|  |
| --- |
| **Title** |
| **National Data Science Network for Health, Agriculture, Environment and Education: A Flemish–Kenyan Partnership for Evidence-Driven Public Policy.**  ***Lead institution (Global south): Kenya Institute of Primate Research (KIPRE), Kenya***  ***Flemish partner Institution: Hasselt University (UHasselt)***  ***Duration: 5 years  Indicative budget: €300,000*** |
| **Context analysis** |
| ****Sustainable Development Context: Scope, Causes, and Interlinkages**** The One Health approach, introduced in 2004, highlights the interconnectedness of human, animal, and environmental health. While globally recognized as crucial for disease prevention, outbreak preparedness, and ecosystem sustainability, its implementation remains a persistent challenge. In Kenya, institutions across health, agriculture, livestock, environment, and education continue to collect and archive critical datasets independently, resulting in fragmentation, lack of interoperability, and underutilization of valuable information.  This fragmentation limits the ability of policymakers to make timely, evidence-based decisions during epidemics, zoonotic outbreaks, or other public health emergencies. Public health, agricultural, and environmental interventions often lack predictive insights due to insufficient data integration and advanced analytical capacity. Researchers and government agencies face obstacles in conducting complex analyses, simulations, and early warning modeling because of gaps in infrastructure, data governance, and skilled personnel.  The root causes of this problem include the absence of a national framework for centralized data management, limited human capacity in biostatistics, bioinformatics, epidemiology, and data science, siloed and non-standardized datasets, and weak governance systems for data curation, sharing, and ethical use. These constraints hinder cross-sectoral collaboration, knowledge sharing, and generation of policy-relevant evidence, reducing Kenya’s preparedness for pandemics and other disease outbreaks.  This project aligns with Agenda 2030, contributing to SDG 3 on Good Health and Well-Being by enhancing disease surveillance, early warning, predictive analytics, and evidence-based interventions. It supports SDG 9 on Industry, Innovation, and Infrastructure by establishing a digital, interoperable infrastructure for national data integration. It also advances SDG 17 on Partnerships for the Goals by fostering multi-institutional, multi-stakeholder collaborations across government ministries, research institutions, and international partners. ****Partner Institutional Context: Strengths, Weaknesses, and Added Value**** Kenya’s national institutions bring significant strengths to this initiative. The Ministry of Health (via KIPRE and NPHI) provides expertise in epidemiology, biomedical research, and public health surveillance. The Ministry of Agriculture and Livestock development (through DVS) contributes data on zoonotic diseases, livestock health, and environmental interactions. The Ministry of Environment, climate change and Forestry (via NEMA) provides environmental monitoring and ecological datasets, while the Ministry of Education, through CEMA, supports capacity-building initiatives and data literacy.  Despite these strengths, institutional capacity for advanced analytics, data integration, and predictive modeling remains limited. Datasets are fragmented, standards for interoperability are inconsistent, and human capacity for multi-sectoral data science is insufficient. Coordination between sectors is ad hoc, limiting the adoption of integrated, evidence-based policy approaches.  This project will establish a **National Data Science Network** through a centralized Data Analytics Centre. The Centre will provide modern digital infrastructure, standardize and integrate datasets, and develop high-level competencies among data scientists across these ministries. Flemish expertise from UHasselt will support training, mentorship, and governance frameworks, ensuring that analytical outputs inform national and regional policy decisions. By connecting previously siloed institutions into a coherent network, the project will foster cross-sectoral collaboration, strengthen regional benchmarking, and embed a culture of evidence-based policymaking. ****Organizational Capacity-Building Context**** Current capacity constraints include limited exposure to advanced analytics, high-performance computing, integrated data pipelines, and standardized governance procedures. Many agencies lack personnel trained in multi-sectoral BEET analytics and opportunities for structured mentorship and international exchange are limited.  The project will address these gaps by providing structured mentorship, hands-on training, and academic exchanges through the Data Analytics Centre. It will develop a national cadre of data scientists capable of integrating, analyzing, and interpreting BEET datasets across ministries. The Centre will ensure long-term sustainability by embedding best practices in data governance, ethical use, and analytical workflows, enabling evidence-driven interventions, early warnings, and preparedness strategies for health, agricultural, and environmental challenges. |
| **Impact statement, ToC (Theory of Change) and Strategy** |
| 1. Impact Statement (The dream), primary beneficiaries, where change happens 2. Areas of change to realise the dream (Why?), institutional capacity, human capacity and skills, integrated and interoperatble data systems, strong data governance and ethical frameworks, improve research quality and uptake outputs, sustainable multi stakeholder partnerships. 3. Priority key areas of change the project can realistically influence (What?), functional DSAC, human capacity, integrated, interoprable data systems, improved dsac generated evidence, rationale for prioritisation 4. How the project will deliver change ( operational and methodological approaches)- Education, research, outreach and dissemination/uptake strategy, service to society, infrastructure and institutional strengthening, multi stakeholder partnership and networking   Integration of gender and environment   1. Coherence check (Summary narrative)   The **impact statement** of this initiative envisions the establishment of a **national network of data scientists** in Kenya capable of integrating and analyzing data across the biomedical, epidemiological, ecological, and translational life sciences. This network will serve as a cornerstone for evidence-driven public policy, enhancing Kenya’s capacity for early warning, preparedness, and response to disease outbreaks, pandemics, and other health, agricultural, and environmental challenges. The primary beneficiaries of the project are the national institutions generating and using data: the Ministry of Health (KIPRE, NPHI), the Ministry of Agriculture (DVS), the Ministry of Environment (NEMA), and the Ministry of Education through CEMA. Indirect beneficiaries include policymakers, researchers, underserved communities, and the general public, who will gain from improved decision-making, timely interventions, and evidence-based resource allocation. The change will occur nationally, with hubs in participating ministries and research institutions forming the backbone of the network, while outreach and training activities will extend impact regionally and nationally.  To realize this dream, several **areas of change** need to be addressed. Institutional capacity must be strengthened through the creation of a functional Data Science and Analytics Centre (DSAC) that serves as a national hub for data integration, management, and analysis. Human capacity and skills development are crucial; this includes equipping data scientists and analysts with advanced competencies in biostatistics, bioinformatics, epidemiology, data management, and predictive modeling. Integrated and interoperable data systems are essential for consolidating fragmented datasets across ministries and research institutions, ensuring that analyses are consistent, reproducible, and timely. Strong data governance and ethical frameworks are necessary to guide responsible use, sharing, and protection of sensitive data. Improved research quality and enhanced uptake of DSAC-generated evidence will ensure that outputs translate into actionable policies and interventions. Finally, sustainable multi-stakeholder partnerships are required to facilitate collaboration, knowledge exchange, and long-term network sustainability across government, research institutions, and international partners.  Within these areas, the **priority key areas of change** that the project can realistically influence include establishing a fully functional DSAC, building human capacity, integrating and standardizing data systems, and improving the generation and use of evidence from the DSAC. These priorities were selected based on feasibility within the project’s five-year horizon, alignment with available expertise and resources, and the potential to deliver high-value outcomes. By focusing on these areas, the project can generate tangible results that form the foundation for a sustainable, nationwide data science ecosystem while leaving room for continued expansion and innovation beyond the project’s lifecycle.  The project will **deliver change** through operational and methodological approaches structured around six interlinked domains. In the domain of **education**, the project will implement structured training programs, mentorship, and international academic exchanges to build the next generation of Kenyan data scientists. Under **research**, the DSAC will enable collaborative studies applying advanced analytics, predictive modeling, and One Health approaches to priority challenges in health, agriculture, environment, and education. **Outreach and dissemination** strategies will include policy briefs, dashboards, workshops, and stakeholder forums to ensure evidence generated by the DSAC informs decision-making. Through **service to society**, data-driven interventions will be translated into actionable public health, environmental, and agricultural policies. In **infrastructure and institutional strengthening**, the DSAC will provide secure, interoperable digital infrastructure, standardized data repositories, and governance frameworks. Finally, **multi-stakeholder partnerships and networking** will facilitate collaboration between national institutions, regional partners, and UHasselt, strengthening North-South and South-South exchange and ensuring long-term sustainability.  Integration of **gender and environmental considerations** is central to the project. All training, leadership, and decision-making opportunities within the network will adopt a gender-sensitive approach, ensuring equitable access for women and underrepresented groups. Environmental data will be incorporated into analyses and modeling, enabling policies that balance human health, livestock, agricultural productivity, and ecological sustainability. The project’s holistic approach captures co-benefits and trade-offs across sectors, supporting evidence-based decision-making that accounts for socio-economic and environmental interlinkages.  As a summary, this project establishes a coherent, practical, and ambitious Theory of Change. By addressing fragmented datasets, building institutional and human capacity, and generating actionable evidence, the initiative will create a sustainable national network of data scientists. This network will strengthen the integration of health, agricultural, environmental, and educational datasets, support One Health implementation, and enable evidence-driven policies for early warning, preparedness, and interventions. The project aligns with SDG 3 (Good Health and Well-being), SDG 9 (Industry, Innovation, and Infrastructure), and SDG 17 (Partnerships for the Goals), ensuring inclusivity, interoperability, and sustainability. Its design balances feasibility, impact, and long-term vision, setting the stage for a transformative contribution to public policy and national resilience in Kenya. |
| **Organisations** |
| The implementation of the National Data Science Network will be overseen through a **multi-institutional and multi-stakeholder organizational structure**, ensuring balanced participation of all partner institutions and enabling strong collaboration between Flemish and Kenyan partners. The project will be led by **KIPRE**, serving as the national coordination hub, while **Hasselt University (UHasselt)** provides technical guidance, mentorship, and capacity-building expertise. Other Kenyan institutions involved include the National Public Health Institute (NPHI), the Directorate of Veterinary Services (DVS), Wildlife Research institutions (WRTI or KWS), the National Environment Management Authority (NEMA), and the Centre for Mathematics and Applied Statistics (CEMA) under the Ministry of Education. These partners will provide domain-specific expertise, access to datasets, and engagement with policy stakeholders to ensure the network produces actionable insights.  The **core project team** will include a project director at KIPRE, responsible for overall project oversight, coordination, and strategic engagement with ministries and partners. A deputy director will manage day-to-day operations and liaise with the Flemish promoter and other partner institutions. Each partner institution will nominate a **focal point**, who will coordinate data contributions, supervise local staff, and ensure adherence to data governance and ethical standards. UHasselt will designate a **technical lead** responsible for designing training curricula, mentoring local data scientists, and supporting the implementation of interoperable data systems. Supporting staff will include data engineers, analysts, IT specialists, and administrative personnel responsible for communications, reporting, and logistical management.  The **project’s organisational structure** will follow a hub-and-spoke model. KIPRE will serve as the central hub for coordination, data integration, and monitoring of DSAC activities. Spokes will represent the partner institutions, each contributing specialized data, technical knowledge, and human resources. Decision-making will be collaborative, with regular steering committee meetings involving all focal points and the Flemish promoter to review progress, address challenges, and guide strategic priorities. An **advisory board** including representatives from government ministries, civil society, and gender and equity experts will ensure that perspectives of underserved and marginalized groups are incorporated into decision-making processes.  **Roles and responsibilities** are defined to promote shared ownership and balanced partnerships. KIPRE will coordinate the network, manage infrastructure, and ensure alignment with national policy objectives. UHasselt will provide capacity-building, technical guidance, and co-supervise the development of research outputs. Partner ministries and institutions will contribute domain-specific datasets, facilitate policy engagement, and ensure sustainability of training outcomes. The DSAC focal points will oversee local implementation, including mentoring junior staff, managing workflows, and monitoring compliance with ethical and governance standards. Gender and environmental experts will ensure that project outputs and processes respect inclusivity, leave no one behind, and integrate environmental considerations.  The **selection of participants** for training and network activities will be based on predefined criteria, prioritizing diversity, representation from all relevant institutions, technical aptitude, and inclusion of women and other underrepresented groups. Recruitment will follow a transparent and competitive process, guided by both LNOB and gender principles. Capacity-building initiatives will combine workshops, mentorship, and hands-on projects, enabling participants to gain practical skills in data integration, predictive modeling, and policy translation.  This organisational model ensures **interconnectedness and interdisciplinarity**, bringing together data scientists, epidemiologists, ecologists, and statisticians to address challenges across biomedical, agricultural, environmental, and educational domains. Multi-institutional and multi-stakeholder partnerships strengthen the project’s legitimacy, enhance resource sharing, and facilitate uptake of DSAC outputs into policy and practice. The structure provides mechanisms for continuous communication, decision-making, and accountability, ensuring that all partners have voice and influence in project implementation. Ultimately, this framework guarantees that the network is inclusive, collaborative, and sustainable, while supporting evidence-driven interventions for national health, agriculture, environment, and education policy. |
| **Stakeholders and coherence** |
| The success of the National Data Science Network hinges on engaging a diverse set of stakeholders who are directly or indirectly involved in the generation, management, and use of biomedical, epidemiological, ecological, and translational life sciences data. Internal stakeholders include staff and researchers within the partner institutions such as KIPRE, NPHI, DVS, NEMA, and CEMA, who will serve as focal points for data integration, analytics, and training. These individuals will be actively involved in implementing the project, contributing technical expertise, coordinating data workflows, and participating in structured capacity-building activities. Trainees selected for specialized courses in biostatistics, bioinformatics, epidemiology, and data science will also be internal stakeholders, as they will directly benefit from skill development and mentoring, while contributing to the functional network of data scientists.  External stakeholders comprise national policymakers in the Ministries of Health, Agriculture, Livestock, Environment, and Education, who will benefit from evidence-based insights generated by the network to inform pandemic preparedness, early warning systems, and intervention strategies. Other external stakeholders include civil society organizations, local communities, and research users who depend on actionable data for decision-making. Private sector entities, including technology and analytics service providers, may provide complementary expertise, infrastructure, or platforms to support data integration and dissemination. Belgian actors, primarily UHasselt and other Flemish universities, will contribute technical guidance, mentorship, and co-supervise research activities, ensuring North-South knowledge transfer and sustainable capacity building.  Stakeholder engagement will be multi-tiered and tailored according to influence and interest. Critical stakeholders—such as national policymakers, ministry representatives, and institutional focal points—will be engaged closely through advisory boards, joint decision-making processes, and continuous dialogue to ensure alignment with national priorities. High-interest but lower-influence stakeholders, including junior researchers and civil society representatives, will be informed and consulted through workshops, training sessions, and participatory platforms, building their capacity to influence the network effectively over time. Lower-priority stakeholders will be monitored and periodically updated to maintain awareness and identify potential opportunities for engagement. The network will employ participatory methods, community-based research, and co-creation of outputs to ensure inclusivity, transparency, and responsiveness to stakeholder needs.  To facilitate uptake of results, the project will implement a continuous engagement strategy rather than one-off dissemination. Knowledge will be communicated through policy briefs, webinars, workshops, interactive dashboards, and targeted training modules, ensuring accessibility to non-specialist audiences. Stakeholders will be supported in using the outputs for decision-making, planning, and interventions, thereby strengthening evidence-based policy development. Trust, mutual respect, and long-term interaction will underpin relationships, with mechanisms for feedback and adaptation as stakeholder priorities evolve. Gender-sensitive approaches will ensure equitable participation, while specific attention will be given to including underserved communities and marginalized groups in consultation and decision-making processes.  External coherence will be ensured by mapping the project against existing Belgian and international initiatives in data science, public health, One Health, and environmental analytics. Opportunities for collaboration, synergy, and complementarity will be actively pursued with Belgian development actors, other VLIR-UOS projects, and regional networks, avoiding duplication and leveraging existing expertise. Internal coherence will be achieved through alignment with other VLIR-UOS programs (ITP, TEAM/SI, ICP) to maximize learning, resource sharing, and methodological harmonization across projects. By embedding stakeholder engagement within an integrated, multi-institutional framework, the project will foster sustained collaboration, institutional capacity building, and the creation of a national network of data scientists capable of generating high-quality, actionable evidence to inform public policy. |
| **Planning and Budgeting** |
| The planning and budgeting of the National Data Science Network is designed to ensure that all project activities are aligned with the overarching goals of building institutional and human capacity, establishing integrated and interoperable data systems, and enhancing evidence-based policy-making. Activities are structured around six core domains: capacity building, infrastructure and institutional strengthening, research and methodological development, multi-stakeholder engagement, knowledge dissemination and uptake, and project management. Each domain includes specific, time-bound activities, assigned responsibilities, and associated financial resources to guarantee smooth and efficient implementation.  Capacity-building activities will include structured training programs in biostatistics, bioinformatics, epidemiology, and data science, complemented by mentorship, exchange programs with UHasselt, and the creation of a national network of data scientists across key ministries and research institutions. Training sessions will be scheduled annually, with continuous follow-up through virtual platforms to ensure knowledge retention and application. Particular attention will be paid to gender balance, equitable inclusion of left-behind groups, and representation from underserved regions, with sufficient budget allocated for inclusive participation, such as childcare support or travel stipends.  Infrastructure and institutional strengthening will focus on establishing the Data Science and Analytics Centre (DSAC), including hardware, software, and secure cloud-based data management systems. This will enable the integration of fragmented biomedical, epidemiological, ecological, and translational datasets from partner institutions. Dedicated technical teams will oversee the setup, configuration, and maintenance of infrastructure, with clear responsibilities assigned to IT staff within KIPRE, NPHI, DVS, WRTI/KWS, NEMA, and CEMA, supported by technical guidance from UHasselt. Budget allocations reflect the high cost of hardware, software licenses, and infrastructure setup to ensure the DSAC is fully functional and sustainable.  Research and methodological development will involve collaborative projects leveraging the DSAC, focusing on predictive analytics, outbreak modelling, and One Health data integration. Teams from partner institutions will co-lead research initiatives, ensuring knowledge transfer, skill development, and generation of high-quality evidence. Multi-stakeholder engagement activities will ensure continuous dialogue with policymakers, civil society, and private sector actors to align research priorities with national development goals, facilitate adoption of results, and strengthen collaborative networks.  Knowledge dissemination and uptake will involve workshops, webinars, policy briefs, interactive dashboards, and participatory forums. Activities will be designed to translate complex data insights into actionable policy recommendations, with continuous feedback loops to enhance relevance. A portion of the budget is dedicated to communication and outreach to ensure broad accessibility and engagement, especially with marginalized groups and communities historically left behind.  Project management will include dedicated coordination, monitoring, and reporting teams to oversee the timely implementation of all activities, manage budgets, and ensure compliance with VLIR-UOS requirements. Clear accountability mechanisms, regular progress reporting, and adaptive management processes will maintain operational efficiency and alignment with the Theory of Change. Budgeting is designed to maximize value for money, with resources allocated proportionally across activities to ensure the DSAC establishment, capacity building, and research outputs receive sufficient funding while maintaining flexibility for emerging needs.  By integrating SDG principles, gender considerations, and equitable inclusion in all activities, the planning and budgeting framework ensures that the project contributes to sustainable institutional and human capacity, fosters multi-institutional collaboration, and strengthens evidence-based public policy for pandemic preparedness, environmental monitoring, and translational research in Kenya. |
| **Risk management** |
| The implementation of a national data science network for health, agriculture, environment, and education involves several uncertainties and potential risks, which need to be proactively identified and managed. Key risks include delays in establishing the Data Science and Analytics Centre (DSAC) due to procurement or logistical challenges, insufficient engagement or buy-in from partner institutions, and limited human capacity in data science across partner organizations. Additionally, risks may arise from fragmented or inaccessible datasets, cybersecurity threats, and potential misalignment between research outputs and policy needs. External risks include changes in government priorities, regulatory constraints, or funding delays, which could impact project continuity and outcomes.  The probability and impact of these risks vary. For instance, logistical delays or procurement issues are moderately probable but can have a high impact on project timelines if not addressed. Limited human capacity is highly probable initially but can be mitigated through structured training and mentorship programs. Fragmented data or interoperability challenges are likely given the historical context but can be reduced by establishing clear data governance frameworks, standardized procedures, and technical support. Policy misalignment or insufficient stakeholder engagement is moderate in probability but could significantly affect the uptake and impact of project outputs.  Risk mitigation strategies include proactive planning and scheduling for infrastructure and procurement activities, early and continuous stakeholder engagement to ensure ownership and commitment, and the establishment of a dedicated project management and technical support team to monitor progress. Standardized training programs, mentorship, and exchange visits with UHasselt will strengthen human capacity, while robust cybersecurity measures and data governance policies will protect data integrity and facilitate interoperability. Regular monitoring, reporting, and adaptive management will ensure that emerging risks are addressed promptly.  Responsibilities for risk monitoring will be clearly assigned: the project coordination team will oversee overall risk management, while each domain lead (capacity building, infrastructure, research, stakeholder engagement) will monitor domain-specific risks. Monthly progress reviews and quarterly risk assessment meetings will provide opportunities to update risk status, implement corrective actions, and document lessons learned. By adopting this proactive and structured approach, the project aims to minimize uncertainties and maximize the likelihood of achieving its objectives, ensuring sustainable institutional and national impact. |
| **Monitoring and evaluation** |
| Monitoring and reporting are critical components of the National Data Science Network, ensuring that progress towards the project’s goals is systematically tracked, lessons are learned, and outputs are effectively translated into actionable evidence. The project will monitor changes at multiple levels, including institutional capacity, human skills development, data system integration, and uptake of evidence by policy makers. Indicators will be selected to track both quantitative outputs—such as the number of trained data scientists, operational data pipelines, and datasets integrated into the Data Science and Analytics Centre (DSAC)—as well as qualitative outcomes, including stakeholder engagement, co-creation of evidence, and the influence of DSAC-generated insights on policy and practice.  The project will adopt both standard VLIR-UOS indicators for education, research, outreach, infrastructure, and partnership, as well as project-specific indicators tailored to measure progress in data governance, interoperability, and evidence-based decision-making. All indicators will follow the SMART and SPICED principles, ensuring they are specific, measurable, realistic, and meaningful for stakeholders. Special attention will be given to disaggregation by sex, stakeholder type, and underrepresented groups to ensure that the project actively addresses Leaving No One Behind (LNOB) principles.  Data collection will be continuous and integrated into project workflows. For instance, participant performance in training programs will be captured via structured assessments, while DSAC outputs will be documented through data logs, reports, and policy briefs. Stakeholder engagement will be monitored through regular surveys, focus group discussions, and consultation records, ensuring feedback loops inform the ongoing implementation of project activities. Responsibilities for monitoring will be assigned to dedicated project staff, with periodic reviews by the project management team to identify deviations, assess emerging risks, and adapt activities as needed.  To facilitate evidence uptake and impact, the project will establish a central reporting system within the DSAC, allowing real-time tracking of key performance indicators and generating insights for both internal decision-making and external communication with partner institutions, policy makers, and regional stakeholders. Annual reports will summarize progress, lessons learned, and changes implemented, providing accountability while supporting continuous learning. By embedding a robust monitoring and reporting framework, the project ensures that objectives are met efficiently, interventions remain relevant, and the long-term sustainability of the National Data Science Network is achieved. |
| **Learning and steering** |
| Learning and steering are central to ensuring that the National Data Science Network continuously adapts, improves, and maximizes its impact. From the outset, lessons from previous VLIR-UOS projects, as well as insights from local and regional data science initiatives, will inform the project design. This includes best practices for establishing multi-institutional networks, training programs, and data governance frameworks. Recommendations from selection committees and prior evaluations will be incorporated to strengthen approaches to capacity building, stakeholder engagement, and evidence translation, ensuring that the project addresses both known challenges and emerging opportunities.  During implementation, the project will establish a systematic learning framework to track the progress of the Theory of Change. Regular review sessions will identify which activities are producing intended outcomes, highlight effective strategies, and uncover areas needing adjustment. Peer-to-peer learning will be encouraged through interactions with other VLIR-UOS projects and regional data science initiatives, enabling the sharing of methodologies, data management practices, and lessons in policy uptake. Progress reports, stakeholder feedback, and interim evaluations will be used to refine activities, inform decision-making, and document emerging good practices. Mechanisms for capturing lessons learned will include structured reflections, knowledge repositories within the Data Science and Analytics Centre, and dissemination of practical guidance to participating institutions and partner networks.  The project will also ensure that learning processes are inclusive, reflecting the principles of Leaving No One Behind. Feedback from vulnerable and underrepresented groups, including women, minority communities, and less experienced researchers, will be systematically incorporated into monitoring and evaluation. Potential negative impacts will be regularly assessed and mitigated, with remedial actions taken promptly where needed. By embedding these participatory and reflective processes, the project will not only improve internal performance but also strengthen the capacity of partner institutions to manage evidence-based decision-making sustainably.  Interconnectedness and co-benefits will be actively monitored, ensuring that trade-offs between health, environment, education, and agriculture are identified and addressed. The project will create a dynamic learning loop, capturing unforeseen synergies and challenges across sectors, and adjusting strategies to maximize overall impact. Multi-institutional and multi-stakeholder partnerships will be continuously evaluated and, where necessary, strengthened or expanded. This adaptive approach ensures that the project remains responsive, relevant, and capable of achieving long-term systemic change in Kenya’s data science landscape for evidence-driven public policy. |

**Title:** **Project summary**

This initiative will strengthen biomedical research, conservation, and disease outbreak and pandemic preparedness at the Kenya Institute of Primate Research (KIPRE) through institutional capacity building and Flemish–Kenyan collaboration.

By establishing a functional Centre for Data Science and Analytics (DSAC), building human resource capacity, and implementing robust data governance, KIPRE will be able to collect, manage, and analyze high-quality datasets to guide public health policies and interventions locally and internationally.

Over the 3–5 years, the initiative will support applied data analytics in drug and vaccine development, epidemiology, tropical and neglected diseases, and non-communicable disease modelling. Key outputs include staff training, secure and integrated datasets, predictive models, dashboards, and joint research and training exchanges with Flemish institutions.

The initiative will ensure sustainability by embedding DSAC within KIPRE’s operations and fostering a culture of evidence-based decision-making. Ultimately, KIPRE will become a regional hub for biomedical data science, driving innovation and policy advisory in Kenya and beyond.

**Background**

The Kenya Institute of Primate Research (KIPRE) has been conducting pioneering evolutionary, biomedical, and translational research since 1958, contributing significantly to national, regional, and international health priorities. Over the decades, KIPRE has evolved into a regional center of excellence in pre-clinical and translational research, making major contributions in reproductive health and biology, tropical infectious and neglected diseases, and non-communicable diseases, while also advancing environmental and animal conservation through applied research and sustainable management of non-human primate (NHP) populations.

In recent years, KIPRE has broadened its scientific scope to include molecule discovery and drug development, particularly through bioprospecting partnerships with local indigenous communities. These collaborations focus on identifying, validating, and developing bioactive compounds of medicinal value, effectively linking traditional knowledge systems with modern biomedical innovation. This approach advances scientific discovery while promoting community empowerment and benefit sharing, aligning with the global call for inclusive and ethical research practices.

In 2023, through a Government reorganization Order, the institute, formerly known as the Institute of Primate Research (IPR) under the National Museums of Kenya (NMK), was delinked and reconstituted as a Semi-Autonomous Government Agency (SAGA), guided by legal notice 207 of 2018, under the Ministry of Health (MoH). Historically, IPR’s mandate centered on advancing human medicine through studies on NHP evolution and biomedical relevance. Under its new legal and institutional framework, KIPRE now operates within the Department of Public Health Policy and Professional Standards, with an enhanced mandate to generate policy-relevant scientific evidence and provide advisory services to the Ministry on matters of public health policy and professional standards.

This governance transformation has positioned KIPRE at a strategic interface between research, innovation, and policy, providing an opportunity to translate research evidence into actionable national, regional and potentially international health strategies. Its expanded mandate now covers a broad spectrum of translational science activities, including:

* Drug discovery and development, focusing on pre-clinical evaluation of vaccines and therapeutics for neglected and emerging diseases;
* Disease surveillance, mapping, and modelling within a One Health framework integrating human, animal, and environmental health;
* Conservation research on endangered animal species, particularly NHPs essential for biomedical studies; and
* Snakebite research and antivenoming development, addressing a neglected tropical health challenge in sub-Saharan Africa.

Despite this strong scientific and institutional foundation, KIPRE lacks a fully functional Data Science and Analytics Section, a critical gap that limits its ability to harness modern data-driven approaches such as biostatistics, bioinformatics, and machine learning. In the current era of data-intensive research and digital transformation, the absence of an integrated data science infrastructure constrains the institute’s ability to manage complex biomedical datasets, generate predictive insights, and provide high-quality evidence for policy formulation and implementation.

Establishing a Data Science and Analytics Capacity Building Unit (DSA-CBU) within KIPRE would therefore be **t**ransformative. The unit would enable the institute to systematically integrate advanced analytics, biostatistical modelling, and digital infrastructure into its core programs; enhance the reproducibility, transparency, and visibility of its outputs; and strengthen north–south and south–south collaborations aligned with international scientific and public health research standards**.**

Furthermore, this initiative directly supports the aspirations of Kenya’s Vision 2030, the Kenya Universal Health Policy 2020–2030, and the Ministry of Health Digital Health and Innovation Strategy, which all emphasize data governance, technology adoption, and human resource development as drivers of sustainable health outcomes.

Through this proposed investment in data science and analytics capacity strengthening, KIPRE seeks to build institutional resilience, cultivate advanced analytical and computational skills among its scientists, and contribute to a national ecosystem of evidence-based health research and innovation. This initiative also aligns with the VLIR-UOS TEAM 2026 focus on institutional capacity building, equitable partnerships, and locally relevant solutions, thereby positioning KIPRE as a strategic partner in driving data-enabled biomedical innovation in Kenya and the wider East African region.

**Partner institutions and their roles**

|  |  |
| --- | --- |
| **Partner/ Institution** | **Role** |
| **KIPRE** | 1. This will be the lead and project hosting institution 2. Will provide institutional leadership, project coordination, and overall implementation oversight. 3. Will establish and operationalize the DSAC including infrastructure, procurement and staffing. 4. Identify and enroll local scientists and technical staff for training and mentorship under the program. 5. Lead data collection, digitization, and integration of biomedical and environmental datasets. 6. Ensure ethical and regulatory compliance in all data related activities in line with Kenyan and International standards. 7. Translate research outputs into policy-relevant evidence for the MoH and other stakeholders. 8. Sustain project outcomes through institutional policy adoption,, resource allocation, and staff retention. |
| **Hasselt University(Uhasselt)** | 1. Will be the Flemish Partner and facilitate North-South collaborations. 2. Provide technical and academic mentorship in biostatistics, bioinformatics, quantitative epidemiology, machine learning, and AI-driven data analysis through joint training programs, workshops, and short-term exchange initiatives. 3. Co-design the data science and analytics curriculum, training modules, and research frameworks tailored to KIPRE research priorities. 4. Facilitate joint supervision of postgraduate trainees and short-term research fellows under the project. 5. Support development and implementation of data governance protocols, aligning KIPRE practices with FAIR and GDPR principles. 6. Collaborate in joint publications, grant applications, and dissemination of results at international scientific platforms. 7. Contribute to long-term sustainability through capacity transfer, mentorship networks and institutional linkages. |
| **Africa Collaborating Partners(to be confirmed)** | 1. KEMRI: To align project outputs with national health data systems and support integration into Kenya health research ecosystem. 2. APHRC: Provide guidance on ethical data governance, community engagement, and evidence translation. 3. Makerere University/ Ardhi university/University of Rwanda: (i).Participate in co-training and benchmarking regional biomedical data science initiatives. (ii). Facilitate regional knowledge sharing and south-south collaboration to enhance the visibility and sustainability of the project outcomes. |

**Objectives**

The main aim of this project/ initiative is to strengthen the institutional capacity of the KIPRE in data science and analytics to enhance multidisciplinary research, improve evidence generation, and promote the translation of scientific knowledge into public health policy and practice in Kenya and the wider region.

Specific objectives include:

1. Establish a functional DSAC unit at KIPRE.

This will be achieved through developing institutional structures, staffing, and infrastructure (Hardware, software, and secure data repositories) to support data driven research and decision making across all KIPRE programs.

1. Build Human Resource Capacity in data science and analytics

This will be achieved through training a core team of scientists, biostatisticians, bioinformatistician, epidemiologists and ICT officers in advanced data analysis, machine learning, bioinformatics, and geospatial analytics through targeted workshops, mentorships, and academic exchanges with University of Hasselt.

1. Digitize, curate and integrate research data.

This will be done through consolidating legacy datasets from biomedical, environmental, and primatological studies into standardize digital databases, enabling efficient data sharing, and reproducibility and cross-disciplinary analysis.

1. Develop and implement data governance and ethical frameworks.

Through development and establishing institutional policies and protocols on data collection, sharing and stewardship in compliance with FAIR data principles, research ethics and Kenya digital health strategies.

1. Apply data analytics to priority research area

This will be done to demonstrate the utility of data science approaches in selected flagship projects that may include disease modeling, drug discovery or one health surveillance, to generate actionable insights for the MoH and partners.

1. Strengthen collaborative networks and knowledge exchange

This will be achieved through fostering partnerships between KIPRE, Uhasselt and regional data science hubs to co-develop research tools, exchange expertise, and co-supervise trainees in data analytics.

**Expected Medium term outcomes**

* A fully operational DSAC unit at KIPRE supporting research work.
* At least 20 KIPRE staff trained in core and advanced data science skills, with sustainable mentorship program in place.
* Standardized digital databases integrating KIPRE research archives and new data streams for open and reproducible science.
* Operational data governance framework aligned to national and international ethical and legal standards.
* Pilot data driven studies completed and informing MoH policies or research interventions.
* Sustained institutional collaborations with Uhasselt and regional research institutions in East and Central Africa.

**Work Packages (Activities)**

**Management and Governance**

**Budget outline, overview and justification**

**Project timeline (5 years)**

**Sustainability**

**Concept note format**

**VLIR-UOS TEAMs Concept Note (2026 Call)**

**Project title: Strengthening the Centre for Data Science and Analytics for Research, Conservation, and Disease Surveillance: A Flemish-Kenyan Capacity-Building Initiative**

**Lead Institution (Global South): Kenya Institute of Primate Research (KIPRE), Ministry of Health, Kenya**

**Flemish Partner Institution: Hasselt University (UHasselt)**

**Duration: 5 years  Indicative budget: €300,000**

**1. Problem Statement and Context (max. 250 words)**

The Kenya Institute of Primate Research (KIPRE) is a leading biomedical and translational research centre recently restructured as a Semi-Autonomous Government Agency (SAGA) under the Ministry of Health (MoH). Since its establishment in 1958, KIPRE has advanced research in tropical and neglected diseases, non-communicable diseases, reproductive health and biology, and biodiversity conservation. Despite this long-standing scientific contribution, KIPRE currently lacks a dedicated Data Science and Analytics Centre (DSAC) capable of managing, integrating, and analysing the rapidly expanding biomedical and ecological datasets it generates.

This gap limits KIPRE’s ability to harness data-driven approaches including biostatistics, bioinformatics, and machine learning which are indispensable for advancing modern biomedical research, disease mapping and modelling, and strengthening public health surveillance and interventions.

Establishing a functional Data Science and Analytics Centre (DSAC) will bridge this gap by enabling KIPRE to systematically digitize and integrate biomedical and ecological datasets, apply predictive analytics, and generate policy-relevant evidence to guide decision-making within the MoH.

The initiative aligns with Kenya’s Vision 2030, the Universal Health Policy (2020–2030), and the Digital Health and Innovation Strategy, all of which emphasize data governance, capacity strengthening, and digital transformation as key enablers of sustainable health outcomes.

By integrating Flemish expertise from UHasselt in quantitative and computational methods, the project will elevate KIPRE’s analytical capacity, advance equitable North-South and South-South collaborations, and position KIPRE as a regional hub for biomedical data science in East Africa.

**2. Overall and Specific Objectives (max. 150 words)**

**Overall Objective:**

To build a sustainable institutional framework for data science and analytics at KIPRE to support innovative biomedical research, integrated conservation efforts, and data-informed public health policies.

**Specific Objectives:**

1. Establish a fully functional Data Science and Analytics Centre (DSAC) equipped with modern digital infrastructure and staffed with competent data scientists, analysts, and ICT professionals.
2. Enhance human capacity in data science, biostatistics, and bioinformatics through targeted training, mentorship, and academic exchanges between KIPRE and UHasselt.
3. Digitize, curate, and integrate legacy and ongoing research datasets into standardized, interoperable repositories to enable efficient data access and reuse.
4. Develop and institutionalize data governance, ethics, and security frameworks consistent with FAIR (Findable, Accessible, Interoperable, Reusable) and GDPR principles.
5. Apply advanced data science and modelling approaches to priority research domains including disease mapping and modelling, One Health surveillance, and drug discovery.
6. Strengthen collaborative partnerships and networks for knowledge exchange, joint research, co-supervision of trainees, and regional benchmarking in biomedical data science.

**3. Expected Outcomes and Outputs (max. 150 words)**

* A fully operational Data Science and Analytics Centre (DSAC) will be established and institutionalized within KIPRE, serving as a national hub for biomedical and ecological data integration.
* At least 20 scientists and ICT officers will be trained in advanced data analytics, bioinformatics, and machine learning through structured mentorship and exchange programs.
* A comprehensive data governance and ethics frameworkconsistent with FAIR and GDPR principles will be adopted.
* Legacy and ongoing research datasets will be digitized, standardized, and integrated into interoperable repositories.
* The project will generate flagship data-driven studies supporting evidence-based policy and public health interventions for the MoH.
* Sustained collaboration between KIPRE, UHasselt, and African partner institutions (University of Nairobi, Ardhi University, and Makerere University) will enhance regional research visibility and position Kenya as a data-enabled Centre of excellence in health and conservation science.

**4. Added Value of the Flemish–Kenyan Partnership (max. 150 words)**

The partnership leverages UHasselt’s technical excellence in biostatistics, bioinformatics, and AI-driven analytics to strengthen KIPRE’s capacity for biomedical and ecological data integration and advanced research.

UHasselt will co-design the DSAC training curriculum, provide mentorship in data-intensive modelling, and support the development of governance and ethics frameworks aligned with international standards.

KIPRE will host the Centre, coordinate implementation, and ensure that analytical outputs inform national health and conservation policies.

The collaboration promotes equitable capacity strengthening: Flemish researchers gain access to unique African biomedical and ecological datasets, while KIPRE scientists acquire advanced analytical and computational skills.

Through joint supervision, collaborative research, and shared innovation, the partnership will build a sustainable data science ecosystem that enhances evidence generation, informs policy, and positions Kenya as a regional hub for integrated data-driven research and decision-making.

**5. Sustainability (max. 150 words)**

The DSAC will be institutionalized within KIPRE’s organizational structure, with dedicated personnel, infrastructure, and operational budgets integrated into the institute’s long-term strategic plan.

The MoH will reinforce sustainability through policy adoption, budgetary support, and integration of DSAC outputs into national health information systems.

A structured mentorship and training framework with UHasselt will facilitate continuous skill transfer, ensuring technical capacity is retained and expanded beyond the project’s duration.

Regional collaborations with the University of Nairobi (Kenya), Ardhi University (Tanzania), and Makerere University (Uganda) will foster South–South cooperation through joint proposals, shared data infrastructure, and collaborative research.

By embedding digital infrastructure, data governance, and human capital development within KIPRE’s institutional framework, the project will enable the DSAC to evolve into a self-sustaining national and regional Centre of excellence for biomedical and ecological data science.

Modules