



WORLD BANK GROUP



Join Us for an Innovative Discussion on AI and Language Access in Malawi

Date: 25th November 2025 @10am, MUBS ODeL Auditorium

Session Synopsis

How can artificial intelligence serve populations when it doesn't speak their language? This workshop is organised by the World Bank Data Lab and presents a Gates Foundation-supported project addressing a critical gap in global AI development: the near-total absence of Chichewa—spoken by over 20 million people—from the digital world. The "AI for Malawi" project is partnering with local publishers, media houses, and the National Statistical Office to build the first comprehensive Chichewa Digital Language Library for training AI models that can serve Malawian citizens in their own language.

This interactive workshop will feature a 30-minute presentation from the project team, followed by Q&A and—importantly—an opportunity for the project team to gather your input and questions. Whether you're interested in natural language processing, international development, journalism, digital inclusion, data governance, or the ethics of AI deployment in low-resource settings, this pilot project offers insights into how we build equitable AI infrastructure. The session will cover the technical, legal, and governance challenges of curating language datasets, preprocessing data for AI training while protecting intellectual property, and establishing sustainable national stewardship.

Suggested Agenda (1.5 hours)

- Welcome and Introductions (University, 10 min)
- AI for Malawi Project Presentation (Project Team, 30 min)
- Presentation Q and A (All, 15 min)
- Project Team Questions for Participants (All, 30 min)
- Closing Remarks (University, 5 min)

Logistics

Date: 25th November 2025 @ 10 am

Venue: ODeL MUBAS Auditorium

Registration:

To help with logistics we ask all participants to confirm attendance via a registration form

<https://forms.gle/CtLiKwwG5om7PcFu9>.

About the workshop organisers

The World Bank places significant and growing emphasis on investing in data and making it publicly available to drive development. The Bank funds and operates extensive, free, and open-access data portals such as Data360 (with 300 million data points across nearly 10,000 indicators) and DataBank. These tools allow researchers, policymakers, and citizens to analyse economic, social, and sustainability trends. World Bank proposed to leverage its platform to help develop and host language datasets, starting with datasets for Chichewa.

Host: Malawi University of Business and Applied Sciences, through Kuyesera AI Lab and Department of Computer Science and Information Systems. Kuyesera AI Lab has been developing datasets for language, public health and disaster and risk management. Members of the lab have been previously involved in the annotation with Named Entities and Parts-of-Speech for Chichewa datasets.