

# JOY OLUSANYA

☎ +234 09159945712 | ✉ [joynaomi81@gmail.com](mailto:joynaomi81@gmail.com) | 🔗 LinkedIn | 🐙 GitHub | 🌐 Portfolio | 🐦 @Naomi-Joy---

## Experience

### Founder & Lead Organizer, **CLRLC-LLMs** Workshop @NeurIPS 2025

Mexico City

*Centering Low-Resource Languages and Cultures in the Age of LLMs (CLRLC-LLMs)*

December 2025 – Present

- Established and led an international research workshop accepted at NeurIPS 2025, focused on advancing low-resource language technologies and cultural representation in LLMs.
- Coordinated program committee, reviewers, invited speakers, and community partners across academia and industry.
- Designed technical program structure, including call for speakers, lightning talks, and community engagement track.
- Managed end-to-end planning, communications, and logistics for global participation and knowledge-sharing.

### NLP Researcher & Language Lead

Nigeria

*Tonative*

Aug 2025

- Led research and development of machine translation systems for low-resource African languages, with a focus on linguistic diversity and model accuracy.
- Directed language data collection, annotation, and evaluation workflows to improve model performance and cultural relevance.
- Collaborated on the design of multilingual NLP systems aimed at enhancing accessibility, inclusivity, and translation quality across African languages.

### Data Validator

Nigeria

*LyngualLabs*

Aug 2025

- Validated dataset quality using linguistic expertise to ensure accuracy and consistency.
- Reviewed and refined language data to support NLP and machine learning applications.

### Atom Group

Lagos, Nigeria

*AI/LLM Developer*

April 2025

- Developed a commercial AI agent with Instagram scraping pipelines for social media intelligence.
- Built modular, automated data collection and API integration systems to enhance efficiency and reliability.
- Improved scraping accuracy and reduced failure rates through robust exception handling and reusable utilities.

### FriendnPal

Lagos, Nigeria

*AI Developer*

January 2023 – March 2025

- Developed NLP models for a mental health chatbot, improving response relevance and patient engagement.
- Implemented Retrieval-Augmented Generation (RAG) pipelines and preprocessing systems to improve contextual understanding.
- Curated multilingual datasets and conducted data analysis to validate outputs and boost system reliability.

### Cohere For AI

Remote

*Open Source Contributor*

July – August 2024

- Validated 500+ multilingual translations to ensure accuracy and consistency for AI benchmarks.
- Collaborated with international teams on open-source projects and model evaluation.
- Certification of Contribution: [View Certificate](#)

### Rendo AI

Nigeria

*ML/NLP Engineer*

June 2024 – December 2024

- Designed Python-based web scraping pipelines using Requests, Selenium, and BeautifulSoup.
- Automated large-scale data acquisition, improving efficiency and dataset completeness.
- Developed reusable scraping tools to handle structured data from diverse sources.

### Brilla AI Project, Kwame AI

Ghana

*AI Developer*

January 2024 – March 2025

- Developed deep learning speech-to-text models for multilingual ASR with accent robustness.
- Validated and curated diverse datasets to improve transcription accuracy and model reliability.
- Built reproducible evaluation workflows and aligned preprocessing pipelines with training for scalability.

## Education

**Obafemi Awolowo University**

*B.A. Linguistics and African Languages*

**Jan 2021 – July 2025**

*Ile-Ife, Nigeria*

- Relevant Coursework: Computational Linguistics, Phonology, Syntax, Sociolinguistics, Dialectology, Data Annotation, Human Translation, Lexicography, Language Documentation, NLP-focused Linguistics.

## Scholarships and Publications

[1]] **WiML @ NeurIPS 2025 Travel Funding Award**

[2]] **Deep Learning Indaba 2025 Travel Grant**

[P.1] Olusanya Joy and Daud Abolade. (2025). **ÒWE-YOR: Leveraging Transformer-Based Models for Yoruba Proverb Classification**. Poster presented at *Deep Learning Indaba 2025*, Kigali, Rwanda — selected among the **Best Posters**. Also accepted at the *Women in Machine Learning Workshop @ NeurIPS 2025* ([WiML NeurIPS link](#)).

[T.1] Olusanya Joy. (2025). **Tone in Yoruba Automatic Speech Recognition: Evaluating the Impact of Tone Recognition on Transformer-Based ASR Models**. Undergraduate Thesis Summary, *Obafemi Awolowo University*, Ile-Ife, Nigeria.

## Language Proficiency

Yoruba, English, Nigerian Pidgin English, German, French

## Core Skills

- Programming: Python, Julia
- Web/Frameworks: Streamlit, FastAPI, Flask, LangChain, LangGraph, RAG
- Databases/Vector DBs: PostgreSQL, MySQL, Pinecone, Chroma
- APIs: OpenAI API, Gemini API, Tavily API
- Data Science/ML: Pandas, NumPy, Scikit-learn, TensorFlow, Keras, PyTorch, Matplotlib, Seaborn, Tableau, Power BI
- Deep Learning: CNNs, RNNs, LSTMs, Autoencoders, JAX
- NLP: Machine Translation, ASR, Transformers, LLM Evaluation, NLTK, spaCy, Hugging Face Transformers, BeautifulSoup, Language Modeling
- Tools/DevOps: Git, GitHub, GitLab, Docker, Postman, Google Colab, Gradio, Weights & Biases, MLflow, AWS, GCP
- Data: Preprocessing, Annotation, Benchmarking, Dataset Creation, Model Deployment
- Evaluation Methods: Human Annotation, LLM-as-a-Judge, BLEU, ROUGE, BERTScore, COMET, Model Evaluation
- Linguistics Software: Praat, ELAN, FLEEx, Audacity
- Linguistics: Language Documentation, Phonetics, Tone Analysis, Sociolinguistics, Dialectology
- Research: Literature Review, Experimental Design, Data Collection, Dataset Curation, Statistical Analysis

## Selected Projects

**English–Portuguese NMT** – Fine-tuned Helsinki MarianMT for EN–PT; evaluated with BLEU. [GitHub](#)

**English–Kikuyu MT** – Built low-resource MT system using NLLB-200; human evaluation ongoing.

**English–French MT Evaluation** – Tested mBART50 for EN–FR with BLEU/ROUGE. [GitHub](#)

**Language Documentation (Arogbo-Ijaw)** – Fieldwork on syllable structure & phonotactics. [Paper](#)

**Lexicography Project** – Compiled bilingual ecology/climate dictionary (EN–Yorùbá). [Dictionary](#)

**Automatic Essay Scoring** – Fine-tuned DistilBERT for AES; achieved strong RMSE/Pearson scores. [GitHub](#)

**Yoruba Proverb Detection** – Classified proverbs vs non-proverbs with Naive Bayes, BERT, Afrolm.

**Speech Emotion Recognition** – LSTM model on MFCC/chroma features; Streamlit deployment. [GitHub](#)

**Breast Cancer Prediction** – Deep learning model for tumor malignancy classification (95%+). [GitHub](#)

**Toxic Comment Detection** – TF-IDF + Naive Bayes model for toxic vs non-toxic comments. [GitHub](#)

## Leadership Experience

---

**Deep Learning/AI Co-Lead, Data Science Nigeria** – Tutored and guided team members in deep learning concepts, participated in planning and executing deep learning events, and supported bootcamp management.

**AI/Machine Learning Lead, Women in Data Science** – Led AI/ML workshops for women, designed learning paths, created resources, and coordinated speaker sessions for all skill levels.

## Volunteer Experience

---

**Virtual Coach, Django Girls (May 2023)** – Guided participants through coding tasks, enabling them to complete their first web app. **Skills:** Mentorship, Python, Django.

**Protocol Assistant, She Code Africa OAU (May 2024)** – Managed guest reception and event logistics to ensure smooth execution. **Skills:** Communication.

**AI Mentor, LevelUp Techies Bootcamp (Sep 2024 – Nov 2024)** – Taught Python and ML fundamentals, upskilling over 50 participants. **Skills:** Mentorship, Leadership, Communication.

**Deep Learning/AI Tutor, Data Science Nigeria OAU (Mar 2025 – Apr 2025)** – Delivered hands-on sessions on neural networks, CNNs, and model evaluation; empowered 20+ participants to deploy deep learning models. **Skills:** Public Speaking, TensorFlow, Keras.

## Professional Memberships

---

She Code Africa	Data Science Nigeria OAU	Masakhane NLP	Lanfrica	AfriClimate AI	Cohere For AI
-----------------	--------------------------	---------------	----------	----------------	---------------

## Certifications

---

**Udemy:** Natural Language Processing in Python for Beginners 2024 **Cousera:** Data Science Fundamentals2024

**Endangered Languages Projects (Webinar):** Language Documentation Webinar2024