

Zainab Tairu

Machine Learning Engineer

✉ tairuzainab5@gmail.com — [in](#) LinkedIn — [GH](#) GitHub — [Globe](#) Website

Bio

Machine Learning Engineer with a strong focus on NLP and speech technologies, applying language models to scientific and low-resource problems.

Skills

- **Programming & Tools:** Python, SQL, Git, GitHub
- **Machine Learning:** Transformer Models, Deep Learning, Agentic AI, Predictive Modelling
- **NLP:** RAG, LLMs, Prompt Engineering, Rasa, LangChain, spaCy, HuggingFace
- **Data Engineering:** ETL Pipelines, Data Integration, Feature Engineering, Automation
- **Frameworks & Libraries:** PyTorch, TensorFlow, scikit-learn, pandas, NumPy, FastAPI, Streamlit, Gradio
- **Databases:** Neo4j, DynamoDB
- **Cloud & MLOps:** AWS (EC2, S3, Lambda)
- **Mathematical Foundations:** Linear Algebra, Calculus, Probability, Statistics

Experience

- **CoMMA Lab** — Texas (Remote) *Jul 2025 – Jan 2026*
NLP Research Intern
 - Engineered NLP algorithms to model polymer properties and co-authored a survey exploring the intersection of LLMs and polymer modeling.
- **NNPC Upstream Investment Management** — Lagos (On-site) *Oct 2025 – Dec 2025*
Process Intern
 - Designed an ETL pipeline automating daily production reports from field operators, improving data accessibility and reporting efficiency.
 - Explored ML applications for predictive maintenance and reservoir characterization using well log and seismic data.
- **The Innoverse** — South Africa (Remote) *Oct 2024 – Jul 2025*
NLP Intern
 - Built an intelligent chatbot integrating RAG and LLMs to enhance response accuracy and personalization.
 - Developed knowledge graph structures and trained multilingual text-to-speech models for English and Zulu.

Projects

- **RFE Risk Analyzer** — *Jul 2025*
 - Created a USCIS-style risk analyzer using RAG to simulate legal adjudicator feedback on EB-1A petitions.
- **LET ME COOK** — *Nov 2024*
 - Developed a Nigerian recipe dataset and app for efficient pantry-based recipe recommendations.
- **Malaria Blood Smear Classifier** — *Jun 2025*
 - Built a CNN algorithm achieving 99% accuracy for malaria diagnosis from blood smear images.

Publications

Papers and Preprints

- From SMILES to Agents: A Decade of Evolution in Language Models for Polymer Discovery (Expected submission Feb 2026)
- Evaluating Retrieval-Augmented Generation for Medication Question Answering on Nigerian Drug Labels in Yorùbá (Accepted at LoResLM @ EACL 2026)

Poster Presentations

- Hot and Sustainable: AI-Guided Design of Thermostable Lipases — Deep Learning Indaba 2025
- Dual AI Architectures for Accessible Medication Management: OCR-LLM vs. VLM — Deep Learning Indaba 2025
- Reservoir Characterization with Machine Learning — SPE OAU Hackathon 2025

Leadership & Activities

- Lead, DSN OAU — Nov 2025–Present
Coordinated speaker sessions on ethical AI and advanced deep learning research
- Public Relations Officer, Space Club OAU — Jan 2025–Oct 2025
Managed public communications and organized STEM outreach initiatives
- Deep Learning Tutor, DSN OAU — May 2025–Jun 2025
Taught CNN fundamentals and guided students through model development with scikit learn
- Data Science Tutor, Star Girls Africa — Jun 2024–Aug 2024
Taught 100+ students Python programming and supervised their first machine learning models
- Member, SPE OAU — Jan 2022–Present
Participated in hackathons and projects applying ML in petroleum engineering