# KENYA INSTITUTE OF PRIMATE RESEARCH

## Data Science and Analytics Section (DS&AS)

## Strategic Framework and Policy Proposal

September 22, 2025

# Contents

# 1   Introduction and context

Modern biomedical and conservation research increasingly relies on advanced data science. Globally, data-driven approaches are accelerating discoveries, improving health outcomes, and informing policy. For KIPRE—a leading primate research and conservation institute—a dedicated **Data Science and Analytics Section (DS&AS)** will provide the infrastructure and analytical expertise required for evidence-based research, decision-making, and innovation.

KIPRE's 2023–2027 Strategic Plan is aligned with Kenya Vision 2030, the Sustainable Development Goals (SDGs), and Africa Agenda 2063. Within this framework, the DS&AS will support cross-cutting initiatives such as *One Health*, recognising the interdependence of human, animal, and environmental health. Achieving this requires integrated, interoperable data systems capable of supporting disease surveillance, ecological monitoring, and rapid response to emerging threats.

By adhering to international best practices—such as FAIR data principles and open science—and national legislation including Kenya's *Data Protection Act* 2019 (modelled on GDPR), the DS&AS will treat KIPRE's data as a valuable, ethically managed resource. This approach ensures that data are not only secure and compliant but also actively leveraged to maximise research impact, inform policy, and strengthen KIPRE's role as a regional centre of excellence.

# 2   Problem statement and rationale

**Fragmented and inconsistent data governance.** KIPRE collects extensive biomedical, ecological, and primatological data —ranging from zoonotic-disease surveillance to climate–biodiversity studies— but management remains siloed by individual projects. The absence of a central platform for storing, integrating, and curating these datasets leads to duplication of effort, poor data quality, and limited accessibility for national and global stakeholders. Critical resources such as the national non-human primate database cannot achieve their full value without unified governance.

**Gaps in advanced analytics and decision support.** Many of the Directorate's core functions require high-level quantitative expertise: real-time outbreak detection and early-warning systems, spatial modelling of snakebite trends, bioinformatics pipelines for genomics, and predictive models for climate–biodiversity interactions. Without a dedicated Data Science and Analytics Section (DS&AS), KIPRE lacks the statistical, computational, and machine-learning capacity to implement these functions at scale.

**Regulatory, ethical, and reproducibility risks.** Kenya's *Data Protection Act* (2019) demands rigorous safeguards for sensitive human, animal, and environmental data. Fragmented systems complicate compliance and compromise transparent, reproducible analyses as required by international standards such as those of the American Statistical Association and the Royal

Statistical Society.

**Strategic opportunity for evidence-driven leadership.** A unified DS&AS will provide real-time analytics and interactive dashboards to guide resource allocation, policy formulation, and programme evaluation across basic, pre-clinical, and translational research. Centralising data integration and advanced modelling will strengthen One-Health surveillance, accelerate discovery, and enhance KIPRE's national and global leadership in biomedical research and primate conservation.

# 3   Objectives of the DS&AS

The DS&AS will serve as KIPRE's central framework for responsible and effective management and use of biomedical, ecological and primatological data. Its objectives are to:

- **Establish robust data governance and management**
  Formulate and enforce institute-wide policies, metadata standards and secure, interoperable repositories. Support the national non-human primate database by integrating census, geospatial and health-surveillance data, and ensure compliance with Kenya's *Data Protection Act* 2019 and FAIR/open-data principles.

- **Advance analytics, modelling and bioinformatics**
  Apply advanced statistics, machine learning and AI across laboratory, field and clinical datasets. Develop predictive epidemiological and ecological models, early-warning systems for emerging/re-emerging diseases, and bioinformatics pipelines for genomics and drug/vaccine discovery.

- **Enhance surveillance and real-time data integration**
  Design and maintain integrated One-Health platforms that merge human, animal and environmental data for rapid outbreak detection, zoonotic disease tracking, snakebite trend analysis, and climate–biodiversity impact monitoring.

- **Support research impact and performance monitoring**
  Provide key performance indicators (KPIs) and interactive dashboards to track research outputs, funding flows and programme implementation, enabling evidence-driven leadership and efficient resource allocation.

- **Build institutional capacity in data science**
  Offer continuous training and mentorship in statistical computing (R, SAS, Python), spatial analysis, bioinformatics and AI, fostering a culture of reproducible, data-driven decision-making across all directorates.

- **Promote ethical analytics and data security**
  Implement rigorous ethical standards and robust cybersecurity practices to protect sensitive

human, animal and environmental data, while ensuring transparent, reproducible analytical workflows.

- **Translate analytics into policy and innovation**
  Convert data insights into actionable decision-support tools for conservation, health policy, and innovation incubation. Collaborate with government agencies, academic partners and international bodies to maximise national and global impact.

## 4   Scope and core functions

To deliver on its objectives and complement the Directorate of Training and Product Development, the DS&AS will perform the following core functions:

- **Data management and governance**
  Design institute-wide SOPs for data collection, curation and archiving across biomedical, ecological and primatology programmes. Maintain interoperable metadata standards and secure data lakes. Support the national non-human primate database (census, geo-spatial distribution, research affiliations) and ensure that all datasets are FAIR (findable, accessible, interoperable, reusable) and compliant with Kenya's *Data Protection Act* 2019.

- **Advanced analytics, modelling and bioinformatics**
  Develop and apply statistical, machine-learning and AI models to laboratory, field, clinical and genomic data. Build predictive tools for early-warning systems on emerging and re-emerging diseases, spatial models for snakebite trends, and bioinformatics pipelines for drug and vaccine discovery. Provide expert guidance on study design, sampling and meta-analysis in support of pre-clinical, translational and conservation research.

- **Integrated platforms and real-time surveillance**
  Build centralised analytics platforms with dashboards and geospatial mapping tools to integrate human, animal and environmental data. Enable rapid detection of zoonotic outbreaks, biodiversity shifts and climate–biodiversity impacts, thereby strengthening One-Health surveillance.

- **Methodological consultation and research design**
  Advise research teams on experimental design, statistical methods and reproducible workflows. Review analysis plans to ensure they meet ethical, technical and international best-practice standards, and perform rigorous quality checks on data and code.

- **Performance monitoring and evaluation**
  Implement KPIs and automated dashboards to measure research outputs, resource utilisation and impact of Directorate programmes, including early-warning systems, diagnostic development and primate-conservation studies.

- **Capacity building and training**

Lead internal workshops on machine learning, R/Python programming, spatial analytics and data visualisation. Provide training in bioinformatics and AI relevant to drug discovery, vaccine development and primate ecology, in partnership with universities and accredited online platforms.

- **Knowledge dissemination and reporting**
  Produce clear reports, visualisations and interactive dashboards for management, government agencies and international partners. Facilitate timely dissemination of findings from surveillance, bioinformatics and climate-biodiversity studies.

- **Innovation and strategic collaboration**
  Serve as KIPRE's hub for data-science innovation by piloting new techniques (e.g. AI for image analysis in wildlife ecology, advanced molecular design modelling). Build partnerships with universities, government agencies and international consortia to support incubation and acceleration of data-driven health and conservation solutions.

## 5 Organizational structure and leadership

The DS&AS is a specialised section within the **Directorate of Research and Product Development**, ensuring full alignment with the Directorate's scientific strategy and operational priorities. This placement enables direct collaboration with all research units and seamless integration of data-driven decision-making into KIPRE's core research agenda.

### Head of DS&AS

A senior data science professional will head the section and serve as part of the Directorate's leadership team. Key responsibilities include setting institute-wide technical standards, approving study designs and analysis plans, overseeing data-intensive projects, and engaging external partners. The Head will coordinate closely with the Data Protection Officer and report performance metrics directly to the **Director of Research and Product Development**, ensuring that DS&AS initiatives remain consistent with Directorate objectives.

### Team composition

Reporting to the Head, the DS&AS team will comprise:

- **Senior Data Scientist** – Leads complex analytics, advanced modelling, and machine-learning projects aligned with Directorate research goals.

- **Biostatistician** – Provides study design and inferential analysis to support biomedical and ecological research within the Directorate.

- **Data Engineer** – Develops and maintains secure databases, data pipelines, and metadata catalogues that integrate with Directorate systems.

- **Research Data Analysts (1–2)** – Perform data cleaning, visualisation, and dashboard development for Directorate projects.

- **Training and Outreach Officer** – Coordinates internal workshops, documentation, and external stakeholder engagement in support of Directorate priorities.

## Reporting lines

The Head of DS&AS reports directly to the **Director of Research and Product Development**. All section staff report to the Head. For cross-institutional governance and compliance, the Head collaborates with other Directorate section leads and the Data Protection Officer.

# 6   Resources and capacity building

## Human resources

All positions require relevant qualifications and experience in research settings to ensure a strong blend of technical and domain expertise aligned with the Directorate's research agenda.

## Technical infrastructure

The section will procure both open-source and licensed tools: R and Python for statistical computing, SAS for specialised analytics, and ArcGIS/QGIS for spatial analysis. A secure relational database system (PostgreSQL or MySQL), high-performance servers (on-premise or cloud), and redundant storage will support large-scale data processing for Directorate projects. Robust backup routines and cybersecurity controls will protect sensitive data in line with institutional and national policies.

## Cloud and collaboration services

Cloud computing will provide scalable analytics and storage capacity for Directorate-wide research initiatives. Collaborative repositories and real-time analytics platforms will be established for authorised KIPRE staff and partners, ensuring full compliance with Kenya's privacy legislation and the Directorate's IT governance.

## Capacity-building plan

To strengthen data science capacity within the Directorate of Research and Product Development, the DS&AS will:

- Publish an annual training schedule covering advanced analytics, machine learning, and data governance.

- Sponsor staff attendance at external courses, conferences, and professional certification programmes.

- Host targeted workshops and mentorship sessions led by invited experts.

- Maintain an internal knowledge portal containing best-practice templates, code libraries, and analysis protocols for Directorate researchers.

**Budget and partnerships**

Funding will be drawn from KIPRE's core allocation to the **Directorate of Research and Product Development** and from competitive grants secured by the Directorate. Strategic collaborations with universities, government research agencies, and international partners will share costs, contribute technical expertise, and ensure that DS&AS activities complement broader national and global research initiatives led by the Directorate.

# 7 Implementation plan

The DS&AS will be established in phases under the oversight of the **Directorate of Research and Product Development**:

1. Draft a DS&AS policy detailing goals, scope, and procedures. Integrate KIPRE's Strategic Plan priorities, Kenya's *Data Protection Act* 2019, and international data-governance guidelines.

2. Present the policy draft and resource requirements to the Directorate's leadership and the KIPRE Board for review and approval. Secure budgetary support for staffing, infrastructure, and software procurement.

3. Appoint the Head of DS&AS and competitively recruit the core team in consultation with the Directorate's HR and scientific leads. Provide structured onboarding within the Directorate framework.

4. Procure and install required hardware and software through the Directorate's procurement channels. Set up secure servers, relational databases, and backup systems with Directorate IT support. Configure user-access controls and network security in compliance with institutional policies.

5. Develop standard operating procedures, data-governance manuals, and analysis-plan templates consistent with Directorate-wide practices. Establish metadata standards and reporting formats. Conduct training sessions for researchers across the Directorate.

6. Launch pilot projects demonstrating the unit's value (e.g. disease-surveillance dashboards or integrated wildlife–health datasets) that directly support Directorate research priorities. Use pilot results to refine workflows and tools.

7. Scale DS&AS services to all research sections under the Directorate, incorporating user feedback and updating training and policies as the unit matures.

The Head of DS&AS will coordinate each phase while maintaining close engagement with the **Director of Research and Product Development** and other key stakeholders.

# 8   Monitoring and evaluation

A robust M&E framework—embedded within the Directorate's performance management system—will track DS&AS outputs and impact.

## Performance metrics

Key indicators will include:

- **Analytical output:** number and complexity of analyses and models delivered to Directorate research teams.

- **Data quality:** measurable improvements in data completeness, consistency, and standardisation across Directorate projects.

- **Efficiency gains:** reductions in project turnaround times and resource utilisation.

- **Research impact:** publications, policy briefs, and grant proposals citing DS&AS support within Directorate programmes.

- **User satisfaction:** survey feedback from researchers, Directorate section heads, and external partners.

- **Capacity metrics:** staff trained, workshops held, and professional certifications achieved within the Directorate.

- **Strategic alignment:** demonstrated linkage of DS&AS outputs to KIPRE's strategic objectives and national health and conservation goals.

## Reporting and review

The Head of DS&AS will submit an annual report to the **Director of Research and Product Development** summarising M&E findings, achievements, and challenges. Mid-year reviews and stakeholder consultations within the Directorate will capture lessons learned and inform adjustments.

## Continuous improvement

M&E results will guide iterative enhancements. If certain metrics lag, processes or tools will be revised in consultation with the Directorate. This feedback loop ensures that DS&AS continually improves its services and remains fully aligned with both Directorate and institutional best practices.

# 9   Qualifications of the proposed DS&AS lead

The Lead of the DS&AS will report directly to the **Director of Research and Product Development**, ensuring that all activities align with the Directorate's scientific and strategic priorities. The current KIPRE career guidelines do not explicitly define the position, job group, reporting structure, or responsibilities for this role. Given that the DS&AS is a full section within the Directorate, it is appropriate to designate the Lead as an **Assistant Director**.

**Required qualifications** should therefore mirror those of other Assistant Directors within the Directorate, with the following additional expectations:

- Advanced degree (MSc or PhD) in data science, biostatistics, bioinformatics, or a closely related field.

- Demonstrated expertise in advanced analytics, statistical modelling, machine learning, and data-governance frameworks.

- Experience managing multidisciplinary research projects and leading high-performing technical teams.

- Familiarity with Kenya's *Data Protection Act* 2019 and international standards for data security, ethics, and reproducibility.

- Strong collaboration skills for engaging with other KIPRE directorates, national agencies, and international research partners.

# 10   Conclusion and recommendations

Establishing the DS&AS as a section under the **Directorate of Research and Product Development** will transform KIPRE into a truly data-driven institution. With robust analytics and governance, leadership will gain a comprehensive, real-time view of research activities to inform strategy and policy.

A well-supported DS&AS will:

- Implement strong data governance and protection systems aligned with Kenya's *Data Protection Act* 2019 and global norms.

- Deliver advanced analytics, modelling, and visualisation capabilities in support of the Directorate's biomedical, ecological, and conservation mandates.

- Build internal expertise through continuous training in modern data-science tools and methods, strengthening capacity across all research sections.

- Foster collaboration across directorates and with external partners, amplifying research impact and innovation.

**Recommendations**

1. Develop and adopt a DS&AS Policy Framework that embeds the Section within the **Directorate of Research and Product Development** and defines its mandate, reporting lines, and resource needs.

2. Allocate a dedicated budget for DS&AS staffing, software, hardware, and capacity-building activities as part of the Directorate's annual planning cycle.

By implementing these recommendations, KIPRE will align with international best practices, safeguard the integrity of its data assets, and position the DS&AS as a regional centre of excellence for data-driven biomedical and conservation research.