

# Kidist Amde Mekonnen

PhD Researcher | Generative Retrieval, Recommendation & Large-Scale Machine Learning

[in](#) kidistamde [G](#) kidist-amde [G](#) Google Scholar [@](#) k.a.mekonnen@uva.nl [📍](#) Amsterdam, NL

## SUMMARY

PhD researcher at **University of Amsterdam (IRLab)** working on generative retrieval, recommendation, neural IR, and large-scale applied machine learning. Publications at **SIGIR 2025**, **SIGIR 2026**, and **ACL Findings 2025**, including work on continual generative retrieval, decoding robustness, and multilingual retrieval. Experienced in end-to-end model development, large-scale experimentation, and retrieval systems using **PyTorch**, **FAISS**, **Hugging Face**, **Docker**, and **Slurm**. Seeking research or applied scientist internship roles in retrieval, recommendation, LLMs, or large-scale ML.

## RESEARCH EXPERIENCE

**PhD Researcher (IR & ML)** Oct 2023 – Present  
**University of Amsterdam (IRLab)** Amsterdam, NL

- Continual GenIR:** developed a parametric memory-head method for updating generative retrievers over dynamic corpora while mitigating catastrophic forgetting; accepted as a **SIGIR 2026 Full Paper**.
- Decoding robustness:** reproduced and stress-tested look-ahead decoding for generative retrieval under query variation and cross-lingual shift; accepted in the **SIGIR 2026 Reproducibility Track**.
- DDRO:** designed a pairwise relevance optimization objective for generative retrieval, improving **MRR@10** by **+7.4%** on **MS MARCO** and **+19.9%** on **Natural Questions** over strong baselines.
- Built end-to-end **Python/PyTorch** training and evaluation pipelines for large-scale retrieval experiments, including ablations, benchmarking, and error analysis.
- Multilingual retrieval:** led dataset construction, model training, and benchmarking for Amharic passage retrieval; improved dense retrieval **MRR@10** by **+5.01%** over multilingual baselines.

## INDUSTRY EXPERIENCE

**Data Science Research Intern** Apr 2022 – Oct 2022  
**Nanoverly Limited** Newcastle upon Tyne, UK

- Implemented and evaluated **GAN-based** models for synthetic DNA strand generation for nanorobot probes, covering preprocessing, model training, and experimental evaluation.
- Developed a **predictive model** for forecasting sensor behavior across lab environments and communicated findings to research stakeholders.

## TECHNICAL SKILLS

- Languages:** Python, Java, C++, R
- ML/IR:** PyTorch, TensorFlow, Hugging Face, neural IR, dense/sparse retrieval, learning-to-rank, recommender systems, LLMs
- Systems:** FAISS, Elasticsearch, Docker, Slurm, Git, Linux/Unix
- Methods:** large-scale evaluation, ablation studies, error analysis, reproducible experimentation

## EDUCATION

**PhD Candidate** Oct 2023 – Present  
**University of Amsterdam** Amsterdam, NL

**MSc in Data Science** 2020 – 2023  
**University of Trento** Trento, Italy

**MSc in Machine Intelligence** 2019 – 2020  
**AIMS-AMMI** Kigali, Rwanda

**BSc in Computer Science** 2013 – 2017  
**University of Gondar** Gondar, Ethiopia

## SELECTED PUBLICATIONS

- Mekonnen, K.A., Tang, Y., and de Rijke, M.** *A Parametric Memory Head for Continual Generative Retrieval*. **SIGIR 2026 (Full Paper Track)** – to appear. doi.
- Mekonnen, K.A., Li, Y., Tang, Y., Lupart, S., and de Rijke, M.** *Lost in Decoding? Producing and Stress-Testing the Look-Ahead Prior in Generative Retrieval*. **SIGIR 2026 (Reproducibility Track)** – to appear. doi.
- Mekonnen, K.A., Tang, Y., and de Rijke, M.** *Lightweight and Direct Document Relevance Optimization for Generative Information Retrieval*. **SIGIR 2025 (Full Paper Track)** – presented. doi.
- Mekonnen, K.A., Alemneh, E., and de Rijke, M.** *Optimized Text Embeddings & Benchmarks for Amharic Passage Retrieval*. **ACL Findings 2025** – presented. link.

## ADDITIONAL PUBLICATIONS

- Mekonnen, K.A. et al.** *Adv-KD: Adversarial Knowledge Distillation for Faster Diffusion Sampling*. **arXiv (2024)** – poster: **ELLIS Doctoral Symposium 2023**. arXiv.
- Mekonnen, K.A.** *Balanced Face Dataset: Guiding StyleGAN for Labeled Synthetic Face Image Dataset for Underrepresented Group*. **arXiv (2023)** – talk: **AMLD Africa 2021**. arXiv.
- Mekonnen, K.A.** *Conditioning GAN Without Training Dataset*. **arXiv (2024)**. arXiv.

## SELECTED ENGINEERING PROJECTS

**Spatiotemporal Mobility Analytics** 2023  
**MSc Course Project**

- Built a spatiotemporal analytics pipeline using GPS and time-diary data, including trajectory cleaning, temporal aggregation, subgroup analysis, and interactive **space-time cube** visualization.
- Produced reproducible demographic comparisons and visual analytics outputs. [🔗](#) Code

**Birth-rate Estimation Pipeline** 2023  
**Course Project**

- Developed an end-to-end ML pipeline using Twitter and Eurostat data: collection, labeling UI with **Flask**, **BERT**-based tweet classification, and regression-based birth-rate estimation.
- Packaged code and experiments for reproducible execution. [🔗](#) Code

## TEACHING & SERVICE

**Teaching Assistant** 2024 – Present  
**University of Amsterdam** TA for Information Retrieval & Recommender Systems

**Assistant Lecturer** 2017 – 2019  
**University of Gondar** Lectured CS courses; supervised labs and assignments

**Hackathon Mentor** 2024  
**ELLIS ESSIR'24** [🔗](#) Mentored teams on IR tasks and reproducibility

## VOLUNTEERING

**Teaching Assistant** 2019, 2024  
**AddisCoder** [🔗](#) Programming and algorithms outreach

**Volunteer** Sep 2020 – Apr 2022  
**Women in AI & Robotics** [🔗](#) Mentoring and community support

**Mentor** 2019 – 2023  
**MeetMentors.org** [🔗](#) Mentored students in AI/ML careers