



# A NEW GENERATION OF AFRICAN LANGUAGE TECHNOLOGISTS

## A Virtual Lab for Virtuous Goals

Kamusi Labs **works with students** and hackers from around the world to build **cutting-edge tools** and resources for **numerous languages**, which we then **share with the public for free**. Our origin and special focus is Africa, where more than a billion people speak more than 2000 languages. These languages remain almost entirely outside the digital realm, thereby excluding their speakers from access to revolutionary engines of knowledge and prosperity. We aim to change that, and to foster **a cadre of young women and men** who gain experience creating tools that their communities can use immediately, as well as the confidence to produce **market-worthy technologies** to lead their languages into the future.

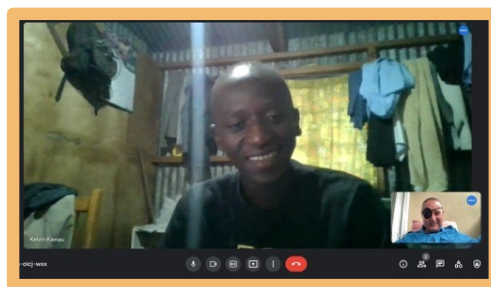
Lab members maintain a steady dialog with the director as they dig into their projects. They also have frequent group meetings to help each other and work through common challenges. Video conferences often **bring brains together** from several countries, across multiple time zones and continents. We recruit especially among passionate young technologists spread **across Africa** in the grassroots Masakhane community ([kamu.si/masakhane](http://kamu.si/masakhane)).

Kamusi Project International • Rue de la Blancherie 22 •  
Chavannes-près-Renens 1022 Switzerland • +41.78.904.94.05 •  
fax +41.21.888.16.06 • [kpi@kamusi.org](mailto:kpi@kamusi.org)

## African Language Empowerment

Kamusi is partnering with the African Academy of Languages (ACALAN), the intergovernmental language body of the African Union, to produce the tools to gather and share the linguistic data for an anticipated technology platform for African languages. Our pathbreaking Kam4D linguistic knowledge graph (<http://kamu.si/kam4d>) will support many further initiatives to place African languages at the forefront of next-generation technologies. The first step is an as-yet unfunded prototype for Kiswahili that we are pursuing with the University of Nairobi, to demonstrate what we are capable of achieving. When we succeed, speakers of African languages will have:

- Rich and comprehensive dictionaries and learning resources
- Translation tools that far surpass Google's accuracy
- User interfaces for the internet, mobile devices, and other technology in myriad local languages
- Bedrock data for speech recognition and other advanced natural language processing
- Access to language-restricted services such as health care, courts, and higher education



## The Big Picture Playbook

We have many projects in the pipeline, from entry-level to potential topics for a Masters or PhD. Look through <http://kamu.si/big-picture-playbook> to see what's coming down the pike. Broadly, lab opportunities include:

- Games for collecting linguistic knowledge from the public
- Tools for collecting linguistic knowledge from experts
- Data preparation – for example, converting an existing dictionary for your language into a useable format
- Graph database development (Neo4j)
- Web user interface and experience
- Mobile apps using Ionic
- Educational applications
- Social service applications
- Natural language processing, such as machine translation and machine learning

## The Bounty Model

It is **almost impossible** to find funding for the development of African languages from foundations, philanthropists, or public agencies – donors direct their limited funds toward other priorities. Kamusi Labs relies largely on volunteer labor to claw toward our objectives. Interns often start with enthusiasm, but evaporate midway when schoolwork or paying jobs demand their attention. **We have learned** that rewarding a **small bounty** upon delivery, or certification for university course credit, **makes everyone happy**, ensuring that projects are completed, with large leaps forward on a **shoestring budget**.

Kamusi Project USA • 605 41st St, #H • Brooklyn, NY •  
11232 USA +1.203.599.1920 • [kpusa@kamusi.org](mailto:kpusa@kamusi.org)