REPUBLIC OF KENYA

National Commission for Science, Technology and Innovation

POLICY BRIEF

Strengthening Genomic Data Governance in Kenya



NACOSTI Policy Brief No. 1/ 2025-2026

JUNE 2026

info@nacosti.go.ke | www.nacosti.go.ke

1. INTRODUCTION

Genomic data is defined as the entirety of the genetic information contained in an organism's deoxyribonucleic acid (DNA). It is a digital blueprint that provides the ultimate instructions for how a living organism develops, functions, and reproduces among other characteristics. It is also defined as genetic data relating to the inherited or acquired genomic characteristics of a natural person which can give unique information about the physiology, characteristics or the health of that person and which result, in particular, from an analysis of the interplay of the genetic information of a person and their environment (WHO, 2024; European Union, 2016).

World Health Organisation in its guidance on human genome data collection, access, use and sharing outlines principles to ensure legal, ethical and equitable use of genomic data. It outlines the ethical commitments in managing human genome data focusing on autonomy, inclusivity, and fairness. It calls for empowering individuals and communities to participate in decision-making, and deliberate effort to achieve social justice. Further it emphasises solidarity and equity to guide the fair sharing of data and benefits, supported by strong partnerships across national, regional and international levels. It stresses responsible stewardship, adherence to laws and ethical standards, and minimizing risks. Transparency in policies and open access to research findings, combined with accountability systems, are essential to ensuring trust and preventing misuse of genomic data (WHO, 2024).

Genomic data generation in Africa has expanded significantly over the past decade. Several Pan African initiatives such as the Human Heredity and Health in Africa (H3Africa) consortium have played a crucial role supporting and training researchers, developing biobanks and bioinformatics networks. Despite the gains, Africa's genome is underrepresented in the global genomic datasets limits the applicability of genomic discoveries to local populations and exacerbating global health inequities. Additionally, structural challenges like limited infrastructure, shortage of skilled personnel, and inadequate governance structures further complicate Africa's effort to global competitiveness. (H3Africa Consortium, 2014; Mulder et al., 2018).

Despite risks involved in genomics research, Yakubu et al. (2018) and Kabata and Thaldar (2023) caution against state-centric "genomic sovereignty" models that treat genomic resources as national assets, approaches shown to hinder international collaboration and violate individual rights without delivering the national benefits they promise. Kenya should, therefore, be focused on building a coherent, rights-anchored framework capable of governing a research environment that is institutionalised but structurally allow controlled access for the public good.

Kenya is one of the countries in Africa that has participated actively in the genomic data initiatives in the continent. While most genomic studies are funded by external donors, institutions like Kenya medical Research Institute (KEMRI), Kenya Agricultural and Livestock Research Organization (KALRO), Kenyatta National Hospital (KNH), University of Nairobi (UON) and others have contributed immensely to genomic research in the country. However, national governance and infrastructure structures remain inadequate.

2. RATIONALE FOR ENHANCED GENOMIC DATA GOVERNANCE

Genomic data is becoming increasingly critical to advancements in medicine, security and the economy with the scale of genetic data generation increasing at a staggering rate since 2001 (Ramanan et al., 2025). WHO warns that as genomic data use expands, so too do the ethical and logistical challenges surround privacy, equitable access and responsible data management.

Kenya's Data Protection Act, 2019 establishes principles for governing the collection, processing, and transfer of personal data, including sensitive health and genetic information. However, genomic data presents unique challenges beyond conventional data governance owing to the specific characteristics related to genomic data such as the volume of data carried by a single genome, genomic data ownership beyond the individual (a person's genomic information could be representative of entire subsets of a population), and the potential for inclusion or exclusion of whole populations in health and other sectors. Genomic data governance requires a comprehensive approach that addresses the whole data lifecycle. To address this interconnectedness and unique challenges, a specific national or institutional, policy, and regulatory framework for genomic data is recommended.

The transformative impact of the application of genomic data and genomic science to transform human health, food security and drug development makes it a strategic and critical area for investment for Kenya.

This Policy brief aims to provide guidance on strengthening governance of genomic data in Kenya to ensure its ethical, legal, equitable and fair collection, access, use and sharing.

3. SITUATION ANALYSIS

Kenya's genomics research governance landscape is characterized by a constellation of institutions operating across intersecting but largely siloed mandates. The Science, Technology and Innovation Act (Cap. 511) vests NACOSTI with overarching responsibility for research licensing, ethics oversight, and approval of research involving biological samples. The Environmental Management and Coordination Act and its 2025 Access and Benefit Sharing (ABS) Regulations task NEMA with governing access to genetic resources under the Nagoya Protocol. The Data Protection Act (Cap. 411C) mandates the Office of the Data Protection Commissioner to oversee personal data, while the Biosafety Act (Cap. 320) governs genetically modified organisms through the National Biosafety Authority. No single body holds comprehensive authority over the genomic data lifecycle.

Kenya holds substantial genomic research capacity. Institutions including KEMRI, KALRO, the Kenya Animal Genetic Resources Centre, Kenya Wildlife Service, and the International Livestock Research Institute (ILRI) generate significant volumes of genomic data across health, agriculture, and biodiversity domains. The genomic data generated from Kenyan and regional biological resources flows into global repositories governed by donor priorities, not by Kenyan regulators. This creates a striking sovereignty gap: advanced genomics infrastructure exists on Kenyan soil but remains outside the reach of national governance.

This gap is compounded by the risk of what Yakubu et al. (2018) term "parachute research", where international teams collect samples and data from Kenya, conduct research with limited local participation, and deposit results in international databases without triggering national benefit-sharing or oversight obligations. Once genomic data enters global repositories such as NCBI, traceability of its Kenyan origin and enforcement of any associated obligations or benefits become practically impossible (H3Africa Working Group on Ethics, 2017).

Kenya's ethics review architecture presents a further structural weakness. The National Scientific Ethics Committee and Institutional Scientific Ethics Review Committees were designed for clinical and social research without a distinction

for governance challenges of genomics, which include broad consent for secondary data use, population-level stigma risks, and the re-use of samples by investigators not involved in original collection. The H3Africa framework (2017) notes that Africa's ethical review infrastructure does not yet accommodate sharing and secondary use, a deficit that Kenya's own system reflects.

Kenya's Constitution (Article 42) guarantees environmental rights, and the STI Act promotes scientific advancement, but no overarching governance philosophy reconciles these with individual genomic data rights, obligations to research participants and communities, and international benefit-sharing norms.

4. METHODOLOGY

This policy brief was developed from mixed-methods research combining primary and secondary data collection.

The National Commission for science technology and Innovation partnered with health Security Partnered to assess the genomic data governance in Kenya. A tool was developed and deployed to 110 institutions mapped to be conducting life science research in Kenya, however, 31 institutions responded to the survey tool. i. Respondents comprised of universities and academic laboratories, government research institutes, private sector laboratories, diagnostic laboratories and public health laboratories.

The study explored practices and governance structures for genomic data including easy and equitable access and use, compliance with existing laws, appropriate protection, and benefit sharing, and privacy management. The assessment focused on the full lifecycle of genomic research and data.

Table 1: Study Respondents

Universities /Academia	Private Labs	Government Research Institutions	Diagnostic Labs
19	4	5	1

The secondary data sources explored are included in the reference section and primarily included the Science, Technology and Innovation Act, the Data Protection Act, the Biosafety Act, WHO's guidance on human genomic data collection, access, use and sharing, scientific journals and selected scholarly publications on genomic data governance.

FINDINGS

Below is the summary of the key findings from both primary and secondary sources relating to Infrastructure and human resource capacity, the governance framework at the national and institutional level, and data and material transfer practices;

Generation or Handing of Genomic Data in Kenyan Institutions

Evidence from the NACOSTI survey underpinning this policy shows that 71% of institutions are actively engaged in genomic data handling, whereas 29% are not currently involved, although they may possess future capacity or contribute to supporting roles across the genomic data value chain (Figure1). The findings suggests that Kenya already has a substantial operational base in genomic data generation.

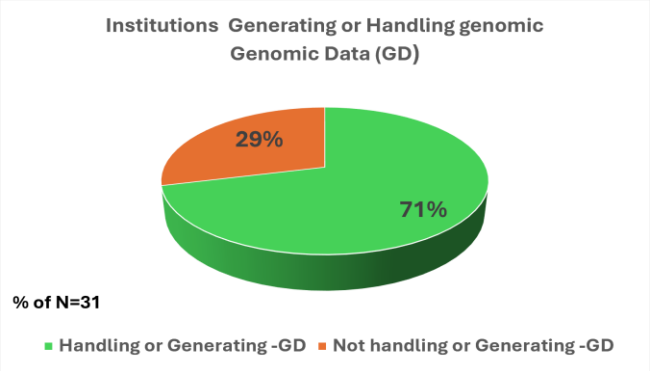


Figure 1: Percentage of institutions generating or managing genomic data

Compliance with Existing Governance Structures

The study explored the level of awareness of the data protection Act of 2019, registration as data controllers or processors and appointment of data protection officer. 81% were fully aware of the Act while 65% of the surveyed institutions were registered with the office of Data protection commission with 71% having a designated data protection officers. However, 19% of respondents who were somewhat aware of the Act can be categorized as non-compliant, highlighting the need for targeted capacity-building interventions. Additionally, 10% of institutions were in the process of registering with the office of data protection commissioner, while 13% were not registered, indicating potential legal exposure. A further 13% were uncertain of the next steps and are therefore likely to be non-compliant. The proportion of institutions without a dedicated Data Protection Officer was 19%, while 10% were uncertain about their status.

Easy and equitable access to genomic data

One of the noticeable observations was that most of the generated genomic data was in the hands of scientists and researchers (71%) or institutional scientific ethical review committees (52%) (Figure 2). Some of the institutions had ethics officers or data protection officers responsible for such data. This implied that most of the institutions did not have a well-structured governance mechanism for the genomic data being generated or handled within the institutions. There was no guarantee for easy access or equitable use of data for the public good.

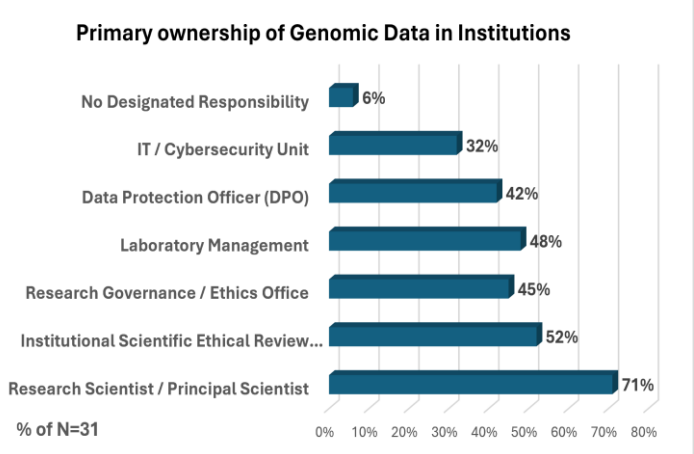


Figure 2: The graph shows a multi-selected ownership of genomic data in institutions

Data Transfer and Cross Border Sharing

The survey findings indicated that 23% of institutions frequently share genomic data with external collaborators, while the majority (52%) do so on an occasional basis. A smaller proportion (16%) reported no external data sharing. Notably, 65% of institutions utilize Data Sharing Agreements (DSAs) to govern such exchanges; however, 19% reported inconsistent use of DSAs, highlighting a significant gap in standardized data governance and compliance mechanisms. Specifically in cross-border data transfer, the results further show that 68% of surveyed institutions transfer genomic data beyond national borders, either within Africa or internationally. In contrast, 19% reported no cross-border data transfer. Of particular concern, 6% of institutions indicated that they share data outside Africa, presenting heightened risks related to data sovereignty and regulatory oversight. These findings underscore the need to strengthen regulatory frameworks governing genomic data sharing and cross-border transfer, including the mandatory and consistent use of Data Sharing Agreements, as well as enhanced safeguards to address data sovereignty and security risks associated with international data exchange.

Infrastructure Capacity

The evidence from the study indicates that Kenya is actively generating genomic data across multiple sectors. However, the findings further reveal that such data is not centrally stored within the institutions surveyed, pointing to fragmentation in data management systems. Institutions engaged in genomic data generation are equipped with high-throughput DNA sequencing platforms, including Illumina, Ion Torrent, and Oxford Nanopore technologies. Despite this capacity, the large volumes of genomic data produced by these platforms exceed existing institutional infrastructure. Most institutions lack adequate high-performance computing systems and dedicated servers with storage capacity to effectively manage, process, and store genomic data. This gap highlights a critical need for investment in centralized data repositories, with high-performance computing infrastructure. Consequently, private companies with fairly adequate infrastructure for GD management are slowly expanding. Kenya lacks a national genome biobank/repository.

KEY CHALLENGES/GAPS IDENTIFIED

The main challenges in genomic data governance identified in the study are summarized in the following below:

i. Regulatory and awareness gaps were identified in data governance and non-compliance to existing laws. Non-compliance indicated either non-compliance between policy design and implementation realities or limited awareness,

ii. The challenge identified was the absence of a structured mechanism for genomic data repositories in most institutions. As a result, most of the data was held at the individual level rather than within institutional systems. The institution does not retain ownership of any genomic data generated by researchers, notwithstanding the fact that such data is produced within its facilities.

iii. Although the Office of the Data Protection Commissioner has developed guidelines on data sharing domestically and across border, institutional policies remain largely misaligned with these requirements. The absence of comprehensive institutional data management policies and standard operating procedures may be attributed to limited awareness of these guidelines at the institutional level. This gap points to broader systemic weaknesses in enforcement frameworks, further evidenced by the lack of consistent monitoring and verification of compliance across institutions.

iv. Although sequencing platforms exist, the critical infrastructure such as high-Performance Computing systems to support GD management including storage and analysis was cited as one of the challenges. This may relate to the high resources required to maintain such a system and the specialised manpower required. There is no national repository or Biobanks which are critical for ensuring equitable access to available data for research.

v. There is a significant skills gap for experts especially in Bioinformatics which constitutes the most critical skill for genomic data management and analysis, as it underpins the ability to process, store, analyse, and interpret large-scale datasets. Although the H3Africa initiative has attempted to build capacity, this is not enough to raise the critical mass required. contributed immensely in the recent past to capacity building through post graduate training for scientists in this discipline

vi. Finance and budget for genomic research and data governance Inadequate funding allocation to genomic research due to lack of a clear definition and classification as ‘special data’.

vii. Data security. The assessment reveals a critical shortage of both physical and digital security safeguards across institutions and agencies engaged in the genomic data lifecycle. Most lack secure storage facilities, access controls, encryption protocols, and incident response procedures, leaving sensitive genomic data exposed to breaches, unauthorised access, and potential misuse.

POLICY RECOMMENDATIONS

To strengthen the governance of genomic data, the Kenyan government and stakeholders will need to;

i. **Strengthen the National genomic data governance frameworks** – This should include comprehensive genomic data law/regulation to address the governance of this special category of data including collection, access, storage, use, ownership and sharing. The law should be anchored on principles including ethical data collection, equitable access, responsible use, and fair benefit sharing, whilst promoting transparency, social justice, and accountability. The scope of the provisions in the Data Protection Act comprehensively cover the lifecycle of genomic data including transport, transfer, storage, biobanking, equitable access and benefit sharing.

ii. **Build National Genomic Data Infrastructure (Centers of Excellence for genomic data repositories and biobanks) to enhance data sovereignty**. This should include the establishment of government funded Centres of Excellence for genomic data, biobanks, national genomic data repositories with fully installed sequencing platforms and High-Performance Computing systems to efficiently manage this kind of data and governance structures that support greater sovereignty.

iii. **Develop Local Human Capacity and Increase Awareness on genomic data**. Enhance access to resources, training, capacity building and strengthening for researchers, healthcare professionals, genomic data administrators, policymakers, individuals, communities and other stakeholders involved in genome data collection, access, use and sharing. Fulfilment of this commitment requires greater effort in some countries and contexts than others due to existing inequities (WHO, 2024). This is achievable through targeted training, policy dialogues, and integration into national research priorities.

iv. **Strengthening Oversight, Coordination, and Accountability Mechanisms**. Enhanced inter-agency collaboration and alignment for instance, through

developing standardized procedures for ethical review practices, as well as clear guidelines for data access and use. Include regular quality audits to enhance compliance and continuous improvement.

v. **Increase Government Funding towards genomic research and governance**. The national government should increase budgetary allocations to support data governance systems that include legal frameworks, oversight systems, physical infrastructure, and enforcement. Adequate funding is critical for sustaining regulatory institutions, repositories, and oversight mechanisms.

vi. **Strengthen informed consent models to dynamic tiered consent.** Traditional consent processes are insufficient for genomic data owing to its potential thus requiring live/continuous consenting for current, broader and any unknown future use. The consent process is further complicated by the non-personal nature, i.e., genomic data of a member of a population can still reveal traits and characteristics of a community/population segment. Communal consent systems are therefore necessary.

vii. **Genomic data security.** Institutions should implement comprehensive cybersecurity measures including regular data backups for laboratory systems, encryption of sensitive data at rest and in transit, formalized cybersecurity training for all staff handling biological and patient data, and the adoption of network-integrated security architectures that bring laboratory devices under centralized monitoring, access control, incident response protocols

CONCLUSION

This policy is informed by primary data collected through a recent survey on genomic data governance, which highlights both existing capacities and critical gaps. The science and application of genomics offer transformative potential to improve individual and population health, as well as to strengthen agricultural systems through advancements in breeding, diagnostics, and disease management.

To fully realize these benefits, it is imperative to establish a robust and coordinated framework that enables the secure collection, access, use, and sharing of genomic data across health, research, and agricultural sectors. This requires urgent and sustained action to strengthen ethical, legal, social, and cultural safeguards, alongside investment in institutional capacity, infrastructure, and regulatory enforcement mechanisms.

Accordingly, this policy calls for the harmonization of data governance frameworks, mandatory compliance with data sharing standards, and the development of national systems for oversight, monitoring, and accountability. Without such measures, the country risks underutilizing genomic innovation and exposing itself to data governance, biosecurity, and sovereignty vulnerabilities

Yakubu, A., et al. (2018). Managing genomic research and biobanking in Africa. BMC Medical Ethics, 19(Suppl 1).

ACKNOWLEDGEMENT

The development of this policy brief was undertaken under the guidance and support of **Dr. David Ngigi, Ag. Director General, NACOSTI**.

This policy brief was prepared through the collective efforts and valuable contributions of the following NACOSTI officers:

Dr. Benson Mburu, Godfrey Kalerwa, Dr. Mary Onsarigo, Julius Mwangi, Carolyne Nekesa, Teresia Nyawira, Ezekiah Gatheru, Lilian Awala, Njuguna Kiuna, Lourine Auma, and Mitchell Gakii.

Their expertise and commitment in the development of this policy brief are acknowledged and appreciated.

REFERENCES

European Union. (2016). General Data Protection Regulation (GDPR).

H3Africa Consortium. (2014). Research capacity. Nature, 517, S1-S3.

H3Africa Ethics Working Group. (2017). Ethical frameworks for genomic research in Africa.

Kaye, J., et al. (2015). Dynamic consent: A patient interface for twenty-first century research networks. European Journal of Human Genetics, 23(2), 141–146.

Mittelstadt, B. (2017). Ethics of biomedical big data. Philosophy & Technology, 30(4), 419–436.

Mulder, N., et al. (2018). H3Africa bioinformatics network. Genome Medicine, 10(1).

Ramanan, P., et al. (2025). Growth of genomic data and global implications.

Ramsay, M., & Choudhury, A. (2026). Global inequalities in genomics.

Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). Missing diversity in human genetic studies. Cell, 177(1), 26–31.

World Health Organization (WHO). (2024). Guidance on human genome data collection, access, use and sharing.