

Leveraging Artificial Intelligence to Advance Bioinformatics in Africa: Opportunities, Challenges, and Ethical Considerations in Combating Antimicrobial Resistance

Beatus Lyimo¹ 

Abstract

Africa continues to bear a disproportionate burden of infectious diseases, particularly antimicrobial-resistant (AMR) infections, which significantly affect public health and socio-economic development. Addressing these complex health threats requires innovative approaches to data analysis, pathogen surveillance, and intervention design. The emergence of advanced computational tools especially artificial intelligence (AI) is expected to reduce turnaround times for AMR prediction from days to hours by leveraging whole-genome sequencing (WGS)-based models. This article explores the synergistic integration of AI and bioinformatics, focusing on their application in combating AMR in Africa. It details how AI techniques, particularly machine learning (ML) and deep learning (DL) algorithms, can enhance genomic research by automating the analysis of large-scale sequence datasets, predicting resistance patterns, and modeling infections transmission dynamics. In regions with limited laboratory capacity, AI models can detect resistance genes rapidly and assist clinicians in selecting appropriate antibiotics, offering a faster and more scalable alternative to traditional diagnostics. Tools such as convolutional neural networks (CNNs) and support vector machines (SVMs) are examples of models capable of classifying pathogen strains based on genetic data. Furthermore, the article highlights the emerging role of large language models (LLMs) in supporting bioinformatics workflows. These tools aid researchers by generating analysis scripts, interpreting complex outputs, troubleshooting code errors, summarizing literature, and preparing manuscripts or grant proposals particularly benefiting early-career scientists who may lack access to advanced training or mentorship. Despite notable progress, significant challenges remain, including limited infrastructure, barriers to data sharing, and the urgent need for ethical guidelines and policies to govern AI integration. Ultimately, this article underscores the transformative potential of AI in advancing bioinformatics across Africa and advocates for sustained investment in infrastructure, capacity-building, and responsible policy frameworks to harness AI for improved health research and disease control outcomes. We propose 3 priority actions: building African AMR genomic datasets, investing in AI-ready infrastructure, and developing responsible data-governance frameworks.

Keywords

Artificial intelligence, bioinformatics, AMR, genomic surveillance, genomic AMR prediction, AI-enabled diagnostics

Received: 30 September 2025; accepted: 9 February 2026

Introduction

The African continent bears a disproportionate burden of infectious diseases, including antimicrobial-resistant (AMR) infections, which continue to significantly affect public health systems and hinder socio-economic development across the region.¹ These infections not only cause immense human suffering but also strain healthcare systems, limit productivity, and place substantial financial burdens on already resource-constrained economies.² AMR has emerged as a major challenge, undermining decades of

medical progress in the prevention and control of infectious diseases. The WHO 2024 Strategic and Technical Advisory

¹School of Life Science and Bio-Engineering (LiSBE), Nelson Mandela African Institution of Science and Technology (NM-AIST), Arusha, Tanzania

Corresponding author:

Beatus Lyimo, Nelson Mandela African Institution of Science and Technology (NM-AIST), P.O. Box 447, Arusha, Tanzania.
Email: beatus.lyimo@nm-aist.ac.tz



Group on AMR has further highlighted the situation in Africa as an escalating crisis and a growing public health emergency.¹ In sub-Saharan Africa, high rates of multidrug-resistance (MDR) have been documented in key bacterial pathogens, including *Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *Salmonella enterica*. These pathogens are increasingly resistant to commonly used antibiotics such as third-generation cephalosporins, fluoroquinolones, and carbapenems, severely limiting treatment options.¹ Alarming, AMR is now estimated to cause more than 250 000 direct deaths annually in Africa, with even greater indirect mortality due to complications from untreatable infections and longer hospital stays.³ By 2050, AMR is projected to cause a cumulative global economic loss of up to \$100 trillion due to rising healthcare costs, reduced productivity, and the loss of effective treatments for life-threatening infections.⁴

The lack of diagnostic capacity, inadequate laboratory infrastructure, and weak antimicrobial stewardship programs have led to widespread reliance on empirical treatments, which are often ineffective and contribute to further resistance. More also, the absence of comprehensive genomic surveillance frameworks in most African countries has hindered efforts to understand and mitigate AMR dynamics. Traditional phenotypic methods, while essential, are time-consuming and provide limited insight into the genetic mechanisms underlying resistance. In contrast, whole-genome sequencing (WGS) and bioinformatics analyses allow high-resolution mapping of resistance determinants, virulence factors, and mobile genetic elements, enabling more precise monitoring of pathogen evolution and transmission pathways.⁵ Genomic surveillance can also support early detection of emerging resistance variants, thereby guiding targeted interventions and informing antimicrobial stewardship programs.⁶

Recent advances in AI, particularly ML and DL,^{7,8} have introduced new opportunities to enhance AMR genomic surveillance in Africa. These technologies enable more efficient AMR data analysis, predictive modeling, and real-time surveillance, thereby supporting timely and evidence-based public health responses across the continent. In this context bioinformatics, the application of computational tools to analyze and interpret biological data plays a critical role in characterizing resistance genes, tracking the evolution of drug-resistant pathogens, and elucidating transmission dynamics of antimicrobial-resistant strains. It enables high-throughput analysis of genomic data to detect mutations and resistance genes, characterize population structure, and guide the design of targeted interventions, including vaccines and therapeutics. However, despite its growing importance, many African countries lack the infrastructure, skilled workforce, and institutional capacity required to fully harness the potential of bioinformatics.⁹ Across sub-Saharan Africa, bioinformatics training remains fragmented, with few universities offering standalone undergraduate or graduate programs and most capacity building occurring through short-term courses or integration within broader life sciences curricula.¹⁰ Survey

evidence from Tanzania reveals that 96.4 % of researchers perform bioinformatics analyses on personal computers, with only about 10 % reporting access to high-performance computing facilities (HPC), cloud, or institutional servers, underscoring weak computational capacity at the institutional level.¹¹

This gap significantly constrains the region's ability to conduct local genomic analyses, interpret high-throughput sequencing data, and participate fully in global genomic surveillance initiatives.⁷ Similar challenges, have been reported across the region, including limited access to HPC, insufficient funding for genomic research, and a shortage of trained bioinformaticians.¹² As a result, many African researcher groups continue to rely heavily on collaborations with institutions in Europe or North America for data processing and interpretation an approach that can delay responses to local public health threats and limit research autonomy.

In this context, AI, particularly ML and deep learning, offers a transformative opportunity for strengthening regional bioinformatics capacity. Artificial intelligence can automate the analysis of large-scale genomic datasets, enhance AMR prediction, and support real-time outbreak detection.¹³⁻¹⁶ When integrated with bioinformatics pipelines, AI enables faster, more accurate, and scalable analysis of complex biological datasets, even in resource-constrained settings. For example, AI-driven tools can support researchers and clinicians across Africa in identifying emerging resistant bacterial strains, forecasting disease hotspots, and informing targeted treatment and control strategies.^{17,18} Strategic investment in bioinformatics education, regional computing infrastructure, and AI-enabled analytical frameworks could therefore help close existing capacity gaps and promote locally driven, sustainable genomic and public health solutions across Africa.

The Promise of Artificial Intelligence in Bioinformatics

Bioinformaticians are increasingly integrating AI into AMR analytics to improve efficiency and scalability, particularly in resource-constrained settings such as sub-Saharan Africa. Rather than functioning as isolated tools, AI applications can be conceptualized as accelerators across 4 key stages of the bioinformatics workflow: (1) data preprocessing, (2) predictive modeling, (3) interpretation and validation, and (4) dissemination and decision support (Figure 1).^{7,8} This framework provides a structured lens for understanding how AI contributes to AMR surveillance and research. At the data preprocessing stage, AI-assisted methods can support automated quality control, normalization, and feature extraction from genomic and phenotypic datasets. Large language models and code-generation tools can assist in drafting scripts for common preprocessing tasks such as sequence filtering, adapter trimming, and data imputation.¹⁹ However, these tools must be used cautiously, as they are prone to hallucinated outputs, coding errors, and reproducibility issues, necessitating mandatory expert review. Furthermore, hosted

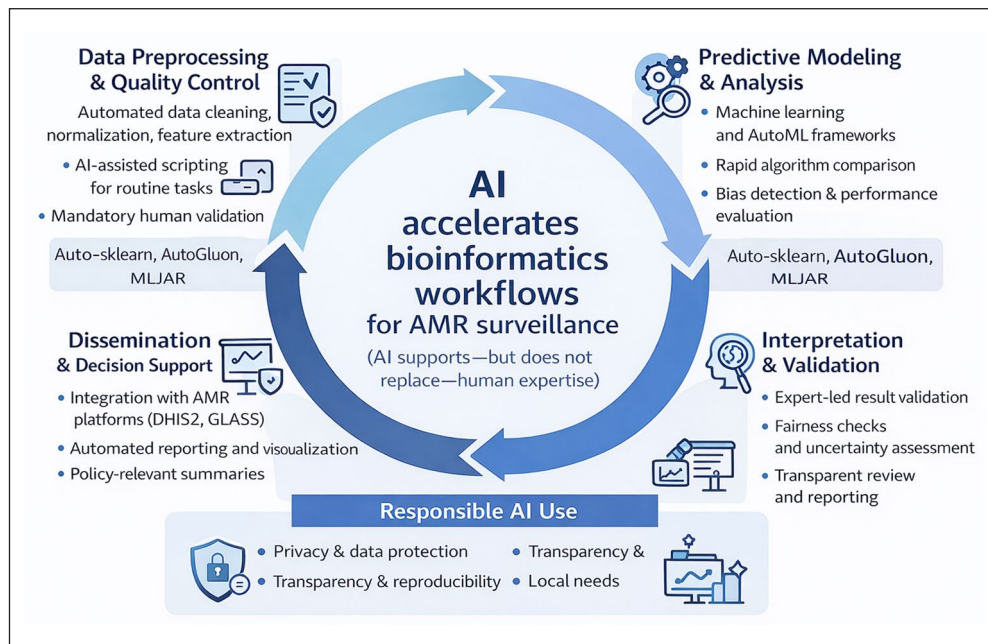


Figure 1. AI-assisted support for bioinformaticians, illustrating the major ways in which artificial intelligence enhances bioinformatics workflows.

LLM platforms should not be used with identifiable patient or genomic data due to data-privacy and security concerns, particularly in health research contexts.²⁰

During the modeling stage AutoML tools (eg, Google AutoML) frameworks enable rapid comparison of algorithms and hyperparameters for AMR prediction tasks.²¹ While several commercial AutoML platforms exist, their adoption in African settings may be constrained by licensing costs, Internet connectivity, and data-sovereignty requirements.²² Open-source alternatives such as AutoGluon²³ provide more accessible and locally deployable solutions. Importantly, AutoML does not automatically eliminate bias; rather, it can expose performance disparities across subgroups, which must then be addressed through explicit fairness assessments, stratified validation, and transparent reporting.¹²

Predictive Modeling of Antimicrobial Resistance

Artificial intelligence–based predictive modeling is reshaping how AMR is detected, characterized, and monitored across clinical, veterinary, and environmental settings. ML algorithms such as support vector machines (SVMs), decision trees, and deep neural networks (DNNs), are increasingly used to predict AMR profiles directly from genomic data. In practice, these models rely on specific genomic feature representations, most commonly k-mer frequencies, single-nucleotide polymorphism (SNP) profiles, and gene presence/absence matrices derived from whole-genome sequencing data.^{24,25} The choice of representation is critical, particularly for African datasets where sequencing depth, assembly quality, and metadata completeness can be highly variable.²⁶ For example, k-mer-based approaches are relatively robust to fragmented assemblies, whereas

SNP-based models often require high-quality reference genomes and consistent variant calling pipelines.²⁴

In sub-Saharan Africa, where routine antimicrobial susceptibility testing (AST), culture capacity, and molecular diagnostics remain limited in many settings,²⁷ ML-based genomic prediction offers a complementary approach to infer resistance patterns from sequencing data.²⁸ Emerging African studies, including ML-driven AMR prediction in *Escherichia coli* and *Salmonella* isolates from Kenya, Nigeria, and South Africa, demonstrate the feasibility of these approaches, although they remain far fewer than comparable studies from Europe and North America.^{15,29-32} This imbalance highlights both the promise of ML for African AMR surveillance and the current underrepresentation of African genomic data in global training datasets, which may affect model generalizability.³³⁻³⁵ ML tools can be trained on curated AMR databases (eg, the Comprehensive Antibiotic Resistance Database [CARD], ResFinder) to automatically detect and classify resistance determinants, track the spread of high-risk clones, and distinguish chromosomal from plasmid-borne resistance.³⁶ ChatGPT further supports this process by assisting in the generation of model training scripts, feature selection strategies, and evaluation pipelines.^{16,21} Together, these ML-enabled approaches can significantly strengthen AMR surveillance and response, particularly in resource-constrained settings. However, translating these predictive models into clinical decision support systems (CDSS) is non-trivial. Minimum requirements include standardized AST data formats, reliable laboratory information systems (LIS), interoperability with electronic health records (EHRs), and rigorous local validation within clinical workflows. Without these technical and organizational foundations, ML outputs risk remaining academic artifacts rather than actionable tools.

Biological Interpretation and Hypothesis Generation

Natural language processing (NLP) tools, powered by advanced AI models, are playing an increasingly critical role in bridging the gap between raw genomic data and actionable biological insights in AMR research. Beyond general literature summarization, concrete AMR-relevant applications include the automated extraction of resistance gene drug relationships from surveillance reports and published studies, as well as the prioritization of putative resistance-associated mutations in locally circulating pathogens, such as *Mycobacterium tuberculosis* strains in high burden African settings.^{37,38} By mining structured and unstructured sources including PubMed, CARD, and national surveillance reports, NLP systems can rapidly contextualize variants within known resistance mechanisms, treatment outcomes, and epidemiological settings. As a concrete example, AI models can analyze SNPs in AMR-associated genes to probabilistically estimate the likely functional impact of amino acid substitutions, helping to prioritize putative resistance-conferring mutations for experimental validation or epidemiological surveillance, rather than directly informing prescribing decisions.^{18,34} By integrating variant analysis with literature mining and functional prediction,³³ AI-driven NLP systems support a more holistic approach to biological interpretation by enabling researchers particularly in resource-limited settings to systematically synthesize evidence, prioritize high impact variants for further investigation, and generate testable, evidence-based hypotheses, thereby informing the design of targeted AMR surveillance and research interventions rather than direct clinical decision-making.

However, the performance and generalizability of NLP models in African contexts are constrained by data quality and representation challenges, including limited availability of full-text articles, underrepresentation of Africa-based research in indexed databases, and the prevalence of multilingual gray literature and non-standardized surveillance reports. Models trained primarily on English-language corpora from high-income settings may therefore miss locally relevant resistance patterns or contextual nuances. Addressing these gaps will require expanded digitization of African AMR reports, inclusion of regional languages and gray literature in training corpora, and close collaboration between computational scientists, microbiologists, and public health practitioners to ensure context-aware interpretation.

Visualization and Reporting

Artificial intelligence tools are playing an increasingly vital role in facilitating the visualization, interpretation, and communication of complex genomic and epidemiological data. Automated machine learning (AutoML) frameworks such as H2O.ai and Google AutoML streamline the development of predictive models by automatically selecting, tuning, and validating machine learning algorithms with minimal human intervention.²¹ The AutoML-based models trained on routine laboratory and genomic surveillance data

can be used to predict the probability of resistance in *E coli* or *Klebsiella pneumoniae* isolates. For example, susceptibility data exported from WHO Network for antimicrobial resistance surveillance software (WHONET)³⁹ and linked with basic epidemiological metadata (facility, specimen type, patient age, and location) can be used within AutoML frameworks to generate predictive resistance risk maps, enabling health officers to anticipate emerging resistance trends and adjust empirical treatment or stewardship strategies accordingly. In resource-constrained African settings, effective AMR dashboards should prioritize low-bandwidth performance, offline or asynchronous functionality, and role-specific views. Clinicians may require simplified summaries of dominant resistance patterns to guide treatment decisions, while policymakers benefit from aggregated trends, geographic comparisons, and early warning indicators. Interactive but lightweight visual elements such as temporal trend plots or district-level resistance maps are often more valuable than highly complex interfaces.

Artificial intelligence workflows also support automated generation of surveillance reports that synthesize genomic findings, visual analytics, and model outputs for diverse stakeholders.^{40,41} However, automated reporting should avoid “black-box” presentation of results.⁴² Reports should explicitly include model uncertainty, data completeness indicators, and clear explanations of analytical limitations to prevent over-interpretation of AI-derived outputs. When designed with these principles, AI-supported visualization and reporting can enhance transparency, support evidence-based decision-making, and improve the uptake of genomic surveillance insights into routine AMR control activities.⁴³

These reports can be tailored to specific audiences such as policymakers, epidemiologists, or clinicians and may include summarized results, visual analytics, model performance metrics, and evidence-based recommendations.⁴⁴ By automating this process, researchers can ensure consistency, reduce manual errors, and accelerate the translation of research findings into health interventions and policy actions.

AI-Assisted Workflow Optimization and Applications in AMR

Artificial intelligence-enabled platforms are increasingly being used to support real-time troubleshooting and optimization of complex bioinformatics workflows, including workflow managers such as Snakemake and Nextflow, as well as scripting and command-line environments. Large language models (LLMs) can assist researchers by interpreting error messages, suggesting code modifications, and providing optimized commands for commonly used tools such as BLAST, Prokka, or BWA. This form of real-time support is particularly valuable in settings such as Tanzania and other parts of sub-Saharan Africa, where access to advanced bioinformatics expertise remains limited.^{42,45} By enabling researchers to independently resolve technical challenges, AI-assisted tools help bridge critical human resource gaps and improve the efficiency of genomic data analysis.

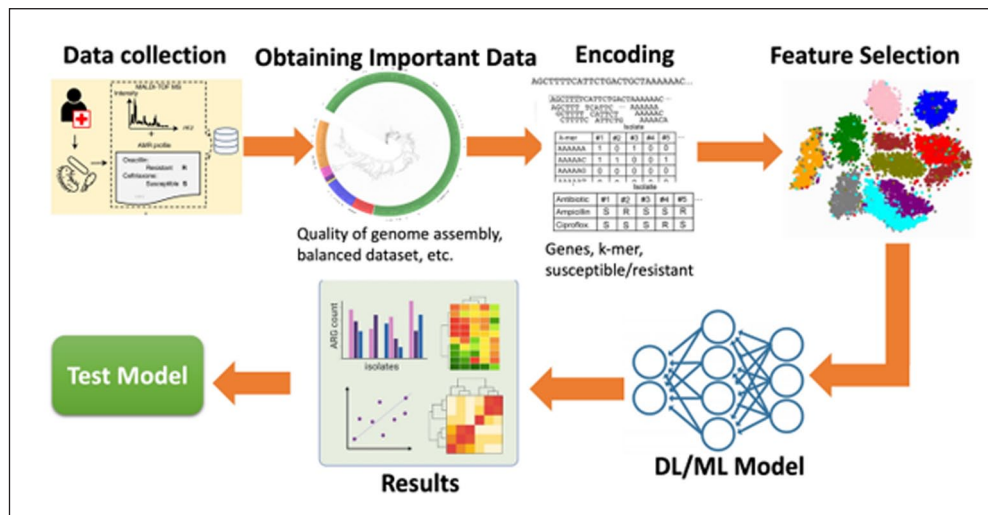


Figure 2. Overall process of applying machine-learning/deep-learning models in AMR identification. Source: This image was adopted from Ali et al.⁴⁹

Beyond troubleshooting, the integration of AI into bioinformatics workflows enhances the ability to process large-scale genomic datasets and extract actionable insights relevant to AMR. Bioinformatics pipelines transform raw sequencing data into interpretable outputs through genome assembly, annotation, and variant calling, enabling the identification of resistance genes, point mutations, and mobile genetic elements such as plasmids and integrons. When combined with machine learning approaches including support vector machines (SVMs), decision trees, random forests, and deep neural networks these features can be used to predict antimicrobial resistance phenotypes directly from genomic sequences (Figure 2). Such AI-driven AMR prediction frameworks, implemented in tools and platforms such as DeepVariant and SeekDeep, offer faster and potentially more scalable alternatives to conventional culture-based phenotypic testing, particularly for pathogens such as *E. coli*, *Salmonella*, *Klebsiella*, and *Staphylococcus aureus*.⁴⁶⁻⁴⁹

Nevertheless, reliance on AI-assisted troubleshooting and automated analysis introduces important risks and dependencies that must be explicitly acknowledged. LLMs are probabilistic systems and may generate plausible but incorrect code suggestions, propagate subtle workflow errors, or obscure underlying methodological flaws if used uncritically as “black-box tutors.”⁴² In addition, sharing pipeline logs, configuration files, or partial datasets with hosted AI services raises data security and confidentiality concerns, especially when human genomic data or sensitive clinical metadata are involved. To mitigate these risks, AI-assisted workflow optimization should be embedded within human-in-the-loop review processes, with mandatory expert validation of code changes, use of version control systems, and reproducibility checks.⁵⁰

To ensure sustainable impact, AI-assisted coding and AMR analytics should be positioned as capacity-building tools rather than substitutes for domain expertise. Practical recommendations include integrating AI-supported debugging and workflow optimization into formal bioinformatics

curricula and short courses, alongside training in best practices for code review, reproducible research, and responsible AI use. This approach supports both improved AMR surveillance and the long-term development of local bioinformatics capacity in resource-constrained settings.

A Mini-Framework for AI-Enabled AMR Applications in Sub-Saharan Africa

In sub-Saharan Africa, where sequencing capacity remain uneven, the integration of AI with bioinformatics can support AMR control across 4 interlinked application domains: (1) detection, (2) surveillance, (3) treatment optimization, and (4) stewardship monitoring. This framing clarifies how AI tools may be operationalized beyond standalone predictive modeling.

Detection and characterization

Artificial intelligence–assisted bioinformatics workflows enable the systematic detection of antimicrobial resistance genes, resistance-associated mutations, and mobile genetic elements from WGS data, supporting faster and more standardized interpretation of pathogen genotypes. When integrated with curated AMR reference databases (eg, CARD, ResFinder), these approaches improve consistency in resistance annotation and reduce inter-laboratory variability.^{51,52} When used appropriately, such workflows can support early resistance triage, prioritization of isolates for confirmatory testing, and enhanced surveillance while awaiting phenotypic results.

Surveillance and trend analysis

When applied to aggregated WGS and AST datasets, AI-enabled analytic frameworks can support the monitoring of temporal trends, evolutionary dynamics, and geographic spread of antimicrobial-resistant strains across

hospitals, districts, and regions.^{53,54} By integrating genomic relatedness, resistance determinants, and epidemiological metadata, these models facilitate early detection of emerging resistance lineages and support outbreak investigation and situational awareness for public health authorities.⁵⁴ When embedded within routine surveillance systems, such approaches can enhance regional trend analysis and inform targeted infection prevention and antimicrobial stewardship interventions.⁵⁵

Treatment optimization

Artificial intelligence-driven predictive models can support clinical decision-support systems (CDSS) by estimating the probability of antimicrobial resistance to specific agents based on pathogen genomic features, local resistance prevalence, and patient- or setting-level context.⁵⁶ When trained and validated on regionally representative datasets, such models may assist clinicians in prioritizing or de-escalating empirical therapy while awaiting AST results. For example, in suspected *Salmonella* or *Klebsiella pneumoniae* infections, genomic-ML models incorporating resistance determinants and regional surveillance data can help assess the likelihood of reduced susceptibility to fluoroquinolones or third-generation cephalosporins, thereby informing empiric treatment choices and stewardship decisions.^{57,58}

Stewardship monitoring

At the health-system level, AI-enabled analytics can integrate antimicrobial prescribing data, resistance patterns, and patient outcomes to support antimicrobial stewardship programs (ASPs). When embedded within routine surveillance and reporting systems, these tools can help identify deviations from treatment guidelines, detect inappropriate or prolonged antibiotic use, and monitor temporal changes in resistance and clinical outcomes following stewardship interventions.^{14,59} By enabling longitudinal analysis across wards, facilities, or districts, AI-supported dashboards and models can assist stewardship teams in prioritizing high-risk settings, evaluating the effectiveness of policy or behavioral interventions, and informing targeted feedback to prescribers.^{14,60} Importantly, the utility of such approaches depends on the availability of standardized prescribing and AST data, transparent model outputs, and close integration with existing stewardship governance structures to ensure that AI insights translate into actionable and context-appropriate decisions.

Data Prerequisites for Safe and Effective Deployment

Although AI-enabled bioinformatics is frequently framed as a means of “democratizing access” to advanced AMR diagnostics, independent and routine deployment in African health systems depends fundamentally on robust local data ecosystems. This is because performance, fairness, and

clinical safety of ML models are critically determined by the availability of regionally representative genomic and phenotypic datasets; in their absence, model outputs may be biased, poorly generalizable, or unsafe for clinical and public-health decision-making.⁶¹ Across many African settings, AMR data remain fragmented across laboratories, hospitals, and surveillance programs, with substantial heterogeneity in AST methodologies, incomplete epidemiological metadata, and limited longitudinal coverage. These structural constraints directly limit model training, external validation, and post-deployment performance monitoring, reinforcing that AI cannot compensate for gaps in core laboratory infrastructure or standardized surveillance systems.^{62,63} Without alignment to internationally recognized AST standards (eg, Clinical and Laboratory Standards Institute [CLSI] or European Committee on Antimicrobial Susceptibility Testing [EUCAST]) and surveillance frameworks such as the WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS), AI-derived predictions risk reinforcing existing data inequities rather than strengthening AMR control efforts.

AI-Driven Tools to Strengthen AMR Diagnostics Across Different Levels of Health Care

Primary care and peripheral laboratories

At the primary-care level, where laboratory infrastructure is minimal, AI-enabled approaches should build on tools that add intelligence to methods laboratories already use, rather than investing in expensive new platforms, low-cost AI-augmented diagnostic pipelines can enable rapid AMR detection in resource-limited hospitals.⁶⁴ A practical approach is to start with routine disk-diffusion AST and layer AI on top: smartphone applications using computer vision can automatically measure inhibition zones on antibiogram plates and interpret S/I/R status according to CLSI breakpoints,⁶⁵ achieving expert-level accuracy at minimal cost and without Internet connectivity, as shown in AI-based mobile AST readers.⁶⁴ These tools operate offline, require minimal training, and reduce human error and turnaround time. Limitations at this level include restricted pathogen coverage, reliance on culture-based methods, and inability to detect specific resistance mechanisms.

Secondary care and district hospitals

Secondary-level facilities with basic molecular capacity can complement phenotypic AST with targeted, rapid assays for high-risk resistance determinants. Isothermal amplification methods (eg, LAMP or RPA) and CRISPR-based assays coupled to lateral-flow strips or simple fluorescence readers allow faster detection of priority resistance genes. AI-assisted mobile applications can guide workflows, interpret results, and digitally capture data for surveillance. However, these approaches are limited to predefined targets and require reliable reagent supply chains, basic biosafety practices, and

periodic staff retraining,⁶⁶ where a phone app guides the workflow, reads bands or signals, and logs results for surveillance.

Tertiary care and referral or teaching hospitals

At tertiary-care level, portable nanopore sequencers (eg, MinION) combined with offline AI/ML tools such as Mykrobe, TB-Profiler, or DeepARG offer a relatively affordable way to generate same-day genomic AMR profiles from priority pathogens and feed those data into local decision support and surveillance systems.⁶⁷ Together, these tiered pipelines smartphone-assisted phenotypic AST at the periphery, rapid gene-targeted assays for high-risk resistance, and focused sequencing with AI prediction at referral centers provide a realistic, scalable roadmap for low- and middle-income African hospitals to deploy AI-enabled AMR diagnostics without prohibitive capital investment. However, any molecular or AI model must be validated against local isolates and phenotypic AST to capture local variants and avoid false predictions.

Cross-cutting considerations

Across all tiers of care, successful implementation of AI-enabled AMR diagnostics depends on enabling health-system components, including reliable supply chains for reagents and consumables, routine equipment maintenance and calibration, continuous staff training, and participation in external quality assurance (EQA) and proficiency-testing programs. These elements are central to laboratory quality management systems and are repeatedly identified as prerequisites for sustainable AMR surveillance and diagnostics in low- and middle-income settings.^{68,69} Importantly, AI-derived outputs should remain advisory and be interpreted alongside phenotypic AST results and clinical judgment, with systematic local validation to ensure analytical accuracy, clinical safety, and contextual relevance before routine deployment.⁶¹

Code Generation and Pipeline Development

Artificial intelligence tools are increasingly reshaping the support ecosystem for bioinformatics, particularly in settings where access to advanced training, expert mentorship, and computational support is limited. These platforms offer interactive and intuitive assistance, making complex computational biology tasks more accessible to researchers in Africa and beyond.⁴⁴⁻⁴⁶ These tools can be grouped by function, reflecting enduring use cases that are likely to persist as specific products evolve. General purpose LLMs provide conversational, task-oriented assistance across the bioinformatics workflow. They can translate natural-language prompts into executable code, outline analytical pipelines, explain methodological concepts, and assist with troubleshooting. Their primary value lies in lowering the barrier to entry for complex computational tasks and supporting

iterative problem-solving, rather than replacing domain expertise. Integrated code-assistant tools operate within programming environments and support real-time code completion, syntax correction, and debugging. By accelerating routine scripting in languages such as Python, R, and Bash, and by assisting with workflow managers such as Snakemake or Nextflow, these tools reduce technical friction while keeping analytical control with the user.

Domain-specific biomedical LLMs are trained on curated biological and medical corpora and are particularly suited to tasks such as variant annotation, gene-disease association exploration, pathway analysis, and interpretation of genomics or AMR outputs. Their strength lies in contextualizing computational results within biomedical knowledge, complementing rather than substituting experimental validation and expert review.

Literature-triage and evidence-synthesis tools support rapid screening, summarization, and comparison of large bodies of scientific literature. These tools are especially valuable for hypothesis generation, guideline appraisal, and keeping pace with rapidly expanding AMR and genomics research, provided outputs are cross-checked against primary sources. Across all categories, the enduring principle is that AI tools function best as assistive systems, not autonomous decision-makers. While they can substantially reduce the learning curve and improve productivity, they may also mask underlying conceptual gaps if used uncritically. Their effective use therefore depends on pairing AI-assisted workflows with foundational training in bioinformatics, microbiology, and statistics, alongside human oversight, validation, and reproducibility checks.

Current Status of African Genomic Diversity in Public Repositories for AMR Research

The effective application of AI in healthcare critically depends on the availability of large, high-quality, and region-specific datasets. Artificial intelligence models trained on genomic or clinical data from one population may not perform accurately when applied to another due to differences in genetic diversity, pathogen strains, and epidemiological patterns. In Africa, despite increasing sequencing efforts, the representation of both human and microbial genomes in public repositories such as NCBI/SRA remains limited. Pathogen genomic data from Africa currently represent only a small fraction of global sequence repositories (estimated at approximately 4.4 terabytes) (Figure 3), highlighting persistent underrepresentation in the datasets used to develop and validate genomic and AI-driven models.⁷⁰⁻⁷² Moreover, for bacterial pathogens in the East African Community, nearly 97% of genome assemblies are analyzed externally, highlighting capacity and data ownership challenges.⁷³ For *Mycobacterium tuberculosis*, curated datasets now include more than 17 000 strains from African countries, reflecting progress but also illustrating that a substantial portion of regional diversity remains unrepresented.⁷⁴ These gaps underscore the need

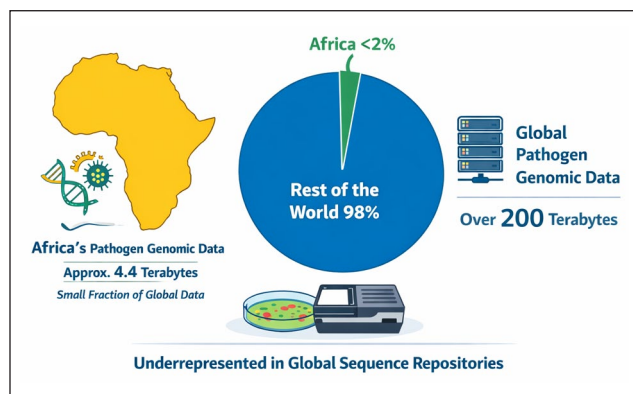


Figure 3. Underrepresentation of African pathogen genomic data in global sequence repositories, illustrating the disproportion between Africa's contribution and global sequencing outputs.

for locally generated, curated, and accessible datasets to train AI models capable of supporting antimicrobial resistance surveillance research tailored to African populations.

Efforts From the African Government to Include AI-Based Healthcare Solutions

Recognizing that AI-driven healthcare depends on large volumes of well-governed data, African governments and regional bodies have launched a series of mutually reinforcing initiatives to build and manage these data ecosystems. At the continental level, the African Union's Continental AI Strategy, Data Policy Framework, and Africa CDC's Health Data Governance Initiative position health data as a strategic asset and promote harmonized standards, data sovereignty, and ethical data sharing.⁷⁵⁻⁷⁸ These frameworks provide an enabling basis for regionally coordinated AMR genomics, including the development of continental or sub-regional pathogen genomic data hubs, alignment with WHO GLASS reporting requirements, and the establishment of AMR-specific data standards that integrate WGS, phenotypic AST, and epidemiological metadata.

Nationally, countries such as Kenya, Nigeria, Rwanda, Ghana, and South Africa are adopting AI strategies, digital health acts, and eHealth roadmaps that emphasize electronic health records, interoperability, and privacy-preserving use of patient data for AI-driven services.⁷⁸⁻⁸⁰ These policies are reinforced by investments in digital health infrastructure, open and Africa-relevant datasets, legal instruments on cybersecurity and personal data protection, and large-scale AI capacity-building programs in partnership with industry and academia. Together, these efforts aim to ensure that AI-based healthcare solutions are built on secure, high-quality African health data and deliver benefits to local populations rather than merely exporting data value.

However, there is a risk that such strategies remain aspirational ("policy-washing") if not accompanied by sustained financing, implementation guidance, and accountability

mechanisms. Experience from digital health and laboratory strengthening initiatives shows that uneven rollout can concentrate benefits in a small number of better-resourced countries or flagship institutions, potentially widening regional inequities in AMR surveillance capacity.

African Government Hospitals With Automated Electronic Health Record Systems for Data Collection

The adoption of automated electronic health record (EHR) systems in African government hospitals is gradually increasing, offering significant potential to strengthen health data management and surveillance. Countries such as Kenya, Uganda, Zambia, South Africa, and Tanzania have established varying levels of hospital-based EHR infrastructure (Figure 4).^{78,81,82} As shown in Figure 4, variation in EHR system strength across African countries has direct implications for readiness to deploy AI-enabled AMR surveillance and clinical decision-support tools. Countries with relatively strong EHR infrastructures are better positioned to integrate AI models that rely on structured clinical data, longitudinal patient records, and routine reporting of antimicrobial use and outcomes. Although implementation has often been concentrated in disease-specific programs, platforms like OpenMRS have enabled digital recording of patient information, laboratory results, and reporting in low-resource settings. In Tanzania, government-supported platforms such as GoT-HoMIS and eHMS are increasingly used across major hospitals, improving clinical workflow and enabling aggregation of data at regional and national levels. Despite these advances, challenges persist, including high costs of procurement and maintenance, limited interoperability between systems, insufficient infrastructure, and inadequate training of healthcare personnel, often resulting in parallel paper-based and electronic records.^{82,83} Integrating EHR systems with AMR surveillance represents a critical opportunity for improving public health outcomes. In Tanzania, AMR surveillance is coordinated across sentinel hospitals, with data collected at the National Health Laboratory and reported to global platforms such as GLASS.⁸⁴ Automated EHRs can facilitate near-real-time monitoring of resistant pathogens, enable accurate linkage of laboratory results with clinical and demographic data, and support evidence-based prescribing practices. However, the lack of standardized data-sharing protocols and limited interoperability between hospital systems hampers the effective use of EHRs for AMR surveillance.

Challenges and Opportunities in AI-Driven Bioinformatics in Africa

Despite notable progress AI into bioinformatics workflows across Africa, several persistent challenges continue to hinder its full potential. Beyond the well-documented barriers

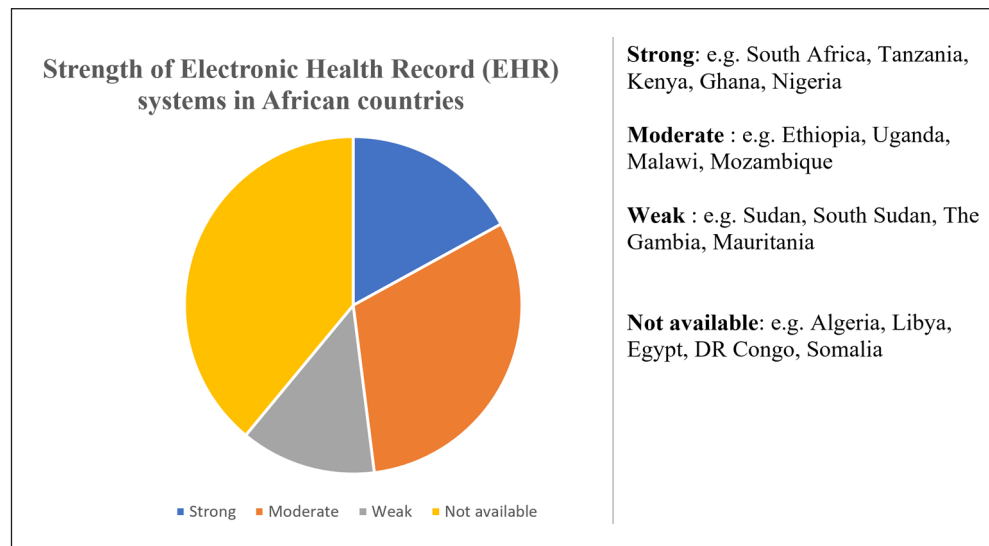


Figure 4. Strength of electronic health record (EHR) systems in African countries, source of the data.⁷⁸

of limited Internet connectivity and HPC infrastructure, institutional disincentives also constrain progress. For example, researchers often lack recognition for data stewardship activities, depend heavily on short-term, project-based funding, and face brain drain as skilled personnel move to better-resourced institutions abroad. In countries such as Tanzania, public universities and regional research centers experience frequent network disruptions and insufficient computing capacity, which impede training and deployment of computationally intensive AI models, accessing cloud-based platforms, and processing large-scale genomic datasets. These constraints significantly impede researchers' ability to train and deploy computationally intensive AI models, use cloud-based platforms, or access and process large-scale genomic datasets essential for modern bioinformatics research. Without robust digital infrastructure, the scalability and reproducibility of AI-driven genomic analyses remain limited, exacerbating disparities in global research capacity and slowing progress in areas such as genomic antimicrobial resistance surveillance.^{85,86}

To mitigate these challenges, targeted and actionable interventions are required at both institutional and regional levels. Universities and research institutes could formally allocate protected time for data curation, annotation, and sharing, recognizing these activities within promotion and performance evaluation frameworks to incentivize high-quality data stewardship. Reducing dependence on short-term project funding by embedding data management roles into core institutional budgets would further enhance continuity and sustainability. At a regional scale, the establishment of shared HPC consortia—coordinated across neighboring countries and linked to national research and education networks—could lower costs, pool technical expertise, and provide equitable access to computational resources. Such consortia would enable routine training and deployment of AI models for genomics, improving

timeliness of outbreak detection and the calibration of risk prediction models for African populations.

Responsible AI: Data Sharing and Ethical Considerations

The integration of AI into health-related bioinformatics in Africa raises not only ethical questions but also critical cybersecurity and data governance challenges. Issues such as data ownership, patient confidentiality, and cross-border data sharing are compounded by the lack of robust regulatory and digital security frameworks across many African countries. These gaps heighten the risks of data breaches, misuse, and inequitable benefit-sharing, particularly when sensitive genomic and clinical data are processed using AI tools hosted on third-party platforms or cloud environments. For instance, sharing TB genomic data across borders on third-party cloud providers may expose populations to unintended data exploitation or loss of control over genomic resources. Furthermore, over-reliance on AI without sufficient human oversight in clinical or research decision-making can result in biased, opaque, or unaccountable outcomes.

To ensure safe and ethical deployment, AI-driven bioinformatics systems must adhere to recognized responsible AI principles. Based on WHO guidance on AI in health,⁸⁷ these principles can be grouped into 4 key areas:

1. **Transparency:** Algorithms and predictive models should be interpretable and auditable, with clear documentation of data sources, limitations, and intended use.
2. **Accountability:** Human oversight should be maintained for all AI-supported decisions, and clear responsibility assigned for errors or unintended outcomes.

3. **Equity and data sovereignty:** Data sharing and AI deployment must safeguard the interests of local populations, ensuring fair benefit-sharing and that sensitive data remain under appropriate national or institutional control.
4. **Cybersecurity and privacy:** Robust digital security measures, privacy-preserving data sharing protocols, and secure storage and transmission systems are essential to protect sensitive genomic and clinical data.

By explicitly considering these principles in realistic African scenarios, policymakers and researchers can foster trust, ensure safe deployment, and maximize the public health benefits of AI-driven bioinformatics while minimizing ethical and security risks.⁸⁸⁻⁹⁰

Lack of Locally Trained AI Models

Most AI models are developed using datasets from Europe, North America, or Asia, which may not accurately reflect the genetic diversity of African pathogens, human populations, or local epidemiological patterns. This mismatch can lead to biased predictions, reduced diagnostic accuracy, and poorly informed public health interventions across African settings. For example, models trained on Ugandan *M. tuberculosis* genomic data showed good predictive performance on the Uganda testing dataset, but the logistic regression model for rifampicin and streptomycin resistance did not generalize well when applied to an independent South African dataset.¹³ This highlights how models trained in one region may degrade when applied to data from other regions with different genomic variations. To address this, local calibration studies should be mandatory before deployment, ensuring that AI models are validated and adapted to the African context. To support this, greater emphasis must be placed on empowering African research institutions to share data ethically and securely, fostering inclusive model development that addresses regional needs.⁹¹⁻⁹³ Concrete governance mechanisms can help achieve this, and we specifically propose the following:

1. African-hosted federated learning infrastructures, which allow AI models to be trained across multiple institutions without transferring sensitive genomic or clinical data across borders, thereby preserving privacy and data sovereignty.
2. Data-access committees with community representation, ensuring that decisions about data use are transparent, ethically reviewed, and aligned with local values and participant consent.
3. Regional AMR data trusts, which coordinate equitable data sharing across institutions, establish clear rules for data access, and facilitate collaborative model development while safeguarding sensitive information.

By implementing these mechanisms, African research institutions can actively participate in AI-driven bioinformatics, ensure equitable benefit-sharing, and maintain high ethical standards in line with local and international guidance.

Policy Development for Responsible and Collaborative AI Use in African Academia

As AI becomes increasingly embedded in research and education, African countries and academic institutions must urgently develop robust policies to guide the ethical, professional, and equitable use of AI technologies. These policies should explicitly frame AI as a collaborative tool that enhances, rather than replaces, human expertise. In practical terms, this includes developing clear guidelines on acceptable AI use in coursework, theses, and grant writing, as well as institutional statements defining appropriate AI assistance in coding, data analysis, and scientific writing, with requirements for disclosure and human oversight. Such policies can help maintain academic integrity while enabling responsible innovation.

Promoting multidisciplinary collaboration across bioinformatics, clinical research, data science, and public health is essential for driving impactful AI innovations. Rather than creating entirely new structures, existing research networks and consortia such as Human Heredity and Health in Africa (H3Africa), Africa CDC surveillance networks, and regional university consortia could be strengthened by explicitly incorporating AI-bioinformatics working groups, shared training programs, and joint infrastructure initiatives. Embedding AI capacity within these established platforms would accelerate sustainable adoption, facilitate knowledge exchange, and reduce duplication of effort. Furthermore, investing in African-centric AI models and datasets that reflect local health systems, pathogen diversity, and socio-cultural contexts is critical to creating globally relevant yet locally tailored bioinformatics solutions.^{27,94}

Conclusion and Recommendations

AI offers important opportunities to strengthen bioinformatics and AMR surveillance in sub-Saharan Africa, particularly in settings where laboratory capacity and specialist expertise remain limited. However, AI depends on the availability of high-quality data, robust governance frameworks, and sustained investment in human and institutional capacity. Without these foundations, AI risks reinforcing existing inequities or diverting attention from essential laboratory and surveillance infrastructure.

To translate AI-enabled bioinformatics from promise into practice, a focused set of priority actions is required (Box 1):

Box 1. Priority actions for AI-enabled bioinformatics and AMR surveillance in Africa.

1. Strengthen digital and laboratory infrastructure (Governments, Funders): Invest in reliable internet connectivity, scalable computing resources (including cloud and regional HPC), and routine microbiology and sequencing capacity to ensure AI models are built on robust data.
2. Build sustainable human capacity (Universities, Research Institutes): Establish formal degree programs and regional centers of excellence in bioinformatics, data science, and AI, complemented by short courses for clinicians and public health professionals.
3. Develop governance and ethical frameworks (Governments, Regulators): Implement clear policies on data sharing, privacy, model transparency, and responsible AI use to support trustworthy deployment in biomedical research and surveillance.
4. Promote context-specific AI development (Universities, Hospitals, Industry; ongoing): Prioritize local data generation and model training that reflect Africa's pathogen diversity and health system realities, reducing dependence on externally trained models.
5. Foster cross-sector collaboration (All stakeholders; ongoing): Encourage partnerships between academia, healthcare systems, public health agencies, and technology developers to ensure AI tools align with real-world needs.

In summary, AI can meaningfully enhance bioinformatics and AMR surveillance in Africa only when embedded within strong data ecosystems, laboratory capacity, and governance structures. Strategic, coordinated investment rather than technological solutionism will be essential to ensure that AI contributes to equitable, sustainable health system strengthening across the continent.

Acknowledgments

Not applicable.

Ethical Considerations

This study is conceptual in nature and did not involve the collection or analysis of primary human or animal data; therefore, formal ethical approval was not required. Nevertheless, the work adheres to principles of responsible research practice, including the ethical use of secondary data sources and artificial intelligence tools. No identifiable human, clinical, or genomic data were uploaded to third-party platforms, and all examples discussed are illustrative and based on publicly available information.

Author Contributions

BL conceived and designed the study, wrote the manuscript, and submitted it for publication.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Beatus Lyimo  <https://orcid.org/0000-0002-6743-4546>

References

1. WHO. *WHO Strategic and Technical Advisory Group for Antimicrobial Resistance (STAG-AMR)*. WHO; 2024.
2. World Bank. *Drug-Resistant Infections a Threat to Our Economic Future*. World Bank; 2017. www.worldbank.org
3. Antimicrobial Resistance Collaborators. The burden of bacterial antimicrobial resistance in the WHO African region in 2019: a cross-country systematic analysis. *Lancet Glob Health*. 2024;12:e201-e216. doi:10.1016/S2214-109X(23)00539-9
4. Ahmed SK, Hussein S, Qurbani K, et al. Antimicrobial resistance: impacts, challenges, and future prospects. *J Med Surg Public Health*. 2024;2:100081. doi:10.1016/j.glmedi.2024.100081
5. Chigozie VU, Aniokete CU, Ogbonna PI, Iroha RI. Transforming antimicrobial resistance mitigation: the genomic revolution in one health and public health. *Discov Appl Sci*. 2025;7:1187. doi:10.1007/s42452-025-07053-7
6. Sherry NL, Lee JYH, Giulieri SG, et al. Genomics for antimicrobial resistance—progress and future directions. *Antimicrob Agents Chemother*. 2025;69:e0108224. doi:10.1128/aac.01082-24
7. Deo RC. Machine learning in medicine. *Circulation*. 2015;132:1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593
8. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet*. 2015;16:321-332. doi:10.1038/nrg3920
9. Lyimo BM, Popkin-Hall ZR, Giesbrecht DJ, et al. Potential opportunities and challenges of deploying next generation sequencing and CRISPR-Cas systems to support diagnostics and surveillance towards malaria control and elimination in Africa. *Front Cell Infect Microbiol*. 2022;12:757844. doi:10.3389/fcimb.2022.757844
10. Kiosia A, Boylan S, Retford M, et al. Current data science capacity building initiatives for health researchers in LMICs: global & regional efforts. *Front Public Health*. 2024;12:1418382. doi:10.3389/fpubh.2024.1418382
11. Sangeda RZ, Mwakilili AD, Masamu U, et al. A Baseline evaluation of bioinformatics capacity in Tanzania reveals areas for training. *Front Educ (Lausanne)*. 2021;6:665313. doi:10.3389/educ.2021.665313
12. Ochola R. The case for genomic surveillance in Africa. *Trop Med Infect Dis*. 2025;10:129. doi:10.3390/tropicalmed10050129
13. Babirye SR, Nsubuga M, Mboowa G, Batte C, Galiwango R, Kateete DP. Machine learning-based prediction of antibiotic resistance in Mycobacterium tuberculosis clinical isolates from Uganda. *BMC Infect Dis*. 2024;24:1391. doi:10.1186/s12879-024-10282-7
14. Peiffer-Smadja N, Rawson TM, Ahmad R, et al. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clin Microbiol Infect*. 2020;26:584-595. doi:10.1016/j.cmi.2019.09.009
15. Xu Y, Mao Y, Hua X, et al. Machine learning-based prediction of antimicrobial resistance and identification of AMR-related SNPs in Mycobacterium tuberculosis. *BMC Genom Data*. 2025;26:48. doi:10.1186/s12863-025-01338-x
16. Nsubuga M, Galiwango R, Jjingio D, Mboowa G. Generalizability of machine learning in predicting antimicrobial

- resistance in *E. Coli*: a multi-country case study in Africa. *BMC Genomics*. 2024;25:287. doi:10.1186/s12864-024-10214-4
17. Tanui CK, Ndembu N, Kebede Y, Tessema SK. Artificial intelligence to transform public health in Africa. *Lancet Infect Dis*. 2024;24:e542. doi:10.1016/S1473-3099(24)00435-3
 18. Kasse GE, Cosh SM, Humphries J, Islam MS. Leveraging artificial intelligence for One Health: opportunities and challenges in tackling antimicrobial resistance—scoping review. *One Health Outlook*. 2025;7:51. doi:10.1186/s42522-025-00170-8
 19. Smit D, Smuts H. The impact of GitHub copilot on developer productivity from a software engineering body of knowledge perspective, 2024. <https://aisel.aisnet.org/amcis2024>
 20. Yu H, Fan L, Li L, et al. Large language models in biomedical and health informatics: a review with bibliometric analysis. *J Healthc Inform Res*. 2024;8:658-711. doi:10.1007/s41666-024-00171-8
 21. Bisong E. *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*. Apress Media LLC; 2019. doi:10.1007/978-1-4842-4470-8
 22. Abdelwanis M, Simsekler MCE, Gabor AF, Sleptchenko A, Omar M. Artificial intelligence adoption challenges from healthcare providers' perspectives: a comprehensive review and future directions. *Saf Sci*. 2026;193:107028. doi:10.1016/j.ssci.2025.107028
 23. Erickson N, Mueller J, Shirkov A, et al. AutoGluon-tabular: robust and accurate AutoML for structured data. Published Online March 13, 2020. <http://arxiv.org/abs/2003.06505>
 24. Jaillard M, Lima L, Tournoud M, et al. A fast and agnostic method for bacterial genome-wide association studies: bridging the gap between k-mers and genetic events. *PLoS Genet*. 2018;14:e1007758. doi:10.1371/journal.pgen.1007758
 25. Drouin A, Giguère S, Déraspe M, et al. Predictive computational phenotyping and biomarker discovery using reference-free genome comparisons. *BMC Genomics*. 2016;17:754. doi:10.1186/s12864-016-2889-6
 26. Vaseghi H, Akrami SM, Rashidi-Nezhad A. The challenges in the interpretation of genetic variants detected by genomics techniques in patients with congenital anomalies. *J Clin Lab Anal*. 2023;37:e24967. doi:10.1002/jcla.24967
 27. Serge Andigema A, Tania Cyrielle NN, Ekwelle E. Artificial intelligence in African healthcare: catalyzing innovation while confronting structural challenges. Preprint Posted Online June 23, 2025. doi:10.20944/preprints202506.1824.v1.
 28. Tadesse BT, Ashley EA, Ongarello S, et al. Antimicrobial resistance in Africa: a systematic review. *BMC Infect Dis*. 2017;17:616. doi:10.1186/s12879-017-2713-1
 29. Park SE, Pham DT, Pak GD, et al. The genomic epidemiology of multi-drug resistant invasive non-typhoidal *Salmonella* in selected sub-Saharan African countries. *BMJ Glob Health*. 2021;6:e005659. doi:10.1136/bmjgh-2021-005659
 30. Moradigaravand D, Palm M, Farewell A, Mustonen V, Warringer J, Parts L. Prediction of antibiotic resistance in *Escherichia coli* from large-scale pan-genome data. *PLoS Comput Biol*. 2018;14:e1006258. doi:10.1371/journal.pcbi.1006258
 31. Anyaegbunam ZKG, Mba IE, Doowuese Y, et al. Antimicrobial resistance containment in Africa: moving beyond surveillance. *Biosaf Health*. 2024;6:50-58. doi:10.1016/j.bshealth.2023.12.003
 32. Pesesky MW, Hussain T, Wallace M, et al. Evaluation of machine learning and rules-based approaches for predicting antimicrobial resistance profiles in gram-negative bacilli from whole genome sequence data. *Front Microbiol*. 2016;7:1887. doi:10.3389/fmicb.2016.01887
 33. Wheeler NE, Price V, Cunningham-Oakes E, et al. Innovations in genomic antimicrobial resistance surveillance. *Lancet Microbe*. 2023;4:e1063-e1070. doi:10.1016/S2666-5247(23)00285-9
 34. Massé J, Lardé H, Archambault M, et al. Conventional and unsupervised artificial intelligence analyses identified risk factors for antimicrobial resistance on dairy farms in the province of Québec, Canada. *J Dairy Sci*. 2024;107:11398-11414. doi:10.3168/jds.2024-25088
 35. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell*. 2019;177:26-31. doi:10.1016/j.cell.2019.02.048
 36. Djordjevic SP, Jarocki VM, Seemann T, et al. Genomic surveillance for antimicrobial resistance — a One Health perspective. *Nat Rev Genet*. 2024;25:142-157. doi:10.1038/s41576-023-00649-y
 37. Oikonomou ED, Karvelis P, Giannakeas N, Vrachatis A, Glavas E, Tzallas AT. How natural language processing derived techniques are used on biological data: a systematic review. *Netw Model Anal Health Inform Bioinform*. 2024;13:23. doi:10.1007/s13721-024-00458-1
 38. Cheng S, Wei Y, Zhou Y, et al. Deciphering genomic codes using advanced natural language processing techniques: a scoping review. *J Am Med Inform Assoc*. 2025;32:761-772. doi:10.1093/jamia/ocaf029
 39. WHO. Implementation Status of Integrated AMR Surveillance within a One Health Approach in the WHO African Region. Accessed January 8, 2026. <https://iris.who.int/server/api/core/bitstreams/32863215-b739-4a20-bbb4-1b3e07c81ebb/content>
 40. Lim C, Miliya T, Chansamouth V, et al. Automating the generation of antimicrobial resistance surveillance reports: proof-of-concept study involving seven hospitals in seven countries. *J Med Internet Res*. 2020;22:e19762. doi:10.2196/19762
 41. Argimón S, Abudahab K, Goater RJE, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom*. 2016;2:e000093. doi:10.1099/mgen.0.000093
 42. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019;1:206-215. doi:10.1038/s42256-019-0048-x
 43. Tacconelli E, Sifakis F, Harbarth S, et al. Surveillance for control of antimicrobial resistance. *Lancet Infect Dis*. 2018;18:e99-e106. doi:10.1016/S1473-3099(17)30485-1
 44. Wang L, Ge X, Liu L, Hu G. Code interpreter for bioinformatics: are we there yet. *Ann Biomed Eng*. 2024;52:754-756. doi:10.1007/s10439-023-03324-9
 45. Shue E, Liu L, Li B, Feng Z, Li X, Hu G. Empowering beginners in bioinformatics with ChatGPT. *Quant Biol*. 2023;11:105-108. doi:10.15302/J-QB-023-0327
 46. Merow C, Serra-Diaz JM, Enquist BJ, Wilson AM. AI chatbots can boost scientific coding. *Nat Ecol Evol*. 2023;7:960-962. doi:10.1038/s41559-023-02063-3
 47. Jia H, Li X, Zhuang Y, et al. Neural network-based predictions of antimicrobial resistance phenotypes in multidrug-resistant *Acinetobacter baumannii* from whole genome sequenc-

- ing and gene expression. *Antimicrob Agents Chemother.* 2024;68:e0144624. doi:10.1128/aac.01446-24
48. Gao Y, Li H, Zhao C, Li S, Yin G, Wang H. Machine learning and feature extraction for rapid antimicrobial resistance prediction of *Acinetobacter baumannii* from whole-genome sequencing data. *Front Microbiol.* 2024;14:1320312. doi:10.3389/fmicb.2023.1320312
 49. Ali T, Ahmed S, Aslam M. Artificial intelligence for antimicrobial resistance prediction: challenges and opportunities towards practical implementation. *Antibiotics.* 2023;12:523. doi:10.3390/antibiotics12030523
 50. Sandve GK, Nekrutenko A, Taylor J, Hovig E. Ten simple rules for reproducible computational research. *PLoS Comput Biol.* 2013;9:e1003285. doi:10.1371/journal.pcbi.1003285
 51. Bortolaia V, Kaas RS, Ruppe E, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother.* 2020;75:3491-3500. doi:10.1093/jac/dkaa345
 52. Alcock BP, Raphenya AR, Lau TTY, et al. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 2019;48:D517-D525. doi:10.1093/nar/gkz935
 53. Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat Rev Genet.* 2018;19:9-20. doi:10.1038/nrg.2017.88
 54. Hendriksen RS, Bortolaia V, Tate H, Tyson GH, Aarestrup FM, McDermott PF. Using genomics to track global antimicrobial resistance. *Front Public Health.* 2019;7:242. doi:10.3389/fpubh.2019.00242
 55. Seale AC, Gordon NC, Islam J, Peacock SJ, Scott JAG. AMR Surveillance in low and middle-income settings—a roadmap for participation in the Global Antimicrobial Surveillance System (GLASS). *Wellcome Open Res.* 2017;2:92. doi:10.12688/wellcomeopenres.12527.1
 56. Rawson TM, Ahmad R, Toumazou C, Georgiou P, Holmes AH. Artificial intelligence can improve decision-making in infection management. *Nat Hum Behav.* 2019;3:543-545. doi:10.1038/s41562-019-0583-9
 57. Nguyen M, Brettin T, Long SW, et al. Developing an in silico minimum inhibitory concentration panel test for *Klebsiella pneumoniae*. *Sci Rep.* 2018;8:421. doi:10.1038/s41598-017-18972-w
 58. Chanamé Pinedo LE, Franz E, Dallman TJ, et al. Genomic epidemiology of *Salmonella* Enteritidis human infections in the Netherlands, 2019 to 2023. *Microb Genom.* 2025;11:001394. doi:10.1099/mgen.0.001394
 59. Rabaan AA, Alhumaid S, Al Mutair A, et al. Application of artificial intelligence in combating high antimicrobial resistance rates. *Antibiotics.* 2022;11:784. doi:10.3390/antibiotics11060784
 60. Harandi H, Shafaati M, Salehi M, et al. Artificial intelligence-driven approaches in antibiotic stewardship programs and optimizing prescription practices: a systematic review. *Artif Intell Med.* 2025;162:103089. doi:10.1016/j.artmed.2025.103089
 61. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med.* 2019;25:1337-1340. doi:10.1038/s41591-019-0548-6
 62. Wertheim HFL, Huang VTL, Kuijper EJ. Clinical microbiology laboratories in low-resource settings, it is not only about equipment and reagents, but also good governance for sustainability. *Clin Microbiol Infect.* 2021;27:1389-1390. doi:10.1016/j.cmi.2021.07.027
 63. Argimón S, Yeats CA, Goater RJ, et al. A global resource for genomic predictions of antimicrobial resistance and surveillance of *Salmonella* Typhi at pathogenwatch. *Nat Commun.* 2021;12:2879. doi:10.1038/s41467-021-23091-2
 64. Pascucci M, Royer G, Adamek J, et al. AI-based mobile application to fight antibiotic resistance. *Nat Commun.* 2021;12:1173. doi:10.1038/s41467-021-21187-3
 65. CLSI. *M100-S24 Performance Standards for Antimicrobial*. CLSI; 2014.
 66. Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat Protoc.* 2019;14:2986-3012. doi:10.1038/s41596-019-0210-2
 67. Chan WS, Au CH, Chung Y, et al. Rapid and economical drug resistance profiling with Nanopore MinION for clinical specimens with low bacillary burden of *Mycobacterium tuberculosis*. *BMC Res Notes.* 2020;13:444. doi:10.1186/s13104-020-05287-9
 68. *Global Antimicrobial Resistance Surveillance System (GLASS) Report: Early Implementation*. World Health Organization; 2020.
 69. Ombelet S, Ronat JB, Walsh T, et al. Clinical bacteriology in low-resource settings: today's solutions. *Lancet Infect Dis.* 2018;18:e248-e258. doi:10.1016/S1473-3099(18)30093-8
 70. Fatumo S, Yakubu A, Oyedele O, et al. Promoting the genomic revolution in Africa through the Nigerian 100K Genome Project. *Nat Genet.* 2022;54:531-536. doi:10.1038/s41588-022-01071-6
 71. Omotoso OE, Teibo JO, Atiba FA, Oladimeji T, Adebisin AO, Babalghith AO. Bridging the genomic data gap in Africa: implications for global disease burdens. *Global Health.* 2022;18:103. doi:10.1186/s12992-022-00898-2
 72. Mboowa G, Kakooza F, Egesa M, et al. The rise of pathogen genomics in Africa. *F1000Res.* 2024;13:468. doi:10.12688/f1000research.147114.2
 73. Nguinkal JA, Zoclanclounon YAB, Molina A, et al. Assessment of the pathogen genomics landscape highlights disparities and challenges for effective AMR surveillance and outbreak response in the East African community. *BMC Public Health.* 2024;24:1500. doi:10.1186/s12889-024-18990-0
 74. Laamarti M, El Fathi Lalaoui Y, Elfermi R, Daoud R, El Allali A. Afro-TB dataset as a large scale genomic data of *Mycobacterium tuberculosis* in Africa. *Sci Data.* 2023;10:212. doi:10.1038/s41597-023-02112-3
 75. African Union. Continental artificial intelligence strategy harnessing AI for Africa's development and prosperity. Published July 2024. Accessed November 19, 2025. https://au.int/sites/default/files/documents/44004-doc-EN-Continental_AI_Strategy_July_2024.pdf
 76. The Centre for Intellectual Property and Information Technology Law (CIPIT). The state of AI in Africa report. Published 2025. Accessed November 19, 2025. <https://aiconference.cipit.org/documents/the-state-of-ai-in-africa-report.pdf>
 77. African Union. *AUDA-NEPAD Champions AI Integration in Healthcare Regulation at 4th African Medicines Regulatory Harmonisation Week*. African Union; 2024.
 78. Olufadewa II, Iyiola OP, Nnatus J, et al. National eHealth strategy frameworks in Africa: a comprehensive assessment using the WHO-ITU eHealth strategy toolkit and FAIR guidelines. *Oxf Open Digit Health.* 2024;2:oqae047. doi:10.1093/oodh/oqae047
 79. Ministry of Communications and Digitalisation with Smart Africa. Republic of Ghana national artificial intelligence strategy: 2023-2033. Published October 2022. Accessed

- November 19, 2025. <https://www.africadataprotection.org/Ghana-AI-Strat.pdf>
80. ICTworks. Introducing the National Artificial Intelligence Policy for Rwanda. Published December 20, 2023. Accessed November 19, 2025. <https://www.ictworks.org/national-artificial-intelligence-policy-rwanda/>
 81. Mwogosi A. Revolutionizing primary health care in Tanzania: unravelling the contextual factors on electronic health record systems implementation for effective decision support [published online ahead of print August 5, 2024]. *J Sci Technol Policy Manag*. doi:10.1108/JSTPM-11-2023-0205
 82. Ddamba A, Nsubuga B, Kamabare M, et al. Factors influencing the availability and use of electronic medical records systems in public health facilities in Uganda: a cross-sectional assessment. *BMC Med Inform Decis Mak*. 2025;25:372. doi:10.1186/s12911-025-03190-6
 83. Akanbi MO, Ocheke AN, Agaba PA, et al. Use of electronic health records in Sub-Saharan Africa: progress and challenges. *J Med Trop*. 2012;14:1-6.
 84. WHO. *Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2021*. World Health Organization; 2021.
 85. Joseph M. Challenges of educational digital infrastructure in Africa: a tale of hope and disillusionment. *J Afr Stud Dev*. 2019;11:59-63. doi:10.5897/jasdev2019.0539
 86. Navaux POA, Lorenzon AF, Serpa M, da S. Challenges in high-performance computing. *J Braz Comput Soc*. 2023;29:51-62. doi:10.5753/jbcs.2023.2219
 87. *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*. World Health Organization; 2021.
 88. Dhali A. Using artificial intelligence in healthcare—some ethical and legal considerations. *S Afr J Bioeth Law*. 2024;17:2. Accessed September 13, 2025. https://hdl.handle.net/10520/ejc-m_sajbl_v17_n3_a2
 89. Solaiman B. Legal and ethical considerations of artificial intelligence for residents in post-acute and long-term care. *J Am Med Dir Assoc*. 2024;25:105105. doi:10.1016/j.jamda.2024.105105
 90. Ratti E, Morrison M, Jakab I. Ethical and social considerations of applying artificial intelligence in healthcare—a two-pronged scoping review. *BMC Med Ethics*. 2025;26:68. doi:10.1186/s12910-025-01198-1
 91. Fatumo S, Inouye M. African genomes hold the key to accurate genetic risk prediction. *Nat Hum Behav*. 2023;7:295-296. doi:10.1038/s41562-023-01549-1
 92. Taddese AA, Addis AC, Tam BT. Data stewardship and curation practices in AI-based genomics and automated microscopy image analysis for high-throughput screening studies: promoting robust and ethical AI applications. *Hum Genomics*. 2025;19:16. doi:10.1186/s40246-025-00716-x
 93. Chikowore T, Läll K, Micklesfield LK, et al. Variability of polygenic prediction for body mass index in Africa. *Genome Med*. 2024;16:74. doi:10.1186/s13073-024-01348-x
 94. Akinwalere SN, Ivanov V. Artificial intelligence in higher education: challenges and opportunities. *Border Cross*. 2022;12:1-15. doi:10.33182/bc.v12i1.2015