Centering Low-Resource Languages and Cultures in the Age of Large Language Models (CLRLC-LLMs).

NeurIPS 2025 Workshop.

"Enshrined in a language is the whole of a community's history and a large part of its cultural identity. The world is a mosaic of visions. To lose even one piece of this mosaic is a loss for all of us."

— David Crystal, British Linguist.

Workshop Goals

Center Low-Resource Languages and Cultures in AI Innovation

Reframe the dominant narratives in language technology by placing low-resource languages and cultural contexts at the heart of LLM development.

Advance Research and Scalable Approaches for Low-Resource LLMs

Showcase cutting-edge research, methods, and tools that improve data creation, model alignment, evaluation, and deployment for low-resource languages.

Foster Global Collaboration and Cross-Disciplinary Exchange

Bring together researchers, linguists, technologists, developers, community leaders, and policy thinkers to share practical insights, challenges, and strategies.

Ideate on Ethnically Culturally Grounded Data Principles

Create space to critically examine ethical, legal, and socio-cultural challenges in dataset collection for low-resource languages.

Accepted Works at the CLRLC-LLMs Workshop, NeurIPS 2025

Program at a Glance

Time	Session	Details
09:00–09:30	Workshop Opening & Registration	Don Alberto 1
09:30–09:45	Welcome & Opening Remarks	Workshop Organizer- Oluchi
09:45–10:30	Keynote Speaker	Towards Representative, Trustworthy and Scalable Multilingual, Multi-cultural Evaluation. ***Sunayana Sitaram (Microsoft India)***
10:30–10:45		**Coffee Break**
10:45–11:00	Spotlight Presentation	Community-in-the-Loop: A Framework for Participatory Dataset Creation for Indigenous and African Languages. ***Balogun David Taiwo***
11:00–11:15	Spotlight Presentation	Culturally aware instruction and alignment methods for low-resource Indian-language LLMs. ***Praveen Tangarajan & Madhvi Sharma***
11:15–11:30	Spotlight Presentation	ENLIVEN-1000: A Comprehensive Revitalization Framework for 1000+Endangered Languages via Broad-Coverage LID and LLM-Augmented MT. ***Philip Meng***
11:30–11:45	Spotlight Presentation	AfriStereo: A Culturally Grounded Dataset for Evaluating Stereotypical Bias in Large Language Models. ***Oche Ankeli***
11:45–12:00	Spotlight Presentation	Developing a Non-Western Culture Aware RAG Application. ***Ariana Mondiri***
12:00–13:00		**Lunch Break**

Program at a Glance

Time	Session	Details
13:00–13:45	Keynote Speaker	Endangered Languages in the Age of LLM. ***Emily Prud'hommeaux (Boston College)***
13:45–14:00	Spotlight Presentation	CL-PDE: A Heterogeneous Graph Framework for Cross-Lingual Mental Health Ontologies. ***Isabel Michelle***
14:00–14:45	Keynote Speaker	Active Learning for Low-resource Efficient NLP. ***Bonaventure F. P. Dossou (McGill University | Mila Québec)***
14:45–15:00	Spotlight Presentation	Beyond Surface Text: Revealing Distinctive Personas in LLMs using Cognitive Bridging. **Jongwon Ryu**
15:00–15:15	**Afternoon Break**	
15:15- 15:30	Spotlight Presentation	HealthAlign-Agents: Self-Play Reflective Prompting for Culturally Aligned Health Communication in Low-Resource Languages. **Aura Arefeh Yavary**
15:30– 16:30	**Special Session**	**Ethical Dataset Collection and Representation in Low-Resource Language Context**
16:30– 17:30	Closing Remarks	Workshop Chair – Joy Olusanya

Keynote Speakers

Sunayana Sitaram
Microsoft Research India
9:45-10:30

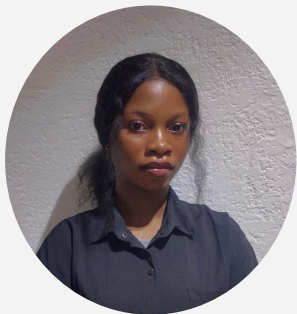**Emily Prud'hommeaux**
Boston College
13:00–13:45

Bonaventure F. P. Dossou
McGill University | Mila Quebec AI Institute
14:00–14:45

Organizers

Joy Olusanya
Workshop Chair
Obafemi Awolowo University

Mary Salami
University of California, Santa Barbara

Cynthia Jayne Amol
Maseno University | Tonative Community

Yann Le Beux
YUX Cultural AI Lab

Sharon Ibejih
Tonative

Flora Oladipupo
Tolaram Group

Oluchi Audu
YUX Cultural AI Lab