

Email jeffreyotoibhi@gmail.com

Mobile +234 906 155 4618

Assets

Al Research & Development

ML System Design

Big Data

CI/CD & MLOps

Linear Algebra & Probability

Statistics & Optimization

Professional Affiliations

Masakhane member Contributor to African NLP advancement.

DataFest Africa

Igbo Al

Grants

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

Social networks

- 🕊 pauljeffrey
- 🗲 jeffreypaul

in jeffreypaul

Languages

English German

Jeffrey Otoibhi Medical Doctor | Al Researcher, Engineer

Seasoned AI Engineer and Medical Doctor, with five years of experience in machine learning and innovation. Led the development of Nigeria's first foundational language model and have contributed to 5+ AI research projects. Delivered scalable AI solutions across healthcare and business. Passionate about addressing societal challenges such as cancer treatment and medical rehabilitation. Committed to leveraging AI and robotics to transform healthcare and drive global technological advancement.

Address Lagos, Nigeria. Website https://www.aletheia.com.ng/jeffreyotoibhi

Work experience

Coptic Multi-specialist Hospital
General Physician

Victoria Island, Lagos Since April 2024

- Provided comprehensive medical care, including emergency management, patient consultations, and surgical procedures.
- Collaborated with interdisciplinary teams to optimize patient care pathways.

Freelance

AI Engineer/Researcher

Remote Since September 2019

- AI Mental Health Support Companion
 - Developed, deployed an AI chatbot companion with context-aware interactions, robust dialogue management, documentation systems and progress tracking.
- RAG-Powered Customer Support Chatbot
 - An Information Retrieval enhanced GPT4All-based chatbot. Enhanced response relevance and accuracy by 30%, resolved 10% of inquiries in real-time.
- AI-Driven Traffic and Accident Event Analysis Engine
 - Using ResNet-based video segmentation, CNN-LSTM, VideoViT, and LLMs for accident classification, weather analysis, and narrative generation. Accuracy -71%.
- Brazilian Appliance Repair Recommendation System
 - Created an API-based ML system for appliance repair recommendations, reducing human support needs and improving first-time fix rates.
- Neuronal Responses Prediction to Excitable Images (Macaque v4)
 - Developed a custom model (CNN core + attention readout) to predict neuronal responses in macaque V4 visual areas. Correlation score -0.27.
- SignSynth: AI-Powered Pose-to-Sign Language Generator
 - Built a pose sequence generator, GAN and a U-Net-based model for translating pose vectors into sign language images. MAE: 0.04, SSIM: 0.9478.
- Multi-Modal Clinical Depression Detection Model (DAIC-WOZ dataset)
 - Detection using speech and text. Accuracy: 82%, Precision: 0.69, Recall: 0.7, Fl-score: 0.63.

Aletheia Research Lab

Chief AI Scientist-SabiYarn-125M Foundational LM Si

Lagos Since November 2023

Lagos

- 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 average across Nigerian language benchmarks for downstream tasks.

Sienna Analytics Consulting

Al Engineer, LLM & Generative Al

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%.

İnterests

Al Research

Explainable AI

Robotics

Biomedical Engineering

Education

 College of Medicine, University Of Lagos, MBBS

Certifications

Deep Learning Specialization (Andrew Ng)

Google TensorFlow Developer (Google ML 2022 Cohort)

AI

- AI Safety: RLHF, Explainable AI, adversarial robustness.
- Programming: Python, Javascript, Typescript, c++.
- Libraries & Frameworks: Pytorch, Tensorflow, Numpy, Jax, Scikit-learn, OpenCV
- Tools & Platforms: MLflow, Docker, Git, AWS.
- Data Engineering: Pyspark, Hadoop.
- Database: MongoDB, Postgresql, Redis.

Research Publications

SabiYarn-125M: Advancing Underrepresented Languages with Multitask NLP Pretraining Otoibhi J., et al.

Research contributions

PEPLER with reparameterization for explainable recommendations

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

Automatic Post-Edit (APE) Translator with Online Adaptation

Designed and trained APE model. BLEU-4 -0.43, TER-0.66.

Adapting Large Language Models for Collaborative Semantic Recommendations

Modified original implementation to support unique user index generation.

JHU++ Image Crowd Counting (open source implementation)

Implemented a confidence-guided deep residual crowd counting model using PyTorch.

Alzheimer's disease diagnosis

Using patient speech analysis

Personal projects

MediVault Al

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

Object Detection & Tracking Systems (computer vision)

- built a real-Time face, hand, object detection and tracking system.
- Real Time AI Virtual Mouse and Painter.

Drug Information Retriever

- Developed an <u>NER</u> based system to extract drug names and properties from a database based on user's query in natural language.
- 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

Cart Pole Game (Reinforcement Learning)

 Designed, Implemented a <u>Deep Q network</u> to solve the cart pole game problem using reinforcement learning (RL).