

# Mahi Aminu Aliyu

Abuja, Nigeria

✉ mahigital@gmail.com | 🌐 mahigital.me | 🌐 Mahi Aminu Aliyu | 🎧 A-I-Mahi

## RESEARCH INTERESTS

My research focuses on causal representation learning and inductive biases — specifically, how incorporating causal structure into learned representations can improve robustness and out-of-distribution generalization, with applications in low-resource NLP and public health.

## EDUCATION

**M.Sc., Computer Science**, Abubakar Tafawa Balewa University, Nigeria 2023 – 2026 (Expected)

*Advisors:* Dr Badamsi Imam Ya'u 🌐

**B.Sc., Computer Science**, Al-Qalam University Katsina, Nigeria 2018 – 2023

*Thesis:* Water Quality Prediction Model

*Advisor:* Dr Mansir Abubakar 🌐

## EXPERIENCE

**Artificial Intelligence and Social Informatics Associate** 2025 – Present

*Africa Hub for Innovation and Development*

- Refactored WellBridge (a Tuberculosis chatbot) into class-based architecture with GPU-accelerated STT, TTS, and LLM components, improving maintainability, scalability, and reducing latency by 30%.
- Expanded its reach through WhatsApp integration and established a structured feedback loop with end-users to guide system improvements.
- Scraped and analyzed 2,000+ public health-related social media posts across four platforms to build a dataset for sentiment analysis on vaccine misinformation.

**Research Collaborator** 2024 – Present

*ML Collective (Remote)*

- Led the ML Theory and Optimization group (5 members), conducting literature-driven research on generalization techniques with a focus on causal representation learning, including experimental design, execution, and research reporting.
- Contributed to multimodal research by curating, cleaning, and organizing datasets to support reliable experimentation and analysis.
- Presented research ideas and findings at conferences and workshops, engaging with the research community and integrating feedback into ongoing work.
- Facilitated weekly Saturday reading sessions, guiding technical paper discussions and fostering critical analysis of research papers across the community.

**Data Engineer & Analyst** 2023 – 2025

*Horizon IT Services*

- Curated a low-resource Hausa dataset by extracting and cleaning 500+ text entries from Al-Barka radio station's news archives
- Applied advanced data cleaning techniques to transform 2,000+ raw Hausa news entries into a structured CSV dataset
- Collaborated on dataset enhancement by introducing categorization and optional labeling, enabling downstream use in Kaggle competitions and applications in NER, sentiment analysis, and machine translation for Hausa language

## PUBLICATIONS & MANUSCRIPTS

[1] **Mahi Aliyu**, Mansir Abubakar, Armaya'u Z. Umar, Alwatben Batoul Rashed, and Rukayya Musa Ismail. Analysis of machine learning models for efficient water quality prediction. In *Journal of Information and Technology Management (JISTM)*, 10(40), Sep. 2025. 🌐

[2] Mardiyah Oduwale, Prince Mireku, Fatimo Adebajo, Oluwatosin Olajide, **Mahi Aminu Aliyu**, Jekaterina Novikova. AfriCaption: Establishing a New Paradigm for Image Captioning in African Languages. In **7th Workshop on African Natural Language Processing (AfricaNLP)**, Jan 2026 🌐

[3] **Mahi Aminu Aliyu**, Chisom Chibuike, Fatimo Adebajo, Omokolade Awosanya, Samuel Oyeneye. Towards Robust Generalization in African AI: Causal Inference and Domain Shift Mitigation. *arXiv preprint* 🌐

CONFERENCES, SUMMER SCHOOLS & WORKSHOPS

|                                                                                                                                                                                                                                                                                                            |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <b>Foundations of Causal Inference Workshop, Isaac Newton Institute, Cambridge</b> – Virtual Participant<br>Selected to attend workshop on causal inference foundations led by top global researchers from EPFL, Cambridge, Harvard, and ETH Zürich.                                                       | 2026 |
| <b>Machine Learning Summer School - Senegal (MLSS)</b> – Participant<br>Selected to attend advanced ML training covering NLP, optimization, reinforcement learning, optimal transport, diffusion models, federated learning, and foundation models, led by researchers from Google, NYU, Apple, and KAUST. | 2025 |
| <b>Deep Learning Indaba 2025, Rwanda</b> – Attendee<br>Selected to attend Africa’s largest AI conference with travel grant acquired from ML Collective.                                                                                                                                                    | 2025 |

SELECTED PROJECTS

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |  |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <b>Cross-Lingual Aspect-Based Sentiment Analysis in African Languages</b> <ul style="list-style-type: none"><li>- Translated SemEval-2014 to <b>17 African languages</b> (afri-SemEval-2014) with aspect alignment heuristics.</li><li>- <b>Modified</b> and <b>fine-tuned</b> pretrained transformer architectures (e.g., XLM-R, AfroXLM-R) for <b>causal reasoning</b> for aspect-based sentiment analysis tasks.</li><li>- Causal models showed superior out-of-distribution (OOD) performance with fewer training steps.</li></ul> |  |
| <b>SageBot</b> <ul style="list-style-type: none"><li>- Built AI tool that summarizes technical docs for developers, cutting review time by 15%.</li><li>- Leveraged LLMs (ChatGPT, LLaMA-3), Hugging Face, and Langchain for high summarization accuracy.</li><li>- Improved developer workflow efficiency <b>by enabling 25%</b> faster comprehension of complex documents.</li></ul>                                                                                                                                                 |  |
| <b>Dono Health</b> <ul style="list-style-type: none"><li>- Led development of inclusive insurance MVP with Health Center, Vendor, and Insured modules.</li><li>- Designed serverless AWS backend with Lambda, API Gateway, and CDK.</li><li>- Integrated Flutterwave for payments and Africa’s Talking USSD API for voucher-based funding.</li></ul>                                                                                                                                                                                   |  |

TALKS & PRESENTATIONS

|                                                                                                                                                     |      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <b>Deep Learning Indaba ’25</b><br>Poster presentation: "Towards Robust Generalization in African AI: Causal Inference and Domain Shift Mitigation" | 2025 |
| <b>MLSS-Senegal ’25</b><br>Poster presentation: "Towards Robust Generalization in African AI: Causal Inference and Domain Shift Mitigation"         | 2025 |
| <b>ML Collective #25 Research Jam</b><br>Oral Presentation: "Pretrained Models for Water-based Applications"                                        | 2025 |

COMMUNITY SERVICE

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| <b>City Lead</b> , Data Science Nigeria <ul style="list-style-type: none"><li>- Organize AI workshops, hands-on sessions, and reading groups for 150+ community members</li><li>- Promote AI adoption and capacity building across underserved sectors</li></ul>                                                                                                                                                                                                                      | 2024 – Present |
| <b>DeepLearning.ai Ambassador</b> <ul style="list-style-type: none"><li>- Facilitated knowledge sharing and engaged 100+ peers as an active community ambassador</li><li>- Hosted "Growing an AI Community" event in Bauchi with 50+ participants</li><li>- Organized and participated in 5 cross-community events to strengthen the regional AI ecosystem</li></ul>                                                                                                                  | 2023 – Present |
| <b>Lead</b> , Google Developer Student Clubs <ul style="list-style-type: none"><li>- Led the Google Developer Student Club (GDSC) at Al-Qalam University, Katsina. Organized and facilitated 10+ technical workshops and hands-on sessions with core-team members, reaching and training over 350 students in data science, web development, and Android development.</li><li>- Applied Google technologies to address UN SDGs, delivering solutions to 5 local communities</li></ul> | 2021 – 2022    |

## TECHNICAL SKILLS

---

**Programming Languages:** Python, C++, R

**ML/AI Frameworks:** PyTorch, TensorFlow, Hugging Face, Scikit-learn

**Causal Inference & Experiment Tools:** DoWhy, EconML, Weights & Biases

**Web Frameworks & Tools:** Django, Flask, Git/GitHub

**Cloud Platforms:** Microsoft Azure, AWS & GCP

## REFERENCES

---

**Dr. Badamasi Imam Ya'u**

Senior Lecturer, Abubakar Tafawa Balewa University

Email: biyau@atbu.edu.ng

**Dr. Mansir Abubakar**

Senior Lecturer, Universiti Teknologi MARA (UiTM),

Email: mansir@uitm.edu.my

**Riccardo Cadei**

PhD Candidate, Institute of Science and Technology Austria (ISTA)

Email: riccardo.cadei@ist.ac.at