Reccomendations on the inclusion sub-Saharan Africa in Global AI Ethics

- **Artificial Intelligence (AI) and algorithmic decision-making (ADM) systems can improve efficiency in society, in business, in public authorities and the in public sector, where it can help to reduce fraud, corruption and cut unnecessary bureaucracy.**
- **AI has the potential to both perpetuate and amplify the deep-rooted social exclusion of Africans for three key reasons:**
  - **AI can reinforce or amplify long-standing societal biases and prejudices.**
  - **Africans may be marginalised from the economic and social benefits of AI due lack of control over and ownership of AI technologies.**
  - **It is the perspectives of the Global North that predominantly inform the current discussions on inclusion and ethics, with limited commitment to addressing historical social and economic injustices. This imbalance carries risks, particularly where the ethical norms and values designed into these technologies collide with those of the African communities in which they are deployed.**
- **There is an urgent need to address the ethical challenges of AI and to develop and deploy AI in Africa in a manner that is inclusive, socially beneficial, and that adds economic value (AI for good).**
  - **In addition to the current AI ethics policy initiatives, multi-stakeholder principles, codes of ethics and industry toolkits, the following is needed:**
  - **Political commitment and a strong formulation of the value proposition of AI at the policy level in order to generate public interest in AI technologies, particularly in sub-Saharan Africa, where interest within policy circles is still embryonic.**
  - **Application of universal human values and international standards that take into account Africa's historical peculiarities, while achieving development objectives. This includes treating social, economic and cultural rights on a par with civil and political rights.**
  - **Inviting a cross section of Africans working on different aspects of AI to articulate their views on AI ethics and the inclusion of Africa.**
- **The common themes laid out in this brief cluster around four key principles:**
  - **the need to introduce safeguards to balance AI's risks and opportunities;**
  - **the need to protect personal and collective privacy rights in cross-border data flows;**
  - **the need to define African values and align AI with them; and**
  - **the need for fair, inclusive and socially responsible AI development.**

Introduction

Ethics is the philosophy of human conduct. Addressing the moral issues underpinning human action, it can mean either: being engaged in the philosophical sub-discipline of ethics; or acting in a way that is ethical.[1] This policy brief mostly uses the latter definition.

Recent advances in artificial intelligence (AI),[2] in particular the deepening of 'machine learning' methods and the growth of algorithm assisted and automated decision-making, have brought to the fore the question of ethics, often under the slogan of 'AI for Good'.[3] AI has the potential to dramatically transform both the economy and society. Governments, for example, are already implementing AI in their operations and service delivery, to improve efficiency, save time and money, and deliver better quality public services. On the other hand, businesses can employ automated decision-making (ADM) in a growing range of use-cases.

AI and machine learning algorithms are replacing humans in making many of the decisions that affect our daily lives, raising a number of key ethical and moral questions.

Machine learning algorithms are replacing humans in making many of the decisions that affect our everyday lives, raising a number of questions. How can we decide how machine learning algorithms and their designers should act? What is the ethics of today and what will it be in the future? This has invigorated the debate on the ethical issues surrounding their use. The debate is underpinned by further ethical issues that are more acute in the Global South as a result of a number of complex issues which this brief will address below.

Despite the global nature of the ethical implications of artificial intelligence, attention to date has focused primarily on the US and the EU, with growing awareness of China, especially its increasing AI capabilities, its impact on the Global South and the global geopolitical order.[4] Despite the clear need to understand how AI affects people around the world, a truly global perspective remains a critical blind spot in the ethics conversation.[5]

The United Nations, national legislators and industrial bodies in developed countries are asking these questions and are already acting to protect their constituents from some potentially negative effects of AI, such as: algorithmic discrimination;[6] voter manipulation;[7]

---

[1] Mireille Hildebrandt, 'Closure: On Ethics, Code and Law', Law for Computer Scientists, 2 June 2019, https://lawforcomputerscientists.pubpub.org/pub/nx5zv2ux.

[2] Cf: David Kaye, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression', New York, Human Rights Council, 29 August, 2018, https://un.org/ga/search/view_doc.asp?symbol=A/73/348.

[3] See, for example, the 'AI for Good Global Summit' of the International Telecommunication Union, Geneva, https://aiforgood.itu.int/.

[4] Arthur Gwagwa, 'The Geopolitical impact of China's Artificial Intelligence Programmes on sovereign and foreign policy decisions', GAC Geopolitics of AI, Canada, March 2019.

[5] Alexa Hagerty and Igor Rubinov 'Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence', 2019, arxiv.org/abs/1907.07892.

[6] Pro Publica, 'Machine Bias: Investigating Algorithmic Injustice', *Pro Publica*, https://www.propublica.org/series/machine-bias.

[7] Nicholas Confessore, 'Cambridge Analytica and Facebook: The Scandal and the Fallout So Far', *New York Times*, 4 April 2018, http://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

disappearance of jobs due to automation;[8] lethal results of AI failures; gender bias in job recruitment tools; discriminatory credit algorithms; and 'dirty data' in predictive policing that results in people of colour being unfairly targeted, especially in the US.[9]

More importantly, Western governments and businesses are implementing domestic measures. For instance, the UK, France, and Canada have AI frameworks, while lobbying at UN platforms such as the World Trade Organisation (WTO) for positions that protect their economic interests.[10]

Despite the possible negative impact on Africa, however, the value proposition and public interest in AI technologies in sub-Saharan Africa is weak, while the policy discourse is embryonic and mostly focused on AI for development through the lens of international development. It is important for Africa to be included in AI discussions if it is to fully benefit from the so-called 'Fourth Industrial Revolution', while at the same time mitigating some of the potential harms of AI.

As Africa adopts AI, these technologies will be used in imperfect and unequal environments where they may be used by governments for non-democratic ends under political corruption and authoritarian rule. Misused, AI may threaten basic human rights and civil liberties, such as the right to privacy and freedom of expression. Further, the corporate use of algorithms in modern data-processing techniques raises important human rights and ethical issues, including privacy and the protection of personal data. The unwanted social and economic impacts of AI may be felt most immediately by historically marginalised groups. AI's impact on labour disruption and digital divides will likely be more pronounced in low- and middle-income countries, the majority of which are in Africa.

The adoption of AI in Africa may be used by governments for non-democratic ends to entrench political corruption and authoritarian rule.

It is therefore necessary to assess the extent to which Africa has been included in the AI ethics discussions to date. To that end, this paper examines current global AI ethics initiatives and evaluates the extent to which sub-Saharan Africa is being included. This is done through the lens of social inclusion: seeking to improve the terms on which individuals and groups take part in society, to improve the ability, opportunity, and dignity of those disadvantaged on the basis of their identity.[11] The research assesses the degree of social inclusion (or exclusion) of Africans in global initiatives around AI. It also provides some recommendations for improvements, based on emerging African views.

This policy brief addresses a number of key questions in relation to the AI ethics debate:

- To what extent is sub-Saharan Africa included in emerging global initiatives dealing with artificial intelligence ethics?

---

[8] Moshe Y Vardi, 'Are Robots Going to Steal Your Job? Probably', *The Guardian*, 13 September 2018, http://www.theguardian.com/commentisfree/2016/apr/07/robots-replacing-jobs-luddites-economics-labor

[9] Alexa Hagerty and Igor Rubinov, 'Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence,' 2019, arxiv.org/abs/1907.07892.

[10] Paul Twomey, 'Building on the Hamburg Statement and The G20 Roadmap For Digitalization – Towards a G20 Framework For Artificial Intelligence in the Workplace', Kiel, Centre for International Governance Innovation, http://www.economics-ejournal.org/economics/discussionpapers/2018-632018.

[11] A. Saloojee & N. Saloojee, 'From Social Exclusion to Social Inclusion: Theory and Practice over Two Continents', *African Journal of Public Affairs,* 4(2), 2011.

- Is Africa included in the processes of the emerging global AI ethics initiatives, including the nascent ones in Africa? And are issues that are relevant to Africa being addressed in such initiatives?

Despite the fact that there is no agreed international legal framework to guide development and implementation of AI on a global scale, initiatives focusing on its ethical design and implementation are evolving apace.[12] In this vacuum, legislators and policy-makers are well aware that they are struggling to understand how these technologies work and what their ethical implications are. But, despite this uncertainty, a recent meta-analysis reveals that AI Ethics has seemingly converged on a set of principles that closely resemble the four classic principles of medical ethics.[13]

Nevertheless, there are still some reasons to be concerned about the future impact of AI development and its governance reasons, including a lack of accountability, a lack of specificity (as most of the codes are more aspirational than prescriptive), and a number of ongoing political and normative disagreements. [14] "As political initiatives and policy instruments, some of the existing guidelines remain insufficient to ensure that the design and deployment of AI is human-centric and respects individual rights".[15]

Policy-makers and legislators are struggling to understand how these technologies work and what their ethical implications are.

There are also additional problems, such as value alignment, ensuring that AI aligns with human values, and, consequently, whose sets of values should be embedded in machine learning.[16] Some remain sceptical of the efficacy of codes of ethics. A recent study, for example, showed that such codes had no impact on the decision-making of computer scientists, who rather based decisions on personal ethic.[17]

Social inclusion of Africa in the 4IR

At a social level, some groups in Africa may be excluded from the benefits of AI due to specific characteristics or attributes, such as skin colour, language, culture, racial categorisation. A complex set of issues thus exists around the divide between the Global North and the Global South when it comes to the development, design, and application of AI-based technologies, algorithms and data. For example, non-representative or biased data can further entrench existing inequities, as AI systems reproduce the representational gaps and biases of the data sets on which they are trained. This will lead to AI amplifying

[12] UNESCO Courier, Artificial Intelligence: The Promises and the Threats, en.unesco.org/courier/2018-3.

[13] Brent Mittelstadt, AI Ethics – Too Principled to Fail? Oxford Internet Institute. May 2019 papers.ssrn.com/sol3/papers.cfm?abstract_id=3391293.

[14] Brent Mittelstadt, 'AI Ethics – Too Principled to Fail?' Oxford Internet Institute.

[15] Access Now. 'Laying down the law on AI: ethics done, now the EU must focus on human rights' https://www.accessnow.org/laying-down-the-law-on-ai-ethics-done-now-the-eu-must-focus-on-human-rights/

[16] Jessica Fjeld, et al, A Map of Ethical and Rights Based Approach, Principled Artificial Intelligence July 4 2019 ai-hr.cyber.harvard.edu/primp-viz.html

[17] Matt Shipman, Code of Ethics Doesn't Influence Decisions of Software Developers, NC State University October 8, 2018. news.ncsu.edu/2018/10/software-developer-ethics/

and entrenching long-standing societal biases and prejudices, mostly on the grounds of characteristics like race.[18]

A lack of diversity and inclusion has resulted in flawed AI systems that amplify existing gender and racial biases. Because the AI field is mainly 'white' and 'male', this has contributed to many failed solutions such as "image recognition services making offensive classifications of minorities, chatbots adopting hate speech, and technology failing to recognize users with darker skin colours."[19]

Artificial intelligence stands to generate vast wealth for the corporations and countries that develop it, but uneven access may lead to the exclusion of Africa from its social and economic benefits.[20] Such exclusion may also include the very processes themselves, such that the Global North may lead the social inclusion discourse and take decisions on how African civil society should be included. The current inability of Western thinkers to link the inclusion discourse to the universality of solidarity is highly problematic, and demonstrates the acute need for increased interactions among intellectuals globally. The discourse cannot be a Western discourse exclusively, particularly when the reality of global exclusion is felt most substantially in the developing world.[21]

Lack of access to the dominant AI ethical discourse may lead to the exclusion of African individuals, communities and countries from its social and economic benefits.

Data equity and privacy

The increasing adoption of AI brings with it certain challenges and constraints. In particular, AI (and the associated methods of machine learning, deep learning, data science and the like) relies on access to vast amounts of data that can help train and develop new systems. Not only is quality (unbiased and fully representative) data often unavailable in emerging economies, but the relevant stakeholders may also lack the capacity (technical and otherwise) to make use of it.[22]

Ownership and control of data is another issue. Currently, many in the Global South are denied access to and control over their own country's data, which is locked under contract rules. Data collection, management and distribution frameworks are not built to ensure healthy partnerships between industry and government, thereby preventing countries from

[18] Julia Powles & Helen Nissenbaum, 'The Seductive Diversion of 'Solving' Bias in Artificial Intelligence', Medium, 7 December 2018, http://www.medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53.

[19] Kari Paul, 'Disastrous' lack of diversity in AI industry perpetuates bias, study finds', The Guardian, 17 April 2019, https://www.theguardian.com/technology/2019/apr/16/artificial-intelligence-lack-diversity-new-york-university-study.

[20] Dan Robitzski, 'Microsoft's President Met With the Pope to Talk About Ethical AI', *Saintly Robotics*, 14 February 2019 http://www.futurism.com/microsoft-pope-ethical-ai.

[21] Anver Saloojee, 'Social Inclusion, Anti-Racism and Democratic Citizenship', Laidlaw Foundation, 2003, http://library.bsl.org.au/jspui/bitstream/1/1796/1/Social_Inclusion_antiracism__democratic_citizenship (2).pdf.

[22] Stefaan Verhulst, 'Data Collaboratives as an enabling infrastructure for AI for Good', *Medium*, 16 April 2019 http://www.medium.com/data-stewards-network/data-collaboratives-as-an-enabling-infrastructure-for-ai-for-good-99aeb1192c10

realising their potential. The absence of regulations that protect developing countries' control of their own collected data simply prolongs global inequities.[23]

The question of asymmetric access to data is closely linked to the right to privacy. Privacy ultimately governs the set of responsible policy outcomes that arise in response to the data equity problem. Firms will, and already do, invoke consumer privacy as a rationale for not permitting access to their data.

Currently, the interests of those deploying advanced data systems (for example, Uber, Tencent, Amazon, and Facebook), can overshadow public interest, acting in ways that are contrary to individual autonomy and collective welfare, often working in ways that are invisible, unquantifiable, and using a business model where the industry controls the data it collects without reward, let alone user transparency.[24]

Negotiation for regulatory specifications to make data a shared resource between data subjects and data controllers should be considered, to protect the interest of individuals.[25] In this regard, the AI policy community must maintain a healthy attitude of scepticism towards 'ethical codes of conduct developed by industry.[26] Such codes are likely to be self-serving, and to contain a formulation of privacy that, unless carefully crafted, operates to help shield the company from an obligation to share aggregated and anonymised data with other stakeholders.

The challenge is also coming from governments. In their policy interventions at the United Nations, the governments of the Global North, with some exceptions, mostly approach the ownership and protection of data simply from a personal privacy angle, without considering the economic value of processed and redacted data, whilst those in the Global South are only beginning to see such datasets as a valuable collective informational resource.[27]

Regulation should specify that data is a shared resource between data subjects and data controllers in a way that protects the rights and privacy of individuals.

Uneven access to AI

The uneven access to AI and related technologies[28] and their data often has the greatest impact on those who sit at the intersection of marginalised groups within the least developed countries (LDCs). This increases the risk of amplifying digital inequalities between and within nations. Data on populations can be seen through multiple frames: the frame of the uncounted (those who 'don't exist' because they are not included in any sort of database), the unaccounted (portrayals of people and groups with less inclusion into the digital world, who are, therefore not adequately represented, possibly due to reasons of economic or social exclusion) and the discounted (those who exist and are in the system,

[23] Yasodara Cordova, 'Artificial Intelligence and the need for data fairness in the global south', Medium, 21 March 2018, https://thelivinglib.org/artificial-intelligence-and-the-need-for-data-fairness-in-the-global-south/

[24] Franklin Foer, F, *World Without Mind: The Existential Threat of Big Tech*, Penguin, 2017.

[25] Yasodara Cordova, 'Artificial Intelligence and the need for data fairness in the global south'

[26] Ryan Calo, 'Artificial Intelligence Policy: A Primer and Roadmap', University of California, Davis Law 51(399), 2017.

[27] Cf: TWN, 'South nations take firm stand on geo-economics of data, Third World Network, 17 April 2019, https://www.twn.my/title2/wto.info/2019/ti190413.htm.

[28] The RIO Global Symposium on AI, 8 - 10 November, 2017. http://cortexlogic.com/2017/11/18/global-symposium-on-artificial-intelligence-inclusion-rio-de-janeiro/

but who are not subjects of interest to the controllers of the data – for example, individuals unable to afford the services being marketed by companies, or outside the parameters of government service delivery). Data is expensive and hard to come by at scale, and AI algorithmic training relies on available data sets, rather than complete data sets. This type of data can easily privilege socio-economically advantaged populations, those with greater access to connected devices and online services.[29] In aggregate, data-driven systems and probabilistic models, inclusive of AI, have a pattern of disproportionately harming disadvantaged, marginalised and vulnerable groups.[30]

As the technologies of the so-called Fourth Industrial Revolution make themselves felt, it is important to assess the extent to which the sub-Saharan Africa region is being included in the course and outcomes of the various social, economic and political processes underpinning the current changes - all the more so since the United Nations has included the importance of social inclusion in a number of their sustainable development goals.[31]

Findings from case studies of the global AI initiatives

As noted above, three broad categories of global initiatives that include policy and governance, multi-stakeholder involvement and industrial models and toolkits were chosen and examined[32]. Although the findings from the case studies, which are presented below, inform the recommendations of this paper, the details of the cases fall outside the purview of this brief. They are, however, in an extended policy brief that is available upon request. An examination of the primary and secondary literature on the above projects was aided by the researcher's involvement in AI policy and governance initiatives, including in the IEEE Standard P7003 and the UN Global Pulse Africa Task Force on AI.

Although existing global AI initiatives address established universal values, engagement with AI and its ethics is only now emerging in Africa.

[29] Tandem Research, 'AI for All: 10 Social Conundrums for India', Tandem Research, January 2019, https://tandemresearch.org/publications/ai-for-all-10-social-conundrums-for-india-working-paper

[30] Alexa Hagerty & Igor Rubinov, 'Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence', 2019, arxiv.org/abs/1907.07892.

[31] Hilary Silver, 'The Contexts of Social Inclusion', New York, UN Department of Economic and Social Affairs, https://www.un.org/esa/desa/papers/2015/wp144_2015.pdf.

[32] This involved analysing context, processes and stakeholders as well as the rationale for such projects. By relying on a collective intelligence approach, where gaps existed in literature, clarification was sought from the project leaders who include Mila Romanoff and Tim Engelhardt of Global Pulse, Knowledge 4 All Foundation AI lead Davor Orlic, IEEE Standard P70003 lead Ansgar Koene, the Africa IBM team, especially Reginald Bryant who works on AI bias, and Joana Bryson, who was in the Google Ethics Board.

Policy and Governance Initiatives

Evidence on the case studies was collected from the African participants[33] who attended the AI4D initiative[34] and the United Nations' AI ethics workshop run by the UN Global Pulse,[35] and was analysed using the collective intelligence (CI) approach[36] Whereas the AI4D workshop focussed more on development issues, the latter dealt specifically with AI ethics. For this reason, recommendations from the latter are covered extensively.[37]

With a limited value proposition for AI in Africa, engagement remains low and African issues are just emerging.[38] On the other hand, global initiatives cover universal values that are established evolving issues.[39] Issues for Africa include artificial intelligence for development (AI4D), AI Readiness and the much-hyped Fourth Industrial Revolution (4IR). While a number of African countries have data protection laws, unlike the GDPR, these do not explicitly address AI and ADM. However, recently, South Africa[40] and the UN Global Pulse have started bringing this issue up for discussion on the continent.

The following overarching principle emerges from the workshops referred to above:

Ethics should come before, during and after the law and underpin AI policy and implementation.

Ethics that align to African values should be the foundations and default of AI development and application in Africa. Africa should not rush to pass laws and regulations, but follow an incremental approach based on strong ethical foundations since ethical initiatives help develop a shared language with which to discuss and debate social and political concerns. Future regulatory frameworks should not merely be imported from the West as policy transfer, but engaged with and adapted to the African context in a process of policy learning. A learning approach will ensure that regulation learns from and responds to advances in AI

[33] The majority of the participants of the two workshops were Africans living in Africa, some Africans from the diaspora or working for global technology companies such as Google. Attendees comprised academics, policy-makers, civil society, and people from different sectors including start-ups, education and finance, as well as UN Special Rapporteur on the right to privacy, Joe Cannataci.

[34] Artificial Intelligence for Development in sub-Saharan Africa network, coordinated by the Knowledge 4 All Foundation and IDRC, in close collaboration with UNESCO Chair in AI.

[35] This is an innovation initiative of the United Nations Secretary-General. See https://www.unglobalpulse.org/.

[36] John Danaher, et al. 'Algorithmic Governance: Developing a Research Agenda through the Power of Collective Intelligence.' *Big Data & Society*, (Dec. 2017).

[37] The Workshops sought views on the opportunities, challenges and risks of AI. Specific views were also gathered on Africa's emerging thinking on AI ethics and AI development in general within Africa's own political and economic trajectory.

[38] For instance at the workshop Toward a Network of Excellence in Artificial Intelligence for Development (AI4D) in sub-Saharan Africa, 3-5 April 2019, https://ai4d.ai/event/ssa-network/.

[39] See: The RIO Global Symposium on AI, 8-10 November 2017. http://cortexlogic.com/2017/11/18/global-symposium-on-artificial-intelligence-inclusion-rio-de-janeiro/.

[40] Alison Gillwald, 'South Africa is caught in the global hype of the fourth industrial revolution', *The Conversation*, 20 August 2019, http://theconversation.com/south-africa-is-caught-in-the-global-hype-of-the-fourth-industrial-revolution-121189.

so that regulation does not stifle the advancement of AI[41]. A recent global AI ethics assessment recommends the advancement of legislative efforts with iterative development based on research of social impacts[42].

From this, the following specific principles emerge:

❖ *Principle 1: Introduce safeguards to balance AI opportunities and risks*

A harmonised approach is necessary to protect individual and collective privacy rights in the cross-border flows of AI data.

AI offers opportunities in many areas, including: research and innovation, smart automation in core areas such as health (e.g. disease diagnostics for malaria, TB and more), clerical and office processing; agriculture (which employs 70 per cent of Africa's labour), energy, and tourism. AI also presents opportunities for more efficient public sector decision-making and resource allocation especially in social protection schemes, improved business analytics through increased data intelligence.

Regulatory safeguards to balance to opportunities and risks of AI, including its implications for employment and role of digital ID.

However, safeguards are needed to minimise risk, including:

- The **threat to employment**. For example, AI will increase the capital to labour ratio in diverse spheres of the economy, leading to job losses and shifts in labour market supply and demand.[43] It will also threaten taxation. For example, since labour currently counts for a significant portion of the tax base, the shift from labour to capital threatens the capacity of governments to fund an appropriate social safety net[44].
- The use increasing of **digital identity**, especially in social protection schemes, presents risks, especially as 'single source of truth' approaches threaten privacy and the citizen-government social compact. The social contract may need to be re-negotiated in the AI era, since digital ID changes how individuals and communities are governed in the African context.

❖ *Principle 2: Protect individual and collective privacy rights in cross-border data flows*

The collective rights of peoples and communities must be protected - in addition to the more standard provisions covering personal privacy. This can be achieved through any of the following ways:

Harmonise data and AI frameworks. The adoption of AI in Africa comes at a stage of implementation when data is more important than the technology itself. African countries have already been working on data issues and facing data protection challenges for some

---

[41] AI4D, 'A roadmap for artificial intelligence for development in Africa', 8 May 2019, https://ai4d.ai/blog-africa-roadmap/.

[42] Alexa Hagerty & Igor Rubinov, 'Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence'.

[43] Gerogios Petropoulos, ' Do we understand the impact of artificial intelligence on employment?', 27 April 2017, https://bruegel.org/2017/04/do-we-understand-the-impact-of-artificial-intelligence-on-employment/.

[44] Arthur Gwagwa, 'Ethical Approaches and processing for the global good -'a focus on the African Region', 2019, https://www.academia.edu/39703546/Ethical_Approaches_and_processing_for_the_global_good_-_a_focus_on_the_African_Region.

time, and a number are in the process of coming up with legal frameworks. To ensure better harmonisation, AI frameworks should cover all the constituent aspects of AI, from core technology through to data and domains. Further, in cases where a country has existing data or domain frameworks, these should be updated so as to harmonise with any subsequent continental-level AI framework rather than allow contradictions to be created.

Build the capacity to generate and use data. Frameworks should explore ways to make data available to those who need it most - especially businesses, since good aggregate data is often unavailable in emerging economies. In addition, the relevant stakeholders may need to be supported to develop the capacity (both technical and analytical) to make use of the data.

Protect personal data. However, caution should be exercised when merging personal and non-personal data. Security, confidentiality and integrity are essential requirements. Safeguards must be built in, especially when data is used for other purposes other than the primary one for which it was collected.

Safeguard cross-border data. Where such data crosses international borders, safeguards must be inbuilt, taking into account all implications of cross border data flows. Datasets need to be accompanied with a datasheet that documents their motivation, collection process, composition, recommended uses. Where possible, dataset be kept free and publicly available (with the necessary ethical provisions having been taken into account).

Protect collective rights in cross border data flows. It is essential for African businesses to extend their radius of trust in respect of cross-border data flows. Governments should also protect collective data rights when entering into agreements with foreign governments, in addition to ensuring informed consent from individual data subjects. This includes ensuring respect for collective and cultural rights.

Balance economic value against protection of open data. Open data can be useful for many purposes, such as diversified data sets, content innovation,[45] and benchmarking fairness. However, there should be safeguards in its use, and compensation required where it generates economic returns to its users. Big open data can be used for nefarious purposes, especially if the technology implementation prioritises profit and functionality ahead of ethical principles.

❖ *Principle 3: Define African values for AI and align AI frameworks with such values*

African countries and regions should define their own ethical values, and rely on these African AI values to inform policy, regulation, development and deployment of AI on the continent, as well to inform contributions to global AI ethics initiatives. **African ethical values for AI** adopt and contextualise those elements of best practice from global initiatives that accord with their own ethical values and cultural contexts.

[45] Arthur Gwagwa, 'Better algorithms and open data frameworks needed for inclusive and diverse online content,' *LinkedIn*, 8 November 2018, https://www.linkedin.com/pulse/better-algorithms-open-data-frameworks-needed-inclusive-arthur-gwagwa/.

AI project implementations should be preceded by **values and risk assessments**, along the lines of regulatory impact assessment practice.

Ethical values determine the type of AI that is created. And, what ethics means to Africa may differ from other regions. For instance, African cultures, despite their diversity, share certain cultural commonalities, such as the notion of 'Ubuntu', which encompasses a collective approach to life, along with a number of sentimental values and beliefs. Western and oriental values[46] may not accord with, or may even collide with African interests in the context of AI.

❖ *Principle 4: Practise fair and socially-responsible AI*

AI should be both fair and inclusive, taking into account the variations and granularities of the continent. This can be achieved in several ways.

Use open data sets for benchmarking fairness. The use of open data sets - with safeguards - offers the benefits of benchmarking fairness.

Include young people especially women and girls. For AI to be inclusive and socially beneficial, young people, especially women and girls, should be included in AI development - especially since Africa's population is very young and that girls have been historically disadvantaged under patriarchy.

Build environmentally-friendly and sustainable innovations. Fair and inclusive AI should be friendly to the environment and conscious of climate change for it to be sustainable.

Fair and socially-responsible deployment of AI needs to include and embody African values.

❖ *Principle 5: Build inclusive partnerships based on community and co-creation*

Build equal and beneficial partnerships. To achieve AI for good in Africa, partnerships between global technology companies and local stakeholders, such as developers, communities and governments, should include ethical safeguards. Once an ethical framework from a stakeholder process is in place, it needs to be evaluated and tested. Despite the benefits of partnerships, stand-alone African AI start-ups should also be encouraged, as this will allow Africans to decide and apply their own ethical standards and define their own AI future.

Promote multidisciplinary teams based on co-creation. Task forces of social scientists, statisticians, domain experts need to be formed to work together in creating ethical AI applications, and in evaluating the impact of AI on African communities.

Engage local communities. The ethics of co-creation needs to go beyond formal adherence to legal requirements, Community engagement needs to be more than a mere tick-box exercise. It involves understanding the specific needs of the people in Africa, adopting an approach based on dialogue with communities to make sure they are included in research,

[46] Cf: Meredith Somers, 'Are we encouraging the wrong AI?', Cambridge MA, Massachusetts Institute of Technology, 8 May 2019, https://mitsloan.mit.edu/ideas-made-to-matter/are-we-encouraging-wrong-ai.

and undertaking public engagement and meaningful consultation during which communities can inform the experts of their unique needs.[47]

❖ *Principle 6: Adopt an adaptive, open minded and humble approach*

The workshops from which this set of principles is derived, should be seen as a preliminary needs analysis based on early conversations. Further discussions on the issues are needed, especially to improve the understanding African concerns in global AI ethics discussions. The AI debate should also be approached with humility, and an open-minded attitude, and a willingness to learn. The ethics of AI in Africa is still a work in progress: only four African countries have nascent AI policies – Ghana, Kenya, Tunisia and Uganda. Some global technology companies are also adopting a cautious approach, acknowledging that this area is dynamic and evolving. For example, in its AI values statement, Google states that they will "approach [their] work with humility, a commitment to internal and external engagement, and a willingness to adapt [their] approach as [they] learn over time".[48]

Key Recommendations

- The existing codes of ethics and toolkits for AI may offer a good starting point. However, as important as algorithmic fairness is, it is crucial to avoid reducing ethics to a fairness problem alone. There is a need for inter-disciplinary approaches, including social science approaches, that consider issues beyond technical robustness and legal compliance – including the impact of AI on fundamental human rights and on collective social and ethical values[49] - to ensure that AI adheres to ethical principles and values[50].
- There is a need for political commitment and the development of a strong value proposition at public policy level in order to generate public interest in AI technologies, for example, through awareness raising, education, and training.
- Ethical AI requires the application of universal human values and international standards. However, it also needs to take into account Africa's historical peculiarities. It needs to use a development paradigm lens to treat Africa's social, economic and cultural rights on a par with more generalised civil and political rights. This requires a respectful and honest dialogue between the Global North and the Global South, especially at norm-setting forums such as the WTO.

We need further discussions on the ethics of AI in the context of Africa, along with a willingness to incorporate African concerns in global AI ethics discussions.

---

[47] Alexandra Hagerty, & Igor Rubinov, 'Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence', 2019, arxiv.org/abs/1907.07892.

[48] Google, 'AI at Google: our principles', https://www.blog.google/technology/ai/ai-principles/.

[49] Alessandro Mantelero, 'AI and Big Data: A Blueprint for Human Rights, Social and Ethical Impact Assessment,' Computer Law & Security Review, 34(4), 2018.

[50] Cf: EC, 'European Guidelines on Trustworthy AI', Brussels, European Commission, 2019, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

- AI policy formulation needs to include broader ethical and socio-cultural perspectives so to inform the current discussions on inclusion. This will assist to mitigate the risks, particularly where the ethical norms and values embodied in AI technologies collide with those of the communities in which they are delivered and deployed. In particular, there is a need to invite African views, some of which have been articulated above.

_____

For more RIA updates, sign up here and download full South Africa report Hyperlink

Author

Arthur Gwagwa: arthurgwagwa@gmail.com

Enquiries

info@researchictafrica.net

409 The Studios, Old Castle Brewery, Beach Road, Woodstock, Cape Town

T: +27 214476332

W: www.researchictafrica.net

This is policy brief was based on independent research and funded by Microsoft. The views and recommendations in this paper do not necessarily present the views of Microsoft on AI.