Discovering HPC Resources in Africa

Empowering Collaborative Research Opportunities

Bryan Johnston
CHPC
CSIR
Cape Town, South Africa
bjohnston@csir.co.za

Katherine Johnston
University of Cape Town
Cape Town, South Africa
katherine.johnston@uct.ac.za

Suresh Maslamoney
University of Cape Town
Cape Town, South Africa
suresh.maslamoney@uct.ac.za

Abstract

Africa faces significant challenges in accessing advanced cyberinfrastructure due to resource constraints. Consequently, the African scientific community must explore innovative approaches, such as fostering collaboration and leveraging shared resources, to overcome financial barriers to cyberinfrastructure adoption. We outline the methodology and results of a preliminary discovery survey aimed at mapping High-Performance Computing (HPC) resources across Africa to support scientific computing research. The initiative stemmed from the recognition of the critical importance of an African HPC resource catalogue in fostering research and scientific collaboration. The survey is a result of the collective efforts of diverse stakeholders to promote scientific advancement throughout the continent by generating a preliminary overview of available computational resources in Africa. The survey gathered a total of 51 completed submissions from 23 African countries, establishing a solid foundation for further exploration of existing HPC resources across Africa.

CCS Concepts

• **Computer systems organization** → **Pipeline computing**.

Keywords

HPC discovery survey in Africa, HPC resource catalogue, HPC Resource Constrained Environments, Research collaboration, Research capacity development

ACM Reference Format:

Bryan Johnston, Katherine Johnston, and Suresh Maslamoney. 2025. Discovering HPC Resources in Africa: Empowering Collaborative Research Opportunities. In *Practice and Experience in Advanced Research Computing (PEARC '25), July 20–24, 2025, Columbus, OH, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3708035.3736027

1 Introduction

Africa faces notable limitations in scientific computing resources, yet its demand for advanced research computing infrastructure is comparable to that of more resource-rich regions [5, 6].

Africa is home to only 1 percent of global data centre capacity despite being home to 18 percent of the world's population. The continent has limited high-performance computing services, cloud service challenges, concentrated server infrastructure and high workstation costs. Bridging these resource gaps is essential and requires an alternative approach to relying on heavy financial investments [7]. Africa might better tackle its challenges through enhanced collaboration and resource sharing across the continent. A crucial first step in this process involves identifying existing resources to understand which gaps require attention. While previous efforts have been applied to assessing scientific computing capacity in Africa, they have not achieved the scale, reach, scope or detail required to establish an HPC collaboration catalogue [2, 9]. By uniting several regional communities' efforts and leveraging shared contacts, this effort allowed for a broader reach and higher response rate.

The HPC Discovery Survey partnership comprises stakeholders from several regional scientific communities:

- **AfricaBP (African BioGenome Project)**[1]: A coordinated pan-African initiative aimed at building capacity and infrastructure to generate, analyse, and utilise genomic data from organisms indigenous to Africa.
- **DS-I Africa (Data Science for Health Discovery and Innovation in Africa)**[2]: Aims to leverage data science technologies to transform biomedical and public health research, leading to improved health for individuals and populations.
- **H3ABioNet (Human Heredity and Health in Africa Bioinformatics Network)**[3]: A pan-African bioinformatics network supporting genomics research and capacity building under the H3Africa initiative [8].
- **eLwazi**[4]: An African-led open data science platform designed to support health discoveries by enabling data sharing, analysis, and collaboration with flexible, scalable tools.
- **HPC Ecosystems Project**[5]: A national initiative to build HPC readiness in African partner countries through system deployment and workforce development [5, 6].

PEARC '25, Columbus, OH, USA

ACM ISBN 979-8-4007-1398-9/25/07
https://doi.org/10.1145/3708035.3736027

[1] https://africanbiogenome.org/
[2] https://dsi-africa.org/
[3] https://h3abionet.org/
[4] https://www.elwazi.org/
[5] https://ecosystems.nicis.ac.za/

Table 1: Dimensions Captured in Resource Survey

Dimension	Description
Organisation Type	Academia/Research • Private/Industry • National/Government
Type of Resource	Datacentre or Cloud-Based • Computer Lab/Small-Scale Cluster • Workstation • Other • None
Placement	On Premise • Remote (Same entity, different location) • Remote (Partner site) • Cloud • Other
Age	Based on deployment year of the current HPC system, irrespective of hardware age.
Collaboration	Indicates whether the resource is shared or exclusive - includes sharing practices and scope (local, regional, international).
Skills	Describes skill sets available at the site and how they might be shared collaboratively.
Consent	Indicates site consent for future contact, public listing, private repository inclusion, or survey participation.

The goal of the preliminary survey is to establish an initial catalog identifying willing collaboration partners, outlining available HPC systems and resources across the continent. The results are expected to enable collaborative scientific computing partnerships, highlight gaps in technology and human resources that require attention, and support initiatives to secure funding and upskill the scientific computing workforce in Africa.

2 Execution

In the absence of a comprehensive understanding of HPC resources in Africa, limiting the survey to targeted questions directed only at known HPC centres within the partnership would have introduced selection bias and left significant coverage gaps. To mitigate this, a preliminary REDCap-based survey [3, 4] was broadly distributed via the contact networks of collaborating partners. It aimed to reach a diverse set of respondents — including group leaders, system administrators, and active HPC users — by offering broad categorical options that promoted data consistency while accommodating varied site contexts. Table 1 provides an overview of the survey design.

The survey distribution extended to contact networks beyond African participants, allowing international users of HPC systems in Africa to contribute their input. Acknowledging that the survey's wide reach could lead to duplicate responses from members of the same groups submitting separate responses on their group's behalf, as well as incomplete or inaccurate submissions, the partnership carefully designed the questions to mitigate these challenges. Additionally, once the survey was completed, the data was reviewed to validate and ensure its accuracy, also noting the obvious missing respondents.

The main goal was to identify a wide range of stakeholders to enable future targeted engagement. While inaccuracies or contradictions in the data were expected, the primary focus of the preliminary discovery survey was to collect contact information to support follow-ups for further data validation. Put simply, false positives were considered more acceptable than false negatives because they could be addressed and eliminated during future engagements. In contrast, false negatives represented missed opportunities, as they would remain unidentified and excluded from subsequent efforts.

A summarized overview of the responses is available in a data archive [1]. Upcoming targeted surveys will focus on addressing specific inquiries raised by partnership members. The aggregated data is stored in a central repository, which will contribute to the development of an HPC resource catalogue. This repository will be made publicly accessible after follow-up surveys are conducted to verify and validate the initial findings. Considering the exploratory nature of the preliminary survey, releasing its results prematurely may lead to inaccurate or misleading conclusions, posing risks to the cooperation of the scientific community in future efforts to build a central HPC systems catalogue.

3 Findings

As summarized in Table 2, the initial response demonstrates promising geographic and institutional diversity. However, despite a significant number of HPC resources being reported, the partnership noted a low response rate from its own community, suggesting potential non-response bias. Notably, the survey revealed HPC systems previously unknown to the survey's stakeholders, highlighting a disconnect between policy oversight and on-the-ground infrastructure awareness.

Table 2: Preliminary Survey Response Summary

Metric	Count
Total responses received	72
Complete responses	51
Sites represented	46
Countries represented	23

Since the survey was distributed broadly without restricting respondent roles, it had to be designed to elicit meaningful input regardless of technical background. As a result, the quality and depth of responses varied considerably. Some participants were HPC users or prospective users, while others were technical administrators — occasionally leading to conflicting submissions from the same site, such as differences in system specification details. As noted earlier, this variability was anticipated and did not undermine the overarching goal of discovery.

4 Future Work and Resources

The survey will be refined and distributed to a wider network, leveraging existing contacts to broaden outreach whilst also seeking to include identified non-respondents and investigate the reasons for

their non-response. To contribute towards a more comprehensive HPC resource catalogue, targeted subgroup surveys will address specific community needs. Following completion of the refined follow-up survey, the catalogue will be populated and made available for general reference, with privacy measures in place for sites opting out of public repository inclusion.

5 Conclusion

A promising foundation has been established, with survey responses validating not only the value of the initiative but also an expressed interest in collaboration, reinforcing that this approach is key to addressing resource constraints in Africa and enabling impactful science. This initiative offers significant momentum towards establishing a clear overview of available African HPC resources and fosters collaboration with a stronger emphasis on regional integration. The recognition that many parties repeatedly seek the same information about HPC resources in Africa underscores the need for a unified approach. By working together to prevent survey fatigue and publish collected knowledge, the community can take meaningful steps forward.

Acknowledgments

The authors gratefully acknowledge the broader community of volunteers in the scientific computing research field for their invaluable contributions. We also thank Prof. Nicola Mulder, affiliated with both the eLwazi Open Data Science Platform and the DS-I Africa Consortium, and Dr. ThankGod Ebenezer from the African BioGenome Project, for their support. Minor editorial assistance, including grammar and phrasing refinements, was provided with the assistance of Microsoft Copilot. This work expands upon the OSF preprint *Summarisation of the HPC Resources in Africa – Discovery Survey: 2024* [1].

References

[1] 2024. Summarisation of the HPC Resources in Africa - Discovery Survey: 2024. https://osf.io/6vkuh. Version 1, OSF Preprint.

[2] George O. Amolo. 2018. The Growth of High-Performance Computing in Africa. *Computing in Science and Engineering* 20, 3 (5 2018), 21–24. doi:10.1109/MCSE.2018.03221926

[3] Paul A. Harris, Robert Taylor, Brenda L. Minor, Vicki Elliott, Michelle Fernandez, Lindsay O'Neal, Gabrielle Delacqua, Francis Delacqua, Jennifer Kirby, and Stephanie N. Duda. 2019. The REDCap consortium: Building an international community of software partners. *Journal of Biomedical Informatics* 95 (2019), 103208. doi:10.1016/j.jbi.2019.103208

[4] Paul A. Harris, Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, and Jose G. Conde. 2009. Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* 42, 2 (2009), 377–381. doi:10.1016/j.jbi.2008.08.010

[5] Bryan Johnston. 2019. HPC Ecosystems Project: Facilitating Advanced Research Computing in Africa. In *Proceedings of the Practice and Experience in Advanced Research Computing (PEARC '19)*. Association for Computing Machinery, New York, NY, USA, Article No. 107, 1–3. doi:10.1145/3332186.3333264

[6] Bryan Johnston, Lara Timm, David Macleod, and John Poole. 2024. Ten Years of the HPC Ecosystems Project: Transforming HPC in Africa for the Past Decade. In *Proceedings of the Practice and Experience in Advanced Research Computing (PEARC '24)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3626203.3670537

[7] Robin Miller, Alim Ladha, Stephanie Mambo, Victor Kogo, and Alex Tsado. 2025. Unlocking Compute in Africa. https://cdn.prod.website-files.com/66e31d90ea60e260f5ea025f/68546eacc565927bae162dc5_COMPUTE%20PAPER%20-%20ONLINE%20-%20TO%20PUBLISH.pdf

[8] Nicola J. Mulder, Ezekiel Adebiyi, Rachid Alami, Alia Benkahla, James Brandful, Seydou Doumbia, Dean Everett, Faisal M. Fadlelmola, Farida Gaboun, Simani Gaseitsiwe, Hassan Ghazal, Scott Hazelhurst, Winston Hide, Azeddine Ibrahimi, Yuven Jaufeerally Fakim, Claude V. Jongeneel, Fourie Joubert, Sadia Kassim, Julius Kayondo, Judit Kumuthini, Sulphina Lyantagaye, Julie Makani, Abdullah Mansour Alzohairy, Daniel Masiga, Abdramane Moussa, Ori Nash, Odile Ouwe Missi Oukem-Boyer, Ellis Owusu-Dabo, Soraya Panji, Hugh Patterton, Fouad Radouani, Khalid Sadki, Fatima Seghrouchni, Özlem Tastan Bishop, Nicki Tiffin, Naftali Ulenga, and H3ABioNet Consortium. 2016. H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. *Genome Research* 26, 2 (Feb 2016), 271–277. doi:10.1101/gr.196295.115 Epub 2015 Dec 1.

[9] Ghita Rahal. 2022. Status of the Computing for Research in Africa. doi:10.48550/arXiv.2206.05306 arXiv:2206.05306 [cs.CY]