Breaking Barriers:

AI Generative Tools for Inclusive Stakeholder Engagement in Sub-Saharan Africa's Education Sector

Beatrice Fagan, Cem Ayanoglu and Ozlem Ozel[1]

Artificial Intelligence (AI) is increasingly driving technological progress and the digital revolution. This paper explores how recent advancements in AI can enhance representation for vulnerable communities, in particular in the sub-Saharan African context, during the environmental and social (E&S) scoping process.

Its widespread application worldwide highlights how rapidly it can streamline processes and enhance not only routine work activities but also broader aspects of daily life. However, as technology advances, a significant digital gap persists between developing countries and the rest of the world. This disparity is also evident in the development of AI generative tools, while the top 100 languages in natural language processing (NLP) cover 96% of global GDP, they represent less than 60% of the world's population[2].

Several factors contribute to this gap, including financial constraints, a shortage of skilled professionals, and limited documentation and data, particularly in African countries. NLP and AI generative tools are closely linked to socio-economic inclusion, as they play a crucial role in improving representation for vulnerable communities. Yet, as long as the digital divide remains unaddressed, inequalities in access to NLP solutions will continue to grow.

By leveraging AI-generative tools and NLP for African languages, the paper will shed spotlight on the Artificial Intelligence for Development (AI4D) Framework and its current implementation at Malawi University of Business and Applied Sciences (MUBAS) University for a Chichewa (official language of Malawi and Zimbabwe) dataset creation. The aim of the paper is to show how common social risks can be addressed through AI contribution on meaningful stakeholder engagement and the upholding of the right to informed participation.

Taking as example a case study of university campus projects in Malawi, where RINA conducted environmental and social (E&S) scoping studies in 2023 and 2024, key social sensitivities identified during stakeholder engagement will be presented. Subsequently, the "AI4D framework" program will be introduced, as well as the ongoing implementation at the University of Malawi – MUBAS, for the development of a dataset in Chichewa Language, to show how hypothetically these tools can positively affect stakeholder engagement process.

---

[1] Beatrice Fagan is a Social Impact Consultant specializing in African studies and human rights, with a strong background in the socio-economic inclusion of vulnerable communities. Currently, she works at RINA Consulting, assessing social compliance of financed projects with international standards across various African regions. Cem Ayanoglu is a Senior Environmental Consultant at RINA. Experienced in environmental and social impact assessment on a variety of sectors in Turkey and Africa. Ozelm Ozel is Middle East Technical University Sociology graduate and Social Impact Consultant at RINA. Experienced in working with vulnerable groups, leads stakeholder engagement activities through focus group discussions, in-depth interviews and participatory research methods.

[2] African languages falling behind voice technology innovations

1. Key Social Risks in the Scoping Process of Development Projects in Sub-Saharan Africa

1.1.   The Role of the Scoping Phase in Impact Assessment

The scoping phase is widely recognized as a foundational step in the impact assessment (IA) process. According to the Social Impact Assessment Guidance (IAIA, 2015):

"Scoping can be defined as the process of identifying the main issues of concern as well as determining the interested and affected parties for a particular planned intervention. It is a preliminary process that produces an interim list of issues to be considered that are later properly assessed"[3]

A robust scoping process serves multiple interconnected purposes. It confirms the scope and extent of the proposed project, and identifies the legislative, regulatory, and policy frameworks that will guide acceptable outcomes (IAIA, 2018). It also facilitates the early identification of interested and affected parties including right-holders, regulatory authorities, and most important, it helps identify individuals and groups that may be directly and differentially or disproportionately affected by the project because of their disadvantaged or vulnerable status[4].

Therefore, scoping phase and the preliminary identification of stakeholders are critical steps in shaping the project's social performance, as they contribute to the prioritization of social risks and define how stakeholders will be meaningfully involved throughout the impact assessment process, outlines the required specialist studies and their terms of reference, and agrees on the methodologies for assessing impact significance (IAIA, 2018). Furthermore, it establishes mitigation approaches, typically guided by the mitigation hierarchy.

When key stakeholders are excluded from the scoping phase, they may only be consulted after specialist studies and assessments are completed, limiting their ability to meaningfully influence the process. As such, their late-stage involvement may reduce the legitimacy and effectiveness of the IA. It is therefore vital to document how stakeholder inputs during the scoping phase shape the design and scope of the assessment itself, ensuring transparency and accountability in the process.

Effective stakeholder engagement hinges on the prior disclosure of relevant and comprehensive project information to ensure stakeholders fully understand the risks, impacts, and opportunities of the project. The objectives of such engagement include fostering understanding through culturally appropriate and transparent communication, actively involving stakeholders in identifying mitigation measures, optimizing local benefits, building relationships that foster a

---

[3] Guidance for assessing and managing the social impacts of projects, IAIA, 2015, p.41
[4] The International Finance Corporation (IFC) Performance Standards (PSs) define this disadvantaged or vulnerable status may stem from an individual's or group's race, color, sex, language, religion, political or other opinion, national or social origin, property, birth, or other status. The client should also consider factors such as gender, age, ethnicity, culture, literacy, sickness, physical or mental disability, poverty or economic disadvantage, and dependence on unique natural resources (IFC Performance Standards on Environmental and Social Sustainability, p.9)

social license to operate (SLO), and managing expectations through clear and accessible information[5].

However, challenges such as language variability, especially in remote regions, and low literacy levels can hinder meaningful engagement. It is essential that these obstacles are addressed to ensure the process is truly inclusive and effective. Meaningful consultation not only enhances the quality and relevance of the impact assessment but also mitigates the risk of opposition or delays.

1.2 Common Social Risks and Barriers in Sub-Saharan Africa

Sub-Saharan Africa is characterized by a large rural population, with an average about 57% of people residing in rural areas[6]. This demographic factor presents significant challenges during the early stages of development projects, particularly in the scoping phase.

A key issue is the prevalence of multiple languages spoken across the region. According to the Survey Report for Bantu Languages (Nurse, 2001) there are 2,058, approximately languages spoken in Africa (30% of the estimated languages spoken in the world), many of which are indigenous or local languages. Experts estimates almost 440 Bantu "varieties"[7], while some of these languages are widely spoken, others are limited to small communities. In rural areas, non-dominant language speakers often struggle to participate meaningfully in consultations due to a lack of access to information in their languages.

Furthermore, literacy rates in many sub-Saharan African countries remain low. For instance, the adult literacy rate is often below 70%, with some countries having rates as low as 50% in rural regions[8]. This poses significant barriers to effective engagement, as illiterate communities may struggle to understand written project materials and participate in consultations. Additionally, rural areas face infrastructural challenges such as poor road conditions, limited access to electricity, and unreliable internet connectivity, which further hinder the flow of information and participation in the consultation process. These factors contribute to the exclusion of vulnerable groups and marginalize their voices in decision-making processes, making it critical to ensure that communication is accessible and inclusive during the scoping phase of development projects.

In addition, in several sub-Saharan African countries, including Malawi, weak national frameworks for disclosure and consultation often hinder the systematic inclusion of vulnerable and marginalized groups. If not proactively addressed, these gaps risk undermining the legitimacy, transparency, and overall quality of the assessment process.

In Malawi, the legal foundation for Environmental and Social Impact Assessment (ESIA) is established by the Environmental Management Act (EMA) of 2017, which provides the overarching framework for environmental protection and project oversight. Complementing this,

---

[5] Stakeholder Engagement Good Practice Handbook (2007)

[6] https://www.macrotrends.net/global-metrics/countries/SSF/sub-saharan-africa-/rural-population?utm

[7] Communities speaking Bantu languages are indigenous to twenty-seven African countries: Angola, Botswana, Burundi, Cameroon, CAR, Comoros, Congo, DRC, Equatorial Guinea, Gabon, Kenya, Lesotho, Madagascar, Malawi, Mayotte, Mozambique, Namibia, Nigeria, Rwanda, Somalia, South Africa, Sudan, Swaziland, Tanzania, Uganda, Zambia, and Zimbabwe. Non-Bantu Niger-Congo languages are spoken north and mainly west of Bantu.

[8] https://databank.worldbank.org/embed/Adult-literacy-rates-in-Sub-Saharan-Africa/id/3aa273b7

the Environmental Impact Assessment Guidelines (originally published in 1997 under a previous version of the Act) continue to guide project disclosure and stakeholder engagement practices.

While the EMA 2017 mandates stakeholder engagement as part of the ESIA process (as outlined in Section G of the EIA Guidelines), the scope of these requirements remains limited. Engagement is only formally required during the scoping phase and again prior to the submission of the final draft to the Malawi Environmental Protection Authority (MEPA). This narrow approach falls short of international standards such as those established by the IFC, which call for continuous, iterative consultation and disclosure throughout the project life cycle, including implementation and monitoring stages.

The combination of inadequate stakeholder consultation and limited disclosure presents a dual social risk: on one hand, affected communities are excluded from key decision-making processes, and on the other, developers may miss critical insights that could improve project design and social outcomes. These regulatory and structural shortcomings often result in stakeholder engagement that is superficial and reactive, rather than inclusive and forward-looking.

1.3 Case Study of RINA: Scoping Study in Malawi for MUBAS University

Since 2023, RINA Consulting is conducting an E&S Scoping Study in Malawi as part of the due diligence and project structuring phase of a Purpose-Built Student Accommodation (PBSA) project supported by the IFC and the Government of Malawi.

The objective of the study is to inform Project stakeholders and future bidders of the potential E&S risks and impacts related to the development and operation of the PBSA on the proposed site, providing guidance to align the project with World Bank Performance Standards and national regulatory requirements. The Project Government of Malawi (GoM), represented by the Ministry of Finance, in collaboration with the Malawi University of Business and Applied Sciences (MUBAS), has appointed the IFC as lead transaction advisor for a Public Private Partnership (PPP) aimed at delivering PBSA. To ensure compliance with national environmental and social (E&S) regulations and the World Bank Group's performance standards, IFC engaged RINA Consulting as the E&S consultant.

Between social risks identified during the scoping phase, the risk of exclusion of vulnerable groups was assessed.

In particular, the potential exclusion of vulnerable groups from meaningful consultation processes. This risk is particularly relevant given that most local communities speak Chichewa, a Bantu language estimated to be spoken by over 10 million people in Malawi, with a wide distribution in Southern Africa[9]. The language is the national language of Malawi and is also spoken in neighboring countries such as Mozambique, Zambia and Zimbabwer [10]. Project-related information and disclosure documents are typically produced in English, creating significant language barriers for many residents.

In addition, Malawi has an adult illiteracy rate of approximately 31%, with clear disparities based on location: urban youth literacy reaches 95.6%, while rural youth literacy drops to 80.7%[11]. These

---

[9] https://www.worlddata.info/languages/chichewa.php#google_vignette
[10] Taylor A., Using Machine Learning to Detect Fraudulent SMSs in Chichewa.
[11] Malawi Population and Housing Census 2018

factors combine to reduce access to information and limit community members' ability to understand project implications and express concerns.

Geographically, MUBAS is located in Blantyre District, where 98.2% of the population lives in rural areas, despite Blantyre City being entirely urban[12]. Rural communities in this area face infrastructural challenges, including poor road networks and limited internet access, further impeding their ability to participate in stakeholder engagement processes.

Altogether, these linguistic, educational, and infrastructural barriers reinforce the marginalization of rural populations and increase the risk of their exclusion from critical stages of project planning and decision-making.

According to the above, withing the scoping study one of the core challenges identified was conducting meaningful stakeholder engagement in a context marked by language diversity, low literacy, and limited access to project information. Weak national consultation frameworks, power asymmetries between stakeholders, and community mistrust further complicated engagement efforts. These barriers often hinder the participation of non-dominant language speakers, women, youth, and other marginalized groups.

The scoping phase demonstrated the importance of early engagement, particularly for identifying risks that could influence the design, timeline, and financial viability of the project.

2. AI4D and the Development of Chichewa NLP Tools

2.1 What is AI4D?

The AI4D initiative is a promotion program that provides inclusive and context-sensitive artificial intelligence research and application. A major component of the initiative is the AI4D – African Language Program, which specifically focuses on the low-resource status of African languages in NLP. The program addresses structural challenges such as the scarcity of digital language data, lack of standardized orthography, and limited access to expert linguistic resources, all of which hinder the development of AI tools that are linguistically inclusive and culturally relevant.[13]

AI4D has three key phases:

- Data Collection Challenges: Open competitions were held to crowdsource African language datasets. These resulted in 52 submissions covering 15 languages, evaluated both qualitatively and quantitatively.[14]
- Language Dataset Fellowships: Selected teams received fellowships (3–4 months) to build higher quality datasets for various NLP tasks, including machine translation, text-to-speech, sentiment analysis, and document classification. Among these, a Chichewa news classification dataset was developed in collaboration with MUBAS.[15]

---

[12] Ibid
[13] Siminyu, K. et al. (2021). AI4D – African Language Program. In Proceedings of Africa NLP 2021, pp. 1–12.
[14] Ibid., pp. 5–6
[15] Ibid., p. 9.

- Machine Learning Competitions: Datasets were used to host Mahine Learning competitions for benchmarking AI models and generating reusable tools for African contexts.[16]

AI4D supports social inclusion, stakeholder participation, and sustainable development, especially in sectors like education, health, governance, communication and public services, while using local-language NLP development.

2.2 The Chichewa NLP Project

The Chichewa NLP Project is one of several initiatives under the AI4D framework which supports the development of AI tools' usage in African languages and aim to increase access to essential services and rights such as health, education, and financial security.

In Malawi, a team from the MUBAS was part of the AI4D African Language Dataset Challenge and created annotated datasets in Chichewa, for tasks such as machine translation and named entity recognition[17]. Another team from Kuyesera AI Lab at MUBAS focused specifically on detecting fraudulent SMS messages. They developed a labeled dataset distinguishing between normal and fraudulent texts, using crowdsourced data from students and the local community[18].

To improve the dataset's size and quality, the team applied text augmentation and translated the messages into English using both human and machine translation methods and they achieved up to 98% accuracy while highlighting the importance of developing AI tools in native languages[19].

This work has a positive social impact on social inclusion and sustainable development of under representative groups. Many SMS scams in Malawi involve impersonating government or NGO programs, such as the well-known Mtukula Pakhomo cash transfer scheme or fake job offers related to COVID-19[20].

Based on real examples in the local language, AI models can help identify and prevent these scams, which mostly target vulnerable people with limited access to formal support systems.

Additionally, students and community members gained digital awareness and technical knowledge within participation in data collection and annotation processes. In a country where mobile communication is more widespread than internet access, being able to process and understand Chichewa messages with AI tools is a major step toward inclusive and safer digital services[21].

3. Implications for Stakeholder Engagement in Impact Assessment

It is almost certain that AI tools will improve the impact assessment processes by making stakeholder engagement more inclusive and effective. As anticipated above, in many development projects, community participation might be hindered by barriers such as language differences, low

---

[16] Ibid., pp. 11–12.
[17] Siminyu K. et al., AI4D – African Language Dataset Challenge, 2021. pp. 1–12.
[18] Amelia Taylor & Amoss Robert, Using Machine Learning to Detect Fraudulent SMSs in Chichewa
[19] Ibid., pp. 13–14.
[20] Ibid., p. 11.
[21] Ibid., pp. 2–3.

literacy levels, and limited access to data or communication channels. By leveraging AI tools, IA experts can break down language barriers, present information in accessible formats, and analyze feedback at scale, thereby upholding the right to informed participation for all groups.

3.1 AI for Language Accessibility and Inclusive Communication

One of the most immediate ways AI can improve inclusion is by overcoming language barriers during stakeholder engagement. In sub-Saharan Africa and other diverse regions, project documents and meetings are often conducted in a dominant language (such as English or French), which many community members may not speak fluently. This leads to the exclusion of non-dominant language speakers from meaningful dialogue.

AI-powered translation tools offer a practical solution by rapidly converting project disclosures, consent forms, and consultation materials into local languages, and vice versa. For instance, an environmental impact report written in English can be automatically translated into the several local languages, enabling villagers to read or hear it in their mother tongue. Conversely, community feedback given in local languages can be translated for project teams, ensuring that all voices are understood regardless of language.

3.2 AI for Visual and Audio-Based Community Engagement

Another common barrier identified in impact assessments is low literacy, which makes it difficult for some community members to understand lengthy written reports or technical jargon. Traditional stakeholder engagement often attempts to address this by using visual aid (like posters or flipcharts) and verbal presentations. However, producing high-quality, engaging multimedia content for every project can be resource intensive. Here, generative AI can fill the gap by rapidly creating visual and audio materials that convey key information in more intuitive ways.

These AI-enabled visual and audio tools make engagement more inclusive for people with limited reading ability or formal education. They also cater to oral cultural preferences, as many communities transmit knowledge through spoken words. This closes the information gap and builds trust, since community members see that efforts are being made to communicate with them rather than at them.

3.3 AI for Data Collection and Analysis in Remote Areas

Reaching and understanding communities in remote or under-resourced areas poses another significant challenge for stakeholder engagement. Villages far from urban centers may have limited internet connectivity, poor roads, or simply fall outside the usual reach of consultation efforts. As a result, vital feedback from these areas can be under sampled. On the data collection side, AI chatbots can extend the reach of engagement. For example, a simple SMS-based chatbot could automatically send out questions in a local language (even on basic cell phones) to solicit community input on a project's impacts, translating and storing replies for analysis. Similarly, an AI voice assistant system could be used, letting people in remote villages call a number and verbally submit concerns or ideas in their own dialect. These approaches reduce the need for constant physical presence of survey teams, yet still gather qualitative data from afar.

Moreover, AI can tap into indirect data sources: social media posts, messaging apps, or community radio transcripts that discuss the project. By scraping and transcribing these channels (with permission), practitioners can capture sentiments that locals express in day-to-day conversations.

Even traditional feedback mechanisms like grievance hotlines or radio call-in shows can be leveraged and their content can be transcribed and fed into AI systems so that no community voices are overlooked.

AI excels at making sense of the volume. NLP algorithms can sort and categorize thousands of comments or transcripts much faster than manual reviewers. One valuable application is sentiment analysis: AI can detect whether stakeholder feedback in each village skews positive, negative, or neutral about the project, and flag specific statements that indicate strong concerns. By identifying trends in the data, AI helps pinpoint which issues are emerging repeatedly across communities.

Conclusion

The aim of this paper was to analyze the potential of generative AI to significantly reshape the way IA processes are conducted, particularly in terms of enhancing the inclusivity and effectiveness of stakeholder engagement. It was discussed how such tools can play a crucial role in breaking down communication barriers and supporting social inclusion not only by translating information, but by bridging vocabulary gaps, advancing knowledge in vernacular languages, and simplifying technical content into accessible, culturally appropriate formats.

AI-powered translation tools ensure that project materials and community feedback are rapidly and accurately converted into multiple local languages, making participation more informed and equitable. Also, generative AI enables the creation of visual and audio formats that resonate with orally oriented and low-literacy communities, promoting meaningful engagement beyond the limits of written communication. In remote or underserved areas, AI facilitates broader outreach through mobile-based chatbots, voice-enabled platforms, and the analysis of indirect channels like social media or radio transcripts. Sentiment analysis and natural language processing further allow project teams to monitor feedback in real time, identifying emerging risks and concerns early in the process.

From a consultancy perspective, this shift has significant implications: it enhances the quality and depth of early social risk identification and contributes to stronger, trust-based relationships with communities.

The AI4D framework is already being implemented through various initiatives, one of which is built on partnerships between universities, local innovators, and international experts. These collaborations foster the development of inclusive technological solutions that are contextually appropriate and driven by the very communities they aim to serve.

Programs like AI4D have already demonstrated that students and academic experts are key drivers of innovation, capable of developing inclusive, community-centered technologies. Universities serve as vital hubs for this innovation, where knowledge, creativity, and collaboration intersect. For instance, at MUBAS, a program is currently underway to explore how AI can meaningfully enhance stakeholder engagement in impact assessments. Supporting such academic initiatives is crucial to ensuring that emerging technologies empower vulnerable groups, promote equitable development, and contribute to rights-based, participatory governance.

Bibliography

AI4D – African Language Program (Siminyu, Kathleen, et al.). 2021. AI4D – African Language Program. In Proceedings of Africa NLP.

International Association for Impact Assessment. 2018. FastTips: Scoping.

International Finance Corporation (IFC). 2012. Performance Standards on Environmental and Social Sustainability. Washington, DC: World Bank Group.

International Finance Corporation (IFC). 2007. Stakeholder Engagement: A Good Practice Handbook for Companies Doing Business in Emerging Markets. Washington, DC: World Bank Group.

Malawi National Statistical Office. 2018. Malawi Population and Housing Census 2018. Zomba, Malawi.

Nurse, Derek. 2001. A Survey Report for the Bantu Languages. Cartographer: Irene Tucker. SIL International.

Taylor, Amelia, and Amoss Robert. 2021. "Using Machine Learning to Detect Fraudulent SMSs in Chichewa." AI4D – African Language Program.

Vanclay, Frank, Ana Maria Esteves, Ilse Aucamp, and Daniel M. Franks. 2015. Social Impact Assessment: Guidance for Assessing and Managing the Social Impacts of Projects. Fargo, ND: International Association for Impact Assessment.

Website consulted

International Association for Impact Assessment. 2018. "FastTips: Scoping." Accessed April 17, 2025. https://www.iaia.org/uploads/pdf/Fastips_18%20Scoping.pdf.

MacroTrends. n.d. "Sub-Saharan Africa Rural Population 1960–2023." Accessed April 17, 2025. https://www.macrotrends.net/global-metrics/countries/SSF/sub-saharan-africa-/rural-population?utm.

World Bank. n.d. "Adult Literacy Rates in Sub-Saharan Africa." *World Development Indicators (Databank)*. Accessed April 17, 2025. https://databank.worldbank.org/embed/Adult-literacy-rates-in-Sub-Saharan-Africa/id/3aa273b7.

World Data Info. n.d. "Languages in Malawi: Chichewa." Accessed April 17, 2025. https://www.worlddata.info/languages/chichewa.php#google_vignette.

"African Languages Falling Behind Voice Technology Innovations."