Ubuntu as African theory for social work practice. Soc. Work 55, 439–450.
Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M., Fuso Nerini, F., 2020. The role of artificial intelligence in achieving the Sustainable Development Goals. Nat. Commun. 11, 233. https://doi.org/10.1038/s41467-019-14108-y

Index

D. O. Eke et al. (eds.), *Responsible AI in Africa*, Social and Cultural Studies of Robots and AI, https://doi.org/10.1007/978-3-031-08215-3