# The importance of funding and investment to strengthen data science in Africa

Hazal Kayalioglu, John Pateña, Raphael Z. Sangeda, Upendo Masamu, Bruno Mmbando, Belinda Njiro, Conrad Iyegbe, Joyce Gyamfi, Dorice Vieira & Emmanuel Peprah

Check for updates

Currently Africa is not fully realizing the potential of data science (DS) to improve health outcomes. In this comment we advocate for policymakers to make concerted and strategic efforts that complement existing strategies to enhance DS and propel Africa into a prominent role in the global DS arena.

Data science (DS) plays a transformative role in advancing scientific discovery and informing policy by enabling the systematic analysis of large and complex data sets. In the health sector, it supports the development of predictive models, optimizes intervention strategies, and informs evidence-based decision-making, particularly in resource-constrained settings. Data science can be used to develop tools for diagnosis, treatment, and prevention of diseases that disproportionately overburden Africa to improve health outcomes. Harnessing the full potential of data science (DS) capacity is challenging because there is a lack of financial investment in DS in Africa. There are compelling reasons to channel more resources into African DS infrastructure. An opportunity exists to intensify investments to train and retain data scientists in Africa, build and scale-up data science (DS) infrastructure, and coordinate efforts between governments to sustain DS initiatives. To achieve these outcomes, sustained funding and investment for DS is needed. This commentary discusses the critical role of DS in addressing health challenges in Africa, emphasizing the need for greater investment in DS infrastructure, talent development, and research funding. It highlights the shortage of DS professionals across the continent, and the limited financial resources allocated to DS. The article advocates for strategic investments through collaboration between governments, the private sector, and international partners, while exploring successful funding models and the necessity for improved infrastructure to support DS growth across Africa.

**Current status of Data Science research in Africa.** A recent analysis of the global distribution of individuals trained in DS identified 11,400 professionals in North America, Europe, and Asia[1]. [1]In contrast, Africa had a limited number of data scientists, with South Africa having a few hundred[1]. Furthermore, information on the distribution of trained DS beyond South Africa is limited. This information gap underscores the urgency of documenting and mapping current DS capacity, sources and funding distributions. The African Union published the Digital Transformation Strategy for Africa (2020–2030), which aims for "an integrated and inclusive digital society and economy"[2]. However, there is a scarcity of local financial resources resulting in a dependency on donor funding, limited funding through government initiatives, and private foundations spearheading key investments. With the exception of South Africa, there is a considerable variation in research investment between most African governments and a failure to meet the agreed-upon 1% GDP target for expenditure on scientific research[2]. The limited overall research expenditure hinders scientific progress, especially in emerging fields including DS, despite its recognition by the African Union as a critical area of investment.

We explored funding channeled to Africa for DS compared to other continents reported by the National Institutes of Health (NIH) World Reporter (Figs. 1 and 2)[3]. The NIH is one of the major science funders in Africa. Whilst acknowledging that the NIH data does not include all funding for DS in Africa, it highlights that gaps in investment in DS exist. NIH World Reporter shows that about 70% of the funding was from the NIH and 30% from other funders. Only 1.54% of this funding was allocated to Africa for DS, to ~100 unique projects, mainly in South Africa, Kenya, and Nigeria (Fig. 2).

The NIH allocate millions of dollars through initiatives such as Human Hereditary and Health in Africa (H3Africa) and Data Science for Health Discovery and Innovation in Africa (DS-I Africa) to support the emerging field of DS on the African continent. Additional funding efforts exist from non-governmental organizations (NGOs), venture capitalists, and non-profit organizations (Table 1).

Moreover, selected African governments also continue to play a pivotal role in advancing DS within and outside their borders. Academic institutions and educational programs are beginning to receive funding to facilitate collaborative endeavors between academia and data scientists based elsewhere, such as those in industry and private organizations, supported by funders and local governments. Despite these positive developments, reducing gaps in funding is imperative. Robust financial support from diverse sources (e.g., private philanthropy, private sector, and public institutions) is essential to unlock its full potential and drive positive change across Africa[4]. Exploring diverse funding models is critical for sustained development of DS. Funding from additional sources helps to ensure programs are less vulnerable to changes in funding priorities and attracts various levels of resources, expertise, and perspectives to address challenges in DS. For example, the Developing Excellence in Leadership, Training, and Science in Africa (DELTAS Africa)[5] program uses the hub-and-spoke model. This model is designed to create a hub of resources and expertise that act as a focal point for training and innovation. Smaller or less resourced partners can benefit from this hub, enabling collaboration, enhancing capacities and creating opportunities for shared growth. This approach could also be utilized for funding DS. Another approach is the Grand Challenges model, designed to stimulate innovation to solve key global

**Fig. 1 | Annual trends and total investments in funding of data science in Africa.** Main graph shows data sourced from the NIH World Report (https://worldreport.nih.gov/) which indicates that total global investment from 2017 to 2022 amounted to $2,683.81 million. The (**A**) insert shows the relative share of global funding allocated to different regions, with just 1.54% invested in data science in Africa. The (**B**) insert shows the share provided by each funding organization listed to fund DS in Africa. Abbreviations: BMGF Bill Melinda Gates Foundation, CIHR Canadian Institutes of Health Research, EC European Commission, EDCTP European & Developing Countries Clinical Trials Partnership, MRC Medical Research Council, KHIDI Korea Health Industry Development Institute, NHMRC National Health and Medical Research Council, NIH National Institute for Health, Pasteur Pasteur Institute, Wellcome Wellcome Trust.
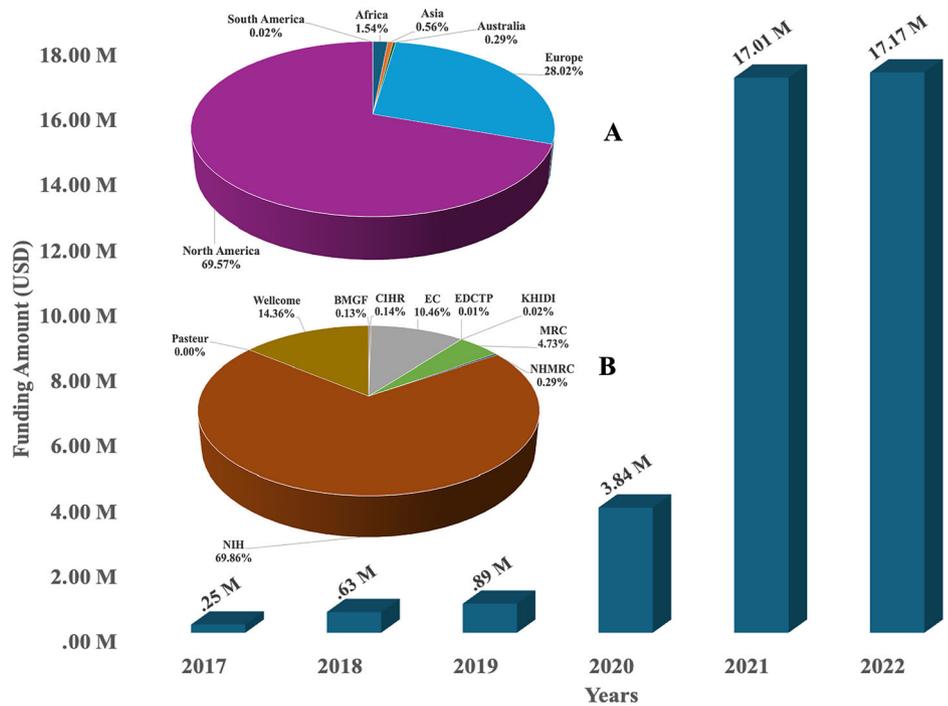


**Fig. 2 | Distribution of funded projects by agencies across Africa.** The size of each circle indicates the number of funded projects in a particular location. Abbreviations: BMGF Bill Melinda Gates Foundation,EC European Commission, EDCTP European & Developing Countries Clinical Trials Partnership, MRC Medical Research Council, NIH National Institute for Health, Pasteur Pasteur Institute, Wellcome Wellcome Trust. The map was generated using Microsoft Power BI software.
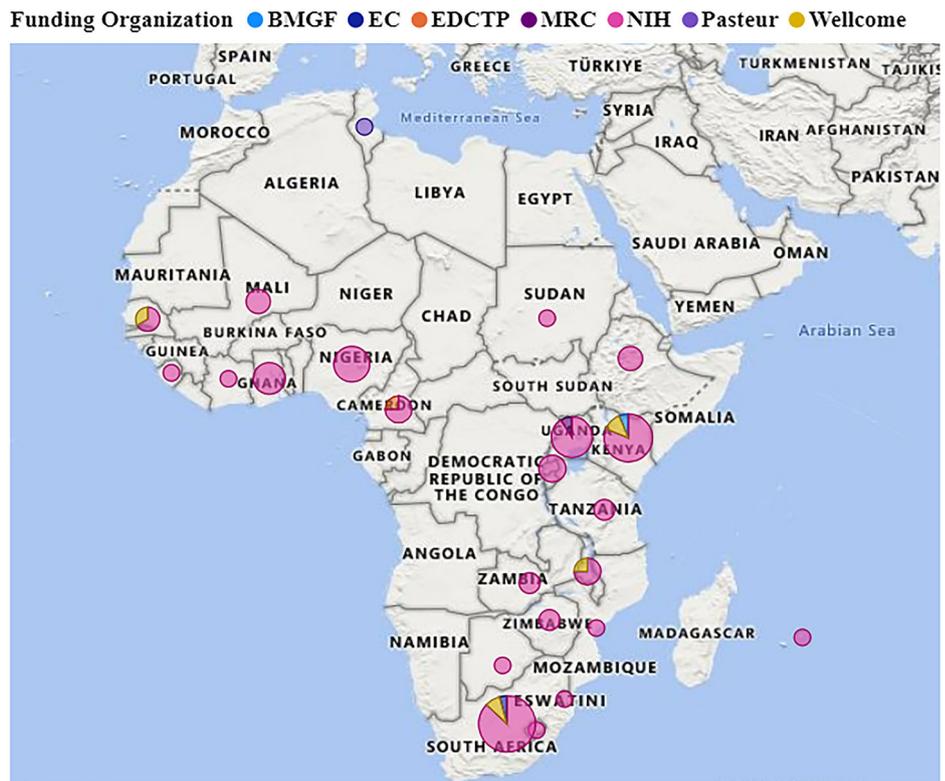
**Table 1 | African DS Funding Efforts**

| Funding Source | Initiative/ Program | Link (if applicable) | Purpose/Contribution |
|---|---|---|---|
| NIH (U.S. National Institutes of Health) | Human Heredity and Health in Africa (H3Africa) | https://h3africa.org/ | Supports genomics and health research in Africa |
| NIH (U.S. National Institutes of Health) | Data Science for Health Discovery and Innovation in Africa (DS-I Africa) | https://dsi-africa.org/ | Builds data science capacity for health innovation in Africa |
| NGOs / Non-profit / Private Entities | The Rockefeller Foundation | https://www.rockefellerfoundation.org/ | Supports global health and data science initiatives |
| NGOs / Non-profit / Private Entities | American Heart Association | https://www.heart.org/ | Supports health-related data science initiatives |
| NGOs / Non-profit / Private Entities | The World Data Science Initiative | https://www.worlddatascience.org/ | Promotes global data science collaborations |
| NGOs / Non-profit / Private Entities | Global Grand Challenges | https://gcgh.grandchallenges.org/ | Funds innovative global health solutions |
| NGOs / Non-profit / Private Entities | Grand Challenges Africa | https://www.grandchallenges.org/ | Funds African-led health and development innovations |
| NGOs / Non-profit / Private Entities | The Bill and Melinda Gates Foundation | https://www.gatesfoundation.org/ | Major funder of global and African data science and health research |
| NGOs / Non-profit / Private Entities | Global Partnership for Sustainable Development Data | https://www.data4sdgs.org/ | Accelerates data-driven decision-making for sustainable development |
| NGOs / Non-profit / Private Entities | Ingressive For Good | https://ingressive.org/ | Provides training and resources for African tech talent |
| NGOs / Non-profit / Private Entities | Science for Africa Foundation | https://scienceforafrica.foundation/ | Supports research and data science across Africa |
| Capacity-Building / Research Centers | Sub-Saharan African Consortium for Advanced Biostatistics (SSACAB) | https://ssacab.co.za/ | Builds biostatistics expertise in Africa |
| Capacity-Building / Research Centers | South African Centre for Epidemiological Modeling and Analysis (SACEMA) | https://www.sacema.org/ | Focuses on epidemiological modeling in Africa |
| Capacity-Building / Research Centers | Mathematics in South Africa (MASAMU) Program | https://samsa-math.org/ | Enhances math research collaboration in Southern Africa |
| Capacity-Building / Research Centers | Harmonizing and Improving Statistics in West and Central Africa Project (HISWACA) | https://projects.worldbank.org/en/projects-operations/project-detail/P178497 | Improves statistical capacity and infrastructure |

health problems through DS[6]. Both models share a common strength: they empower African leadership in DS by using established frameworks. However, additional models, such as the Paradigmatic Principal-Agent Model (which emphasizes performance-based accountability), Sector-Differential Model (tailoring support to the unique needs of sectors like health or agriculture), Embedded Principal-Agent Model (which promotes co-design and shared ownership), and Multiple Principal-Agents Model[7] (fostering coordination among diverse funders and implementers), could further expand DS capacity in Africa. These models are particularly well suited to the African context as they promote context-specific solutions, foster multi-sectoral collaboration, and mitigate risks associated with reliance on single-source funding, thereby contributing to long-term sustainability of DS initiatives. When effectively nurtured, these models have the potential to significantly boost DS output by enhancing both capacity and training.

**Aligning Funders' Strategies and Priorities.** The demand for human and technological resources is escalating, and strategic investments in DS can advance healthcare systems and enhance the well-being of the population. Creating and maintaining comprehensive datasets, biobanks, and electronic health record systems that depict the health status of populations is pivotal. Such data infrastructure empowers nations to better inform healthcare policies and interventions, fostering more effective and targeted approaches[8]. Within the global landscape, data-driven methodologies enable prediction, management, and mitigation of the impact of epidemics and disease outbreaks, contributing to public health resilience[9–11]. Investing in building talent capacity, innovation and retaining talented researchers in Africa is imperative for elevating DS capabilities in health research. These investments will also be beneficial beyond the healthcare sector, catalyzing innovation and fueling economic growth. This ripple effect will lead to a more sustainable and self-reliant DS ecosystem, fostering a landscape in which Africa can harness the full potential of DS.

Impactful DS initiatives require substantial investments in technology, infrastructure, and talent. However, tight budgets within African governments and the need to prioritize other critical needs present significant obstacles to developing DS on the continent. It has been argued that digital transformation in Africa will be central to maintaining sustainable growth over the next few decades[2]. Infrastructure improvement has emerged as a crucial factor in bridging the DS gap. High-speed internet and cutting-edge computing technology are the foundation for successful DS endeavors[12,13]. Lack of reliable internet connections and modern computing facilities across some African countries impedes data collection, analysis, and dissemination[14]. Substantial disparities in internet access exist across and within African regions, with Southern Africa reporting 60% access compared to only 25% in Central Africa[15]. Investing in high-speed internet access and establishing well-equipped data centers can significantly enhance the DS landscape. This infrastructure development would empower researchers and data scientists to work more efficiently, resulting in better data-driven solutions and informed decision-making. Aligning government infrastructure development with funders priorities is essential for advancing

and sustaining DS in Africa[16]. DS progress cannot occur without the necessary infrastructure, making it essential to synchronize government investments with funders' strategic goals. Similar collaborative partnerships and long-term strategic planning have been successfully employed to tackle communicable diseases[17]. This approach is equally vital for DS. Moreover, donor support for infrastructure development is feasible when embedded within a broader vision for DS that promotes transparency, accountability and long-term sustainability.

**Overcoming challenges in financing and nurturing data science initiatives.** The shortage of skilled data scientists in Africa poses a significant challenge. While numerous DS programs are available globally, accessibility challenges such as geographical proximity, travel constraints, and limited Internet availability hinder many Africans from benefiting from these opportunities[18]. There is a compelling need to fund programs that can be developed and expanded within Africa, which requires additional sponsors and partners for their execution. Education and training programs in Africa have recently improved by offering DS skills and training. This surge in development is characterized by the establishment of fellowship programs within higher learning institutions and collaborative initiatives with advanced DS institutions[19]. Educational institutions should collaborate with industry experts to design relevant, up-to-date curriculum aligned with industry demands. Scholarship programs and internships can be pivotal in developing local talent, making DS a more accessible career choice for Africans. This approach promotes skill development and strengthens workforces in other STEM-related fields. In addition, there remains a need to retain talent in African institutions. Investments also need to fund positions post-training so that the skill sets developed can support the growth of local capacity building.

As DS continues to exert its transformative influence on our world, strategically allocating research, innovation, and educational resources in this dynamic field is an indispensable catalyst for progress. While existing global funding models have undoubtedly encouraged the discipline in high-income countries (HIC), there is currently insufficient research, training, and investment to meet the multifaceted demands and opportunities presented by DS in diverse settings, especially in low- and middle-income countries (LMICs).

**The importance of strategic investments in data science for African development.** A robust and sustained commitment to adequate funding is imperative to unlock the transformative power of DS fully. Governments, international organizations, and entities within the private sector each play a pivotal role in providing the financial support necessary for meaningful advances. Investing in DS equips countries with tools to address their unique challenges effectively. It is a gateway to economic growth and enhanced global competitiveness. In addition to their immediate health applications, DS initiatives have the potential to spur innovation, cultivate new industries, and position nations at the forefront of a rapidly evolving digital landscape. As we navigate the data-driven future, strategic and substantial investments in DS are not merely an option but required to foster a more resilient, equitable, and prosperous global society.

**Hazal Kayalioglu** (ID)[1] ✉, **John Pateña**[1], **Raphael Z. Sangeda** (ID)[2], **Upendo Masamu**[2], **Bruno Mmbando**[3], **Belinda Njiro** (ID)[2], **Conrad Iyegbe**[4], **Joyce Gyamfi**[1], **Dorice Vieira**[1] & **Emmanuel Peprah**[1]

[1]Implementing Sustainable Evidence-based interventions through Engagement (ISEE) Lab, Department of Global and Environmental Health, New York University School of Global Public Health, New York, NY, USA. [2]Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania. [3]National Institute for Medical Research (NIMR), Dar es Salaam, Tanzania. [4]Icahn School of Medicine at Mount Sinai, New York, NY, USA. ✉e-mail: hk3570@nyu.edu

### References

1. Stitch. The State of Data Science. Accessed 15 Oct 2023. https://www.stitchdata.com/resources/the-state-of-data-science/.
2. African Union. The Digital Transformation Strategy for Africa (2020–2030). Addis Ababa, Ethiopia: African Union. Retrieved from https://www.au.int (2020).
3. Health Nlo. World Report. https://worldreport.nih.gov/.
4. Arvanitis, R., Mouton, J. & Néron, A. Funding Research in Africa: landscapes of Re-institutionalisation. *Sci. Technol. Soc.* **27**, 351–367 (2022).
5. Gichoga, S. Innovative funding systems are key to fighting inequities in African science. *Nature* https://doi.org/10.1038/d41586-024-00348-6 (2024).
6. Forum W. E. Innovative approaches for unlocking R&D funding in Africa. 2023. https://www.weforum.org/agenda/2023/11/innovative-approaches-for-unlocking-research-and-development-funding-in-africa/.
7. OTT. National Research Funding Mechanisms in West Africa. 2024. https://onthinktanks.org/wp-content/uploads/2024/06/Research-Funding-Mechanisms-in-West-Africa_02-FOR-WEBSITE-NO-NOTES.pptx.pdf.
8. Aung, T., Niyeha, D. & Heidkamp, R. Leveraging data visualization to improve the use of data for global health decision-making. *J Glob. Health* **9**, 020319 (2019).
9. Judson, S. D. et al. COVID-19 data reporting systems in Africa reveal insights for future pandemics. *Epidemiol. Infect* **150**, e119 (2022).
10. Gaye, Y. E., Agbajogu, C. & El Oakley, R. COVID-19 on the Nile: review on the management and outcomes of the COVID-19 Pandemic in the Arab Republic of Egypt from February to August 2020. *Int J Environ Res Public Health*. **18** https://doi.org/10.3390/ijerph18041588 (2021).
11. Ngueilbaye, A., Huang, J. Z., Khan, M. & Wang, H. Data quality model for assessing public COVID-19 big datasets. *J Supercomput*. 1–33. https://doi.org/10.1007/s11227-023-05410-0 (2023).
12. Atayero, A. A. et al. Trends and patterns of broadband Internet access speed in a Nigerian university campus: a robust data exploration. *Data Brief* **23**, 103705 (2019).
13. Adeyemi, O. J. et al. Exploration of daily Internet data traffic generated in a smart university campus. *Data Brief* **20**, 30–52 (2018).
14. Beyene, J. et al. A roadmap for building data science capacity for health discovery and innovation in Africa. *Front. Public Health* **9**, 710961 (2021).
15. Statista. Internet usage in Africa - statistics & facts. accessed 15 October 2023. https://www.statista.com/topics/9813/internet-usage-in-africa/#topicOverview.
16. Green A. Donors learning to align with health priorities of African governments. 2024 https://www.devex.com/news/donors-learning-to-align-with-health-priorities-of-african-governments-107712.
17. Oyebanji, O. Rethinking Donor Partnerships in Africa. 2024. https://www.thinkglobalhealth.org/article/rethinking-donor-partnerships-africa.
18. Munga J. To Close Africa's Digital Divide, Policy Must Address the Usage Gap. https://carnegieendowment.org/research/2022/04/to-close-africas-digital-divide-policy-must-address-the-usage-gap?lang=en.
19. Coker M. O. et al. Data science training needs in sub-Saharan Africa: Implications for biomedical research and therapeutics capacity. *Open Res Africa*. **6** https://doi.org/10.12688/openresafrica.14109.1 (2023).