



Of Guardrails, Sandboxes & Human Centred Artificial Intelligence :

**Reflections, Notes
& Lessons for Africa**

by
Toyosi Akerele-Ogunsiji
Rise Networks, Lagos, Nigeria

risenetworks.org





OVERVIEW

What is Artificial Intelligence Safety?

AI Safety can be referred to as the specific methods and strategies employed in managing inherent risks and vulnerabilities associated with the utilization of Artificial Intelligence.





Introduction

In this age of information, where data reigns supreme, Artificial Intelligence [AI] has emerged as a pivotal force in shaping the future of global economies, societies, and governance. The first Global AI Safety Summit, hosted by the UK Government on the 1st & 2nd of November 2023, was not just a gathering of the world's foremost minds in AI technology and frontier AI Institutions but also a facilitator for international collaboration in charting the ethical, economic, and social contours of AI. Rise Networks, representing the burgeoning AI landscape of Nigeria, participated in this summit with the objective of contributing to and gleaning insights from the global discourse on AI to enrich Nigeria's perspectives and guide the country on the next steps for applicable open and responsible use of AI.

This Document is a combination of my personal notes based on my observations, learning points, and engagements with key actors in addition to the additional comments made by relevant stakeholders and industry leaders who participated in the Summit. It captures the key themes of the deliberations and outcomes from the Summit, offering a synthesis of key discussion points, themes, and emerging concerns as distilled by the Chairs of each roundtable session and the closing remarks made by the UK Secretary of Science, Innovation and Technology, Michelle Donelan. Through the lens of these discussions, my objective for curating these Reflection Notes is to illuminate the path of neo-AI centre points like Nigeria and Africa forward, balancing the deployment of rapid development of AI technologies with the duty of care and imperatives of global safety and ethical integrity. It is a call to action for a concerted, multidisciplinary effort to ensure responsible stewardship of AI, underpinning the advancement of human society.

Toyosi Akerele-Ogunsiji

Founder, Rise Networks

Lagos, Nigeria



SUMMARY OF OUTCOMES

Roundtable 1:

Risks to Global Safety from Frontier AI Misuse

Chaired by François-Philippe Champagne – Minister for Innovation, Science and Industry, Government of Canada, the dialogue scrutinized the safety implications of state-of-the-art AI technologies, acknowledging the escalating cyber and biosecurity threats. The conversation underscored the urgent need for collaborative efforts across all sectors to enforce safeguards and counteract the misuse of AI.

Yi Zeng, from the Chinese Academy of Sciences, led the conversation on the unforeseeable advancements in AI capabilities, stressing the necessity for rigorous development and testing protocols. The roundtable emphasized the dual nature of AI as a harbinger of both ground-breaking benefits and significant risks.

Roundtable 2:

Risks from Unpredictable Advances in Frontier AI Capability



Roundtable 3:

Risks from Loss of Control over Frontier AI

Josephine Teo's [Minister for Communications and Information, Government of Singapore] chairpersonship focused on the potential loss of human control over advanced AI systems, advocating for immediate, concrete measures to forestall such risks and maintaining a cautious approach toward the progression of AI capabilities.

Roundtable 4:

Risks from the Integration of Frontier AI into Society

Under the guidance of Marietje Schaake of the Stanford Cyber Policy Institute, participants evaluated the societal implications of AI integration, touching upon the threats to democratic institutions and global inequities. The discussion highlighted the necessity for continuous, context-aware technical assessments. I was quite apt in my submissions in this Session with my comments centred around championing a technology ethos that prioritizes societal well-being given that I believe that “Nations with lower data availability, broader penetration, lesser technological and human capability and prone to greater risk of the misuse of frontier AI if it ends in the hands of wrong people.



Roundtable 5: What should frontier AI developers do to scale capability responsibly?

Michelle Donelan, the UK Secretary of State for Science, steered the debate towards the responsibilities of AI developers in responsibly scaling AI capabilities, pointing to the convergence of corporate safety policies and the need for external regulatory benchmarks.



Roundtable 6: What should national policymakers do concerning the risks and opportunities of AI?

Rebecca Finlay's session underscored the role of national policymakers in harmonizing the dichotomy between the opportunities and risks posed by AI, advocating for innovative and agile governance models. Rebecca is the CEO for Partnership for AI. Here, I advocated for stakeholder management and human-centered AI specifically stating that "Technology that excludes the majority of its beneficiaries from contributing to its formation is precarious to human society. Hence, we must create safe spaces and identify strategies for inclusion of all classes of citizens to assess and scrutinize emerging technologies like AI to avoid putting humanity at great risk".



Roundtable 7: **What should the International Community do in relation to the risks and opportunities of AI?**

Tino Cuéllar from Carnegie Endowment emphasized the imperatives for international collaboration to address the shared challenges of AI, with a focus on developing a concerted approach to AI safety and equitable benefits distribution.

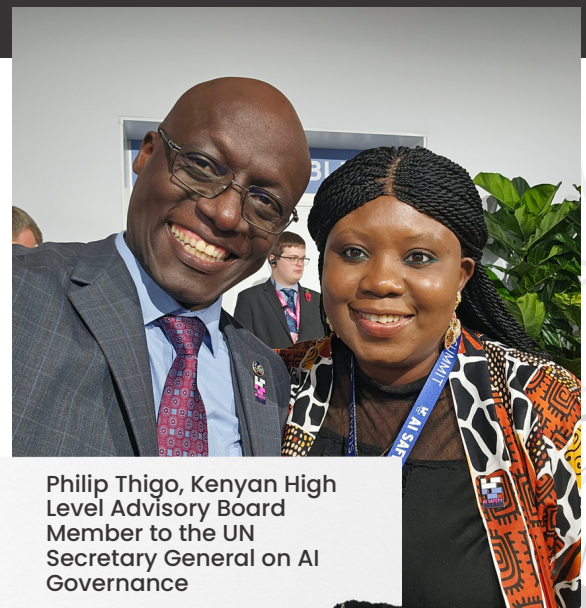
Chaired by Angela McLean, the Government Chief Scientific Adviser of the United Kingdom, emphasized the scientific community's responsibility in pioneering safer AI architectures and conducting thorough safety evaluations, thereby informing effective policy-making.

Roundtable 8:

What should the Scientific Community do in relation to the risks and opportunities of AI?



Yi Zeng, Professor of Brain Inspired AI, AI Ethics and Governance, Chinese Academy of Sciences



Philip Thigo, Kenyan High Level Advisory Board Member to the UN Secretary General on AI Governance



BROAD TOUCHPOINTS OF THE SUMMIT:

The Summit's focus underscored the necessity to understand and leverage AI for the greater good. The UK Government's initiative to host such an event underlines the critical role of AI as a key driver for innovation and economic growth. Rise Networks is grateful for the UK Government's invitation, seeing it as a recognition of the importance of incorporating diverse perspectives from emerging and promising technology nerve centres like Nigeria and other key nations in Africa. It stands as a compelling call to action for countries to pull their intellectual and strategic resources to steward AI toward beneficial ends.



Learning Outcomes and Peer Reviews:

Global Insights and Shared Ambitions

The summit provided an invaluable platform for countries to share their strides in AI. The UK Prime Minister's vision that "the UK is to be a science superpower," reinforced the significance of robust AI policies. His assertion that "AI could add £232 billion to the UK economy by 2030" is not only a projection of economic growth but also a clarion call for all to harness AI responsibly for societal benefits.



Karim Beguir, Co-Founder and CEO of Instadeep



Tino Cuellar, President, Carnegie Endowment



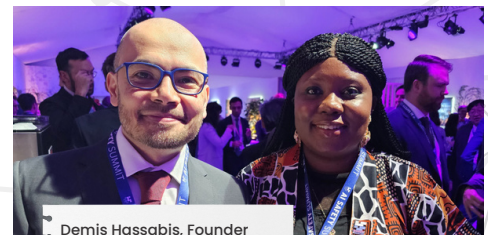
Ahmet Yozgatigil, Deputy Minister of Industry and Technology of Turkey



Josephine Teo, Minister of ICT, Singapore



Elon Musk, Founder of Tesla. Owner of Twitter.



Demis Hassabis, Founder and CEO of Deepmind



Michelle Donelan

the UK Secretary of State
for Science and Tech

we were reminded of the transformative potential of AI with its catalytic ability to create a springboard for opportunities and equality. Her emphasis on "Reducing bureaucracy and paperwork" and introducing an "AI-powered assistant for teachers to help create lesson plans," showcased innovative approaches to empowering educators as well. The multi-stakeholder involvement she advocated for is key to unlocking AI's potential in personalized education.





His Excellency **Omar Sultan Al Olama**

the UAE Minister of State for Artificial Intelligence,
Digital Economy, & Remote Work Applications,

Emphasized the importance of collaboration over competition in the AI landscape. His call to “do more peer-to-peer reviews across countries” is a testament to the UAE's forward-thinking approach to AI development and safety measures. By advocating for a system where “if there are already working models of AI inventions in one country, there’s no need for another country to start the same from scratch, we can learn from each other and localize the solutions for our immediate environments”, he laid out a vision for a more efficient and interconnected global AI community. This approach not only promotes knowledge exchange but also ensures that resources are optimized and advancements in AI are leveraged to their maximum potential across borders.



Ivan John Uy

Philippines Information and
Communication Technology Secretary

The Philippines is carving its niche in the AI sphere as mentioned by Ivan John Uy, the country's Information and Communication Technology Secretary. The country's focus on "accelerating digital skills, upskilling, and reskilling" reflects an understanding of the foundational role that human capital plays in the digital economy. The protection of cyberspace and the pursuit of an all-encompassing strategy for AI integration speak to the Philippines' recognition of the multifaceted nature of AI's impact on society. The Philippines' EUREKA policy model – with its focus on "Empowerment through Education, Universal Access of Digital Access, Responsible AI, Ethical Innovation Design, Knowledge-Based-Society and Agile Governments" – illustrates the Philippines' comprehensive parallel approach between AI policy and human capital growth. It's a blueprint that not only aims to enhance digital infrastructure and emerging technologies but also underpins the necessity of the investment in any Nation's most valuable resource – its people. The model's emphasis on agile governance and the central role of the DICT (Department of ICT) in implementing policy, highlights the dynamic nature of AI strategy and necessitates governments to be as adaptive and innovative as the technologies they seek to regulate and foster.

Ivan John Uy, Minister of ICT,
Philippines





Kamala Harris

Vice President of the
United States of America

The Summit also highlighted the United States' commitment, with Vice President Kamala Harris underscoring the importance of "bridging the digital divide and ensuring that technology serves the people." This statement aligns with the ethos of Rise Networks, where the social implications of AI are paramount. The President of the European Commission, Ursula von der Leyen, brought attention to the EU's approach to AI, focusing on ensuring that it "works for the people," resonating with our Rise Networks' vision and commitment to responsible and human-centered AI. In the field of entrepreneurship and innovation, Elon Musk's insights calls our collective conscience to action on the critical nature of "innovation without sacrificing safety and ethical considerations."



Dr. Bosun Tijani

Minister of Communications,
Innovation, and Digital Economy
of Nigeria

Shared valuable perspectives on the integration of AI into our National Development Strategies, spotlighting the need for infrastructural readiness and policy frameworks conducive to AI growth in African nations. In addition to that, he underscored "Nigeria's role as a nation of innovators and entrepreneurs", emphasizing the significance of harnessing Artificial Intelligence for the benefit of humanity and affirmed our Nation's appreciation to the UK Government for including Nigeria in the UK AI for Development Programme, which will bolster the country's AI research capabilities, the development of inclusive datasets and the National AI Research Scheme".



Lee Jong-ho

South Korean Science
and ICT Minister

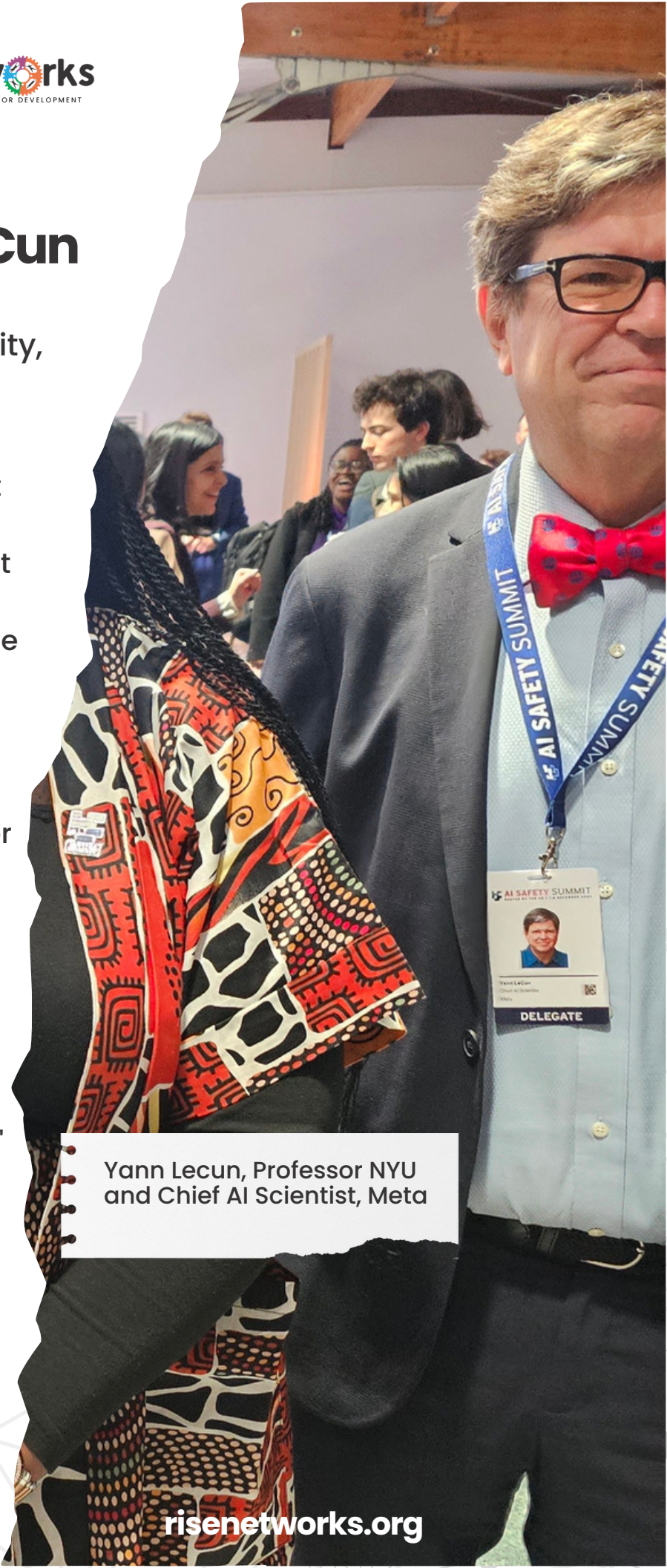
in his closing remarks, highlighted the crucial need for continuity and action. Minister Lee emphasized the "importance of sustaining the momentum generated by the Summit", recognizing that platforms for international dialogue on AI safety and responsible use are vital for progress. His commitment to this cause is evidenced by South Korea's prompt decision to host the next mini AI Safety Summit in the next six months, reflecting a dedication to actively building upon the collaborative efforts of the global community as already initiated by the UK Government. This stance on sustainability and ethical practice in AI underlines a shared global responsibility, and it echoes the need for ongoing engagement, assessment, and reinforcement of the agreements and insights that have emerged. By maintaining the steady pace of these conversations, Minister Lee reaffirms the "collective goal of safe and ethical AI", ensuring that the inaugural summit's pioneering work continues to resonate and evolve in the forthcoming sessions.



Professor. Yann LeCun

Leading figure in AI research,
Professor at New York University,
and Chief AI Scientist at Meta

Amidst the discussions at the Global AI Summit, a salient point made by Yann LeCun, a leading figure in AI Research, Professor at New York University, and Chief AI Scientist at Meta, emphasized the crucial role of open-source AI platforms. In our brief exchange on X (formerly Twitter), LeCun asserted that open-source AI serves as a critical foundation for global technological empowerment, particularly for developing nations. His statement, "The availability of open-source AI platforms is crucial for everyone, but I agree with you that it is particularly crucial for developing countries," highlights a shared understanding that access to these technologies can be a game-changer for countries striving to enhance their AI capabilities especially for social impact and bias control.



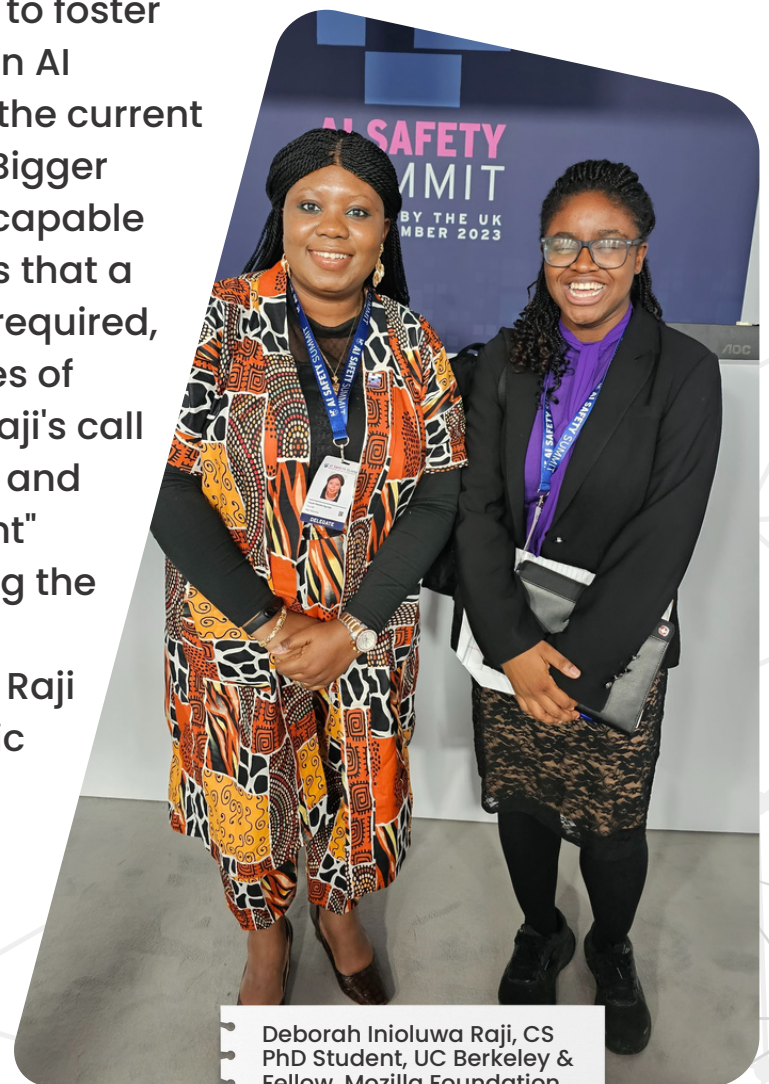
Yann Lecun, Professor NYU
and Chief AI Scientist, Meta



Deb Raji

Fellow of the Mozilla
Foundation

A vital contribution to this discourse came from Deb Raji from her X (formerly Twitter) handle, a Fellow of the Mozilla Foundation, whose stance on open-source AI models adds depth to our understanding of responsible AI deployment. Raji asserts that "open-source models are not antithetical to guardrails and caution." This perspective reinforces the notion that open-source AI can coincide with rigorous evaluation and data governance practices, emphasizing the potential for open-source models to foster transparency and accountability in AI development. He also challenges the current framing of AI regulation, stating, "Bigger models are not necessarily more capable or more dangerous." This suggests that a nuanced approach to AI policy is required, one that does not conflate the sizes of AI models with their risk profiles. Raji's call for consequences for "unsafe use" and "inappropriate model development" while cautioning against restricting the openness of AI models provides a balanced view. Openness in AI, as Raji notes, is "our only avenue for public accountability at the moment."



Deborah Inioluwa Raji, CS
PhD Student, UC Berkeley &
Fellow, Mozilla Foundation



As we synthesize the insights from the Global AI Summit, the Mozilla Foundation's statement on AI safety and openness also provides a foundational lens through which we can evaluate and strategize the future of AI governance. The Foundation asserts that, "To mitigate current and future harms from AI systems, we need to embrace openness, transparency, and broad access," championing this as a global imperative. The stance is clear: openness should not be viewed as a vulnerability but as a catalyst for safety and accountability. The Foundation cautions against the misconceived security offered by proprietary control and underscores the potential dangers of hastily conceived regulations that could centralize power and stifle innovation. Emphasizing the role of openness in enhancing understanding, and scrutiny, and lowering barriers to responsible AI, the statement reinforces our orientation at Rise Networks' vision of a collaborative AI future. The collective signatories of the statement, with their myriad perspectives, united in the belief that an open, responsible, and transparent AI landscape is pivotal for ensuring a secure and equitable AI era. This shared conviction reinforces the direction in which global AI policy and development should move, promoting a balance between innovation and the safeguarding of public welfare in the age of AI.



Megan Smith, Algorithmic Justice League and former 3rd CTO of the United States

Philip Thigo, Kenyan High Level Advisory Board Member to the UN Secretary General on AI Governance



Paula Inagbire, Rwanda's Minister of ICT and Innovation



Rise Networks' recommendations:

Shaping AI from an African perspective

Drawing from the Summit's reservoir of knowledge and the partnerships formed across the board, Rise Networks proposes several key recommendations to drive Nigeria's and Africa's progress within the global AI Landscape:

Enhancing Education with AI:

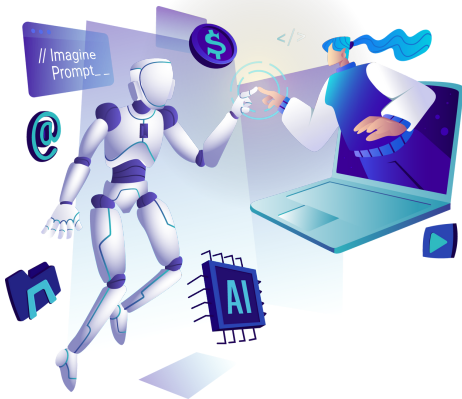
For Nigeria and Africa, the integration of AI in education must extend beyond administrative efficiency to redefine pedagogical frameworks. By leveraging AI, educators can tailor curricula to individual learning styles, predicting and addressing learning challenges before they hinder student development. This transformative approach can democratize education, making quality learning accessible in even the most remote regions. It is a step towards dismantling the barriers of educational inequality, preparing a generation that is AI-savvy and ready to contribute to the continent's growth.





A Robust AI Infrastructure:

Developing a robust infrastructure is pivotal for AI's growth in Nigeria and across Africa. This involves not just physical infrastructure but also the digital architecture that will support the surge in AI-driven solutions. It means creating cloud services, data centers, and high-speed connectivity that can handle the enormous data throughput AI applications demand. Establishing such an ecosystem will act as the backbone for innovation, attracting global AI investments and partnerships.



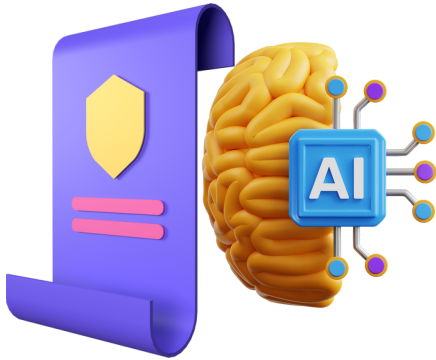
3 Ensuring AI Safety and Ethics:

The imperative of AI safety transcends ensuring that systems do what they are supposed to do—it involves instilling ethical principles at the core of AI development. For Nigeria and Africa, this means establishing a set of governance protocols that align with local cultural values and international human rights standards. These protocols should guide the ethical deployment of AI, ensuring that these systems enhance societal welfare without infringing on individual freedoms.





Creating Inclusive AI Policies:



Inclusivity in AI policy-making must reflect the diverse fabric of Nigerian and African societies. Policies should be designed to enable equitable access to AI technologies, ensuring that all segments of society, particularly the underserved, can benefit from AI advancements. This approach will help avoid the exacerbation of existing inequalities and promote a more harmonious integration of AI into the societal fabric.



Fostering Multi-Stakeholder Collaboration:

AI's potential can only be fully realized through a concerted effort involving all sectors—government, private sector, academia, and civil society. Nigeria and Africa's strategy should be to cultivate an ecosystem where these sectors collaborate on AI projects, from research and development to deployment. This will ensure that AI solutions are well-rounded, addressing the myriad challenges faced by the continent.





Committing to AI for Social Good:

AI should be harnessed to tackle some of Nigeria and Africa's most pressing societal challenges, from healthcare delivery to agricultural productivity. AI could offer predictive analytics for disease outbreaks, personalized learning systems, and optimized resource distribution. The focus should be on deploying AI in ways that make a tangible difference in the everyday lives of citizens.



Human Capital Development in AI Ethics:

To engage effectively with AI, Nigeria, and Africa must develop human capital with a strong understanding of AI ethics. This involves integrating ethics into the core curriculum of STEM education and supporting interdisciplinary research that explores the societal impacts of AI. Professionals trained in this way will be pivotal in guiding AI development in an ethical direction, ensuring that AI serves the common good.





Emphasizing AI Model Transparency & Accountability:

There is a crucial need for transparency and accountability in AI model development, especially pertinent for Nigeria and Africa. This involves creating mechanisms for monitoring AI development processes and ensuring that AI models are transparent in their operations and decision-making. Such transparency is essential not only for building trust among the public but also for enabling experts to assess and mitigate potential risks associated with AI technologies. Accountability mechanisms should be established to hold developers and users of AI systems responsible for their outcomes, particularly in cases where AI decisions significantly impact individual lives or societal norms. This approach aligns with global efforts to foster responsible AI development, ensuring that AI systems are not only advanced in their capabilities but also aligned with ethical standards and societal values.





International Collaboration and Dialogue in AI Safety:

International collaboration and sustained dialogue are vital for advancing AI safety and ethical practices in Nigeria and Africa. This approach entails engaging in global forums, sharing best practices, and learning from the successes and challenges of other nations in AI governance. The continent can benefit significantly from such international exchanges, gaining insights into effective regulatory frameworks, safety protocols, and ethical guidelines that align with the unique African context. By actively participating in these global dialogues, Nigeria and other African countries can ensure that their voices and perspectives are represented in shaping the global AI narrative. This engagement is crucial for developing AI policies and strategies that are not only locally relevant but also globally informed, contributing to the broader discourse on creating a safe and equitable AI future for all.

Fostering AI Literacy:



AI literacy should be a priority, equipping the populace with the knowledge to engage with AI technologies critically. This involves not just understanding how AI systems function but also being able to question their decisions and demand accountability. By fostering AI literacy, Nigeria and Africa can cultivate a society that is not only adept at using AI but also vigilant about its implications.



Conclusion

The Global AI Summit has provided a rich tapestry of insights and recommendations that Rise Networks and Nigeria can leverage to forge a path toward a responsible and equitable AI future.

By embracing these global insights and tailoring them to Nigeria's unique context, we can catalyse an AI-driven transformation that is not only economically beneficial but also socially responsible and inclusive. As we march into the future, the Summit's lessons will serve as beacons to guide our strategies, ensuring that Nigeria plays a pivotal role in the global AI narrative.

Toyosi Akerele-Ogunsiji
Founder, Rise Networks
November, 2023 Lagos,
Nigeria

WITH NOTABLE COMMENTS BY:

Dimeji Sowole
AI Team Lead, Rise Networks

Daniel Ekhaifo
Chief Economist, Nigeria
Liquified Natural Gas [NLNG]





References

1. <https://www.gov.uk/government/publications/ai-safety-summit-1-november-roundtable-chairs-summaries/ai-safety-summit-2023-roundtable-chairs-summaries-1-november--2>
2. <https://www.gov.uk/government/speeches/prime-ministers-speech-at-the-ai-safety-summit-2-november-2023>
3. https://www.linkedin.com/posts/dr-%E2%80%98bosun-tijani-1b027b_this-week-i-had-a-number-of-engagements-activity-7126625532384845824-3ZXK?utm_source=share&utm_medium=member_desktop
4. <https://drive.google.com/file/d/1vpvS6TNQZ3GNBYjtsdAFKFBk5Na6rwAe/view>
5. <https://twitter.com/ylecun/status/1721191000280813714>
6. <https://twitter.com/rajiinio/status/1720367295539638481>
7. <https://open.mozilla.org/letter>
8. <https://www.gettyimages.com.au/editorial-images/news/event/government-hosts-ai-safety-summit-2023/776056543>
9. <https://www.alamy.com/washington-vereinigte-staaten-19th-sep-2023-dr-yann-lecun-vice->





Of Guardrails, Sandboxes & Human Centred Artificial Intelligence : Reflections, Notes & Lessons for Africa

