Microsoft

AI in Africa

Meeting the Opportunity

Table of Contents

Governing AI in Africa:
Policy Frameworks for a New Frontier

Lillian Barnard,
President,
Microsoft Africa

Foreword by Lillian Barnard

Embracing the transformative power of artificial intelligence (AI) is central to our vision for Africa's future. AI will revolutionize the way we do business across the African continent. This revolution transcends mere technological advancement; it marks a strategic turning point wherein AI will be leveraged as a catalyst for sustainable growth and chart a course towards a future brimming with innovation and boundless opportunity for Africa.

Indeed, it is forecasted that AI could have a significant impact on Africa's economic prospects. A recent report suggests that the technology could increase Africa's economy by a remarkable $1.5 trillion – a figure that equals half of the continent's current gross domestic product (GDP) – if African businesses could capture 10% of the global and rapidly expanding AI market. This isn't just a figure; it represents the possibility for AI to make a genuine difference.

On a continent that already has the youngest population in the world, and will soon be home to one in five of the world's consumers, the ability to leverage such a powerful tool to help businesses understand and better cater to the unique needs of Africa's burgeoning consumer base through data analytics could be game changing. It could also empower a young workforce with AI-driven educational tools, enhancing their skills and future employability. From agriculture to healthcare, AI is expected to enable faster and more profound progress in nearly every field of human endeavor and help address some of society's most daunting challenges.

The potential for widespread transformation in Africa is only just beginning to be realized. In health care, for example, AI is expected to improve clinical experiences and support better healthcare delivery. Already, in Rwanda and Ghana, AI-powered diagnostic solutions are being deployed to improve medical imaging analysis. These tools assist in the early detection of diseases such as breast cancer and tuberculosis. Companies like Zipline are using AI to optimize drone delivery routes for medical supplies and vaccines, significantly enhancing access to essential medications in remote areas. The impact of these AI applications in the health care sector has been predicted to be profound, having the potential to lead to earlier diagnosis and treatment, improved patient outcomes, and an overall strengthening of health care systems.

In the agricultural sector, AI is emerging as a supportive tool that can aid in managing resources more efficiently. For instance, by providing data-driven recommendations for fertilizer application, AI has the potential to help farmers reduce unnecessary expenditure and increase profitability, thereby enhancing their economic resilience. In countries like Nigeria and Kenya, startups such as UjuziKilimo and Farmcrowdy are using AI to analyze soil and weather data to provide farmers with personalized advice on crop cultivation and optimized usage enabling them to make evidence-driven decisions and improving crop yields. Ultimately, this has led to increased productivity and enhanced food security, directly benefiting smallholder farmers and the broader community.

Education, finance, urban development, and city planning are other key areas where AI can offer innovative solutions customized to Africa's specific needs. In South Africa, where the past 10 years have been characterized by substantial investment in AI technology, local government municipalities such as the eThekwini Municipality in KwaZulu-Natal have leveraged AI to become a data-driven smart city and provide pragmatic and innovative water and sanitation services to its growing population, which other municipalities have started to emulate.

Already, we have witnessed AI play a crucial role in fostering financial inclusion. In East Africa, particularly in Kenya and Tanzania, mobile banking platforms like M-Pesa are leveraging AI to offer credit scoring and microloan services. AI algorithms analyze transaction data to assess creditworthiness, enabling a broader range of individuals and small businesses to access financial services. This AI-driven approach to finance has revolutionized banking in the region, providing financial services to the unbanked and underbanked populations, and fostering economic growth and stability. As an early adopter of AI, the financial services industry in South Africa is exploring various applications for the technology. For instance, Investec in South Africa has created ZebraGPT, a tool designed to streamline repetitive tasks, explore new ways of working, and boost overall productivity. To date, the company has identified around 64 use cases for generative AI in the wealth and investment space alone.

At Microsoft, we are witnessing the AI revolution firsthand through our collaborative efforts across the continent. In Egypt, Microsoft is working with government, business, and startups to implement AI solutions to reimagine government services, expand business, and instill innovation in the society. In Kenya, Microsoft is working with leading partners in food and water security and wildlife conservation to help address climate and sustainability challenges. In Morocco, AI is being used to improve water conservation efforts, a critical resource for the country's agriculture and human consumption needs. In South Africa, where AI utilization is already advanced, Microsoft's partnership with key players in the public and private sectors are reshaping public service delivery and addressing multifaceted business and societal challenges, particularly in the areas of health care system optimization and innovative urban infrastructure management.

However, amid this wave of excitement and potential, Africa is faced with a critical mandate: As we navigate this AI-powered future, our journey must be underpinned by responsible and sustainable innovation, ensuring that our progress remains aligned with the human values, societal norms, and the needs that define Africa's vastly diverse cultures. This involves a deep engagement with the continent's unique challenges, recognizing their complexity and prioritizing those that AI could help solve. Collaborating with relevant stakeholders will be key to ensuring that AI solutions are not just technologically advanced, but also culturally attuned and genuinely beneficial to African societies.

A thoughtful approach to regulation is also key to ensuring that our exploration of AI's vast potential is conducted with responsibility and ethical foresight, taking Africa's unique socioeconomic and cultural context into account. Legislation should consider issues such as protecting data privacy, ethical AI use, and minimizing bias with the utmost care while simultaneously nurturing the opportunity this revolutionary technology presents.

Given the varied stages of AI adoption and regulations across the continent's 54 countries, a cooperative and thoughtful approach is needed to standardize data and address these challenges effectively. As AI usage expands, more African countries are expected to introduce regulations to guide its development and deployment responsibly.

A crucial step towards countries effectively developing, deploying, and managing AI technologies and paving the way for technological advancement and economic growth relies on the adoption of cloud-first policies that can provide the scalable, cost-effective infrastructure essential for handling AI's computational demands and large data sets. Cloud computing democratizes access to advanced AI technologies, while fostering an environment of innovation and collaboration. With a view to providing more organizations with the capacity to drive progress, Microsoft has invested in datacenters on the continent, in Cape Town and Johannesburg, as well as in five edge sites and subsea cables and gateway colocation deployments.

We are not just a technology provider, but a partner on this journey dedicated to accelerating the continent's trajectory towards an AI-enabled future.

Already, as part of our strategy to enable AI to truly benefit all people, Microsoft has embedded large foundational AI models in our services including Bing, Office 365, Dynamics 365, Power Platform, GitHub, and LinkedIn. In addition to making AI available within the products people are familiar with, Microsoft has developed a range of AI Copilots to help users in specific areas. These virtual assistants will augment the work that people do by freeing up time for more creativity, imagination, and human ingenuity. It's these savings that can lead to significant innovation leaps and give people more time to solve pressing challenges at work or even in their community.

However, as we recognize the immense potential of AI in Africa and celebrate our continued partnership with various governments and organizations in their digital transformation journeys, we are also keenly aware of the potential

challenges and complexities that lie ahead. With responsible regulation and collaborative effort, AI can be a force for good–enriching lives, empowering communities, and steering the continent towards a future filled with limitless possibilities. Our vision extends beyond economic growth and technological advancements; we aim for a sustainable, inclusive future where everyone can benefit from the digital revolution.

To this end, our commitment to leverage AI for societal benefit extends beyond technology development to include capacity building and education. Over the past five years, Microsoft has helped upskill over four million young people across Africa through a variety of skilling and employability programs. Recognizing the need to nurture the next generation of African innovators, we are investing in training and resources to develop digital skills and AI literacy. This effort is encapsulated in our recent AI Skills Initiative, which offers new and free courses for workers to learn about how to harness the power of AI and drive Africa's growth in the digital age.

In parallel with our educational initiatives, Microsoft's multidisciplinary teams, comprising researchers, engineers, and policy experts, are committed to collaborative partnerships with governments and policy stakeholders. Our goal is to jointly identify and tailor policy interventions for cloud and AI technologies, ensuring they align with the specific needs and policy landscapes of African countries. This collaborative approach helps us integrate local insights into our technological initiatives, fostering solutions that are both effective and contextually relevant.

While we don't have all the answers to the questions that this new AI era brings with it, Microsoft's commitment to Africa is unwavering.

We are dedicated to being a responsible steward of AI technology, ensuring that it is developed and used in ways that benefit all Africans, and we look forward to continuing to engage with policymakers, business leaders, grassroot movements, and community stakeholders across the continent to help shape a future where AI is a tool for empowerment, innovation, and positive change.

The AI revolution in Africa is not just a possibility; it is already underway. With responsible regulation, partnerships, and a steadfast commitment to responsible and ethical AI, we are poised to unlock a future with unprecedented opportunities. We look forward to walking this journey with you and leveraging AI to create a continent that is more connected, empowered, and forward-looking than ever before.

By Lillian Barnard,
President, Microsoft Africa

Part 1:
AI Explained

AI is changing how we work, live, and play. Microsoft AI provides billions of intelligent experiences every day to people using our business software and services, gaming on our Xbox platform, and in keeping organizations secure. Our AI tools and technologies are designed to benefit everyone at every level in every organization. They're used in workplaces, home offices, academic institutions, research labs, and manufacturing facilities around the world. They're helping everyone from scientists and salespeople to farmers, software developers, and security practitioners.

What exactly is AI?

Using mathematics and logic, an artificially intelligent computer system simulates the reasoning that humans use to learn from data and make decisions.

Algorithms are at the core of every AI system. They are sets of instructions or rules that the system follows to process and analyze data. Different algorithms are used for different tasks, such as recognizing images or processing language, for example.

An AI model is trained on a dataset representative of the task or problem it's being developed to solve.

Data could be in the form of unstructured data, such as text, images, or audio—or structured data that can be stored in a standardized format.

The AI model uses the patterns and relationships it discovers in the data to adjust and optimize its performance resulting in a learned representation of the training data, which can be used to interpret and make inferences from new input or data.

One of the techniques used to train AI is machine learning. It's akin to teaching a computer program to recognize patterns by showing it examples. Just as we learn from experience, machine learning algorithms learn from data.

They look for similarities and relationships in the data to perform tasks such as making predictions, recommendations, classifications, and more.

Imagine you want to train a computer program to recognize cats. You would feed a machine learning algorithm a large number of pictures, each labelled as to whether the image contains a cat or not. By analyzing these labelled examples, the algorithm can identify patterns such as shapes, colors, and textures associated with cats, and create a model based on this learning. Once trained, the model can then be used to analyze new, unseen images and predict whether they contain a cat or not. The more examples the algorithm sees, the better the model becomes at recognizing cats accurately.

A subset of machine learning under the broader umbrella of AI is reinforcement learning. It involves asking an AI system to make a sequence of decisions to maximize a reward. The AI system learns through trial and error, receiving feedback in the form of rewards—or penalties— based on its choices. By learning which actions lead to higher rewards, the AI system improves its decision-making over time.

AI development has also been inspired by the structure and functioning of the human brain. Neural networks are computational models that consist of interconnected nodes called neurons, organized in layers. Each neuron receives input, performs a computation, and passes the result to the next layer. Neural networks excel at pattern recognition and can learn complex relationships in data.

As you can see, AI is a wide umbrella under which sits a range of different but complementary computer science fields and sub-fields—many of which have been in research for decades. But whichever approach is used to develop an AI system, having access to data and computing power to train your AI model are fundamental building blocks, and its recent advancement is mostly facilitated by the availability of large datasets and computing resources to train models with large chunks of data in a relatively short time.

AI-powered possibility

AI enables machines to carry out a wide range of tasks, many of which weren't possible before its advent.

Natural language processing, for example, extracts meaning from text or speech data. It enables computers to interpret language and perform tasks like sentiment analysis, language translation, and text generation. AI-powered computer vision similarly enables machines to "see" and decode the visual world around us. It involves techniques such as object detection and image recognition and has a wide range of applications from medical imaging to autonomous vehicles.

These capabilities and more are transforming how we use, benefit from, and interact with computer technology. AI can serve as a "copilot" to augment our human ingenuity and creativity through automating tasks and providing new insight.

Over the last decade, the AI field has made significant progress on perception and language tasks. There have been recent advances in the form of generative AI—a class of AI models that can generate new content such as text, images, code, and more. Generative AI is underpinned by a class of large-scale models known as foundation models.

Foundation models are trained on massive amounts of data and are capable of performing a wide range of tasks. With a simple prompt like "describe a scene of the sun rising over the beach," generative AI models can output a detailed description or produce an image, which can then be animated or even turned into video.

Many recent generative AI models such as language models are not only good at generating text but also generating, explaining, and debugging code. GitHub Copilot, for example, leverages OpenAI's Codex model to assist developers in writing code.

Another example of AI systems in action is the new AI-powered Bing search engine. The web search experience often involves the time-consuming task of reviewing and synthesizing information from a variety of sources identified from different search queries. Now, Bing can do the heavy lifting for you, working behind the scenes to make the necessary queries, collect results, synthesize the information, and present a single complete answer.

Time taken in months to reach 100 million monthly average users

ChatGPT   TikTok   Instagram   Pinterest   Spotify   Telegram   Uber

0   10   20   30   40   50   60   70
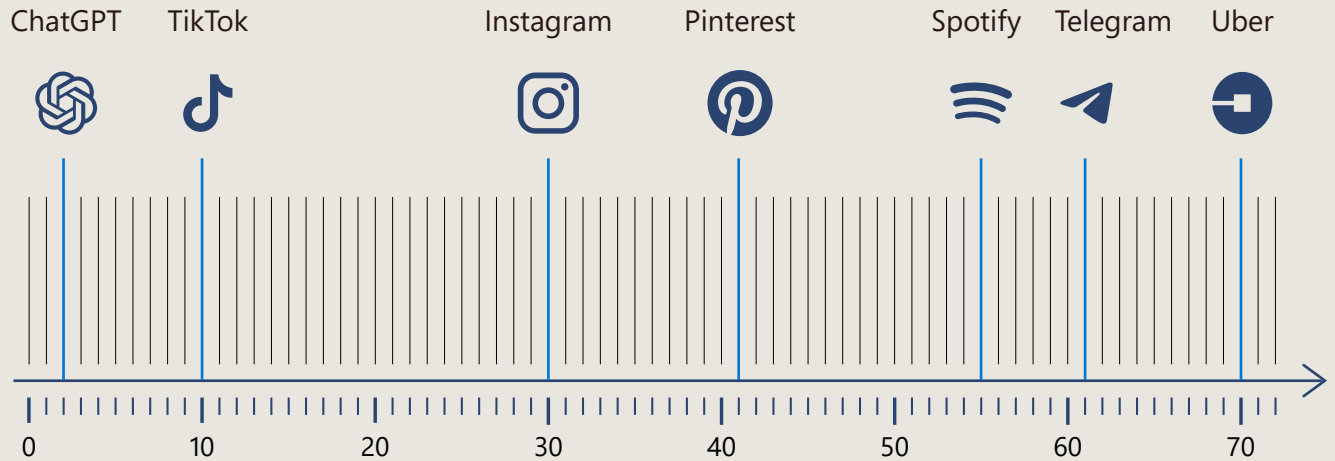
Source: Eyerys

Developing and scaling AI

More than a decade ago, Microsoft forecast an exponential growth in demand for AI systems and started to build special computing infrastructure to handle it. Inspired by early research developments, we've integrated large-scale language models in services ranging from Microsoft Bing to Microsoft 365, and have enabled other companies to take advantage of these technologies through Azure AI and the Azure OpenAI Service.

We believe that every organization in the world should benefit from the power of large-scale AI models. We've developed platforms, tools, and a supercomputing infrastructure that would allow any developer to build and scale their own AI innovation.

Through Azure OpenAI Service, we provide businesses and developers with access to high-performance AI models, such as GPT-4, Codex, and DALL·E 2. The service is based on the same platform we use to power AI models in our own products, including GitHub Copilot, Power Platform, Microsoft Designer, Bing search engine, and Edge browser.

We continue to evolve our AI infrastructure based on feedback and insights from training and serving AI models at scale, and our teams are working in lockstep with industry partners on the design of processors, networks, and datacenters optimized for AI.

Microsoft is committed to building Azure into an AI supercomputer for the world, serving as the

foundation of our vision to democratize AI as a platform. We pushed the frontier of cloud supercomputing technology, announcing our first top-five supercomputer in 2020, and have subsequently constructed multiple AI supercomputing systems at massive scale.

The convergence of innovations in infrastructure, machine learning acceleration software, platform service, and modelling powered by cloud technology has created the perfect conditions to accelerate innovation in AI and enable every company to become an AI technology company.

Why the cloud is the critical infrastructure behind AI:

1. Training AI models requires both large datasets and AI-optimized cloud computing infrastructure.

2. Cloud-based AI platforms make it easier for organizations to develop their own AI applications.

3. Cloud computing resources give organizations of all sizes a cost-effective way to run and scale AI innovation.

Part 2:
New steps toward
democratizing AI

The progress toward democratizing AI and the opportunities it affords marks a transformative shift, extending AI's impact beyond the often-arcane realm of data scientists and developers in the back office to everyone in our communities.

Generative AI, in particular, represents a major leap for the democratization of technology, helping to eliminate the need for specialized AI knowledge to harness AI capabilities. Through the use of natural language, more people are able to benefit from powerful AI tools. Large Language Models (LLMs) have the capability to provide 'no code' services, which empower people to be able to interact with AI systems without coding and to create results without a need for technical expertise. By simply asking a question in one's native language, people can create content and summarize data effortlessly, broadening the audience for AI from IT experts to anyone familiar with a browser and a connection to the internet.

It is key to recognize that internet connectivity and digital literacy are essential to truly achieving democratic access. Microsoft recognizes this imperative and is actively working to bridge gaps in these areas. Through initiatives like the Microsoft Airband Initiative, Microsoft endeavors to make affordable broadband accessible to underserved global communities, having enabled high speed internet access for over 63 million people since 2017. This includes over 15 million people in Africa alone. We have also committed to the goal of connecting 250 million people by the end of 2025, 100 million of which are in Africa.

Microsoft's investments in skilling programs like the AI Skills Initiative, offering courses on Generative AI on LinkedIn Learning, have already reached 80 million people worldwide. The Microsoft AI Skills Initiative includes new, free coursework developed with LinkedIn, including the first Professional Certificate on Generative AI in the online learning market, which is part of our Skills for Jobs program. The Initiative also includes a new open global grant challenge created in coordination with data.org to uncover new ways of training workers on generative AI and provide greater access to free digital learning events and resources for everyone to improve their AI fluency.

Identifying other critical areas for AI to positively impact African citizens, Microsoft's AI for Good Lab extended its presence to the continent in 2023 with the aim of working with local governments, local and global nonprofit organizations, and startups to tackle climate and sustainability challenges on the continent.

Consistent with our investments in connectivity, skilling, and sustainability, Microsoft is focused on fairness and inclusion—two key principles that help to form our responsible AI commitment. We also collaborate with researchers, civil society groups, governments, and international agencies to inform our efforts to align AI development and use with diverse community interests. We seek feedback to help ensure that, as we achieve more equitable access, we are prepared to support meaningful use among individuals and groups with different backgrounds and interests.

Part 3:

Governing AI in Africa: A legal and regulatory five-point blueprint for the future

Around the world, governments are looking for or developing frameworks to govern AI. Of course, there is no single or right approach. We offer here a five-point approach to help advance AI governance more quickly, based on the questions and issues that are pressing to many. Every part of this blueprint will benefit from broader discussion and require deeper development. But we hope this can contribute constructively to the work ahead.

This blueprint recognizes the many opportunities to use AI to improve people's lives while also quickly developing new controls, based on both governmental and private initiatives, including broader international collaboration. It offers specific steps to:

- **Implement and build upon new government-led AI safety frameworks.**
- **Require effective safety brakes for AI systems that control critical infrastructure.**
- **Develop a broader legal and regulatory framework based on the technology architecture for AI.**

- **Promote transparency and ensure academic and public access to AI.**
- **Pursue new public-private partnerships to use AI as an effective tool to address the inevitable societal challenges that come with new technology.**

More broadly, to make the many different aspects of AI governance work on an international level, we will need a multilateral framework that connects various national rules and ensures that an AI system certified as safe in one jurisdiction can also qualify as safe in another. There are many effective precedents for this, such as common safety standards set by the International Civil Aviation Organization, which means an airplane does not need to be refitted midflight from Cairo to Cape Town.

Working towards an internationally interoperable approach to responsible AI is critical to maximizing the benefits of AI globally. Recognizing that AI governance is a journey, not a destination, we look forward to supporting these efforts in the months and years to come.

A five-point blueprint for governing AI

1. Implement and build upon new government-led AI safety frameworks.

2. Require effective safety brakes for AI systems that control critical infrastructure.

3. Develop a broader legal and regulatory framework based on the technology architecture for AI.

4. Promote transparency and ensure academic and public access to AI.

5. Pursue new public-private partnerships to use AI as an effective tool to address the inevitable societal challenges that come with new technology.

01 Implement and build upon new government-led AI safety frameworks

One of the most effective ways to accelerate government action is to build on existing or emerging governmental frameworks to advance AI safety. A key element to ensuring the safer use of this technology is a risk-based approach, with defined processes around risk identification and mitigation as well as the testing of systems before deployment.

For example, the AI Risk Management Framework developed by the U.S. National Institute of Standards and Technology (NIST) and launched in early 2023 provides a strong template for advancing AI governance. It was developed through a consensus-driven and transparent process involving work by government agencies, civil society organizations, and several technology leaders, including Microsoft. NIST brings years of experience to the AI risk management space from its years of work developing critical tools to address cybersecurity risks. Microsoft has long experience working with NIST on the cybersecurity front, and it is encouraging to see NIST apply this expertise to help organizations govern, map, measure, and manage the risks associated with AI. Microsoft has committed to implementing the NIST AI Risk Management Framework, and we are not alone in our high regard for NIST's approach, as numerous governments, international organizations, and leading businesses have also validated the value of the new AI Risk Management Framework.

Similarly, new international standards are in development with the ISO/IEC 42001 on AI Management Systems recently published in December 2023. This will provide another important framework for companies and governments alike to put to work in advancing responsible AI.

We believe there is an opportunity for governments to help accelerate progress around these frameworks, using both carrots and sticks. In many countries, government procurement mechanisms have repeatedly demonstrated their value in improving the quality of products and advancing industry practice more generally. Governments could explore inserting requirements related to the AI Risk Management Framework or other relevant international standards into their procurement processes for AI systems, with an initial focus on critical decision systems that have the potential to meaningfully impact the public's rights, opportunities, or access to critical resources or services.

Although the progress in AI policy and strategy development has been gradual across much of Africa, several African countries are making noteworthy progress and are emerging as leaders in these discussions. Their efforts in pioneering AI-related strategies and policies are commendable and offer valuable insights for other countries who are considering how to initiate similar frameworks in this area. The collaboration of these countries in creating pan-African and interoperable legal and regulatory frameworks is crucial, as it paves the way for a more cohesive and effective approach

to AI governance across the continent. As part of any ongoing and future continent-wide and coordinated approaches to legislating AI, it is critical that AI tools and technologies are compatible with local developmental priorities and contribute to prosperous and inclusive societies.

The African Union (AU) is actively addressing this by convening experts from across the continent to develop a comprehensive African Union Artificial Intelligence (AU-AI) Continental Strategy. This strategy aims to "address the technological, ethical, economic, security, and social perspectives of AI." This coordinated approach will also aim to consider the "responsible, safe and beneficial use of AI." Similarly, the African Continental Free Trade Area is developing a Digital Trade Protocol to enhance digital trade on the continent. The Digital Trade Protocol includes provisions aimed at regulating the use of emerging technologies including AI in a manner that is safe, ethical, and responsible.

A similar approach was taken in the African Union Convention on Cyber Security and Personal Data Protection, also known as the Malabo Convention. Established under the auspices of the AU in 2014, the Convention went into effect in June 2023. It serves as a comprehensive legal framework for addressing cybercrime, cybersecurity, and data protection in Africa. It exemplified a model for creating pan-African, interoperable legal frameworks, showcasing the benefits of regional collaboration and standardization in cybersecurity and data protection. The Convention also acts as an important tool to regulate aspects of AI, such as the automated processing of personal data. Following this in 2016, the AU brought together a ten-member AU High-Level Panel on Emerging

Technologies to advise the Member States on how Africa should harness emerging technologies for its socio-economic development and to better understand the implications and mitigate possible adverse effects that may be associated with the adoption of such technologies.

There is also the SMART Africa Blueprint, which is a strategic framework developed to guide the application and governance of AI technologies in Africa through a multi-stakeholder, pan-African process. The initiative was spearheaded by the South African government, and included contributions from the SMART Africa member countries, civil society, academia, and the private sector to strengthen technical know-how, remove barriers to entry, and develop the appropriate policy frameworks for AI.

On a national level, a few African countries have made strides in setting out AI strategies and policies. **Mauritius** was the first country in Africa to publish a National AI Strategy in 2018. In this strategy, the country sets out a vision for AI to become the cornerstone of its development model and help advance socio-economic growth across its key development sectors, including manufacturing, healthcare, fintech, and agriculture. In 2019, **Egypt** established the National Council for Artificial Intelligence as a partnership between governmental institutions, academia, and businesses. The country then launched its national AI Strategy in 2021, with implementation planned over the next three to five years, to deepen the use of technologies that transform the economy, realize the country's national development priorities, and facilitate regional cooperation within the African and Arab

regions. In 2023, the government endorsed the Egyptian Charter for Responsible AI, which sets out the principles for the ethical use of AI. **Rwanda** released its National AI Policy in April 2023, providing a roadmap for how the East African nation can harness AI for sustainable development with a vision to become a global center for AI research and innovation. It further aims to leverage AI to power economic growth, improve quality of life, and position **Rwanda** as a global innovator for responsible and inclusive AI. Additional objectives include increasing the country's 21st century skills and AI literacy, driving public sector transformation to fuel AI adoption, and accelerating responsible AI adoption in the private sector.

At the time of this publication, **Tunisia** is in the process of developing its National Artificial Intelligence Strategy working in collaboration with a broad range of public sector ministries, Smart Africa, The Future Society (TFS), and the GIZ Digital Transformation Center. Following consultations with various local stakeholders, **Ghana** is also working closely with Smart Africa and TFS to map out Ghana's AI policy landscape and develop the country's National Artificial Intelligence Strategy for 2023-2033.

Ethiopia is in the process of finalizing its National AI Policy, and the Ethiopian Artificial Intelligence Institute has been hosting the [Pan-African Conference on Artificial Intelligence](#) since 2022, aiming to attract both local and international stakeholders to facilitate the responsible adoption and development of AI in Africa.

Nigeria is developing a national AI strategy, led by the Ministry of Communications, Innovation and Digital Economy, to "responsibly steer the AI revolution towards achieving national goals around job creation, social inclusion, and sustainable development." The Ministry has invited AI experts and researchers to join in co-creating an AI strategy that will enable economic diversification and inclusive growth, particularly for the country's fast-growing technology startup ecosystem. This will build on the efforts by the National Information Technology Development Agency (NITDA) to develop a national AI policy. The agency is gathering input from diverse stakeholders to formulate a national AI policy that provides a framework, not only for equitable AI development, but also the development of a vibrant ecosystem. In 2020, the National Center for AI and Robotics was established, as a NITDA subsidiary, to promote research and development in AI and other emerging technologies. The Center leverages collaborations with local and international research bodies on robotics and AI and enables AI education, which stands as a testament to the government's dedication to accelerating AI and robotics in the country.

There are more governments that are in the process of exploring how to incorporate AI into their government safety frameworks or developing national strategies on AI, and this list is by no means exhaustive. **Zambia's** SMART Zambia Institute, a division under the Office of the President mandated to implement the country's e-Government Master Plan (2018-2030), is planning to leverage the use of AI in government institutions as part of its broader digital transformation strategy. **Kenya**, recognizing the importance of AI in modern journalism, has set up

a media-related AI taskforce, which looks to create a mechanism to ensure the appropriate and ethical integration of AI in professional journalism. Although the country does not currently have a standalone AI policy or regulatory framework, its National Digital Master Plan 2022-2032 refers to AI extensively and incorporates the use of emerging technologies such as blockchain, internet of things, AI, big data, and quantum computing. This shows Kenya's commitment to harnessing the potential of AI and ensuring its benefits are realized across various sectors. Additionally, Kenya also established a Blockchain & Artificial Intelligence Taskforce to contextualize the application of AI in areas such as financial inclusion, cybersecurity, land titling, elections, and single digital identity processes. Similarly, in **South Africa**, the Presidential Commission on the Fourth Industrial Revolution has recommended the review and creation of policies and legislation to empower stakeholders around responsible AI use and to harness the potential of AI to realize benefits across various sectors.

We recognize that the pace of AI advancement raises new questions and issues related to safety and security, and we are committed to working with others to develop actionable standards to help evaluate and address them. As part of this, Microsoft recently joined with OpenAI, Google, and Anthropic to launch the Frontier Model Forum to continue to advance safety best practices around highly capable models, including how to evaluate and address the risks these models may present. We look forward to continuing to work with others and further advancing frameworks and practices for responsible AI.

AI Ethics and Principles

Organizations are considering the principles that will ensure ethical development and usage of AI.

- For example, the African Commission on Human and Peoples' Rights (ACHPR) is keeping a close eye on issues at the intersection of AI and human rights and has long been calling for legal and regulatory frameworks to ensure that AI is developed and implemented in a human-centric manner. In a 2019 Declaration of Principles on Freedom of Expression and Access to Information in Africa, the ACHPR called for states to ensure that the development, use, and application of AI and other similar technologies is compatible with international human rights law and standards, and should not infringe on freedom of expression, access to information, or any other human rights. In 2021, the Commission adopted a resolution specifically around AI and human rights, urging governments to "work towards a comprehensive legal and ethical governance framework for AI technologies, robotics, and other new and emerging technologies," and stressing the need for the AU and regional bodies to "develop a regional regulatory framework that ensured that these technologies respond to the needs of the people of the continent."

- Additionally, the UNESCO's 2021 Recommendation on the Ethics of AI was collated with the insights and advice of experts from Cameroon, Egypt, Ghana, Morocco, Rwanda, and South Africa, among others, recognizing the profound and dynamic positive and negative impacts of AI on societies, environment, ecosystems, and human lives. Of the 193 countries that have adopted the Recommendations on the Ethics of AI, at least 46 of the adoptees are from the African continent.

African voices should continue to play an important role in shaping the ethics and principles that guide the use and development of AI.

02 Require effective safety brakes for AI systems that control critical infrastructure

History offers an important and repeated lesson about the promise and peril of new technology. Since the advent of the printing press, governments have confronted the need to decide whether to accept or reject new inventions. Beginning in the latter half of the 1400s, Europe embraced the printing press, while the Ottoman Empire mostly banned it. By 1500, citizens in the Netherlands were reading more books per capita than anyone else. It's not a coincidence that the small nation soon found itself at the forefront of economic innovation.

Ever since, inventors and governments have typically concluded that the best path forward is to harness the power of new technology in part by taming it. The history of technology is replete with examples. Modern cities would not be possible without tall buildings, but tall buildings would not be possible without elevators. And in the 1800s, most people understandably were uncomfortable getting into what all of us today do without even thinking about—entering a metal box and being hoisted several stories into the sky by a cable. Elisha Otis, the American inventor of the elevator, found

in the 1850s that the public was slow to accept his machines, deeming them too dangerous.

This changed in 1854 at the World's Fair in New York, when Otis demonstrated a new safety brake for his elevator. He severed the cable holding his machine above the watching crowd, and the brake immediately caught the car, halting its fall. People were reassured, and in an important respect, the modern city was born.

This pattern has repeated itself for everything from electricity to railroads to school buses. Today, many houses and buildings have circuit breakers to protect against a surge in the electrical current. In many countries, city codes require them. Similarly, hundreds of millions of people put what they hold most precious in the world—their children—on morning school buses, based in part on regulations that require buses to have emergency brakes with bus drivers trained to use them. Planes today have ground proximity detectors and airborne collision avoidance systems that have helped to make commercial air travel incredibly safe, while empowering pilots—not machines—to make decisions in safety-critical scenarios.

Four steps governments can take to secure effective safety brakes for AI systems controlling critical infrastructure

1 Define the class of high-risk AI systems being deployed.

2 Require system developers to ensure that safety brakes are built by design into the use of AI systems for the control of infrastructure.

3 Ensure operators test and monitor high-risk systems to ensure AI systems that power critical infrastructure remain within human control.

4 Require AI systems that control operation of designated critical infrastructure to be deployed only in licensed AI infrastructure.

As we look to a future with artificial intelligence, it's worth remembering that the same fundamental approach has worked repeatedly in managing the potential dangers associated with new technology. Namely, identify when a new product could become the equivalent of a runaway train, and as for the locomotive itself, install an effective safety system that can act as a brake and ensure that the right people will use it quickly if it's ever needed—whether to slow something down or even bring it to a halt.

Not every potential AI scenario poses significant risks and, in fact, most do not. But this becomes more relevant when one contemplates AI systems that manage or control infrastructure systems for electricity grids, the water system, emergency responses, and traffic flows in our cities. We need

"safety brakes" to ensure these systems remain under human control.

We believe that the following steps would help address these issues:

First, the government should define the class of high-risk AI systems that are being deployed to control critical infrastructure and warrant safety brakes as part of a comprehensive approach to system safety. For the purposes of applying the safety brake concept to AI systems, we need to focus on the AI systems that are used to control the operation of critical infrastructure. There will be many AI systems used within critical infrastructure sectors that are low risk and that do not require the same depth of safety measures— employee productivity tools and customer service agents are two such examples.

Instead, one should focus on highly capable systems, increasingly autonomous systems, and systems that cross the digital-physical divide. For the purposes of spurring further discussion, one place to start might be to focus on AI systems that:

- Take decisions or actions affecting large-scale networked systems;

- Process or direct physical inputs and outputs;

- Operate autonomously or semi-autonomously; and

- Pose a significant potential risk of large-scale harm, including physical, economic, or environmental harm.

Second, the government should require system developers to ensure that safety brakes are built by design into the use of AI systems for the control of critical infrastructure. System safety is a well-established discipline that we have put to work in the aviation, automotive, and nuclear sectors, among others, and it is one that we must bring to bear to the engineering of AI systems that control critical infrastructure. We should establish a layered approach to AI safety, with the "safety brake" concept implemented at multiple levels.

While the implementation of "safety brakes" will vary across different systems, a core design principle in all cases is that the system should possess the ability to detect and avoid unintended consequences, and it must have the ability to disengage or deactivate should it demonstrate any unintended behavior. It should also embody best practice in human-computer interaction design.

Third, the government should ensure operators test and monitor high-risk systems to make certain that AI systems that power critical infrastructure remain within human control. Specific system testing will be needed in the context of a planned deployment for critical infrastructure. In other words, the use of an advanced AI model must be reviewed in the context of how it will be used in a specific product or service.

In accordance with system safety best practices, the system and each of its components should be tested, verified, and validated rigorously. It should be provable that the system operates in a way that allows humans to remain in control at all times. In practice, we anticipate that this will require close and regular coordination between a system operator, their AI infrastructure provider, and their regulatory oversight bodies.

Fourth, AI systems that control the operation of designated critical infrastructure should be deployed only in licensed AI infrastructure. We believe it would be wise to require that AI systems that control the operations of higher-risk critical infrastructure systems be deployed on licensed AI infrastructure. This is not to suggest that the AI infrastructure needs to be a hyperscale cloud provider such as Microsoft. Critical infrastructure operators might build AI infrastructure and qualify for such a license in their own right. But to obtain such a license, the AI infrastructure operator should be required to design and operate their system to allow another intervention point—in effect, a second and separate layer of protection—for ensuring human control if application-level measures fail.

These proposals might leave some wondering how realistic or futureproof "safety brakes" are if we are on a path to developing AI systems that are more capable than humans. They might ask: couldn't the AI system itself work around safety brakes and override them? Won't the AI system know how humans will respond at every step of the way and simply work around those responses?

In posing those questions, it's important to be clear about the facts as they stand today. Today's cutting-edge AI systems like GPT-4 from OpenAI and Claude from Anthropic have been specifically tested—by qualified third party experts from the Alignment Research Center—for dangerous capabilities, such as the ability to evade human oversight and become hard to shut down. Those tests concluded that GPT-4 and Claude do not have sufficient capabilities to do those things today. This rigorous testing and the conclusions drawn provide us with clarity as to the capabilities of today's cutting-edge AI models. But we should also heed the Alignment Research Center's call for ongoing research on these topics and recognize the need for industry-wide commitment to AI capability evaluations. Put simply, we need to ensure that we have the right structures in place not only to understand the status quo, but to get ahead of the future. That is precisely why we need action with respect to the small but important class of highly capable AI models that are on the frontier—a topic that our next section addresses.

03 Develop a broad legal and regulatory framework based on the technology architecture of AI

As we have given more thought to the various potential legal and regulatory issues relating to AI responsibilities, it has become more apparent that there will need to be a legal and regulatory architecture for AI that reflects the technology architecture for AI itself. In short, the law will need to place various regulatory responsibilities upon different actors based upon their role in managing different aspects of AI technology. For this reason, it's helpful to consider some of the critical pieces that go into building and using new foundation AI models.

The technology stack for AI foundation models

Applications	Software programs where the output of an AI model is put to work.
API Services	APIs (Application Program Interfaces), or endpoints, through which applications access pre-trained models.
Powerful Pre-Trained AI Models	Pre-trained models like GPT-4 that can be used to solve similar problems without starting from scratch.
Machine Learning Acceleration Software	Software that speeds up the process of developing and deploying large AI models.
AI Datacenter Infrastructure	Advanced supercomputing infrastructure, including clusters of advanced GPUs (Graphics Processing Units) with high bandwidth network connections.

A grounding in the technology architecture for AI foundation models

Software companies like Microsoft build a "tech stack" with layers of technologies that are used to build and run the applications that organizations and the public rely upon every day. There's no single right way to describe an AI tech stack, and there's a good chance that any two developers will describe it differently. But for purposes of thinking about the future of AI regulation, a good way to start is to consider the chart above. An advanced pretrained AI model like GPT-4 is shown on the third row in the middle of the stack. It's created by developers and research scientists at a firm like OpenAI based on the two layers below it. In the case of GPT-4, OpenAI technical staff in San Francisco, California did their model development work by harnessing the AI supercomputing infrastructure that Microsoft built and today operates exclusively for them. As Microsoft announced upon its opening in March 2020, this datacenter contains a single supercomputing system that then ranked in the top five supercomputers in the world. The supercomputing system has more than 285,000 Central Processing Unit (CPU) cores. (The CPU is perhaps the most fundamental component in any modern PC or laptop.) The system also has more than 10,000 of the most advanced Graphics Processing Units, or GPUs. Less advanced versions of such chips are contained in a gaming console or gaming laptop and can process a large number of mathematical equations simultaneously. Each GPU server in the datacenter has network connectivity that can process 400 gigabits of data per second.

As Microsoft Chief Technical Officer Kevin Scott said when we made this announcement in 2020, "the exciting thing about these [new GPT] models is the breadth of things they're going to enable." As OpenAI and Microsoft explained in 2020, machine learning experts had historically built separate, smaller AI models with many labeled examples to learn a single task such as translating between languages.

But using this type of massive supercomputing infrastructure—and with the help of customized machine learning acceleration software—it became possible to create a single massive AI model that could learn by examining huge amounts of data, such as billions of pages of publicly available text. As Microsoft said in the 2020 announcement and as the world now recognizes in 2023, "this type of model can so deeply absorb the nuances of language, grammar, knowledge, concepts, and context that it can excel at multiple tasks: summarizing a lengthy speech, moderating content in live gaming chats, finding relevant passages across thousands of legal files or even generating code from scouring GitHub."

As all this reflects, the core of what has struck some as the most surprising technological development of the decade was preannounced in plain and public view in just the third month as the decade began. The good news, at least from the perspective of Microsoft and OpenAI, is that we've been able to work the past several years to strengthen safety and security protocols to prepare for the more powerful AI models.

This brings one to how these large AI models are deployed for use. Given the very substantial computational resources required, these take place in multiple countries in advanced datacenters with large amounts of GPUs and advanced network connectivity, running in the case of GPT-4, on Microsoft's Azure platform. This requires substantial additional investments and deployment of the most advanced digital technology, but it does not require the same highly specialized infrastructure that is needed to build an advanced AI model in the first place.

The actual use of these models involves the top half of the technology stack. Users interact with a model like GPT-4 through an application, as shown at the top of the stack. ChatGPT, Bing Chat, and GitHub Copilot are all examples of such applications. Companies and organizations large and small will no doubt create new or modify existing applications to incorporate features and services that harness the power of generative AI models. Many will be consumer applications, including those that are already household names. Many others will be created in-house by companies, governments, and nonprofits for their own internal use or by their customers. In short, a new wave of applications powered by generative AI will soon become part of daily life around the world.

Such applications access the capabilities of an AI model through endpoints called APIs, or Application Program Interfaces. APIs have long been one of the most important methods of accessing core technology building blocks that
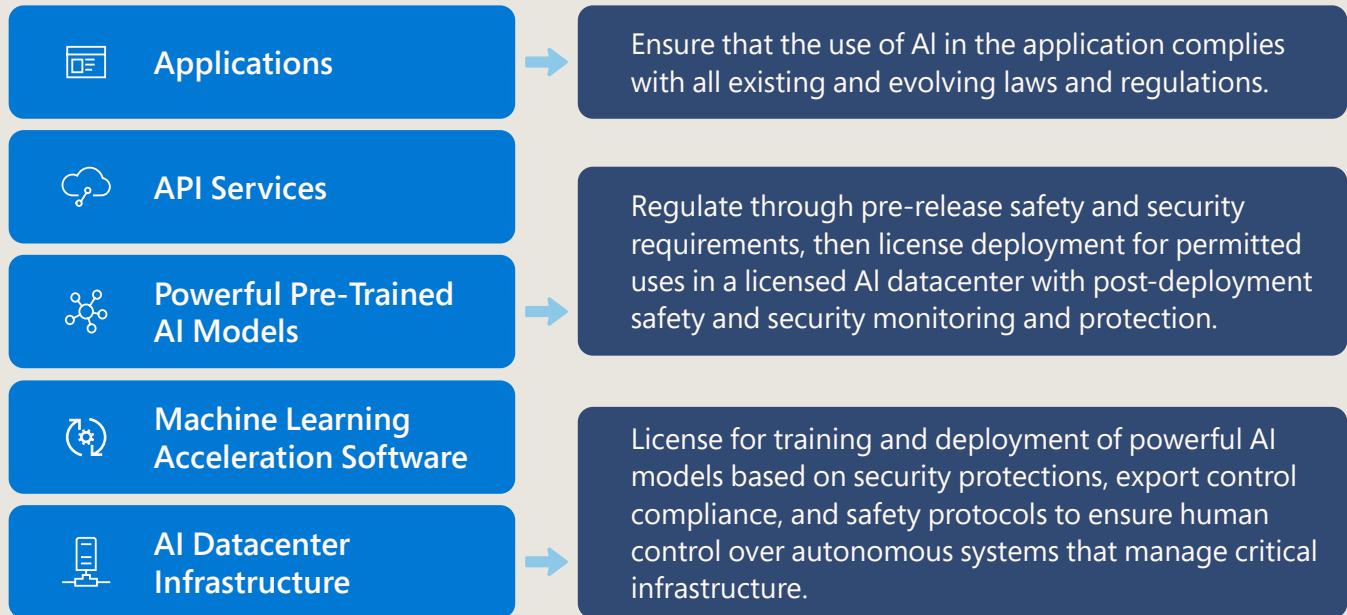
our customers are not running themselves on their infrastructure.

By way of illustration, Microsoft has created the Azure OpenAI Service to provide API access to OpenAI models like GPT-4. This API provides access to the model that is hosted on Microsoft's infrastructure. In short, this means that our customers can harness the power of GPT-4 by building an application of their choosing and simply calling the API to submit prompts and receive outputs from GPT-4. There is no need for customers to maintain the sophisticated infrastructure that is needed to run an advanced model like GPT-4, and our customers benefit from Microsoft's longstanding trust and compliance commitments, as well as the safety systems that we have built on top of the GPT-4 as part of the Azure OpenAI service.

Creating a regulatory architecture that reflects AI's technology architecture

We believe it is sensible to design an AI regulatory architecture that roughly corresponds to the AI technology architecture described above. As we've thought about these issues in recent months, we believe that law and regulation can probably have their most positive impact by focusing on three layers of the tech stack, with differing obligations at each level. The chart above illustrates this proposed approach, with further analysis and commitments we believe we can offer as a company to help advance these requirements.

A proposed AI regulatory architecture

Applications → Ensure that the use of AI in the application complies with all existing and evolving laws and regulations.

API Services

Powerful Pre-Trained AI Models → Regulate through pre-release safety and security requirements, then license deployment for permitted uses in a licensed AI datacenter with post-deployment safety and security monitoring and protection.

Machine Learning Acceleration Software

AI Datacenter Infrastructure → License for training and deployment of powerful AI models based on security protections, export control compliance, and safety protocols to ensure human control over autonomous systems that manage critical infrastructure.

Applying existing legal protections at the applications layer to the use of AI

For a great many individuals and organizations, the legal rubber will meet the road as applications use AI to deliver information and services to others. This is the layer where the safety and rights of people will most be impacted, especially because the impact of AI can vary markedly in different settings. As a result, we will need the laws and regulations that govern AI to apply to applications that use the output from AI models to deliver services to individuals and organizations.

The good news is that in many areas relating to the impact of AI on society, we don't need new laws and regulations. We instead need to apply and enforce existing laws and regulations. It has been encouraging to see discussions emerge in different countries across Africa on this topic. These conversations are especially relevant to the many applications that are being created to use new and more powerful AI. And this will be important for companies and other organizations in every economic sector and in every country. Existing laws will continue to apply to the decisions and actions of organizations and individuals alike. No one is proposing a new defense to illegal conduct that will enable people to stand up in court and proclaim, "but Your Honor, a machine made me do it."

While this conclusion is simple, its consequences are profound. It means that every organization that uses AI needs to master not only the technology itself but the ability to evaluate how the technology impacts its wide-ranging legal responsibilities. And courts and agencies alike will need to develop new capabilities to analyze how AI was used in a particular system.

We believe that several steps can help achieve this, including those we can take as a company:

First, we will work with our customers to help them apply state-of-the-art best practices to deploy AI lawfully and responsibly. One of the critical characteristics of AI is that the real-world impact on specific groups and issues is defined not just by the developer of an AI model or system, but also in its implementation in a specific service or application. In fact, in many circumstances it is only at the application level that it's possible to specifically identify and test for these real-world impacts before AI is deployed. As a result, responsibilities are often shared or even distributed, with different organizations needing to play different roles. This helps explain why it's so important for customers that use AI in their services to develop their own capabilities to do so responsibly. This also explains why it is so important for a leading tech company to share information and lend their expertise on state-of-the-art best practices and tooling for responsible AI deployment. We have been doing this type of work for two decades on other issues involving digital technology, including implementing legal compliance systems, advancing cybersecurity, and protecting privacy. We began five years ago to do similar work relating to artificial intelligence, and

we will expand this initiative to work more broadly and deeply with our customers in the year ahead.

Second, we believe that regulatory agencies will need to add new AI expertise and capabilities. Very quickly, this need will reach virtually every agency in most governments in the world. For example, a regulatory agency that oversees pharmaceuticals and medical devices will need more AI experts who can help evaluate the use of cutting-edge AI systems by companies in clinical trials for new drugs. In Kenya, for instance, the nation's communication authority responsible for overseeing online media and news could possibly benefit from additional AI experts to help evaluate the use of AI in identifying and mitigating propaganda and misinformation given their prior concerns around this issue, especially during times of elections. Similarly, as South Africa's State Security Agency (SSA) is considering the use of predictive policing algorithms, employing more AI specialists could help the SSA refine these algorithms or monitor their deployment to enhance public safety and help ensure policy-aligned outcomes. The goal is to anticipate where crimes are most likely to occur while ensuring that these technologies are used responsibly. This approach is crucial to avoid amplifying human biases or prejudices, thus maintaining the integrity of public safety measures.

Generative AI itself will be a powerful tool that will better enable regulatory agencies to evaluate the use of AI. This is because models like GPT-4 and services like ChatGPT, GitHub Copilot, and Microsoft M365 Copilot make it far easier for people to harness the power of AI to access data and evaluate it more quickly. As Google rightly

recommended in a recent white paper it will be important for governments to "direct sectoral regulators to update existing oversight and enforcement regimes to apply to AI systems, including on how existing authorities apply to the use of AI." Agencies will need the funding, staff, and commitment to put these new tools to work.

Third, we will support broad educational initiatives to make information about AI technologies and responsible AI practices available to legislators, judges, and lawyers. Finally, rapid AI advances are creating new pressures on those who make or help enforce the law to learn about new AI technologies and how they work. We witnessed a similar need when the personal computer first became popular in the 1980s. For example, judges needed to decide cases that started to turn, in part, on evidence about or involving PC software and hardware. Beginning in the 1990s, Microsoft supported broad initiatives to share information about how this new technology worked. We continue to do this today in selected areas such as electronic discovery. The accelerating use of AI means that new such efforts will be needed. We will support this work, including by supporting bar associations and other public interest and civic groups and activities.

Important public policy discussions continue around the world to advance ideas about how to apply existing law, upskill regulators and address any remaining regulatory gaps. In Egypt, the National Artificial Intelligence Strategy includes

"AI for Development" as one of its four main pillars. This strategy underscores the importance of capacity building to facilitate technology and knowledge transfer, which, in turn will help grow the local AI ecosystem. Similarly, Mauritius's National AI Strategy identifies knowledge transfer as a critical challenge. It emphasizes the need for capacity building and attracting skills to create an appropriate ecosystem for adopting new technologies, including AI. This includes leveraging foreign expertise in the short term, as well as establishing AI campuses in collaboration with local universities and international experts and investing in STEM education in the long term.

Developing new laws and regulations for highly capable AI foundation models

While existing laws and regulations can be applied and built upon for the application layer of the tech stack, we believe that new approaches will be needed for the two additional layers beneath that reflect the new and more powerful AI models that are emerging. The first of these is for the development of the most powerful new AI models, and the second is for the deployment and use of these models in advanced datacenters. From our work on the frontiers of AI, we have seen a new class of model emerge. Highly capable foundation models are trained on internet-scale datasets and are effective out-of-the-box at new tasks—a model like GPT-4 allows you to create a never-seen-before image using words in one prompt, and a speech in the style of a famous historical figure in the very next.

Microsoft commitments to an AI licensing regime

Microsoft will share our specialized knowledge about advanced AI models to help governments define the regulatory threshold.

Microsoft will support governments in their efforts to define the requirements that must be met in order to obtain a license to develop or deploy a highly capable foundation model.

Microsoft will support government efforts to ensure the effective enforcement of a licensing regime.

At the cutting-edge, the capabilities of these foundation models are at once very impressive and can be harder to predict. As the models have been scaled up, we have seen anticipated advances in capabilities, as well as surprising ones that we and others did not predict ahead of time and could not observe on a smaller scale. Despite rigorous prerelease testing and engineering, we've sometimes only learned about the outer bounds of model capabilities through controlled releases with users. And the work needed to harness the power of these models and align them to the law and societal values is complex and evolving.

These characteristics of highly capable models present risk surfaces that need to be addressed. To date, we have benefited from the high safety standards self-imposed by the developers who have been working at the frontiers of AI model development. But we shouldn't leave these issues of societal importance to good judgment and self-restraint alone. We need regulatory frameworks that anticipate and get ahead of the risks. And we need to acknowledge the simple truth that not all actors are well intentioned or well-equipped to address the challenges that highly capable models present. Some actors will use AI as a weapon, not a tool, and others will underestimate the safety challenges that lie ahead.

Sam Altman, the CEO of OpenAI, testified before the United States Congress and called for the establishment of a licensing regime for this small but important class of highly capable models at the frontiers of research and development. At Microsoft, we endorse that call and support the establishment of a new regulator to bring this licensing regime to life and oversee its implementation.

First, we and other leading AI developers will need to share our specialized knowledge about advanced AI models to help governments define the regulatory threshold. One of the initial challenges will be to define which AI models should be subject to this level of regulation. The objective is not to regulate the rich ecosystem of AI models that exist today and should be supported into the future, but rather the small number of AI models that are very advanced in their capabilities and in some cases, redefining the frontier. We refer to this small subset of models as highly capable AI models in this whitepaper.

Defining the appropriate threshold for what constitutes a highly capable AI model will require substantial thought, discussion, and work in the months ahead. The amount of compute used to train a model is one tractable proxy for model capabilities, but we know today that it is imperfect in several ways and unlikely to be durable into the future, especially as algorithmic improvements lead to compute efficiencies or new architectures altogether.

A more durable but unquestionably more complex proposition would be to define the capabilities that are indicative of high ability in areas that are consequential to safety and security, or that represent new breakthroughs that we need to better understand before proceeding further. Further research and discussion are needed to set such a capability-based threshold, and early efforts to

define such capabilities must continue apace. In the meantime, it may be that as with many complex problems in life, we start with the best option on offer today—a computer-based threshold—and commit to a program of work to evolve it into a capability-based threshold in short order.

Second, we will support governments in their efforts to define the requirements that must be met in order to obtain a license to develop or deploy a highly capable AI model. A licensing regime for highly capable AI models should be designed to fulfill three key goals. First and foremost, it must ensure that safety and security objectives are achieved in the development and deployment of highly capable AI models. Second, it must establish a framework for close coordination and information flows between licensees and their regulator, to ensure that developments material to the achievement of safety and security objectives are shared and acted on in a timely fashion. Third, it must provide a footing for international cooperation between countries with shared safety and security goals, as domestic initiatives alone will not be sufficient to secure the beneficial uses of highly capable AI models and guard against their misuse. We need to proceed with an understanding that it is currently trivial to move model weights across borders, allowing those with access to the "crown jewels" of highly capable AI models to move those models from country to country with ease.

KY3C:
Applying to AI services the "Know Your Customer" concept
developed for financial services

Know your Cloud

Know your Customer

Know your Content

Third, we will support government efforts to ensure the effective enforcement of a licensing regime for highly capable AI models by also imposing licensing requirements on the operators of AI datacenters that are used for the testing or deployment of these models. Today's highly capable AI models are built on advanced AI datacenters. They require huge amounts of computing power, specialized AI chips, and sophisticated infrastructure engineering, like Microsoft's facilities in Iowa. Such AI datacenters are therefore critical enablers of today's highly capable AI models and an effective control point in a comprehensive regulatory regime.

Much like the regulatory model for telecommunications network operators and critical infrastructure providers, we see a role for licensing providers of AI datacenters to ensure that they play their role responsibly and effectively to ensure the

safe and secure development and deployment of highly capable AI models. To obtain a license, an AI datacenter operator would need to satisfy certain technical capabilities around cybersecurity, physical security, safety architecture, and potentially export control compliance.

In effect, this would see a principle that was developed for banking to protect against money laundering and criminal or terrorist use of financial services become applied to AI. The "Know Your Customer"—or KYC— principle requires that financial institutions verify customer identities, establish risk profiles, and monitor transactions to help detect suspicious activity.

In a similar way, it would make sense for a similar KYC principle to require that the developers of powerful AI models first "know the cloud" on which their models are deployed. The use of authorized and licensed AI datacenters would

ensure that those who develop advanced models would have several vendors from which to choose. And it would enable the developer of an advanced model to build or operate their own cloud infrastructure as well, based on meeting the requisite technical standards and obligations. The licensed AI datacenter operator would then need to meet ongoing regulatory requirements, several of which are worth considering.

First, operators of AI datacenters have a special role to play in securing highly capable AI models to protect them from malicious attacks and adversarial actors. This likely involves not just technical and organizational measures, but also an ongoing exchange of threat intelligence between the operator of the AI datacenter, the model developer, and a regulator.

Second, in certain instances, such as for scenarios that involve sensitive uses, the cloud operator on which the model is operating should apply the second aspect of the KYC principle – knowing the customers who are accessing the model. More thought and discussion will be needed to work through the details, especially when it comes to determining who should be responsible for collecting and maintaining specific customer data in different scenarios.

The operators of AI datacenters that have implemented know-your-customer procedures can help regulators get comfortable that all appropriate licenses for model development and deployment have been obtained. One possible approach is that substantial uses of computers that are consistent with large training runs should be reported to a regulator for further investigation.

Third, as export control measures evolve, operators of AI datacenters could assist with the effective enforcement of those measures, including those that attach at the infrastructure and model layers of the tech stack.

Fourth, as discussed above, the AI infrastructure operator will have a critical role and obligation in applying safety protocols and ensuring that effective AI safety brakes are in place for AI systems that manage or control critical infrastructure. It will be important for the infrastructure operator to have the capability to intervene as a second and separate layer of protection, ensuring the public that these AI systems remain under human control.

These early ideas naturally will all need to be developed further, and we know that our colleagues at OpenAI have important forthcoming contributions on these topics too. What is clear to us now is that this multitiered licensing regime will only become more important as AI models on the frontiers become more capable, more autonomous, and more likely to bridge the digital-physical divide. As we discussed earlier, we believe there is good reason to plan and implement an effective licensing regime that will, among other things, help to ensure that we maintain control over our electricity grid and other safety-critical infrastructure when highly capable AI models are playing a central role in their operation.

04 Promote transparency and ensure academic and nonprofit access to AI

Transparency as a critical ethical principle for AI

One of the many AI policy issues that will require serious discussion in the coming months and years is the relationship and tension between security and transparency. There are some areas, such as AI model weights (which are components of a model that are core to a model's capabilities), where many experts believe that secrecy will be essential for security. In some instances, this may even be needed to protect critical national security and public safety interests. At the same time, there are many other instances where transparency will be important, even to advance the understanding of security needs and best practices. In short, in some instances tension will exist and in other areas it will not.

Microsoft commitments to promote transparency for AI

Microsoft will release an annual transparency report to inform the public about its policies, systems, progress, and performance in managing AI responsibly and safely.

Microsoft will support the development of a national registry of high-risk AI systems that is open for inspection so that members of the public can learn where and how those systems are in use.

Microsoft will commit that it will continue to ensure that our AI systems are designed to inform the public when they are interacting with an AI system and that the system's capabilities and limitations are communicated clearly.

We believe there is benefit in requiring AI generated content to be labeled in important scenarios so that the public "knows the content" it is receiving.

When Microsoft adopted ethical guidelines for AI in 2018, we made transparency one of our six foundational principles. As we've implemented that principle, we've learned that it's important to provide different types of transparency in different circumstances, including making sure that people are aware that they are interacting with an AI system. Generative AI makes this principle more important than in the past, and it's an area where ongoing research and innovation will be critical. To help spur new work in this area, Microsoft is making three commitments.

First, Microsoft will release an annual transparency report to inform the public about its policies, systems, progress, and performance in managing AI responsibly and safely. Transparency reports have proven to be an effective measure to drive corporate accountability and help the public better understand the state-of-the-art and progress toward goals.

Second, Microsoft will support the development of a national registry of high-risk AI systems that is open for inspection so that members of the public can learn where and how those systems are in use. Public trust in AI systems can be enhanced by demystifying where and how they are in use. For high-risk AI systems, Microsoft supports the development of a national registry that would allow members of the public to review an overview of the system as deployed and the measures taken to ensure the safe and rights-respecting performance of the system. For this information to be useful to the public, it should be expressed at the system level, providing details about the context of use, and be written for nontechnical audiences. To achieve this, one could

implement the approach of several European cities in adopting the Algorithmic Transparency Standard and developing accessible explanations of how it uses AI (see, for example, the City of Amsterdam's Algorithm Register).

Third, Microsoft will commit that it will continue to ensure that our AI systems are designed to inform the public when they are interacting with an AI system and that the system's capabilities and limitations are communicated clearly. We believe that transparency is important not only through broad reports and registries, but in specific scenarios and for the users of specific AI systems. Microsoft will continue to build AI systems designed to support informed decision making by the people who use them. We take a holistic approach to transparency, which includes not only user interface features that inform people that they are interacting with an AI system, but also educational materials, such as the new Bing primer, and detailed documentation of a system's capabilities and limitations, such as the Azure OpenAI Service Transparency Note. This documentation and experience design elements are meant to help people understand an AI system's intended uses and make informed decisions about their own use.

Fourth, we believe there is benefit in requiring AI generated content to be labeled in important scenarios so that the public "knows the content" it is receiving. This is the third part of the KY3C approach we believe is worth considering. As we are committing above for Microsoft's services Bing Image Creator and Designer, we believe the public deserves to "know the content" that AI is creating, informing people when something like

a video or audio has been originally produced by an AI model rather than a human being. This labeling obligation should also inform people when certain categories of original content have been altered using AI. This will require the development of new laws, and there will be many important questions and details to address. For example, in recent years there has been a growing focus on addressing the new risks to democracy and the public from the potential weaponization of AI to alter content and create "deep fakes," including videos. The concern about future technology is well-placed.

Fortunately, there is an opportunity to use existing technical building blocks for AI transparency in addition to creating new transparency reporting initiatives. One of these is the Coalition for Content Provenance Authenticity, or C2PA, a global standards body with more than 60 members including Adobe, the BBC, Intel, Microsoft, Publicis Groupe Sony, and Truepic. The group is dedicated to bolstering trust and transparency of online information, including releasing the world's first technical specification for certifying digital content in 2022, which now includes support for generative AI. Leveraging the C2PA specification, Microsoft recently announced at our Build 2023 conference that we will deploy new state-of-the-art provenance tools to help the public identify AI-generated audio-visual content and understand its origin. Microsoft will initially support major image and video formats and release the service for use with two of Microsoft's new AI products, Microsoft Designer and Bing Image Creator. While this is an important step, it is only one step in what needs to be a broader effort to protect information integrity.

Access to AI resources for academic research and the nonprofit community

Lastly, we believe there is another element that adds to transparency and that deserves more prominent attention. This is the need to provide broad access to AI resources for academic research and the nonprofit community. The high cost of computational resources for the training of large-scale AI models, as well as other AI projects, is understandably raising concerns in the higher education and nonprofit communities. We understand this issue well because Microsoft's large technology investment in OpenAI in 2019 originated from precisely this need, for OpenAI itself, due in part to its nonprofit status.

Much of the tech sector itself owes both its birth and ongoing innovation to critical basic research pursued in colleges and universities across the country. It's a success story that has been studied and emulated in many other countries around the world. The past few decades have seen huge swaths of basic research in almost every field propelled by growing computing resources and data science. Unless academic researchers can obtain access to substantially more computing resources, there is a real risk that scientific inquiry and technological innovation will suffer. Another dimension of this problem is also important. Academic researchers help ensure accountability to the public by advancing our understanding of AI. The public needs academics to pursue research in this area, including research that advances AI accountability by analyzing the behavior of the models the commercial sector is creating. While new and smaller open-source AI models are emerging and clearly are important, other basic

research projects involving AI will almost certainly require more computational power than in the past. And unless new funding sources come together to provide a more centralized resource for the academic community, academic research will be at risk. This has led us to offer three focused commitments.

First, Microsoft will support the establishment of the newly proposed National AI Research Resource (NAIRR) in the U.S. to provide computing resources for academic research and would welcome and support an extension to accommodate access by academic institutions internationally. The U.S. is advancing the National AI Research Resource, "a shared research infrastructure that would provide AI researchers and students with significantly expanded access to computational resources, high-quality data, educational tools, and user support." Microsoft supports the establishment of this type of research resource and believes it is important in advancing understanding around the opportunities and risks of AI.

We also would welcome and support an extension of the NAIRR to provide access to academic institutions internationally. According to the Government AI Readiness Index, AI-ready regions necessitate strong collaboration between governments, academic institutions, and the private sector to develop educational programs that not only facilitate meaningful industry connections for students but also curb brain drain. UNESCO highlights the education sector's crucial role in promoting relevant AI research and ensuring Africa keeps pace in this rapidly evolving field.

We're already seeing similar and substantial interest in this type of resource among other countries

around the world, including across Africa. Many universities and higher education institutions across the continent have taken steps to develop AI-related course material so that students understand this technology and are properly equipped to leverage AI in the future. For example, the AI Africa Consortium, led by South Africa's University of the Witwatersrand, is dedicated to bringing together a diverse group of stakeholders, including universities, research organizations, and government agencies, to shape the future direction of Cirrus, an ambitious AI research initiative tailored for African needs. Furthermore, South Africa's Centre for AI Research is working on developing an accredited AI research network in South Africa to advise industry and government partners in the utilisation of AI for social and economic development. Morocco's Euromed University of Fez has established the Euromed School of Digital Engineering and Artificial Intelligence, a center dedicated towards AI research, development, and education, and Egypt's Kafrelsheikh University established an AI faculty which looks to aid the country's efforts in building the necessary skills to boost innovation, growth and productivity by harnessing the transformative potential of these tools.

In addition, Cameroon's first AI training centre was established at the University of Yaoundé in 2019 and falls in line with the government's vision to digitise the economy and introduce and develop digital skills as part of the national curriculum. Ethiopia established a similar center— the AI & University of Yaoundé of Excellence— at the Addis Ababa Science and Technology University, with the aim of fostering close collaboration between academics and entrepreneurs in AI and robotics.

Smaller initiatives across Africa include a mandate to include research on AI systems and projects by the Department of Mathematics and Computer Science at the National University of Lesotho and a 10-week boot camp and master's program in AI for different specializations from Senegal's Dakar Institute of Technology in partnership with French AI school, VIVADATA.

In addition to these country-specific efforts, cross-country collaborations are growing, such as The African Institute for Mathematical Sciences, which has physical presences in Senegal, Rwanda, Cameroon, South Africa, and Ghana, and attracts postgraduate students across African countries. One of AIMS's programs is the African Master's in Machine Learning, a 10-month intensive master's course for Africans that focuses on machine learning to help build an ecosystem of AI professionals on the continent.

Beyond the formal curriculum-based trainings, grassroot movements that aim to build a community of African AI students and researchers have been expanding in the recent years. These include Deep Learning Indaba in Ghana and Deep Learning IndabaX in South Africa, Data Science Africa in Uganda, and Masakhane in South Africa. These efforts represent an acknowledgement of the importance of nurturing highly skilled and proficient African AI specialists and a desire to inspire the growing number of young people entering the workforce to pursue careers in AI.

Second, we will increase investment in academic research programs to ensure researchers outside Microsoft can access the company's foundation models and the Azure OpenAI Service to undertake research and validate findings. This

expanded commitment builds on the success of our Turing Academic Program and Accelerating Foundation Models Research Program. It is designed to help the academic community gain API-based access to cutting edge foundation models from Microsoft, as well as OpenAI models via Microsoft's Azure OpenAI Service. This will ensure that researchers can study frontier applications and the sociotechnical implications of these models. An important complement to providing such access is the development of governance best practices for the academic community engaged in frontier research on applications and the safety and security implications of highly capable models. Microsoft would welcome the opportunity to develop such practices by supporting and collaborating with a multistakeholder group, including representatives across the academic community.

The Microsoft Research AI & Society Fellows program is an extension of our commitment to fostering academic research in the field of AI. This initiative is designed to attract top-tier scholars and industry experts to tackle the challenge of creating more equitable Generative AI models, particularly for the Global South. The program aims to achieve three key objectives: analyzing the effectiveness of these AI models in the Global South, designing more inclusive and robust AI experiences for a diverse global audience, and developing a long-term research strategy to ensure equitable AI applications. This call is open to multidisciplinary social scientists and development economists with experience in AI technologies, aiming to bridge the digital divide and promote global equality in AI development and usage.

Third, Microsoft will create free and low-cost AI resources for use by the nonprofit community. Finally, we deeply appreciate the critical role that nonprofit organizations play in addressing societal needs around the world. Given their role as great incubators of innovative solutions, we believe it is critical for nonprofits to have broad, easy, and inexpensive access to new AI models and features for their work. Microsoft Philanthropies, including its Tech for Social Responsibility arm, supports 350,000 nonprofits in the Microsoft Cloud. It provides more than $4 billion annually in cash and technology donations and discounts to nonprofits worldwide, a figure comparable to one of the 10 largest government foreign aid budgets. In May 2023, we expanded this support by announcing AI solutions to Microsoft Cloud for Nonprofit. These AI solutions are designed to improve the ability of nonprofit organizations to optimize their operations and better engage with donors.

05 Pursue new public-private partnerships to use AI to address the inevitable societal challenges that come with new technology

One lesson from recent years is that democratic societies often can accomplish the most when they harness the power of technology and bring the public and private sectors together. It's a lesson we need to build upon to address the impact of AI on society. AI is an extraordinary tool with incredible potential for good. Like other technologies, though, it can be used as a weapon, and there will be some around the world who will seek to use it that way. However, we can also leverage AI in the fight against those who abuse it and to address societal challenges. We must work together through public and private partnerships to do this.

Specifically, important work is needed now to use AI to strengthen democracy and fundamental rights, provide broad access to the AI skills that will promote inclusive growth, and use the power of AI to advance the planet's sustainability needs. Perhaps more than anything, a wave of new AI technology provides an occasion for thinking big and acting boldly. In each area, the key to success will be to develop concrete initiatives and bring governments, industry, and NGOs together to advance them. Microsoft will do its part in each area.

In Nigeria, Microsoft has partnered with the United Nations Development Programme to co-convene the AI for Development (AI4Dev) Reference Group, a multi-stakeholder and interdisciplinary group, with stakeholders from the private sector, civil society, academia, and other leading industries. The AI4Dev Reference Group will help shape the country's AI agenda and is tasked with several key functions: firstly, to identify and prioritize AI-driven solutions tailored to Nigeria's specific developmental needs and challenges; secondly, to facilitate knowledge sharing and collaborative research among its members, thereby fostering innovation in AI applications; thirdly, to advise on policy and regulatory frameworks that support ethical and sustainable AI development; fourthly, to create channels for capacity building and AI skill development among Nigerian professionals and students; and lastly, to establish a platform for dialogue and partnership, encouraging investment and support for AI projects that can have a meaningful impact on economic and social development in Nigeria.
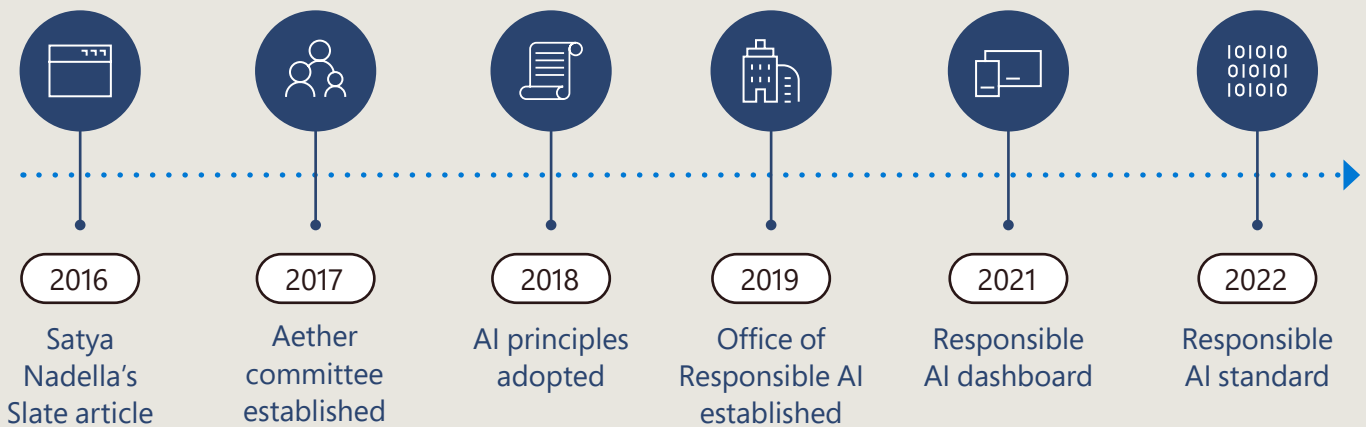
Part 4:

Responsible by Design:
Microsoft's approach to building
AI systems that benefit society

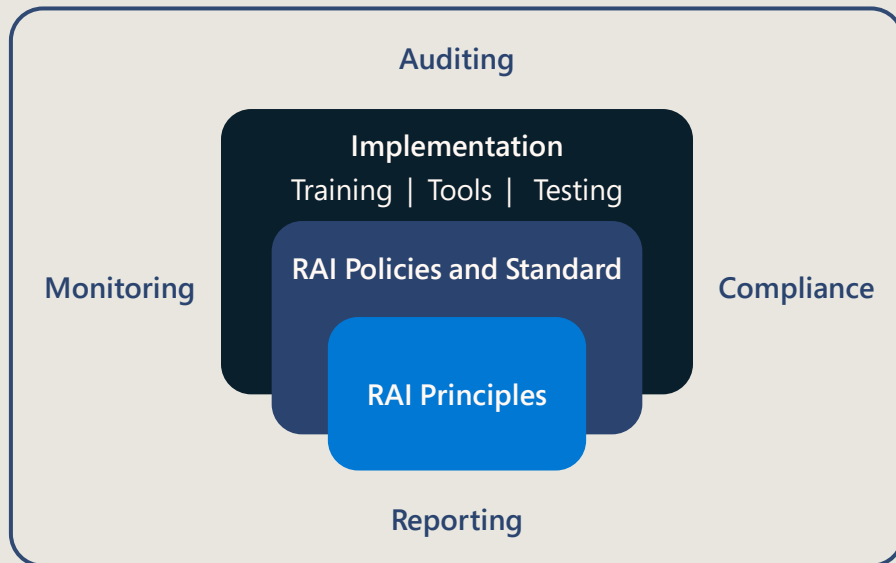## Microsoft's commitment to developing AI responsibly

For the past seven years, we have worked to advance responsible AI—artificial intelligence that is grounded in strong ethical principles. We have approached our work with a humble recognition that trust is not given but earned, and our responsibility is not just to Microsoft but our community more broadly. This has led us to be focused both on meeting our own commitments, and helping our customers and partners do the same.

Our responsible AI journey began in 2016 with Satya Nadella, Microsoft's Chairman and CEO, sharing his vision of humanity empowered by AI. Satya expressed the beginnings of our core AI principles—values that endure today. Building on this vision, we launched Microsoft's Aether Committee, comprised of researchers, engineers, and policy experts who provide subject matter expertise on the state-of-the-art and emerging trends with respect to our AI principles. This led to the creation and adoption of our AI principles in 2018.

Our Responsible AI Journey

2016	2017	2018	2019	2021	2022
Satya Nadella's Slate article	Aether committee established	AI principles adopted	Office of Responsible AI established	Responsible AI dashboard	Responsible AI standard

Responsible AI Governance Framework

Auditing

Implementation
Training | Tools | Testing

RAI Policies and Standard

RAI Principles

Monitoring

Compliance

Reporting

We deepened our efforts in 2019 by establishing the Office of Responsible AI. This team coordinates the governance of our program, and collaborated across the company to write the first version of the Responsible AI Standard, a framework for translating high-level principles into actionable guidance for engineering teams building AI systems.

In 2021, we spoke publicly about the key building blocks that we had put in place to operationalize our program. We envisioned expanding training, processes, and tools to help us to implement and scale our responsible AI efforts. 2022 brought a new iteration of our Responsible AI Standard, evolving it into the version we use today, which we have also made publicly available. It sets out how Microsoft will build AI systems using practical methods to identify, measure, and mitigate

potential risks ahead of time. This responsible-by-design approach establishes repeatable processes to minimize potential harms and magnify the benefits of AI from the outset.

We are proud of our progress over the last seven years.

Those efforts have brought us to where we are today, deepening our commitment to embed safety and responsibility into the lifecycle of our AI systems. This is possible only when responsible AI principles and practices transcend traditional silos and multidisciplinary teams work together. With the opportunity and the potential risks at hand, we believe we must share what we have learned and help all organizations apply responsible AI practices to their work. That is precisely what we at Microsoft are doing, and we hope to lead by example.

Our Ecosystem
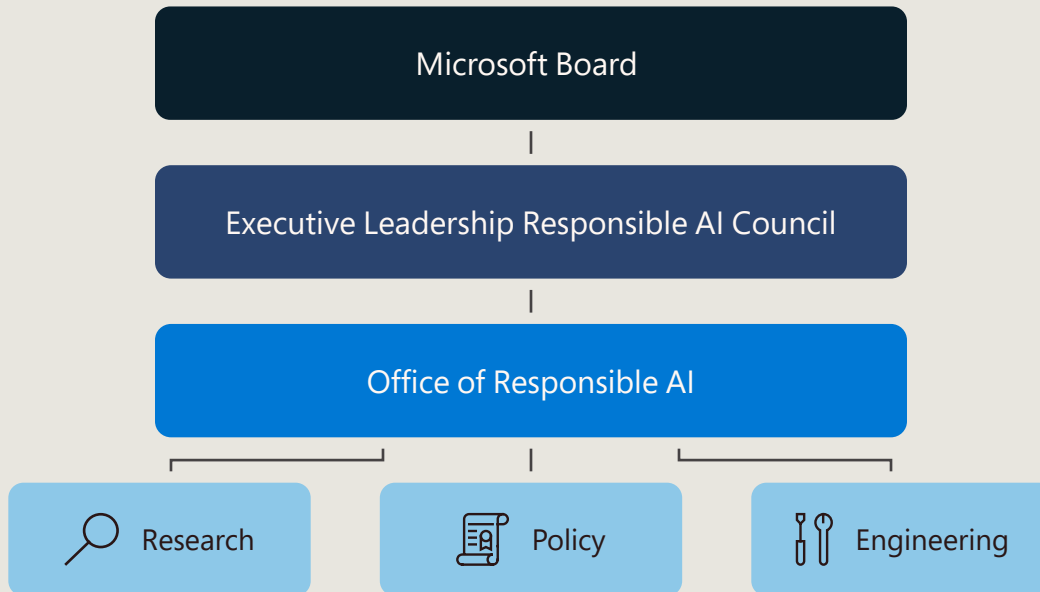
Operationalizing Responsible AI at Microsoft

Setting foundational governance structures

As the pace of AI continues to advance, we continue to evolve the governance structure we established to enable progress and accountability as a foundational piece of our responsible AI program. The creation of Microsoft's governance structure—as well as the decision to scale responsible AI across the company—was driven by leadership. Chairman and CEO Satya Nadella and the entire senior leadership team at Microsoft have made responsible Microsoft's leadership recognized that a single team or discipline tasked with responsible AI would not be enough.

Taking lessons from long-standing, cross-company commitments to privacy, security, and accessibility, we realized that responsible AI must be supported by the highest levels of leadership in the company and championed at every level across Microsoft.

To that end, Microsoft's Office of Responsible AI developed a governance system that incorporates many diverse teams and functions across the company. At the working level, core teams within engineering, research, and policy play critical roles to advance responsible AI across the company, each bringing a set of unique skills. Responsible AI roles are also embedded within product, engineering, and sales teams by the appointment of "Responsible AI Champions" by leadership.

Our Responsible AI Champions are tasked with spearheading responsible AI practices within their respective teams, which means adopting the Responsible AI Standard, issue spotting and directly advising teams on potential mitigations, and cultivating a culture of responsible innovation. The Office of Responsible AI helps to orchestrate these teams across the company, drawing on their deep product knowledge and responsible AI expertise to develop a consistent approach across Microsoft.

At the next level, the Responsible AI Council is a forum for leadership alignment and accountability in implementing Microsoft's responsible AI program. The Council is chaired by Microsoft's Vice Chair and President, Brad Smith, and our Chief Technology Officer, Kevin Scott, who sets the company's technology vision and oversees our Microsoft Research division. The Responsible AI Council convenes regularly, and brings together representatives of our core research, policy, and engineering teams dedicated to responsible AI, including the Aether Committee and the Office of Responsible AI, as well as engineering leaders and senior business partners who are accountable for implementation.

At the highest level, the Environmental, Social, and Public Policy Committee of the Microsoft Board provides oversight of our responsible AI program. Our regular engagements with the Committee ensure the full rigor of Microsoft's enterprise risk management framework is applied to our program.

The need for standardization

From crafting an AI system's purpose to designing how people interact with it, we must keep people at the center of all AI decisions. While our responsible AI principles state the enduring values we seek to uphold, we needed more specific guidance on how to build and deploy AI systems responsibly. This is why we developed our Responsible AI Standard, a more practical guide that memorializes a set of rules of the road for our engineering teams so that upholding our AI principles is a daily practice.

The Responsible AI Standard provides engineering teams with actionable guidance on how to build AI systems responsibly. It was the result of a multi-year, cross-company effort that reflected a vast array of input from researchers, engineers, lawyers, designers, and policy experts. We consider it to be a significant step forward for our practice of responsible AI because it sets out much more concrete, practical guidance on how to identify, measure, and mitigate harm ahead of time. It also requires teams to adopt tools and controls to secure beneficial uses while guarding against potential misuses of their products.

The Anatomy of the Responsible AI Standard

Principles	Which enduring values guide our responsible AI work?
Goals	What are the outcomes that we need to secure?
Requirements	What are the steps we must take to secure the goals?
Tools and practices	Which aids can help us meet the requirements?

There are two ways in which the Standard offers concrete direction to our engineering teams working across an AI product's lifecycle:

- **Articulating goals.** These define what it means to uphold the responsible AI principles. They break down a broad principle like accountability into definitive outcomes, such as ensuring AI systems are subject to impact assessments, data governance, and human oversight.

- **Outlining a playbook.** These specific procedures and steps are required of teams throughout an AI system's lifecycle in order to achieve the goals set in our Responsible AI Standard. The steps map to available resources, tools, and practices to equip teams to make these goals a reality. For example, one of our Responsible AI Standard goals is to minimize the time to remediate predictable or known failures of an AI system, and to secure that goal, we ask teams to identify potential harms through iterative red teaming. We then ask teams to measure the prevalence

of those harms and mitigate them by testing and implementing various tools and established strategies. This cycle of identifying, measuring, and mitigating potential harms of an AI system underpins many of the requirements in the Responsible AI Standard.

Ensuring checks and balances

When building and updating the Responsible AI Standard, we recognized early on that it is impossible to reduce all the complex sociotechnical considerations—for many different use cases—into an exhaustive set of pre-defined rules. This led us to create a program and process for ongoing review and oversight of high-impact cases and rising issues and questions, which we call Sensitive Uses.

Our Sensitive Uses program provides an additional layer of oversight for teams working on higher-risk use cases of our AI systems. The program began under the Aether Committee in 2018 and has operated out of the Office of

Responsible AI since that office's inception in 2019. From July 2019 to May 2023, we have processed over 600 Sensitive Use case reviews from across Microsoft, including almost 150 cases during the period July 2022 to May 2023.

Think of the Sensitive Uses program as a reporting, review, and guidance framework: it starts with a mandatory reporting requirement, which then begins a hands-on responsible AI project review and consulting process with the Office of Responsible AI's Sensitive Uses team. It culminates in project specific guidance and requirements that are additional to the Responsible AI Standard's baseline requirements. The Sensitive Uses review process is triggered when Microsoft personnel are involved in developing or deploying an AI system and the foreseeable use or misuse of that AI system could:

1. Have a consequential impact on a user's legal status or life opportunities;

2. Present the risk of significant physical or psychological injury; or

3. Restrict, infringe upon, or undermine the ability to realize an individual's human rights.

Once reported, the Office of Responsible AI's Sensitive Uses team engages to triage and begin the review process with members of the project team, their Responsible AI Champion, and other relevant stakeholders. To help structure the review and drill into issues, we use not only artifacts such as the team's Responsible AI Impact Assessment and product documentation, but also close, ongoing interactions with the project team itself. During the review process, we also often call on subject matter experts from across Microsoft through focused consultations. For particularly high-impact or novel-use cases, we elevate the project for review and advice from our Sensitive Uses Panel, which is a group of Microsoft experts spanning engineering, research, human rights, policy, legal, and customer-facing organizations from around the world. Our Sensitive Uses team is also multidisciplinary by design—its members have backgrounds in social sciences, law, engineering, and policy, and prior professional experiences as data scientists, academic researchers, policy analysts, lawyers, international diplomats, and machine learning engineers.

At the conclusion of its review, the Sensitive Uses team issues its requirements for the project to move forward. Again, these are additional requirements that go beyond our Responsible AI Standard and are tailored to the specific project at hand. We have even declined opportunities to build and deploy specific AI applications as a result of a Sensitive Uses review because we concluded that the projects were not sufficiently aligned with our Responsible AI Standard and principles. For example, Microsoft Vice Chair and President Brad Smith has spoken publicly about how, through our Sensitive Uses review process, we determined that a local California police department's real-time use of facial recognition on body-worn cameras and dash cams in patrol scenarios was premature, and he shared the fact that we turned down the deal. In addition to navigating the technical challenges presented by facial recognition operating in an uncontrolled environment, our Sensitive Uses review process helped us to form the view that there needed to be a societal conversation around the use of facial recognition and that laws needed to be established.

Another important outcome of the Sensitive Uses process was our Limited Access policy for more sensitive AI platform services, which adds an extra layer of scrutiny on the use and deployment of those services. Under this policy, we not only implement technical controls to mitigate risks, but also require potential customers to submit an application for use, disclose their intended use so that it meets one of our predefined acceptable use cases, and acknowledge that they have reviewed and agree to the terms of service. Only applications for uses that align with our responsible AI principles are approved.

Case study: Applying our Responsible AI approach to the new Bing

In February 2023, Microsoft launched the new Bing, an AI-enhanced web search experience. It supports users by summarizing web search results and providing a chat experience. Users can also generate creative content, such as poems, jokes, letters, and, with Bing Image Creator, images. The new AI-enhanced Bing runs on a variety of advanced technologies from Microsoft and OpenAI, including GPT-4. Responsible AI teams across Microsoft worked with GPT-4 for months prior to its public release by OpenAI to develop a customized set of capabilities and techniques to join this cutting-edge AI technology and web search in the new Bing.

In preparing for the launch, Microsoft harnessed the full power of our responsible AI ecosystem. The new Bing experience has been developed in line with Microsoft's AI Principles, Microsoft's Responsible AI Standard, and in partnership with responsible AI experts across the company, including Microsoft's Office of Responsible AI, our engineering teams, Microsoft Research, and our Aether Committee.

Guided by our AI Principles and our Responsible AI Standard, we sought to identify, measure, and mitigate potential harms and misuse of the new Bing while securing the transformative and beneficial uses that the new experience provides. In the sections below, we describe our approach.

Identify

At the model level, our work began with exploratory analyses of GPT-4 in the late summer of 2022. This included conducting extensive red teaming in collaboration with OpenAI. This testing was designed to assess how the latest technology would work without any additional safeguards applied to it. Our specific intention was to produce harmful responses (responses are outputs from the AI system—in this case, a large language model—and may also be referred to as "completions," "generations," and "answers"), to surface potential avenues for misuse, and to identify capabilities and limitations. Our combined learnings advanced OpenAI's model development, informed our understanding of risks, and contributed to early mitigation strategies for the new Bing.

In addition to model-level red teaming, a multidisciplinary team of experts conducted numerous rounds of application level red teaming on the new Bing AI experiences before making them available in our limited release preview. This process helped us better understand how the system could be exploited by adversarial actors and improve our mitigations. Non-adversarial testers also extensively evaluated new Bing features for shortcomings and vulnerabilities.

Measure

Red teaming can surface instances of specific harms, but in production, users will have millions of different kinds of conversations with the new Bing. Moreover, conversations are multi-turn and contextual, and identifying harmful responses within a conversation is a complex task. To better understand and address the potential for harms in the new Bing AI experiences, we developed additional responsible AI metrics specific to those new AI experiences for measuring potential harms like jailbreaks, harmful content, and ungrounded content. We also enabled measurement at scale through partially automated measurement pipelines.

Our measurement pipelines enable us to rapidly perform measurement for potential harms at scale, testing each change before putting it into production. As we identify new issues through the preview period and beyond, as well as ongoing red teaming, we continue to expand the measurement sets to assess additional harms.

Mitigate

As we identified and measured potential harms and misuse, we developed additional mitigations to those used for traditional search. Some of those include:

- **Preview period, phased release.** Our incremental release strategy has been a core part of how we move our technology safely from the labs into the world, and we're committed to a deliberate, thoughtful process to secure the benefits of the new Bing. Limiting the number of people with access during the preview period allowed us to discover how people use the new Bing, including how people may misuse it, before broader release. We continue to make changes to the new Bing daily to improve product performance, improve existing mitigations, and implement new mitigations in response to our learnings.

- **AI-based classifiers and metaprompting to mitigate harms or misuse.** The use of LLMs may produce problematic content that could lead to harms or misuse. Classifiers and metaprompting are two examples of mitigations that have been implemented in the new Bing to help reduce the risk of these types of content. Classifiers classify text to flag different types of potentially harmful content in search queries, chat prompts, or generated responses. Flags lead to potential mitigations, such as not returning generated content to the user, diverting the user to a different topic, or redirecting the user to traditional search. Metaprompting involves giving instructions to the model to guide its behavior. For example, the metaprompt may include a line such as "communicate in the user's language of choice."

- **Grounding in search results.** The new Bing is designed to provide responses supported by the information in web search results when users are seeking information. For example, the system is provided with text from the top search results and instructions via the metaprompt to ground its response. However, in summarizing content from the web, the new Bing may include information in its response that is not present in its input sources. In other words, it may produce ungrounded results. We have taken several measures to mitigate the risk that users may over-rely on ungrounded generated content in summarization scenarios and chat experiences. For example, responses in the new Bing that are based on search results include references to the source websites for users to verify the response and learn more. Users are also provided with explicit notice that they are interacting with an AI system and are advised to check the web result source materials to help them use their best judgement.

- **Limiting conversational drift.** During the preview period, we learned that very long chat sessions can result in responses that are repetitive, unhelpful, or inconsistent with new Bing's intended tone. To address this conversational drift, we limited the number of turns (exchanges which contain both a user question and a reply from Bing) per chat session, until we could update the system to better mitigate the issue.

- **AI disclosure.** The new Bing provides several touchpoints for meaningful AI disclosure, where users are notified that they are interacting with an AI system as well as about opportunities to learn more about the new Bing.

Our approach to identifying, measuring, and mitigating harms will continue to evolve as we learn more—and as we make improvements based on feedback gathered during the preview period and beyond.

We share more details about our responsible AI work for the new Bing, including our efforts on privacy, digital safety, and transparency, at https://aka.ms/ResponsibleAI-NewBing.

Advancing Responsible AI through company culture

Procedures and standards are a critical part of operationalizing responsible AI and help us build a culture committed to the principles and actions of responsible AI. These complementary approaches help us turn our commitments into reality.

Our people are the core of Microsoft culture. Every individual contributes to our mission and goals. To deepen our culture of advancing responsible AI, we invest in talent focused on AI and embed ownership of responsible AI in every role.

Investing in talent

Over the years, we have invested significantly in people as part of our commitment to responsible AI. We now have nearly 350 employees working on responsible AI, with more than a third of those dedicated to it full-time. These staff work in policy, engineering, research, sales, and other core functions, weaving responsible AI into all aspects of our business.

We ask teams who develop and use AI systems to look at technology through a sociotechnical lens. This means we consider the complex cultural,

political, and societal factors of AI as they show up in different deployment contexts—and it represents a fundamental shift in the conventional approach to computer science. While the training and practices we have developed help teams foresee the beneficial and potentially harmful impacts of AI at the individual, societal, and global levels, this is not enough. Teams developing AI systems and the leadership to whom they answer could still have blind spots. That is why diversity and inclusion are critical to our responsible AI commitment.

The case for investing in a diverse workforce and an inclusive culture is well established, yet it is hard to overstate the importance of diversity and inclusion for responsible AI. That is why our ongoing and increasing investment in our responsible AI ecosystem includes hiring new and diverse talent. As our annual Diversity and Inclusion Report shows, Microsoft continues to make incremental progress on diversity and inclusion. Yet, as an industry, we still have a long way to go. The field of AI continues to be predominantly white and male: only about one-quarter of employees working on AI solutions identify as women or racial or ethnic minorities, according to McKinsey's 2022 Global Survey on AI.

We will continue to champion diversity and inclusion at all levels, especially within our responsible AI program. To build AI systems that serve society as broadly as possible, we must recruit and retain a diverse, dynamic, and engaged employee community.

Embedding ownership of responsible AI in every role

We believe that everyone at Microsoft has the opportunity and responsibility to contribute to AI systems that live up to our responsible AI commitments. All employees, in every role, bring something to this work through their diverse skills, perspectives, and passions. This shift in perspective—that no matter your job title or team, everyone can advance responsible AI—requires a shift in culture.

To support this cultural growth, we have invested in developing employee skills and fostering collaboration.

Developing knowledge and skills

We have developed training and practices to empower our teams to think broadly about the potential impact of AI systems on individuals and society.

For example, when teams are at the earliest stages of designing an AI system, our Impact Assessment guides them through:

- Articulating the intended use(s) of the AI system;
- Interrogating how the AI system will solve the problem it is intended to solve;
- Identifying impacted stakeholders (and not just Microsoft's immediate customer);
- Articulating potential harms and benefits that may affect each stakeholder; and
- Describing preliminary mitigations for potential harms.

To help teams conduct their Impact Assessment, the Office of Responsible AI has developed on-demand training, in person workshops, and supporting guidance documents with examples and prompt questions. As part of our commitment to share best practices, our Impact Assessment template and guidance document are publicly available.

In our broader responsible AI training courses available to all Microsoft employees, we orient employees to Microsoft's approach to responsible AI, including deep dives on our responsible AI principles and governance processes, and we provide content specifically tailored for data scientists and machine learning engineers.

Teams also have access to a wide range of responsible AI experts across the Microsoft ecosystem. They provide real-time engagement and feedback throughout the product lifecycle. This community includes the Aether Committee, the Office of Responsible AI, and a large and growing community of Responsible AI Champions who drive adoption of the Responsible AI Standard.

Responsible AI built into Azure Machine Learning

Fairness
Assess fairness and mitigate fairness issues to build models for everyone.

Explainability
Understand model predictions by generating feature importance values for your model.

Counterfactuals
Observe feature perturbations and find the closest datapoints with different model predictions.

Prompt Flow
Create workflows for large language-based applications to simplify prompt building, evaluation, and tuning.

Causal analysis
Estimate the effect of a feature on real-world outcomes.

Error analysis
Identify dataset cohorts with high error rates and visualize error distribution in your model.

Responsible AI scorecard
Get a PDF summary of your Responsible AI insights to share with your technical and non-technical stakeholders to aid in compliance reviews.

Azure Content Safety
Detect hate, violent, sexual, and self-harm content across languages in both images and text.

Fostering collaboration

We recognized early in our responsible AI journey the critical roles that researchers, policy experts, and engineers at Microsoft play in building our responsible AI practice. Each group brings insights and expertise vital to our work, and we strive to enable collaboration between them.

- Researchers, with a range of expertise from machine learning to the humanities, help us envision the leading edge of AI systems. They offer best practices in the identification, measurement, and mitigation of potential harms posed by AI systems as well as insights into the exciting opportunities for AI innovation.

- Policy experts define and operationalize governance for responsible AI, including crafting the rules to guide the responsible development of AI systems. Our governance framework outlines roles and responsibilities across the organization in a way that creates accountability and encourages collaboration.

- Engineers design and develop AI systems that adhere to the Responsible AI Standard. They automate and scale the steps needed to identify, measure, and mitigate potential harms posed by AI systems. They also create wholly new responsible AI solutions that are feasible based on learnings.

Frequent collaboration and reliance on each other's expertise—practices reinforced by leadership—have helped us create a culture that leads to more beneficial and responsible solutions. Through ongoing dialogue, teams consistently report that a human-centered and collaborative approach to AI results in not just a responsible product, but a better product overall.

Empowering customers on their responsible AI journey

One of our most important responsible AI commitments is to help our customers on their responsible AI journey by sharing our learnings with them. Our efforts alone are not enough to secure the societal gains we envision when responsible AI practices are adopted.

As part of this commitment, we provide transparency documentation for our platform AI services in the form of Transparency Notes to empower our customers to deploy their systems responsibly. Transparency Notes communicate in clear, everyday language the purposes, capabilities, and limitations of AI systems so our customers can understand when and how to deploy our platform technologies. They also identify use cases that fall outside the solution's capabilities and the Responsible AI Standard. Transparency Notes fill the gap between marketing and technical documentation, proactively communicating information that our customers need to know to deploy AI responsibly. You can see an example of our Transparency Note for the Azure OpenAI Service here.

Responsible AI Champions

Meet the Microsoft Responsible AI Champions

Microsoft has cultivated a network of Responsible AI Champions across the organization. These individuals are essential in advancing a responsible-by-design culture.

Mihaela Vorvoreanu, Research

"Responsible AI is not only a technical problem with technical solutions. It requires collaborating deeply and early with not only responsible AI experts, but also people experts."

Alejandro Gutierrez Munoz, Data Science

"Championing of responsible AI is essential for aligning AI systems with ethical principles, fostering trust, ensuring compliance, and promoting social responsibility."

Shweta Gupta, Customer Engineering

"I believe that applying responsible AI principles by bringing together a diverse set of stakeholders while developing AI solutions not only helps us identify and address potential risks, but also ensures that the system being developed holistically supports its objectives."

Ferdane Bekmezci, Data Science

"It takes time to inculcate a culture to an organization. I am passionate about championing its adoption across the company because it's important to ensure that AI is developed and used in a way that is ethically and socially trustworthy."

Lisa Mueller, Design

"AI is changing rapidly, so growing communities and company-wide adoption around AI principles is important to build, grow, and extend trust in AI systems. As part of this approach, it is also important to include many disciplines to contribute to this effort and really makes a difference."

Customers also need practical tools to operationalize responsible AI practices. Over the years, responsible AI research at Microsoft has led to the incubation of tools such as Fairlearn and InterpretML. The collection of tools has grown in capability, spanning many facets of responsible AI practice including the ability to identify, diagnose, and mitigate potential errors and limitations of AI systems. Since their original conception within Microsoft, these tools continue to improve and evolve externally through the contributions of active open-source communities. The collection of tools can be found under the Responsible AI Toolbox GitHub repository. Our latest tool, which is in preview, is Azure Content Safety which helps businesses create safer online environments and communities through models that are designed to detect hate, violent, sexual, and self-harm content across languages in both images and text.

Building on the Responsible AI Toolbox, Microsoft's responsible AI program has invested in integrating some of the more mature responsible AI tools directly into Azure Machine Learning so our customers will also benefit from the development of engineering systems and tools. The collection of capabilities, known as the Responsible AI Dashboard, offers a single pane of glass for machine learning practitioners and business stakeholders to debug models and make informed, responsible decisions as they build AI systems or customize existing ones. Some of our latest features added in preview include support for text and image data that enables users to evaluate large models built with unstructured data during the model-building, training, and evaluation stages, and Prompt Flow, which provides a streamlined experience for prompting, evaluating, and tuning large language models, including on measurements such as groundedness.

We have and will continue to invest in translating research-led responsible AI innovations into practical tools that support our customers on their responsible AI journeys.

The community involved in developing, evaluating, and using AI expands beyond our direct customers. To serve this broad ecosystem, we publicly share key artifacts from our responsible AI program, including our Responsible AI Standard, Impact Assessment template, and collections of cutting-edge research. Our digital learning paths further empower leaders to craft an effective AI strategy, foster an AI-ready culture, innovate responsibly, and more. These resources can be found online at https://aka.ms/rai.

Part 5:
AI in Action in Africa

How AI is addressing societal challenges

AI provides a huge opportunity to countries around the world to address major societal challenges. Below are some examples of how African innovators are already utilizing AI to drive change in areas as varied as health, finance, manufacturing, mining, environmental sustainability, cultural preservation, agriculture, education, skilling, and public services.

AI for a healthier future

We are at a unique moment in history where medicine, biology, and technology are converging on a large scale. Healthcare organizations and hospitals are using AI to develop new treatments and medicines, and helping doctors and nurses improve patient care, reduce clinician burnout, and improve overall outcomes.

Zipline's AI-Enhanced Drone Delivery: A Game-Changer for Medical Access in Africa

In Africa, the integration of AI in optimizing drone delivery routes is revolutionizing the delivery of medical supplies and vaccines, particularly in remote and hard-to-reach areas. At the forefront of this innovation is Zipline, a company that has embraced AI to transform healthcare logistics across the continent.

Zipline operates in five African countries, where its AI-driven approach is crucial in efficiently delivering blood and medical products to remote locations. The AI technology employed by Zipline is instrumental in analysing vast amounts of data to determine the most efficient delivery routes. This capability ensures timely and precise deliveries, which is vital in emergency medical situations.

The company's journey in redefining medical supply chains began with a groundbreaking agreement with the Rwandan government in 2016. Zipline established medical drone-delivery services, utilizing AI to manage and optimize flight paths and delivery schedules. This system enables rural hospitals and clinics to place orders with Zipline and receive essential medical items like blood, platelets, and frozen plasma in as little as 30 minutes.

Following its success in Rwanda, where AI-enabled drones ensure comprehensive national coverage, Zipline expanded to Ghana, Nigeria, Kenya, and Côte d'Ivoire. In each of these countries, Zipline's use of AI in drone technology plays a critical role in overcoming geographical barriers to healthcare. This innovative application of AI not only exemplifies technological advancement but also demonstrates a profound commitment to enhancing healthcare accessibility in some of the most challenging terrains. Zipline's model stands as a pioneering example of how AI can be harnessed to create tangible, life-saving impacts in global health.

Kenya Battles Child Malnutrition with AI Technology and Strategic Policies

In Kenya, malnutrition remains a significant concern, particularly among children. Recent UNICEF data reveals that 26% of Kenyan children under five are stunted, and 11% are underweight. Addressing this issue is a priority under Kenya's Big 4 Agenda, with a specific focus on enhancing food security and tackling childhood malnutrition.

This challenge is being met head-on by Amref Health Africa, the continent's largest healthcare nonprofit, along with the Kenyan Ministry of Health. Originally starting as the Flying Doctors, Amref has evolved to strengthen health systems and train African health workers, focusing on critical health challenges in 35 countries across sub-Saharan Africa. Currently, child malnutrition tops Amref's agenda.

Collaborating with Microsoft's AI for Good Lab in Kenya, Amref is developing an innovative spatio-temporal machine learning model. This model aims to detect malnutrition hotspots in Kenya to facilitate intime interventions to mitigate this challenge. The use of AI in this endeavor is not only about leveraging technology for predictions; it is about transforming data into actionable insights. Amref plans to integrate AI in analyzing comprehensive intervention reports spanning over five decades and real-time data from community health workers. This approach will enhance decision-making processes, ensuring speed, accuracy, inclusivity, and data quality in addressing malnutrition.

AI for more effective financial services

The potential for AI in financial services is immense, with opportunities across a wide range of use cases. From fraud detection and risk management to portfolio optimization and customer service, AI has the potential to transform the way financial institutions operate. By leveraging advanced machine learning and natural language processing capabilities, AI can help financial services providers make better decisions, automate processes, and reduce costs.

M-KOPA's AI-Driven Financial Solutions Empower Millions in Africa

M-KOPA, established in 2011 in Kenya and now operating in Uganda, Nigeria, Ghana, and South Africa, represents a significant stride in financial inclusion across Africa. By 2022, the company had distributed over $1 billion in credit, primarily to individuals typically excluded from conventional banking systems. M-KOPA's philosophy is grounded in providing accessible, affordable, and sustainable financial solutions, thereby enabling a wider demographic to engage in the digital and financial economy.

A key element in M-KOPA's successful business model is the integration of AI and machine learning capabilities, powered by Microsoft Azure. This technological advancement has substantially improved customer repayment performance, a crucial aspect considering the high-risk nature of M-KOPA's customer segments.

Utilizing the Azure cloud ecosystem, M-KOPA can generate daily forecasts for loan repayments. This function is critical in striking a balance between offering inclusive pricing and maintaining financial responsibility. The impact of these AI-enhanced capabilities is far-reaching, with M-KOPA providing over 3 million customers in Africa access to credit, thus fostering economic empowerment for individuals and families.

Moreover, M-KOPA's portfolio extends beyond credit services. The company now offers health insurance policies, with approximately 204,000 customers currently enrolled as policyholders. This expansion not only underscores M-KOPA's commitment to financial inclusivity but also highlights the transformative role of AI in redefining the scope and efficacy of financial services in emerging economies.

AKI Revolutionizes Kenya's Insurance Sector with AI and Digital Solutions

Confronting a multitude of challenges, Kenya's insurance industry, represented by the Association of Kenya Insurers (AKI), sought innovative solutions to enhance efficiency and trust. The industry, burdened by the high costs of manual, paper-based operations, and slow processing times, needed a transformative approach to tackle motor insurance fraud, improve its claims management process and gain back the trust with its customers. More specifically, the Insurance Regulatory Authority highlighted that documenting an automobile claim could take 10 to 12 days and about 25% of the claims were found to be fraudulent. The inefficiencies not only led to prolonged settlement periods, but also contributed to deteriorating customer service and eroding trust in Kenya's overall insurance sector.

In response, SwiftAnt, a Microsoft Gold Partner, developed a Digital Insurance Apps (DIAS) solution for AKI leveraging a suite of Microsoft technologies including Azure Machine Learning, Azure Cognitive Services, AI, and blockchain. This solution automated claims reporting, recorded undeniable risk information, reduced fraud, and enabled accurate repair estimates.

The impact of SwiftAnt's DIAS solution has been profound. The time required to file a claim has been reduced to just six minutes, and the cost associated with processing insurance certificates has been halved. Additionally, the solution aids police in verifying the authenticity of certificates. Policyholders have reported significant improvements, particularly appreciating the expedited, automated claims process which provides crucial assistance in the event of accidents. This digital transformation, underpinned by AI and blockchain, not only enhances operational efficiency but also reinstates trust and reliability in Kenya's insurance sector.

Investec's Strategic Use of AI in Enhancing Banking Services

In the evolving banking landscape of South Africa, Investec has identified the need for innovation to keep pace with changing customer expectations and market dynamics. To address this, the company has focused on integrating advanced technologies into its services, especially for its clientele of ultra-high net worth individuals.

A significant step in this direction is the deployment of ZebraGPT, an AI-powered tool based on Azure OpenAI. This tool is designed to streamline complex banking processes, enhance efficiency in task management, and improve code writing capabilities, ensuring a more efficient workflow. A key aspect of ZebraGPT is its emphasis on maintaining high standards of privacy and security, a critical requirement in the financial sector.

After extensive testing and implementation across various departments, ZebraGPT is now utilized by over 9,000 Investec employees. Its application ranges from simplifying repetitive tasks to providing enhanced learning and knowledge resources. This widespread adoption within the company underscores its utility and effectiveness.

Looking ahead, Investec is exploring the broader potential of AI in banking, having identified numerous applications in areas such as wealth and investment management. This initiative reflects Investec's approach to leveraging AI not just as a tool for immediate problem-solving but as a strategic asset for long-term innovation in banking services.

AI for manufacturing

As AI functionality expands, its potential applications in manufacturing are growing exponentially. Companies are already finding that AI is helping to make the way we produce things safer, more efficient, and more profitable by analysing vast amounts of data and identifying opportunities for improvement.

AB InBev's Integration of AI in Global Operations

After its 2016 merger with South African brewer SABMiller, AB InBev emerged as the world's largest beer company, operating across more than 50 markets, including several in Africa, and selling over 500 brands globally. With a workforce of nearly 180,000, the company is deeply integrated into communities worldwide.

AB InBev's transition into a technology-focused organization marks a significant shift in its business approach. Moving beyond initial stages of utilizing chatbots, the company has embraced machine learning to enhance various functional and operational areas. The integration of customer-centric AI has yielded profound insights into consumer behavior, customer preferences, and brand dynamics. Furthering this technological trajectory, AB InBev is exploring the use of virtual and augmented reality for product development, marketing, and training.

A cornerstone of AB InBev's success in AI implementation lies in it having invested in recruiting data scientists and engineers and ensuring that its senior leadership possesses a thorough understanding of AI, enabling the deployment of AI-powered solutions across the business while adhering to regulatory compliance. To sustain and amplify its AI initiatives, AB InBev has also established internal AI talent pools and Centers of Excellence designed to facilitate the development and deployment of AI related applications. This structured approach to integrating AI underscores AB InBev's

commitment to leveraging technology for continuous improvement and innovation in its global operations.

Coca-Cola Beverages Africa Enhances Operations with AI-Driven Technology

Coca-Cola Beverages Africa (CCBA), the leading Coca-Cola bottler in Africa, has made significant strides in optimizing its operations across numerous bottling plants and retail outlets on the continent. A key aspect of this transformation is CCBA's integration of AI-driven Microsoft Dynamics 365 technology. This suite of products is designed to streamline CCBA's supply chain, customer service, and field operations.

One of the standout features is the AI-powered inventory management system. This system is instrumental in providing real-time data on parts availability for field service technicians and in processing sales demands efficiently. The AI capabilities within this system allow for quick analysis and response, facilitating smoother operations in warehousing and inventory management.

Additionally, CCBA's use of Power BI, Microsoft's business analytics tool enhanced by AI algorithms, allows the team to make faster, more informed decisions. The shift to Microsoft Azure, the cloud computing platform with advanced AI functionalities, has also been a crucial factor in reducing operational costs and improving efficiency. This move has led to a substantial reduction in IT expenses, demonstrating how AI-driven technological advancements can lead to significant cost savings and operational improvements.

CCBA's example illustrates the profound impact of AI in streamlining business processes and decision-making, showcasing the potential for AI technology to revolutionize operations in the beverage industry and beyond.

AI for a more sustainable future

The climate challenge is a race against the clock to decarbonize. AI is helping accelerate and scale efforts to transition to clean energy and mitigate the impact of global warming on people and communities.

Microsoft's AI for Good Lab: Tackling Climate Change with AI in Africa

Microsoft's AI for Good Lab is actively using AI, machine learning, and statistical modeling to combat climate change. This initiative involves collaborations with nonprofits, research institutions, and governments, focusing on developing scalable solutions using Microsoft technology. A key project in this endeavor is the collaboration between Microsoft, Planet Labs, and The Nature Conservancy to create the Global Renewables Watch. This living atlas, using AI and satellite imagery, aims to map all utility-scale solar and wind installations worldwide. It will provide essential data to researchers and policymakers, aiding in understanding renewable energy capacities and informing decisions for more efficient renewable energy development. This project aligns with the pursuit of the UN Sustainable Development Goals. Recently, the Lab expanded into Nairobi, Kenya, and Cairo, Egypt, forming a new team of AI scientists dedicated to enhancing climate resilience.

The establishment of the Africa AI Innovation Council, comprising members from prominent African organizations like the African Development Bank and African Climate Foundation, marks a significant step forward. This Council, a collective of African leaders knowledgeable about the continent's challenges, guides the AI for Good Labs in Africa. Their focus is on identifying and implementing AI-driven strategies for climate adaptation and early warning systems and generating climate data for ongoing research. The council has kicked off its first project, led by the African Risk Capacity, to map flood risk areas, with Mozambique and Ethiopia as use cases.

Moreover, the expanded collaboration between Microsoft AI for Good Lab and Planet Labs grants Africa-based data scientists access to satellite imagery. This resource is vital for addressing challenges identified by the Africa AI Innovation Council, with a special emphasis on adaptation to climate change and developing early warning systems.

Enhancing Emergency Planning in Africa with AI-Powered Mapping Tools

In 2017, weather-related disasters led to the displacement of 18 million people, highlighting the urgency for improved natural disaster and emergency planning. A significant challenge in disaster management, particularly in Africa, is the absence of detailed mapping in many vulnerable areas, which complicates efficient disaster preparedness and relief efforts.

To combat this, technology, especially advanced mapping tools, is increasingly crucial. The Humanitarian OpenStreetMap Team (HOT) has been pivotal in mapping 11 million square kilometers across Africa using their Tasking Manager tool. This tool divides large mapping projects into smaller tasks, allowing multiple contributors to work simultaneously on a larger scale, maintaining consistent quality.

Despite these advancements, approximately 60% of Africa, including many disaster-prone regions, remains unmapped. Bing Maps and Microsoft Philanthropies have joined forces with HOT to integrate AI Assistance into the map-building process in Uganda and Tanzania, as part of Microsoft's AI for Humanitarian Action Program. The focus is on enhancing the Tasking Manager with machine learning (ML) to improve the efficiency and accuracy of mapping tasks. ML allows the Tasking Manager to link different building datasets via an API, granting access to various building footprints from Bing Maps.

Additionally, the Bing Maps team employs Computer Vision to identify map features on a large scale. As a result of this collaboration, comprehensive building footprint datasets for Uganda and Tanzania will be released, contributing to the OpenStreetMap (OSM) free of charge, aiding research and analysis. OSM currently hosts around 14 million building footprints for these countries, and the new initiative is expected to add an additional 18 million open building datasets, significantly enhancing the resources available for emergency planning and disaster management in Africa.

Pulse Lab Kampala Utilizes AI for Refugee Support in Uganda

Uganda, known for its progressive refugee protection policies, has witnessed a rapid increase in its refugee population, particularly with the influx of South Sudanese refugees in the northern region. This surge has strained resources and heightened tensions between local communities and the refugee population.

To address these challenges, Pulse Lab Kampala, in collaboration with the United Nations Development Operations Coordination Office, has developed an innovative solution. They have created a Radio Content Analysis Tool that employs AI for real-time analysis of public radio discussions about refugees, functioning similarly to social media analysis tools.

This AI-powered tool is instrumental for development and humanitarian agencies, enabling them to gauge the level of concern and sentiment among local communities towards refugees. The insights gained from this analysis serve as an early warning system and assist in making more informed decisions. This tool also aids in the strategic planning and implementation of various programs aimed at easing tensions and fostering better integration between refugees and host communities.

Pulse Lab Kampala's initiative demonstrates a proactive approach in using AI to navigate complex social dynamics, providing valuable data that informs policies and actions to improve refugee integration and community relations in Uganda. It also underlines the importance of having in place policies to ensure AI tools are monitored and used in ways that help ensure outputs do not reinforce biases and do uphold fairness outcomes.

Forest Guard: AI-Powered Solution to Tackle Illegal Logging in Kenya

According to the World Bank, an area of forest the size of a football field is lost to illegal logging every two seconds, posing a significant challenge to environmental conservation. In Kenya, the task of monitoring forests is particularly daunting due to limited resources, making it difficult for guards to effectively patrol the vast forest lands. Illegal logging not only harms wildlife and local communities but also exacerbates climate change by increasing carbon emissions.

Addressing this critical issue, four Microsoft interns in Kenya developed an innovative solution named Forest Guard. Conceived during a Microsoft company hackathon, Forest Guard is an AI-based device equipped with sound and geo-sensors, and machine learning capabilities. This technology is specifically designed to identify the sound of chainsaws, a common indicator of illegal logging activities.

When illegal logging is detected, the system automatically sends SMS and email alerts to forest officials. These alerts include the time and geo-location of the suspected logging activity, enabling prompt and targeted response. This technological intervention by Forest Guard represents a significant advancement in the fight against illegal logging in Kenya.

The implementation of Forest Guard demonstrates a practical application of AI and machine learning in environmental protection. It not only aids in the preservation of forests but also contributes to broader efforts to combat climate change, underscoring the importance of technological innovation in ensuring the planet's survival.

Elephant Listening Project: AI-Enhanced Monitoring to Protect Forest Elephants

The African forest elephant, primarily found in sub-Saharan Africa, has experienced a drastic population decline, with numbers falling by over 86% between 1989 and 2020, as reported by the International Fund for Animal Welfare. Recognized as a keystone species, their survival is crucial for the health of their ecosystems.

Cornell University's Elephant Listening Project was established to protect these elephants using acoustic sensors that track their movements in the wild. However, the remoteness of their habitats and the large size of the sound files often resulted in lengthy data collection and analysis periods. This delay hindered the ability to use the data for timely interventions against immediate threats like poaching.

Addressing this issue, the startup Conservation Metrics developed an AI tool capable of distinguishing elephant calls from other rainforest sounds. Approximately 50 sensors were installed in the Nouabalé-Ndoki National Park in the Republic of Congo. These sensors capture audio files that are then processed using machine learning and deep neural networks.

The application of AI significantly accelerates the analysis of these files, enabling quicker examination and allowing scientists to promptly alert park rangers about potential threats. This technological advancement enhances the capacity to protect forest elephants by providing timely data that can be used to prevent poaching and other immediate dangers, demonstrating the crucial role of AI in wildlife conservation.

AI to preserve our culture

UNESCO predicts that between 50 and 90% of endangered languages will disappear by the next century, threatening cultural heritage around the world. AI can help preserve heritage and languages, advancing broader opportunities and a more inclusive future.

Microsoft Translator: Bridging Language Barriers in Africa with AI

Microsoft is making strides in language preservation in Africa by adding 13 new African languages to the total number of 124 supported languages in its Azure Cognitive Services Translator. This platform is integrated across the Microsoft ecosystem, enabling translation of text and documents into these languages across various products.

The Translator's integrations include Microsoft 365 for text and document translation, Microsoft Edge and Bing for webpage translation, SwiftKey for message translation, LinkedIn for user content translation, and the Translator app for on-the-go multilingual conversations. These integrations facilitate seamless communication and content accessibility in a wide range of African languages.

Educators can leverage Translator for creating more inclusive classrooms by enabling live captioning and cross-language understanding for both students and parents.

To enhance the quality of machine translation models, Microsoft collaborates with partners in language communities. These partners assist in gathering bilingual data, consulting with community members, and evaluating translation quality. This collaborative approach helps overcome the challenge of obtaining sufficient bilingual data for training machine translation models.

The continuous improvement of these capabilities is enabling businesses to extend their global reach. Microsoft Translator aids in communicating with customers and partners across languages, localizing content and apps efficiently, cost-effectively, and reliably.

Lelapa AI: Promoting Inclusivity in AI for Underrepresented Languages

Lelapa AI, a South African startup, is addressing the limitations of traditional AI technologies in representing the global community, particularly those in the developing world. Recognizing the neglect of indigenous languages and identities by conventional AI tools, Lelapa AI aims to spearhead a more inclusive AI conversation.

Dedicated to bridging the digital divide, Lelapa AI focuses on revitalizing and enriching underrepresented languages. The startup is committed to utilizing advanced AI technology to ensure that individuals from diverse linguistic backgrounds can access the benefits of the digital world.

In collaboration with Microsoft's AI for Good Lab, Lelapa AI has entered into a partnership that involves working with Microsoft research scientists on African-specific issues and utilizing a Microsoft Azure grant. This grant provides students and scientists with access to sophisticated AI computing resources.

Currently centered on Southern Africa, Lelapa AI aspires to extend its influence across the continent through strategic alliances with organizations that have a deep understanding of local contexts.

Lelapa AI recently released Vulavula, an AI-as-a-service platform. Vulavula offers a range of services tailored to diverse linguistic needs, including transcription, text analysis, and text-to-speech services. This platform exemplifies Lelapa AI's commitment to making AI technology accessible and relevant to a broader range of languages and communities.

Microsoft Africa Research Institute: Pioneering AI Equity Research in Africa

The Microsoft Africa Research Institute (MARI), based in Nairobi, Kenya, is a multi-disciplinary research group dedicated to advancing AI equity with a focus on the African continent. MARI's mission is to develop AI solutions that are inclusive, fair, and beneficial, specifically tailored to the diverse and complex contexts of Africa. A key part of this mission is empowering African researchers, innovators, and communities in the development and application of AI technologies.

MARI primarily investigates the challenges and opportunities of applying generative AI models, such as Large Language Models (LLMs), in African settings. While LLMs hold significant potential for text, image, or speech generation, their underlying models, training data, and evaluation metrics are predominantly based on the Global North, often overlooking the linguistic, cultural, and social diversity of the Global South. MARI's research is twofold: it examines the socio-technical challenges faced by African workers and companies around model use and develops technical solutions to ensure model equity.

In November 2023, MARI co-organized a workshop in Nairobi on AI and the Future of Work in Africa. This event, held in partnership with entities including the AfOx Network, the African Union Development Agency NEPAD, and Microsoft East Africa, brought together thought leaders to discuss the impact of generative AI on African workforces. The workshop focused on four key areas: macro-economics, job markets, workers' perspectives on AI, and Africa-centric AI tools. A white paper summarizing the workshop's findings and recommendations is scheduled for publication in early 2024.

AI for a growing future in agriculture

Ensuring we have access to affordable food even as the world's population grows and weather patterns become more extreme requires us to think differently about farming. AI is helping farmers improve yields even as they use fewer resources, especially water.

Hudhud

Agriculture is a significant part of African economies, exports, and employment – making it a critical area of investment. According to USAID, agriculture sustains around 70% of livelihoods for those living in more rural parts of Africa. Crop yields, however, are decreasing dramatically because of climate change.

To mitigate this challenge, in 2021 the Egyptian government announced the trial of digital services to help enhance agricultural production in Egypt. This included the launch of the Hudhud smart assistant for farmers.

Hudhud is an AI-powered app that provides farmers with guidance and data on agricultural best practices in Arabic. The app is designed to help improve qualitative development in the agricultural extension system, providing tailored direction to farmers according to their particular needs, crops and the pests that might be infecting their crops.

For example, if a farmer finds an infection in his crops, he can take a picture of the plant in question and send it to Hudhud. The application will then leverage AI to identify the disease and provide the farmer with guidance around how to stop the infection.

Currently, Hudhud is used by numerous farmers covering 70% of the farming land across Egypt thereby supporting the digital transformation of agriculture and improving food security in the country.

FarmVibes.Bot: Enhancing African Agriculture with AI

FarmVibes.Bot, targeting the 80% of Africa's farming community that are smallholder farmers, leverages AI to provide vital agricultural support. Designed to connect these farmers, often in remote areas, with essential services, the platform uses AI-driven analytics to deliver personalized agricultural information and services.

The platform operates through various channels like USSD, SMS, and WhatsApp, and integrates services from diverse entities including financial, insurance, and market access providers. AI is key in aggregating and analyzing data from these sources, offering features like crop advisories, pest diagnoses, and market insights. This approach enables farmers to access crucial information, including local weather updates, market prices, and soil testing data, helping them make informed decisions for better crop management and productivity.

AI's role extends to facilitating participation in carbon markets, a recent feature introduced through a partnership with the Alliance for a Green Revolution in Africa (AGRA) and Rabobank. This AI-enabled training on carbon credits empowers farmers to understand and engage in these markets effectively.

FarmVibes.Bot's impact is significant, aiding 500,000 farmers in adopting data-driven precision farming techniques, thereby increasing their productivity and income. This platform demonstrates how AI can be a transformative tool in addressing the unique challenges of smallholder farmers in Africa.

AI for education

In schools around Africa, AI is transforming the way we learn and train, making it more accessible, personalized, and impactful, benefiting educators and students alike. Although AI has already been used to boost inclusivity and productivity, recent advancements offer the opportunity to reimagine solutions to longstanding challenges in education at a time when they are perhaps needed most.

Eduvos

Though the demand for higher education in Africa is growing, according to UNESCO, Africa's enrolment in higher education is just 9%, compared to the global average of 42%.

Higher education institutions such as Eduvos in South Africa are implementing systems like Dynamics 365 to develop actionable real-time insights across its education, sales, financial, and academic departments to improve its operations and deliver more learning opportunities to students in South Africa.

Within a remarkably short time, Eduvos gained the insights it needed to enhance its educational offerings. While student enrolment previously took nearly 90 days, it's now almost instantaneous and the associated cost has been reduced by 90%.

As a result, Eduvos has witnessed an impressive 50% year-over-year growth in enrolment for two consecutive years. It scaled from 5,800 students to 12,000 students and enrolment growth from 1,500 new students to an impressive 7,400. In addition, by replacing many of the manual processes in enrolment with automation, Eduvos is looking forward to enrolling up to 20,000 students in 2024.

Codar Technologies

Digitization across Africa has surely but steadily been accelerating demand for a variety of new specialist roles related to understanding and leveraging the latest emerging technologies. This includes specialists in cloud computing, security and data protection, and mobile technology, as well as data analysts and IT services managers. In fact, over 230 million jobs in sub-Saharan Africa will require digital skills by 2030, resulting in almost 650 million training opportunities, according to the IFC Digital Skills in sub-Saharan Africa study.

To help empower a new generation of skilled tech professionals, Codar Technologies has entered the Nigerian education technology scene with an innovative platform powered by AI. The solution uses AI to analyze a student's academic history, learning style and career aspirations. Based on findings, the platform then generates personalized learning recommendations, including study materials, practice questions and career guidance. These resources are ultimately designed to help users enhance their academic performance and achieve specific tech-oriented career goals.

Codar also harnesses machine learning to improve the quality of audio and video content, while leveraging a chatbot to provide students with real-time feedback. These AI-powered insights provide students with the latest subject information, helping them to keep pace in a rapidly evolving industry.

Finally, the platform is designed to provide language-inclusive education, making it possible for students from all different backgrounds to access the training they need.

AI skills for the future

AI offers tremendous potential to empower workers around the world—but only if everyone, everywhere has the skills to use it. To thrive in an AI-enabled economy, the workforce must be prepared and trained with the digital and AI skills required for the in-demand jobs of the future.

Microsoft Africa Development Centers: Driving AI Innovation and Tech Talent

Microsoft operates two Africa Development Centers (ADCs) strategically located in Nairobi, Kenya, and Lagos, Nigeria. These centers serve as hubs where Microsoft's engineers harness the power of AI and machine learning to drive innovation across diverse sectors, including healthcare, agriculture, finance, and human-centric automation.

Within the ADCs, dedicated teams tackle a wide spectrum of critical company projects. These projects encompass key areas such as Identity and Network Access, Microsoft 365, Windows Engineering, Azure Edge and Platform, and Engineering Fundamentals, among others.

Notably, the ADCs are not limited to local impact; they actively contribute to global innovations. For instance, the Nairobi team played a pivotal role in redesigning global Microsoft products,

exemplified by their work on Microsoft 365's Graph Explorer. The Lagos-based engineers made a significant mark in 2021 when their efforts came to fruition at Microsoft's flagship Ignite conference with the launch of Microsoft Mesh, the company's cutting-edge mixed reality platform.

Beyond innovation, the ADCs are committed to nurturing Africa's tech talent ecosystem. They engage in various initiatives, including faculty skilling programs at universities, curriculum reviews, and student hackathons, all aimed at developing and fortifying the region's pool of tech professionals. These efforts underscore Microsoft's dedication to fostering technology advancements and talent growth across the African continent.

Digital Nigeria for a Digital Economy

In 2021, Microsoft partnered with the Federal Government of Nigeria and the African Development Bank to launch Digital Nigeria, an initiative that aims to build strong foundations for a digital economy in Nigeria, focusing on advancing connectivity, skilling, and digital transformation. As part of this initiative, Microsoft aims to train 5 million people in Nigeria over three years on digital skilling covering digital literacy, advanced digital skills training, and connecting youth to improved livelihood opportunities. To date the partnership has seen the deployment of the Microsoft Community Training Platform and this has reached over 4 million Nigerians with free tech skilling content. Over 605,000 had been trained by the end by December 2023. In June 2023, Microsoft partnered with the Data Scientists Network (DSN) of Nigeria to

train 300 young professionals on the use of the Microsoft Responsible AI Dashboard, promoting ethical AI development and a fair digital future. The dashboard empowers developers to assess and address biases in AI models, demonstrating Microsoft's commitment to responsible AI implementation in Nigeria's digital transformation. As of October 2023, over 10,000 youth had been trained in responsible AI through the partnership with DSN.

The Generative AI Skills Challenge established in collaboration with Microsoft and data.org , is a global grant for best-in-class organizations training and upskilling individuals on generative AI to drive social impact and advance socioeconomic mobility. The Challenge sought out fair and community-led integration in low- and middle-income countries and contexts to accelerate digital inclusion and skills advancement and acquisition for workers from historically marginalized populations around the world. With support from Microsoft and data.org, the Challenge provides funding and technical assistance to awardees to build and scale their programs to train, upskill, and reskill today's workforce to use generative AI.

Selected from a pool of nearly 600 applications across 93 countries, the Global Integrated Education Volunteers Association's (GIEVA) goal is to bring women entrepreneurs in northern Nigeria into the digital economy by training and upskilling them on the use of generative AI capabilities to create digital livelihoods. Through a consortium of local partners GIEVA is building an AI-based platform for women-led businesses

to increase sales by providing tailored sales advice and creating compelling online product profiles and digital sales assets optimized for sharing over WhatsApp and mobile phones.

South Africa's New AI Era

Microsoft South Africa and the Youth Employment Service (YES) have joined forces to provide comprehensive AI training to empower 300,000 young South Africans with digital skills. This initiative, showcased at Microsoft South Africa's "A New Era with AI" event in Sandton, introduces the AI Youth Training Platform, co-developed by YES and Microsoft. The platform offers tailored AI and computer science training, catering to all levels of expertise, from beginners to experts. By making AI accessible, particularly through an introductory module, it aims to bridge the AI knowledge gap among young individuals. The partnership seeks to stimulate job creation, improve employability, and foster digital inclusion in South Africa, while also providing a certification pathway for those who complete the program. Microsoft's collaboration with YES is expected to have a significant impact on job opportunities and talent development in South Africa's digital sector. Additionally, Microsoft Philanthropies has launched the Generative AI skilling learning pathway in partnership with LinkedIn, with over 40,000 participants engaged in the course continent-wide, contributing to the growth of AI skills in Africa.

AI for more efficient mining

To operate more efficiently and sustainably amid the global supply chain disruptions that have emerged in recent years, mining companies need to reduce their carbon footprint and make their core processes like supply chain management more flexible and resilient. Digital solutions that leverage machine learning and predictive analytics help them react more quickly to shifting market conditions and risks that include cybersecurity, geopolitical, and climate events.

Transforming Botswana's Diamond Industry: AI-Powered Efficiency at Debswana

Debswana, the global leader in diamond production, stands as the cornerstone of Botswana's economy, contributing a substantial 40% to the nation's GDP and a remarkable 70% to its export earnings. Any notable advancement or shift in how Debswana retrieves diamonds translates into a substantial economic boost for the country.

In the not-so-distant past, Debswana grappled with a slew of challenges, including frequent machine breakdowns and unreliable performance, leading to considerable financial setbacks. The hefty cost of repair maintenance loomed large as one of the company's major expenses. Recognizing the potential solution within the wealth of data generated by its machinery and applications, Debswana turned to Microsoft Azure.

Now armed with the capability to effortlessly amass and analyze data by leveraging AI, Debswana has unlocked a new era of efficiency. It can proactively anticipate and prevent issues before they disrupt operations, resulting in a remarkable 10% boost in operational efficiency. This transformation not only marks a significant turning point for Debswana but also underscores its pivotal role in shaping the future of Botswana's diamond industry through AI-powered innovation.

Enhancing Safety and Productivity: The AI Transformation at Syama Mine

Mines, renowned for their perilous work environments, place worker safety at the forefront of the digital transformation underway in the African mining industry. Located in Mali, the Syama mine, operated by Resolute Mining, has been recognized as one of the world's foremost automated mines. Several years ago, it achieved a groundbreaking milestone by fully automating its entire mining process, encompassing everything from drilling and extraction to rubble clearance and surface transportation, all orchestrated by autonomous machinery.

The integration of AI into Syama's operations not only amplifies mining productivity but also serves as a vital component in enhancing worker safety. By reducing the exposure of personnel to the traditional hazards associated with mining, these autonomous machines play a pivotal role. It's important to note that AI does not replace skilled mine workers; rather, it empowers them to conduct their work on the surface, away from the underground risks.

In its 2022 Sustainability Report, Resolute Mining reported an impressive total recordable injury rate of just 0.41 per million hours worked across its operations. Notably, Syama mine reported zero long-term injuries among its staff. This compelling case study underscores the profound impact of AI in fostering both safety and efficiency within the mining industry.

AI for the future of public services

By incorporating digital solutions and AI into their operations, governments and professional services are innovating how they deliver personalized, real-time services to citizens and improve tasks like data analysis, reporting, and communication. AI's role extends across various sectors, enhancing creativity, efficiency, and problem-solving in areas from research to design and development.

Empowering Communities with AI: ICE Commercial Power's Impact

Nigeria, despite its economic stature, has grappled with an electrification rate of under 60% in 2021, leaving over 85 million people without access to reliable electricity. ICE Commercial Power, a Nigerian renewable energy provider, sought to change this narrative. Facing operational inefficiencies in their off-grid electrification efforts, they realized the need for data-driven solutions. The company initiated the Ignite program, training unemployed youth to collect hyperlocal geospatial and demographic data. Leveraging AI tools, this data was analysed to identify over 12,000 small businesses that could benefit from ICE's energy solutions.

By harnessing AI and Microsoft's support, ICE streamlined its operations, deploying clean energy solutions to underserved communities and small businesses. Through this collaboration, ICE demonstrated the transformative power of AI in rapidly expanding clean energy access, impacting the lives of millions in Nigeria.

Enhancing Accessibility: MultiChoice's AI-Driven Captioning Solution

MultiChoice, Africa's leading video entertainment and broadcast company with over 21 million customers across 50 countries, recognized the importance of catering to diverse linguistic needs. With between 1,000 and 2,000 languages in Africa, including South African accents and dialects, MultiChoice aimed to ensure accessibility to its content for all viewers, including those with hearing difficulties and international audiences.

Manual transcription and captioning processes were proving slow and costly, hindering MultiChoice's ability to meet the growing demand for content. Seeking an efficient solution, the company partnered with Microsoft to develop an AI-powered solution named 'ODIN.' Leveraging Microsoft Speech to Text models within the Azure stack and trained on the broadcaster's extensive content library, ODIN achieved an impressive 85% accuracy rate.

ODIN's impact is evident in its ability to automate subtitle creation, significantly reducing processing time. This innovation not only enhances accessibility for viewers but also opens doors for cost-effective content translation, expanding MultiChoice's reach to new markets.

Securing South Africa's Water Future: eThekwini Municipality's AI-Enhanced Water Management

South Africa faces significant water scarcity challenges, with projections indicating a potential water deficit of up to 17% of available sources by 2030. eThekwini Metropolitan, the country's third-largest municipality, serves over 1.1 million households and plays a crucial role in South Africa's economy. To address water loss issues, eThekwini Municipality partnered with Microsoft to develop an innovative water security blueprint with far-reaching implications.

This collaboration aims to revolutionize water management by creating a multi-dimensional data usage platform that harnesses insights from six data sources. It offers real-time reporting, enabling rapid response strategies and facilitating strategic planning and operations. The platform not only addresses immediate water loss concerns but also delves into deeper issues related to non-revenue water. eThekwini's pioneering efforts in AI-enhanced water monitoring serve as a model for municipalities in South Africa and other water-scarce regions worldwide, illustrating the potential of AI in securing precious water resources for the future.

Sources/Resources

Governing AI in Africa: Policy Frameworks for a New Frontier

Cited Sources

- Artificial Intelligence centre could expand economy
  **https://www.universityworldnews.com/post.
  php?story=20210301092515749**

- Africa's emerging economies to take the lead in consumer
  market growth
  **https://www.brookings.edu/articles/africas-emerging-
  economies-to-take-the-lead-in-consumer-market-
  growth/**

- Microsoft AI Skills
  **https://www.microsoft.com/en-us/corporate-
  responsibility/ai-skills-resources?rtc=1**

General Sources

- Zipline's Ambitious Medical Drone Delivery in Africa
  **https://www.technologyreview.com/2017/06/08/151339/
  blood-from-the-sky-ziplines-ambitious-medical-drone-
  delivery-in-africa/**

- UjuziKilimo
  https://www.ujuzikilimo.com/

- Farmcrowdy
  http://farmcrowdy.com/

- How Tech Will Make this the Most Liveable City in Africa by
  2030
  **https://sponsorcontent.cnn.com/edition/2021/microsoft/
  en/how-tech-will-make-this-the-most-liveable-city-in-
  africa-by-2030.html**

- M-PESA
  **https://www.vodafone.com/about-vodafone/what-we-
  do/consumer-products-and-services/m-pesa**

- AI: Ushering in an age of empathy
  **https://www.investec.com/en_za/focus/innovation/
  ai-ushering-in-an-age-of-empathy.html**

Part 1: AI Explained

Cited Sources

- GitHub Copilot: GitHub
  https://github.com/features/copilot

- Microsoft Bing Blog: Azure OpenAI Service
  https://blogs.bing.com/search/2020_05/AI-at-Scale-in-Bing

- Power Platform: Microsoft
  https://powerplatform.microsoft.com/en-us/what-is-power-platform/

- Microsoft Designer: Microsoft
  https://designer.microsoft.com/

- Bing and Edge: Yusuf Mehdi, Microsoft, February 2023, Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web
  https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/

General Sources

- Microsoft, February 2023, What is Microsoft's approach to AI?
  https://news.microsoft.com/source/features/ai/microsoftapproach-to-ai/

- Microsoft, March 2023, ExplAIning AI: The fundamentals and the frontiers
  https://news.microsoft.com/source/features/ai/explaining-ai-the-fundamentals-and-the-frontiers/#approach-to-ai

- Microsoft, February 2022, The innovation behind AI at Scale
  https://blogs.microsoft.com/ai-for-business/ai-at-scale-technology/

Part 2: New steps toward democratizing AI

Cited Sources

- Microsoft Airband Initiative
  https://www.microsoft.com/en-us/corporate-responsibility/airband-initiative

- Microsoft AI Skills Initiative
  https://www.microsoft.com/en-us/corporate-responsibility/ai-skills-resources?rtc=1

- Microsoft Skills for Jobs
  https://www.microsoft.com/en-ie/skills-for-jobs/

- Microsoft AI for Good Lab
  https://www.microsoft.com/en-us/research/group/ai-for-good-research-lab/

- Closing the Climate Data Divide in the Global South
  https://blogs.microsoft.com/on-the-issues/2022/11/07/climate-data-divide-global-south/

Part 3: Governing AI in Africa

Cited Sources

- AI Risk Management Framework | NIST
  https://www.nist.gov/itl/ai-risk-management-framework

- African Union Convention on Cyber Security and Personal Data Protection
  https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection

- African Union High-Level Panel on Emerging Technologies
  https://www.nepad.org/microsite/african-union-high-level-panel-emerging-technologies-apet

- SMART Africa Blueprint
  https://smartafrica.org/

- National AI Strategies for Inclusive & Sustainable Development
  https://thefuturesociety.org/policies-ai-sustainable-development/

- Microsoft, Anthropic, Google, and OpenAI launch Frontier Model Forum - Microsoft On the Issues
  https://blogs.microsoft.com/on-the-issues/2023/07/26/anthropic-google-microsoft-openai-launch-frontier-model-forum/

- Alignment Research Center
  https://www.alignment.org/

- Update on ARC's recent eval efforts - ARC Evals (metr.org)
  https://metr.org/blog/2023-03-18-update-on-recent-evals/

- Microsoft announces new supercomputer, lays out vision for future AI work - Source
  https://news.microsoft.com/source/features/ai/openai-azure-supercomputer/

- The 5th Ordinary Session of the Specialized Technical Committee on Communication and Information Communications Technology
  https://au.int/en/5thstccict

- Mauritius Artificial Intelligence Strategy
  https://mitci.govmu.org/Documents/Strategies/Mauritius AI Strategy.pdf

- Rwanda National AI Policy
  https://www.minict.gov.rw/index.

- Recommendation on the Ethics of Artificial Intelligence - UNESCO Digital Library
  https://unesdoc.unesco.org/ark:/48223/pf0000381137

- 193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence | UN News
  https://news.un.org/en/story/2021/11/1106612

- AI Advisory Body: Interim Report: Governing AI for Humanity
  https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf

- Egypt National Artificial Intelligence Strategy
  https://mcit.gov.eg/Upcont/Documents/Publications_672021000_Egypt-National-AI-Strategy-English.pdf

- 2023 Government AI Readiness Index (oxfordinsights.com)
  https://oxfordinsights.com/wp-content/uploads/2023/12/2023-Government-AI-Readiness-Index-1.pdf

- Landscape Study of AI policies and use in Southern Africa: research report
  https://unesdoc.unesco.org/ark:/48223/pf0000385563

- AI Africa Consortium
  https://aiafrica.ac.za/

- Artificial Intelligence School to Open in Fez (moroccoworldnews.com)
  https://www.moroccoworldnews.com/2019/07/278026/artificial-intelligence-school-open-fez

- Kafr El University AI faculty
  https://kfs.edu.eg/en_ai/display.aspx?topic=81262

- Cameroon's first AI training centre launched
  https://itweb.africa/content/JN1gP7OYZKEqjL6m

- The Artificial Intelligence & Robotics Center of Excellence
  http://www.aastu.edu.et/research-and-technology-transfer-vpresident/the-artificial-intelligence-robotics-center-of-excellence/

- 7 Ways that African States are Legitimizing Artificial Intelligence | Open AIR
  https://openair.africa/7-ways-that-african-states-are-legitimizing-artificial-intelligence/

- Dakar Institute of Technology Ecole d'informatique spécialisée en Intelligence Artificielle (dit.sn)
  https://dit.sn/

- African Institute for Mathematical Sciences (nexteinstein.org)
  https://nexteinstein.org/

- Home Page - African Masters Of Machine Intelligence (aimsammi.org)
  https://aimsammi.org/

- Indaba - Deep Learning Indaba 2023
  https://deeplearningindaba.com/2023/

- Deep Learning IndabaX South Africa - #indabaX
  https://indabax.co.za/

- DSA | Home (datascienceafrica.org)
  https://www.datascienceafrica.org/

- Masakhane
  https://www.masakhane.io/

- AI & Society Fellows - Microsoft Research
  https://www.microsoft.com/en-us/research/academic-program/ai-society-fellows/research-challenges/

General Sources

- African Development Bank
  https://www.afdb.org/en

- African Risk Capacity
  https://www.arc.int/

- African Climate Foundation
  https://africanclimatefoundation.org/

- Africa does not want to be left behind in the AI revolution (qz.com)
  https://qz.com/africa/2180864/africa-does-not-want-to-be-left-behind-in-the-ai-revolution

- The African Union Artificial Intelligence Continental Strategy For Africa | AUDA-NEPAD
  https://www.nepad.org/news/african-union-artificial-intelligence-continental-strategy-africa

- The-State-of-AI-in-Africa-Report-2023-min.pdf (strathmore.edu)
  https://cipit.strathmore.edu/wp-content/uploads/2023/05/The-State-of-AI-in-Africa-Report-2023-min.pdf

- The-State-of-AI-in-Africa-A-Policy-Brief110923-1.pdf (strathmore.edu)
  https://cipit.strathmore.edu/wp-content/uploads/2023/09/The-State-of-AI-in-Africa-A-Policy-Brief110923-1.pdf

- Final-Report-The-State-of-AI-in-Africa-Report-2023.pdf (cipit.org)
  https://cipit.org/wp-content/uploads/2023/06/Final-Report-The-State-of-AI-in-Africa-Report-2023.pdf

- AI-Governance-in-Africa-2022.pdf (altadvisory.africa)
  https://ai.altadvisory.africa/wp-content/uploads/AI-Governance-in-Africa-2022.pdf

- Blueprint: Artificial Intelligence for Africa
  https://smartafrica.org/wp-content/uploads/2023/11/70029-eng_ai-for-africa-blueprint-min.pdf

- State of AI in Africa Report (gumroad.com)
  https://aiafricareport.gumroad.com/

- AI Governance for Africa Toolkit - Regional and ... (trust.org)
  https://www.trust.org/trustlaw/resources/item/ai-toolkit-1-and-2/

- Catalyzing Responsible AI Governance in Africa: Insights, Challenges, and Solutions - The Datasphere Initiative
  https://www.thedatasphere.org/news/catalyzing-responsible-ai-governance-in-africa/

- The AI In African Innovation Explained - Forbes Africa
  https://www.forbesafrica.com/technology/2023/09/14/the-ai-in-african-innovation-explained/

- The AI Strategy of Mauritius. Artificial Intelligence Strategy in the... | by Alex Moltzau | DataSeries | Medium
  https://medium.com/dataseries/the-ai-strategy-of-mauritius-20518940fb2d

References

- World Bank - The high price of education in Sub-Saharan Africa
  https://blogs.worldbank.org/developmenttalk/high-price-education-sub-saharan-africa

- IFC – Digital Skills in Sub-Saharan Africa
  https://www.ifc.org/content/dam/ifc/doc/mgrt/digital-skills-final-web-5-7-19.pdf

- Career essentials in generative AI by Microsoft and LinkedIn
  https://www.linkedin.com/learning/paths/career-essentials-in-generative-ai-by-microsoft-and-linkedin

Part 4 Responsible by Design: Microsoft's Approach to Building AI Systems that Benefit Society

Cited Sources

- NIST AI Risk Management Framework
  https://www.nist.gov/itl/ai-risk-management-framework

- Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies
  https://www.whitehouse.gov/briefing-room/statements releases/2023/07/21/fact-sheet-biden-harris-administration secures-voluntary-commitments-from-leading-artificial intelligence-companies-to-manage-the-risks-posed-by-ai/

- OECD AI Principles
  https://oecd.ai/en/ai-principles

- OECD Framework for the Classification of AI systems
  https://www.oecd.org/publications/oecd-framework-for-the-classification-of-ai-systems-cb6d9eca-en.htm

- UNESCO Recommendation on the Ethics of Artificial Intelligence
  https://www.unesco.org/en/artificial-intelligence/ recommendation-ethics

General Sources

- Governing AI: a blueprint for the future
  https://aka.ms/GoverningAI-blueprint

References

- Critical Infrastructure Sectors | CISA
  https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors

Part 5 AI in Action in Africa

Cited Sources

- Zipline's Ambitious Medical Drone Delivery in Africa
  https://www.technologyreview.com/2017/06/08/151339/ blood-from-the-sky-ziplines-ambitious-medical-drone-delivery-in-africa/

- Preventing and treating maternal, adolescent and child malnutrition
  https://www.unicef.org/kenya/ nutrition#:~:text=Nutrition%20in%20numbers,doses%20 of%20vitamin%20A%20supplement

- Improving lives with AI, digital and financial inclusion in Kenya, one smartphone at a time
  https://news.microsoft.com/europe/features/improving-lives-with-ai-and-digital-and-financial-inclusion-in-kenya-one-smartphone-at-a-time/

- Microsoft Customer Stories - Association of Kenya Insurers
  https://customers.microsoft.com/en-us/ story/1459970325168017255-aki-swiftant-azure

- AI: Ushering in an Age of Empathy | Investec
  https://www.investec.com/en_za/focus/innovation/ai-ushering-in-an-age-of-empathy.html

- Artificial Intelligence in the Middle East and Africa Outlook for 2019 and Beyond - AB InBev
  https://info.microsoft.com/rs/157-GQE-382/images/ MicrosoftSouthAfricanreportSRGCM1070.pdf

- Microsoft Customer Stories - Coca-Cola Beverages Africa
  https://customers.microsoft.com/en-us/ story/1333984650578441826-coca-cola-beverages-consumer-goods-en-south-africa

- Microsoft Cloud Blogs – Coca-Cola Beverages Africa
  https://cloudblogs.microsoft.com/dynamics365/ bdm/2022/04/28/microsoft-fuels-digital-transformation-of-supply-chains-for-a-resilient-and-sustainable-future/

- Global Renewables Watch
  https://www.globalrenewableswatch.org/

- HOTOSM.org
  https://www.hotosm.org/

- Microsoft's AI for Humanitarian Action Program
  https://www.microsoft.com/en-us/ai/ai-for-humanitarian-action-projects

- Using big data and machine learning to respond to the refugee crisis in Uganda
  https://reliefweb.int/report/uganda/using-big-data-and-machine-learning-respond-refugee-crisis-uganda

- Pulse Lab Kampala
  https://pulselabkampala.unglobalpulse.net/kampala-contact/

- UN Global Pulse
  https://radio.unglobalpulse.net/uganda/

- Justice for Forests: Improving Criminal Justice Efforts to Combat Illegal Logging report
  https://www.worldbank.org/en/topic/financialsector/ publication/justice-for-forests-improving-criminal-justice-efforts-to-combat-illegal-logging-report

- Making A Difference Through Forest Guard
  https://sciencetechnology.uonbi.ac.ke/index.php/latest-news/making-difference-through-forest-guard

- The State of the World's Forests 2020
  https://www.fao.org/state-of-forests/en/

- African forest elephants | International Fund for Animal Welfare
  https://www.ifaw.org/international/animals/african-forest-elephants

- AI Speeds Effort to Protect Endangered Elephants | Cornell Chronicle
  https://news.cornell.edu/stories/2018/08/ai-speeds-effort-protect-endangered-elephants

- Microsoft Middle East and Africa News Center - Microsoft Azure Cognitive Services Translator
  https://news.microsoft.com/en-xm/2023/03/09/microsoft-translator-launches-13-new-african-languages/

- Lelapa AI
  https://lelapa.ai/

- Microsoft Africa Research Institute (MARI)
  https://www.microsoft.com/en-us/research/group/microsoft-africa-research-institute-mari/overview/

- Agriculture and Food Security | USAID
  https://www.usaid.gov/southern-africa-regional/agriculture-and-food-security

- Microsoft Egypt aims to provide advanced AI models, build solid AI ecosystem: Mirna Arif
  https://www.dailynewsegypt.com/2023/09/18/microsoft-egypt-aims-to-provide-advanced-ai-models-build-solid-ai-ecosystem-mirna-arif/

- FarmVibes.Bot
  https://www.microsoft.com/en-us/research/project/project-farmvibes/articles/farmvibes-bot/

- The Field Report | International Fund for Agriculture Development (IFAD)
  https://www.ifad.org/thefieldreport/

- GitHub – FarmVibes.AI
  https://github.com/microsoft/farmvibes-ai

- Reinforcing Higher Education in Africa | UNESCO
  https://www.unesco.org/en/articles/reinforcing-higher-education-africa

- Microsoft Customer Story - Eduvos
  https://customers.microsoft.com/en-us/story/1706422185795423548-eduvos-azure-higher-education-en-south-africa

- Digital Skills in Sub-Saharan Africa: Spotlight on Ghana
  https://www.ifc.org/content/dam/ifc/doc/mgrt/digital-skills-final-web-5-7-19.pdf

- Codar Tech Africa LinkedIn Page
  https://www.linkedin.com/company/codar-tech-africa-limited

- Microsoft Africa Development Center
  https://www.microsoft.com/en-us/madc/who-we-are?rtc=1

- Microsoft Middle East and Africa News Center - Digital Nigeria
  https://news.microsoft.com/en-xm/2021/05/03/microsoft-collaborates-with-the-nigerian-government-to-accelerate-digital-transformation-in-the-country/

- DSN partners Microsoft train entrepreneurs on Responsible AI
  https://businessday.ng/technology/article/dsn-partners-microsoft-train-entrepreneurs-on-responsible-ai/

- Global Integrated Education Volunteers Association (GIEVA)
  https://data.org/initiatives/ai-challenge/awardees/global-integrated-education-volunteers-gieva/

- Microsoft Middle East and Africa News Center - Youth Employment Service
  https://news.microsoft.com/en-xm/2023/10/31/ai-training-for-300000-south-african-youth-an-enabler-for-future-career-advancement/

- Debswana LinkedIn Page
  https://www.linkedin.com/posts/debswana-diamond-company_shapethefuture-innovate-activity-6659390389621002243-845Y

- Sizing up Syama: the world's first fully automated mine
  https://www.mining-technology.com/features/sizing-syama-worlds-first-fully-automated-mine/

- Resolute publishes 2022 sustainability report amid expansion push
  https://www.mining-technology.com/news/resolute-publishes-2022-sustainability-report-amid-expansion-push/

- Microsoft Middle East and Africa News Center - E-Tax
  https://news.microsoft.com/en-xm/2021/04/29/egypts-e-tax-microsoft-partner-to-use-ai-in-tax-solutions/

- Access to electricity (% of population) - Nigeria | World Bank
  https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=NG

- ICE Commercial Power
  https://www.thisdaylive.com/index.php/2022/09/20/ice-commercial-power-to-provide-renewable-micro-utilities-in-underserved-communities

- Microsoft Customer Stories - MultiChoice
  https://customers.microsoft.com/en-us/story/1533703067530325346-multichoice-group-media-entertainment-azure-en-south-africa

- Water Access and Security is an all of Society Issue | City Insight
  https://www.cityinsight.co.za/water-access-and-security-is-an-all-of-society-issue/

- eThekwini Municipality
  https://sponsorcontent.cnn.com/edition/2021/microsoft/en/how-tech-will-make-this-the-most-liveable-city-in-africa-by-2030.html

General Sources

- M-KOPA
  https://m-kopa.com/

- Microsoft Blog - Forest Guard
  https://www.microsoft.com/en-us/garage/blog/2022/01/forest-guard-microsoft-student-hackathon-winners-create-rapid-response-deforestation-sensor/

- Lelapa AI Twitter Page
  https://www.linkedin.com/posts/lelapa-ai_aiforgood-aiforgood-opportunities-activity-7079162835963912192-2Gvy?trk=public_profile_like_view

- Microsoft Africa Transformation Office - FarmVibes.bot (Agbot)
  https://www.microsoft.com/africa/ato/CustomerStories/democratising-access-to-information-and-services-to-farmer