Only five percent of Africa's AI talent has the compute power it needs

NOVEMBER 12, 2024

This is a guest blog post authored by *Alexander Tsado,* *Co-founder at Alliance4ai and Ahura AI and* ***Celina Lee,*** *Founder, Zindi*

How the AI Hub for Sustainable Development can help transform Africa's AI landscape, by innovating partnerships and unlocking financing with G7, the African private sector and leaders in the global AI space.

This week, the United Nations' leading development agency, the United Nations Development Programme and the Italian G7 Presidency, together with a delegation of African private sector leaders are meeting in San Francisco to build public–private partnerships to boost the AI revolution. Hosted by the Italian Innovation and Culture Hub (Innovit), the conference is

G7 presidency, the AI Hub is centered on inclusive and representative datasets including low resourced languages, local AI talent development across Africa, accessible and affordable green computing infrastructure, and enabling environments for responsible AI adoption.

With only five percent of Africa's AI talent having access to the computational power and resources needed to carry out complex tasks—known as compute— we must act now to change this trajectory.

In the same way history was written in 1945 during the drafting of the United Nations Charter in San Francisco—which gave birth to the United Nations—this week's conference offers fertile ground to anchor sustainable development at the heart of the global AI private sector. Importantly, the discussions will enable us to collectively move beyond goodwill to impactful, scalable partnerships and ensure that Africa's talent is an equal partner in AI innovation.

As founders of Alliance4ai (Powering Africa's AI Revolution) and Zindi, (the largest network of AI talent in Africa), the conference provided us the opportunity to highlight the compute challenge that is hindering the AI revolution on the African continent. This drew on experience working with data centre product teams at Nvidia and interviews with leaders from AI organizations in Africa.

The gap in compute access and its impact on Africa's AI innovators

We previously analysed compute usage data across the largest network of AI talent in Africa, Zindi, comprising a cohort of 11,000 data scientists. This analysis revealed a significant gap in compute access—only five percent of Africa's talent has access to computational power for research and innovation. Within this group, we found many constraints related to meaningful access and usage:

- **Limited on-premise access:** The first one percent have 'on-premise' access (in their offices) to graphics processing units (GPUs)—powerful chips that quickly perform complex calculations and are often used to train AI models. This enables them to run AI training on GPUs as much as needed, as long as they cover the hourly electricity costs.

- **Limited GPU access:** The next four percent are able to pay to access GPUs on the cloud. Their payments are typically made through cloud credits or venture capital funding. However, the majority of this group is only able to afford about US$1,000 per month. This amount for example provides roughly two hours of daily usage for an old Nvidia A100 GPU. This group operates at a much lower capacity than the first group

- **Reliance on Google Colab:** The remaining 95 percent on Zindi either use regular laptops without GPUs or access GPUs through Google Colab—a free cloud-based tool that offers free compute access for educational purposes. Researchers and startups often face limitations and conditions that restrict their usage.

To put this resource gap into context, consider that while a private sector start-up or innovator in one of the G7 countries can iterate on their models every 30 minutes during training, their African peers may have to wait up to six days before they make their next change. This disparity puts African innovators at a significant disadvantage, effectively halting their ability to develop and deploy transformative AI-powered solutions for development.

A wider issue is that the rapid advancement of AI has intensified a critical resource challenge: the compute–energy nexus. As AI's computational demands surge, this puts a strain on global energy systems and exacerbates the digital divide. This divide is significant: the compute market is growing at nine percent annually with 60 percent of the world's top supercomputers concentrated in just three countries (China, Germany and the United States). In Africa, cloud computing penetration lags at 15 percent, compared to Europe's 71 percent, with limited choices and innovation in access and affordability models.

Co-creating solutions

There are several approaches to addressing this compute challenge. Stakeholders can collaborate to create inclusive market paradigms, aiming for an equitable and efficient future of green compute capabilities in Africa. African stakeholders have highlighted the need for innovative compute investment to support small-to-medium sized private sector companies, enabling them to scale and distribute AI opportunities across the continent while addressing the underlying energy challenges.

A few interventions being undertaken to unlock the constraints in cloud credits and on-premise solutions include the following:

THE CLOUD CREDITS APPROACH

This approach involves purchasing or receiving a large pool of cloud credits at a discounted price from large American cloud companies. This option is usually considered by major donor agencies working to support Africa; it is conceptually faster to start with, and easier to solicit co-sponsorships from American cloud companies—initiatives such as the Partnership for Global Inclusivity on AI (PGIAI) launched by the United States of America) is an example. This approach is considered less risky for donor agencies making this kind of investment for the first time. Nevertheless there are drawbacks, which include:

- African researchers rarely get access to newer GPUs, and typically need to wait until users in the United States, or wherever the GPUs are generated from, are not working. Only then can Africa move up the queue; this defeats the purpose of empowering African researchers to shine.

- Previous programmes in the United States have failed to implement recipient selection strategies that ensure the credits are offered to innovators who will use them. To this end, we have worked on new dynamic distribution approaches.

- When leading researchers and innovators are selected, their allotted credits finish in a few months.

In addition, there remains a pressing need to link the cloud credits approach with clean energy solutions, particularly in regions like Africa with the potential to unlock an abundance of renewable resources. Even with these drawbacks, this is a decent approach that may suit certain scenarios, and we welcome donors who want to support the ecosystem with cloud credits, co-designing the process to mitigate the challenges listed above.

THE ON-PREMISE APPROACH

This approach involves buying or receiving tens or hundreds of GPUs, installing them in African countries and offering them at a free or subsidized rate to AI talent across the continent. This is the approach most AI talent in Africa publicly advocates for, because of its incredible benefits and opportunity to catapult Africa's AI trajectory to the next stage. However it carries significant risks which have deterred donors thus far. The table below profiles some of the bigger risks, and how Alliance4ai has worked to mitigate them.

Challenge	Alliance4ai Approach
Expensive to purchase	Exploring grants and equity funding options
Scarce to access	Leveraging strong relationships with Nvidia datacenter team
Expensive import and customs fees	Advocating with four African governments to provide incentives
Lack of local setup and management expertise	Built a network of African supercomputing experts worldwide and a local training program to scale
Expensive electricity bills	Negotiated discounts with datacenter and renewable energy providers
Limited innovator awareness	Fostering a network of AI innovators across the continent
Lack of algorithm development skills	Offering training such as Nvidia Deep Learning Institute training programs

Return on investment: Growing together by co-designing compute approaches

To demonstrate the difference between each approach, consider the following comparisons on a US$1 million investment

- **Cloud credits:** US$1 million could support 10 startups for two years with the startups spending roughly $5,000 a month to access an Nvidia A100 GPU for five hours a day.
- **On-premise subsidy:** The same US$1 million could support 50 percent more startups providing double the daily GPU usage, for five years or more.

An ideal scenario would be to scale this to a US$15 million investment in on-premise GPU clusters, split across several regions in Africa, serving around 100 startups or research teams to sustain their innovation for several years, providing proof points that should catalyse private sector investment in this space.

compute challenge in Africa. With Italy's support, this could be linked to financing initiatives such as Energy for Growth in Africa. This was a flagship initiative adopted by the G7 at the Leaders Summit of Borgo Egnazia together with African partner countries, including Algeria, Republic of the Congo, Côte d'Ivoire, Ethiopia, Kenya, Mozambique, Nigeria, and South Africa and with the support of Türkiye and the United Arab Emirates.

With UNDP's presence in more than 170 countries and territories and deep relationships with both the private and public sectors, the organization could be instrumental in ensuring Africa stays ahead of the compute access and affordability challenge.

As we continue our journey to empower Africa's AI revolution, we will be able to unlock the necessary resources and expertise to harness the potential of the continent's data scientists and innovators. We are confident that through collaborative efforts and strategic partnerships, we can pave the way for a more equitable and sustainable AI-driven future for Africa.