GAINAfrica: Generative Artificial INtelligence for Africa*

Rima Grati[1], Khouloud Boukadi[2], Bassem Bouaziz[2], Amany H.B. Eissa[3,4], Agnes Mirembe[5], Andrea Declich[6], Fabio Feudo[6], Jocelini do Rego[7], Mustapha Mezghani[8], Sanaa Tazi[9], Cyril Boyault[10], Alessandro Zanasi[11], Adriano Rippa[12], Lanfranco Marasso[12], Raniero Romagnoli[13], Ella Morgan[14], Filippo Bianchini[15], Marco Calamo[15], Francesca De Luzi[15], Flavia Monti[15] and Massimo Mecella[15,*]

[1]*Zayed University, Abu Dhabi, AE*
[2]*University of Sfax, Tunisia*
[3]*Egypt University of Informatics, Knowledge City, Cairo Governorate, Egypt*
[4]*Cairo University, Giza, Egypt*
[5]*Action for Rural Women's Empowerment, Kampala, Uganda*
[6]*Conoscenza e Innovazione Srls, Rome, Italy*
[7]*Preventive Healthcare Innovation Hub, Benin*
[8]*Société de gestion de la technopole Sfax, Tunisia*
[9]*African Smart Territories, Casablanca, Morocco*
[10]*Onetreck (Reckonect SAS), Grenoble, France*
[11]*Defence Research Institute, France*
[12]*Almaviva, Rome, Italy*
[13]*Almavawe, Rome, Italy*
[14]*LGI Sustainable Innovation, France*
[14]*Sapienza Università di Roma, Italy*

Abstract

This paper presents the GAINAfrica project, a recently funded Horizon Europe that empower African societies by co-creating and deploying context-sensitive Generative AI solutions that address pressing societal challenges in healthcare, agriculture, city infrastructure, and education. Through a responsible and inclusive innovation framework, the project aims to bridge the technological divide between Europe and Africa by adapting EU-developed GenAI technologies to the realities and opportunities of African regions. The project aims at demonstrating how novel technologies, methods and approaches, coupled with digital innovation approaches fostering strong local ecosystems, comprised of startups, researchers, civil society and its networks, and public institutions, can promote sustainable development, digital equity, and gender-responsive innovation.

Keywords

Generative AI, Socio-technical systems, Technology transfer models

RCIS 2026: Companion Proceedings of the 20th Conference on Research Challenges in Information Science: RCIS Research Projects and Workshops, May 26-29, 2026, Toulouse, France

*GAINAfrica is supported by the European Commission, Horizon Europe proposal 101298559.

*Contact author (Principal Investigator of the project).

rima.grati@zu.ac.ae (R. Grati); khouloud.boukadi@fsegs.usf.tn (K. Boukadi); bassem.bouaziz@isims.usf.tn (B. Bouaziz); amani.eissa@eui.edu.eg (A. H.B. Eissa); aruwe@aruweug.org (A. Mirembe); andrea.declich@knowledge-innovation.org (A. Declich); feudo@knowledge-innovation.org (F. Feudo); jocelini.dorego@phihub.studio (J. d. Rego); m.mezghani@ictsfax.com (M. Mezghani); african.smart.territories@gmail.com (S. Tazi); cyril@reckonect.com (C. Boyault); alessandro.zanasi@zanasi-alessandro.eu (A. Zanasi); a.rippa@almaviva.it (A. Rippa); l.marasso@almaviva.it (L. Marasso); r.romagnoli@almawave.it (R. Romagnoli); ella.morgan@lgi.earth (E. Morgan); bianchini@diag.uniroma1.it (F. Bianchini); calamo@diag.uniroma1.it (M. Calamo); deluzi@diag.uniroma1.it (F. D. Luzi); monti@diag.uniroma1.it (F. Monti); mecella@diag.uniroma1.it (M. Mecella)

0000-0002-6995-465X (R. Grati); 0000-0002-6744-711X (K. Boukadi); 0000-0002-3692-9482 (B. Bouaziz); 0000-0001-8663-2604 (A. H.B. Eissa); 0000-0002-1304-2571 (A. Declich); 0009-0006-3278-9853 (F. Bianchini); 0009-0006-2602-9604 (M. Calamo); 0000-0002-9896-2528 (F. D. Luzi); 0000-0002-9421-8566 (F. Monti); 0000-0002-9730-8882 (M. Mecella)

Table 1
GAINAfrica Project Information

Project full name:	Generative Artificial INtelligence for Africa
Acronym:	GAINAfrica
Funding:	European Commission (Horizon Europe Work Programme 2025 Cluster 4)
Start/end date:	01 Oct 2026 - 30 Sept 2029
Partners:	Sapienza Università di Roma (Italy), Conoscenza e Innovazione Srls (Italy), University of Sfax (Tunisia), Technopole Sfax (Tunisia), Egypt University of Informatics (Egypt), Almaviva/Almavawe (Italy), OneTreck (Reckonect SAS) (France), Action for Rural Women's Empowerment (Uganda), Preventive Healthcare Innovation Hub (Benin), African Smart Territories (Morocco), Defence Research Institute (France), LGI Sustainable Innovation (France)

1. Introduction

Generative Artificial Intelligence (GenAI) is a transformative paradigm across multiple domains, enabling advanced capabilities in natural language processing, decision support, and content generation. In this context, the current GenAI landscape is strongly dominated by large-scale ecosystems developed in the United States and China [1], raising concerns about technological dependency, lack of diversity, and limited adaptability to heterogeneous socio-economic contexts. On the other hand, Europe faces the strategic challenge of developing interoperable, trustworthy, and context-aware GenAI solutions that can be effectively transferred and adapted beyond its borders. This challenge is particularly relevant in African contexts, where infrastructural constraints, linguistic diversity, and socio-cultural heterogeneity require AI systems that are not only technically robust but also locally grounded and inclusive.

GAINAfrica (acronym for Generative Artificial INtelligence for Africa, see Table 1 for details) is a 3-year Horizon Europe project that tackles three key dimensions:

- Research Pillar 1: understand the African demand for GenAI: conduct a structured, evidence-based analysis to identify and characterise the specific societal needs in Africa within four priority sectors – agriculture, healthcare, education, and infrastructure – capturing regional, linguistic, and socio-economic diversity;
- Research Pillar 2: identify African organisations, start-ups, research institutions, clusters , incubators, end users or buyers and civic actors able to test and adapt GenAI-based solutions, mapping ecosystem readiness and absorptive capacity to empower Living lab sites;
- Research Pillar 3: map and evaluate existing EU GenAI solutions, starting with EU-LLM stacks and Small Language Model (SLM) stacks and match them to needs, localisation and deployment conditions.

The topics related to *Research Challenges in Information Science conference* that will be investigated in the project are many, including Information Systems and their Engineering (requirements and software engineering in specific cultural and geographical areas), Domain-specific IS Engineering (e-health, smart cities, smart agriculture), User-Oriented Approaches (user-centred design and human factors in AI-based information systems), Data Science (machine learning and Generative AI applications) and Reflective Research and Practice (impact of information on the enterprise and the individual, action research and case studies).

1.1. Current approach

The GAINAfrica project proposes a double scouting strategy as a new approach to cooperation between stakeholders involved in the development and adoption of GenAI-based solutions to address social issues in local contexts. The approach is structured as an iterative co-creation process, where GenAI solutions are progressively designed, adapted, and validated in collaboration with local stakeholders. Living

labs serve as a vehicle to support co-creation workshops, rapid prototyping, and real-world testing. Requirements are not predefined, but evolve through iterative cycles of design, prototyping, testing, and validation, enabling continuous refinement of usability, performance, and contextual relevance. The project's research and development flow consists of three main phases:

- *Phase 1 - Inception and needs mapping.* The first phase focuses on defining requirements based on African contexts. During this phase, the project identifies key African needs, prioritizes use cases, and maps available European GenAI assets. Co-design activities involve local living labs and incubators to develop a methodological approach for technology adaptation.
- *Phase 2 - Proof of concept and co-creation.* The second phase aims to refine requirements and use cases, select AI adaptation techniques, and set-up initial GenAI pipelines. Solutions are co-created with African startups and local authorities, resulting in prototypes in simplified pilot environments, such as chatbots in local languages or offline-first health assistants. These GenAI services are validated through living lab, demonstrating the feasibility and effectiveness of solutions adapted to local African contexts.
- *Phase 3 - Consolidated pilots and integrated stack.* The final phase focuses on consolidating product-relevant scenarios and integrating AI pipelines into an interoperable EU-Africa GenAI stack. Large-scale pilots are conducted in 3–5 African sites, ensuring compliance with the AI Act, GDPR, and the African Union AI Strategy.

To facilitate technology transfer, GAINAfrica adopts a dual-thread approach. The first model is a *mediated transfer*, where local startups or incubators acts as intermediary, by customising a technology for end-users. The second model is a *direct end-user uptake*, where African organizations adopt the solution directly, with the European provider configuring and handing over the final application. This dual approach enables the project to study the conditions under which each model is most effective, while simultaneously reducing deployment risks. Figure 1 illustrates the GAINAfrica technology transfer process, highlighting the transfer models, the scope and the stakeholders involved.

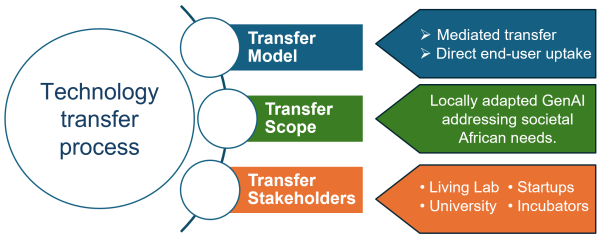

Figure 1: GAINAfrica Technology Transfer Process.

1.2. The technologies and approaches behind GAINAfrica

The GAINAfrica project builds on a set of advanced and complementary technological components to enable the adaptation and deployment of GenAI systems in low-resource and heterogeneous environments. The infrastructure integrates European-developed AI models, lightweight adaptation techniques, and distributed deployment mechanisms into a unified framework. Figure 2 illustrates the high-level framework of GAINAfrica, highlighting the interaction between core components and deployment layers.

At the core of GAINAfrica lies a modular GenAI infrastructure combining: *(i)* European Large Language Models (LLMs) and Small Language Models (SLMs) already available within the consortium; *(ii)* Retrieval-Augmented Generation (RAG) [2] pipelines for grounding model outputs in domain-specific knowledge; and *(iii)* Knowledge Graph (KG) integration to support structured reasoning and contextualisation [3]. These components are orchestrated to support flexible adaptation across multiple domains, such as healthcare, agriculture, digital skills and smart cities.

To address the constraints of African deployment contexts, GAINAfrica adopts a set of lightweight adaptation techniques that avoid costly full model retraining. These include: parameter-efficient fine-tuning [4] (e.g., LoRA-like approaches), model compression techniques (quantization, pruning,

distillation and sparsification), and multilingual and culturally-adaptive GenAI strategies. Preliminary results are already available, cf. [5].

The Data & Deployment layer provides the infrastructure for data management and system execution, ensuring that AI components can be deployed in decentralized, resource-constrained environments typical of many African contexts. In the context of the GAINAfrica project, the development and adaptation of GenAI models require high-quality, representative data. Both EU technology providers and African pilot partners contribute to the creation of these datasets. Specifically, the project integrates heterogeneous data sources, including local datasets (text, speech, and sensor data) and domain-specific repositories, as well as deployment environments spanning cloud, edge, and on-device infrastructures. A key feature of this layer is support for edge and offline-first deployment, enabling AI services to operate under low-connectivity conditions. This is achieved through local storage and caching mechanisms, as well as distributed execution on edge devices. GAINAfrica will also study and incorporate open datasets (e.g., Copernicus, Digital Earth Africa, and open language corpora) and will in turn release a significant portion of its datasets in accordance with the FAIR principles (Findable, Accessible, Interoperable, Reusable) and open science practices.

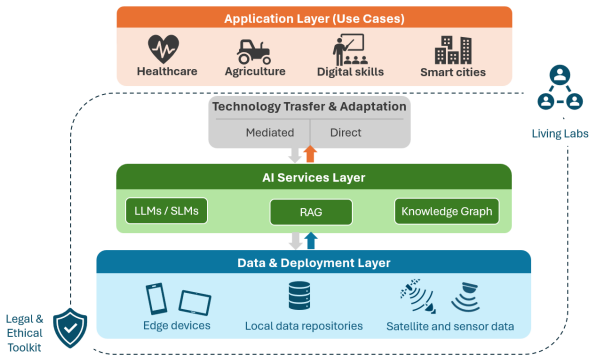

Figure 2: GAINAfrica framework.

To guarantee responsible AI deployment, the project incorporates a compliance-by-design framework aligned with European regulations (e.g., AI Act, GDPR). For example, LLMs may reproduce gender, cultural, or linguistic biases, leading to exclusion of women, rural communities, or minority language speakers. In this context, GAINAfrica applies a participatory co-design approach in living labs. Continuous monitoring and feedback loops will allow iterative correction, ensuring equitable utility across diverse user groups.

The technology infrastructure is closely related to Living Lab environments, in which AI solutions are co-designed and validated with local stakeholders. This enables the iterative refinement of models and interfaces, the integration of local knowledge and data, and, above all, the continuous evaluation of system usability and impact. The living lab approach ensures that the technology remains aligned with real user needs and socio-cultural contexts.

1.3. Use cases

The project aims to apply the two technology-transfer models to four Use Cases (UCs). In addition to UCs, the project uses the interactions with relevant stakeholders within the living labs that will be formed around the UCs.

1.3.1. Use Case 1 - Healthcare

African health systems are characterized by limited human resources for health, high patient loads per doctor, and disparities between urban and rural care. However, the real-world adoption of innovative AI-based solutions is often blocked by by challenges related to cultural adaptation, digital literacy, and workflow integration. This UC aims to support clinicians in their daily practice through *(i)* decision support, by summarising patient history and conditions and suggesting clinical pathways; *(ii)* workflow integration, by reducing time per consultation and automating routine documentation; and *(iii)* patient interaction, by providing clear, culturally adapted health information for improved trust and adherence.

1.3.2. Use Case 2 - Smart cities and urban infrastructure

Several cities (i.e., Casablanca, Rabat, and Sfax) face recurrent peaks during weekends, matches and concerts, leading to: *(i)* congestion and safety risks along pedestrian corridors to venues/old towns; *(ii)* fragmented, last-minute information for visitors across languages; *(iii)* complex multi-actor coordination (municipality, transit, police, organizers) under intermittent connectivity; *(iv)* a need for frugal, replicable solutions that avoid vendor lock-in and comply with AI Act and GDPR-grade safeguards. This UC aims to support citizens in their daily activities through a City Visitor Assistant that provides crowd-aware itineraries, real-time alerts, and accessibility features for women, families and the elderly. It will also support authorities by predicting densities and egress times, and by generating "what-if" scenarios (e.g., gate openings, additional buses, dynamic signage).

1.3.3. Use Case 3 - Smart agriculture

Crop diseases are one of the leading causes of yield losses worldwide. Small and medium-sized farmers, especially in rural regions, often lack immediate access to agronomists or laboratories for timely diagnosis; as a result, delayed detection leads to costly interventions, reduced crop quality, and economic losses. Furthermore, existing solutions are fragmented and not adapted to local languages or farming contexts. Every day, farmers must make complex decisions (e.g, how much to irrigate, when to fertilize, and when to harvest), ofetn based on experience or costly expert consultation. With increasing climate variability, resource scarcity (e.g., water), and volatile market prices, poor decisions can significantly reduce yields and profits. To support farmers in identifying and managing plant diseases, this UC leverages GenAI by integrating computer vision, sensor data, and agronomic expertise. The resulting system generates natural language recommendations tailored to specific crops, climates, and agricultural practices. Preliminary results are already available, cf. [6, 7].

1.3.4. Use Case 4 - Digital skills and learning

Across African and Mediterranean higher education systems, students face several structural barriers: limited faculty-to-student ratios, low levels of digital literacy among first-generation university students, language and cultural barriers in accessing online resources (often English-only), and fragmented e-learning tools that lack personalization and contextual adaptation. These issues result in high dropout rates, uneven skill acquisition and reduced employability. Universities urgently require scalable, low-cost, and inclusive solutions to boost digital literacy, increase retention, and improve learning outcomes. With this UC, GenAI aims to improve digital literacy and personalized learning in university settings by integrating LLMs into digital tutoring platforms to support students in ICT, entrepreneurship, and STEM courses. Through conversational tutoring, adaptive content generation, and multilingual explanations, the system provides academic assistance upon request, filling gaps in access to qualified instructors.

2. Objectives and expected results

The GAINAfrica project aims to foster the development, adaptation, and deployment of human-centric GenAI solutions across African contexts by leveraging European technological assets and promoting inclusive innovation ecosystems. To achieve this vision, the project is structured around five main objectives:

- O1 focuses on identifying and analysing the specific societal and economic needs of African contexts across key sectors such as agriculture, healthcare, education, and infrastructure. The project adopts an evidence-based and participatory approach, combining socio-technical analysis with stakeholder engagement to capture regional, linguistic, and socio-economic diversity. Novel methods for requirement engineering of AI-based information systems, with customization features for regional, socio-economic diversity, are expected to be investigated and validated.

- O2 focuses on identifying and validating EU GenAI solutions suitable for African contexts. It involves mapping the European landscape of projects, startups, and initiatives, and assessing their transferability based on infrastructure, language adaptability, openness, and efficiency. Results are validated through workshops with EU and African stakeholders, leading to the definition of reference architectures and domain-specific use cases, while supporting international cooperation and digital sovereignty.
- O3 aims to establish Living Labs in African regions as hubs for open innovation, co-creation, and testing of GenAI applications. These environments engage local startups, youth, and women-led initiatives, promoting inclusion and skills development. Two transfer models are applied—mediated in North Africa and direct in Sub-Saharan Africa—producing localized innovation blueprints, scale-up plans, and comparative insights to support entrepreneurship and territorial cohesion.
- O4 is dedicated to developing and integrating GenAI models adapted to African contexts. Solutions are co-designed to address constraints such as low-resource environments, limited connectivity, and linguistic diversity. Prototypes in key sectors (agriculture, healthcare, education, infrastructure) are tested locally, ensuring inclusive, sustainable, and human-centric AI technologies. O2–O4 are expected to lead to the development of novel methodologies and tools for GenAI-based information system engineering.
- O5 ensures the ethical, legal, and safe deployment of GenAI solutions. It develops a compliance-by-design framework aligned with the AI Act and GDPR, along with tools for risk and impact assessment. Training, policy guidance, and transfer methodologies are also provided, aiming to build a trustworthy ecosystem based on transparency, accountability, and responsible innovation.

At the core of the project is the development of a unified toolbox, combining LLM tools designed for cloud deployment with lightweight LLMs optimized for edge devices, enabling flexible and context-sensitive AI applications. The project aims to investigate novel language models (large and a small versions) specifically designed to address the linguistic, cultural, and infrastructural realities of African users. Building on these models, GAINAfrica will create four prototype products, co-developed with partner organizations in sectors such as healthcare, smart agriculture, digital skills and learning, and smart cities and urban infrastructure. These prototypes will demonstrate practical, high-impact applications of GenAI across diverse domains.

All tools and models will be integrated into a comprehensive GAINAfrica toolboxBeyond technical outputs, the project will also provide guidelines and a playbook for maximizing societal penetration and socio-economic impact, supporting informed decision-making and sustainable AI deployment. Additionally, GAINAfrica will contribute to shaping the European-African AI regulatory landscape, aligning innovation with ethical and governance standards.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT to: Grammar and spelling check.

References

[1] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang, W. Ye, Y. Zhang, Y. Chang, P. S. Yu, Q. Yang, X. Xie, A survey on evaluation of large language models, ACM Trans. Intell. Syst. Technol. 15 (2024). URL: https://doi.org/10.1145/3641289. doi:10.1145/3641289.

[2] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, et al., Retrieval-augmented generation for knowledge-intensive NLP tasks, Advances in neural information processing systems 33 (2020) 9459–9474.

[3] L. Yang, H. Chen, Z. Li, X. Ding, X. Wu, Give us the facts: Enhancing large language models with knowledge graphs for fact-aware language modeling, IEEE Transactions on Knowledge and Data Engineering 36 (2024) 3091–3110. doi:10.1109/TKDE.2024.3360454.

[4] Z. Han, C. Gao, J. Liu, J. Zhang, S. Q. Zhang, Parameter-efficient fine-tuning for large models: A comprehensive survey, arXiv preprint arXiv:2403.14608 (2024).

[5] A. Puglisi, F. Monti, C. Napoli, M. Mecella, Service-oriented AI model compression for computing continuum environments, in: Service-Oriented Computing - 23rd International Conference, ICSOC 2025, Shenzhen, China, December 1-4, 2025, Proceedings, Part I, Lecture Notes in Computer Science, Springer, 2025, pp. 303–318. doi:10.1007/978-981-95-5012-8_23.

[6] B. A. T. Koné, R. Grati, B. Bouaziz, K. Boukadi, M. Mecella, Optimized and explainable feature selection for soil moisture prediction across sites, in: Proceedings of the 14th International Conference on Data Science, Technology and Applications, DATA 2025, Bilbao, Spain, June 10-12, 2025, SCITEPRESS, 2025, pp. 213–222. doi:10.5220/0013646000003967.

[7] B. A. T. Koné, K. Boukadi, R. Grati, E. B. Abdallah, M. Mecella, Llm-driven semantic explanations for soil moisture prediction models, Smart Agricultural Technology 12 (2025) 101174. doi:https://doi.org/10.1016/j.atech.2025.101174.