



Jeffrey Otoibhi

AI Researcher

AI Researcher with experience developing tailored AI solutions that automate complex, domain-specific tasks in various industries. Skilled in fine-tuning large language models and applying RLHF to build trustworthy, accurate, and compliant AI systems that integrate seamlessly with existing workflows. Committed to advancing ethical, responsible and scalable AI that are robust and reliable.

Email
jeffreyotoibhi@gmail.com

Mobile
+234 906 155 4618

Website
<https://www.aletheia.com.ng/jeffreyotoibhi>

Assets

AI Research & Development
ML System Design
MLOps
Mathematics For AI

Professional Affiliations

Masakhane
Datascience Nigeria
Igbo AI

Grants

ML Collective Compute Grant
Awarded for SabiYarn-125M development.

Social networks

pauljeffrey
 jeffreypaul
 jeffreypaul

Languages

English

Work experience

- Omal** Nigeria
Head of Data Science Since January 2025
 - Responsible for designing and evaluating an AI-powered dermatology diagnostic system tailored to analyze skin images and handle workflows, providing:
 - Diagnosis of dermatological conditions and possible home remedies.
 - scheduling appointments and handling payment systems
- Freelance (Contract)** Remote
AI Engineer Since September 2019
 - RAG-Powered Customer Support Chatbot**
 - Built a GPT4All-based chatbot with Retrieval Augmented Generation (RAG), enhancing response relevance and accuracy by 30% while resolving 10% of inquiries in real-time, reducing support tickets.
 - SignSynth: AI-Powered Pose-to-Sign Language Translator**
 - Built a transformer-based pose sequence generator and a U-Net-based model for translating pose vectors into sign language images. MAE: 0.04, SSIM: 0.9478.
 - Brazilian BERT-Based Appliance Repair Assistant**
 - Created, integrated a BERT-based system for appliance repair recommendations, reducing human support needs and improving first-time fix rates, integrated via a scalable REST API.
 - AI-Driven Traffic and Accident Event Analysis Engine**
 - Using ResNet-based video segmentation, CNN-LSTM, VideoViT, and LLMs for accident classification, weather analysis, and accident narrative generation, improving baseline accuracy by 15%.
- Freelance (Contract)** Remote
AI Researcher Since September 2019
 - Multi-Modal Clinical Depression Detection Model (DAIC-WOZ dataset)**
 - Accuracy: 82%, Precision: 0.69, Recall: 0.7, F1-score: 0.63.
 - Prediction of Visual Neuronal Responses to Excitable Images (Macaque v4)**
 - Designed CNN Core + attention readout model. Correlation score-0.27.
 - Automatic Post-Edit (APE) Translator with Real-time Adaptation**
 - Designed and trained APE model. BLEU-4 -0.43, TER-0.66.
 - Arabic Offensive and Hate Speech Classification:**
 - Designed and trained a custom MARBERT based model achieving 94.6% accuracy for offensive speech and 97.4% for hate speech.
- Aletheia** Lagos
Chief AI Scientist-SabiYarn-125M LM Since November 2023
 - Led the development of Nigeria's first decoder-only foundational language model supporting multiple Nigerian languages (**Train loss- 2.5, Val loss-2.86**).
 - Outperformed GPT-4 by 48% and LLaMA2-13B by 259% on Nigerian language benchmarks for downstream tasks.
- Sienna Analytics Consulting** Lagos
AI Engineer, LLM & Generative AI From October 2023 to February 2024
 - Designed, developed and deployed advanced AI chatbot systems seamlessly integrating them into the business applications for 4 clients.
 - Employed strategic prompt engineering approaches improving AI responses (correctness and hallucinations) by 25%.

Interests

AI Research

Health AI

Explainable AI

Biomedical Engineering

Robotics

AI in Finance

Education

● College of Medicine, University Of Lagos,
MBBS

Nigeria
From November 2011 to July 2018

Certifications

Deep Learning Specialization - Andrew Ng (2022)

Google TensorFlow Developer (2022)

Tech Stack/Skills

AI

- **AI Safety:** RLHF, Explainable AI, adversarial robustness.
- **Programming:** Python, Javascript, Typescript, c++.
- **Libraries & Frameworks:** Pytorch, Tensorflow, Numpy,pydantic ai, Langchain, Scikit-learn, OpenCV, Deepspeed, Ultralytics, Mediapipe, transformers, peft, accelerate, pydub, librosa, pydantic.
- **Tools & Platforms:** MLflow, Docker, Git, AWS, Fastapi.
- **Data Engineering:** Pyspark, Hadoop.
- **Database:** MongoDB, Postgresql, Redis.

Research Publications

SabiYarn-125M: Advancing Low Resource Languages with Multitask NLP Pretraining

- *Otoibhi J., et al (2025)*. Accepted by the AfricaNLP workshop for the ACL Conference 2025

Research contributions

Adapting Large Language Models for Collaborative Semantic Recommendations

- Modified original implementation to support unique user index generation.

Evaluating Bias in Large Language Models For African Languages

- Language bias in LLMs, comparing metrics between African, English and European texts.

PEPLER with reparameterization for explainable recommendations

- Integration of reparameterization to PEPLER's implementation (**DIV-3.55, FCR-0.11, BLEU-4 0.8197**)

JHU++ Image Crowd Counting (open source paper implementation)

- Implemented a confidence-guided deep residual crowd counting model using PyTorch (from paper).

Alzheimer's disease diagnosis

- Using patient speech analysis and Brain MRI scans

Personal projects

Hapi AI (In Progress)

- An AI-enhanced robotic system to digitize paper-based medical health records using research-backed OCR technology.

Ottobiz (In progress)

- a horizontal service that automates marketing, sales and customer support.

AI-Powered Drug Information Retrieval system

- Using BERT **Precision**-0.90, **Recall**- 0.88, **F1 score** - 0.89, **Acc** - 0.93

ECG Anomaly detection Using AutoEncoders

- **Val loss**-0.03, **test accuracy** - 0.942, **Precision** - 0.99 , **Recall** - 0.9