

OFFIONG BASSEY EDET

Machine Learning Researcher | Speech AI | NLP | Low-Resource Language Technologies

Lagos, Nigeria | Email: offiongbassey99@gmail.com | Profile: <https://www.offiongbassey.com.ng> | Google Scholar: <https://scholar.google.com>

PROFESSIONAL PROFILE

Machine learning researcher and NLP/speech-AI practitioner working on multilingual NLP, speech processing, and machine translation for **low-resource African languages**. Founder of EfikAI, building machine translation, ASR, and TTS systems for Efik, an extremely low-resource tonal language, including expanding its speech corpus from a 3-hour baseline to over 300 hours across read and spontaneous speech. Author of accepted papers at **Interspeech**, **AfricaNLP (EACL)**, and **Speakable (LREC)** spanning low-resource TTS, ASR adaptation, and machine translation. Volunteer researcher with **Howard University's NLP Group**, and active in the **Masakhane and Machine Learning Collective** research communities. Broadly interested in data-efficient training strategies for underrepresented languages, with current work examining how TTS performance scales with the quantity and acoustic style of training data.

RESEARCH INTERESTS

Natural Language Processing • Speech Recognition • Speech Synthesis • Machine Translation • Low-Resource Languages • African Language Technology • Multilingual Models • Conversational AI • Language Model Efficiency • Scaling Laws

EDUCATION

B.Sc. Computer Science - *University of Cross River State*

2019 – 2024

PUBLICATIONS

- “Towards Digital Preservation of Efik: TTS for a Low-Resource African Language.” Accepted to **Interspeech 2026**.
Offiong Bassey Edet, Emmanuel Oyo-Ita, Archibong Okon Archibong, David Effanga, Mbuotidem Sunday Awak.
- “Adapting Foundational ASR Models to Efik: An Empirical Study of an Extremely Low-Resource Tonal Language.” Accepted to **Speakable at LREC 2026**.
Offiong Bassey Edet, Stephen Orok Duke, Enoima Essien Umoh, Benjamin Okon Nyong, Andrew Asuquo Nkpanam.
- “Developing an English-Efik Corpus and Machine Translation System for Digitization Inclusion.” Accepted to **AfricaNLP at EACL 2026**.
Offiong Bassey Edet, Mbuotidem Awak, Emmanuel Ubene Oyo-Ita, Benjamin Okon Nyong, Ita Etim Bassey.

RESEARCH EXPERIENCE

Volunteer Research Assistant - *Howard NLP (Howard University Natural Language Processing Group)*

2026 – Present

- Conduct research on LLM-based question answering, multilingual NLP, speech processing, and speech synthesis, with a focus on low-resource language settings.
- Clean and preprocess extremely low-resource language data for multilingual NLP tasks, including question answering, machine translation, audio processing, and speech synthesis.

Machine Learning Researcher - *Machine Learning Collective*

2025 – Present

- Conduct research on multilingual NLP and speech AI, with a focus on low-resource language modeling.
- Lead research on Efik language technologies, including machine translation, automatic speech recognition, and text-to-speech.
- Collaborate with an open research community on reproducible, community-driven AI initiatives.

Research Community Member - *Masakhane Research Foundation*

2025 – Present

- Member of the pan-African NLP research community focused on advancing language technology for African languages.
- Participate in community research discussions, workshops, and knowledge-sharing on multilingual NLP and speech technologies.

Founder & Machine Learning Researcher - *EfikAI*

July 2025 – Present

- Founded EfikAI, a community-based research initiative building the first publicly documented machine translation, ASR, and TTS systems for Efik, an extremely low-resource tonal African language.
- Curated 13,865 parallel English-Efik sentence pairs in collaboration with linguists and native speakers.
- Fine-tuned state-of-the-art machine translation models (mT5-base, M2M-100, NLLB) using PyTorch and Hugging Face.
- Developed natural-sounding ASR and TTS systems for Efik using Whisper, XLS-R, VITS, and diffusion-based models.

RESEARCH & INDUSTRY EXPERIENCE

Machine Learning Engineer - NLP -- *Plotweaver*

Jan 2026 – Present

- Led development and deployment of multilingual text-to-speech systems for Yoruba, Swahili, Hausa, and other African languages.
- Fine-tuned and evaluated modern TTS architectures including MMS-TTS, Orpheus-TTS, F5-TTS, and Qwen3-TTS.
- Built and deployed a real-time multilingual speech pipeline (ASR → MT → TTS) for live football commentary and event broadcasting.
- Reduced end-to-end latency from 24 seconds to 3 seconds (87% improvement) through Faster-Whisper, CTranslate2, and inference optimization.
- Deployed scalable inference services and APIs using AWS SageMaker.

AI Engineer — *RapidMind AI (Remote)*

Jun 2025 – Oct 2025

- Developed conversational AI systems and Rasa NLU chatbots for fintech customer support automation.

Full Stack Engineer — *Klas (Remote, Toronto)*

May 2024 – Present

- Led development of an AI Note Taker for live classes and online sessions.
- Built real-time transcription pipelines for lectures and meetings using speech recognition systems.
- Integrated LLMs / generative AI for automatic lecture summarization, key-point extraction, and structured study notes.
- Leveraged AWS services (EC2, Lambda, DynamoDB, S3) to optimize image processing, file uploads, and cloud infrastructure.

Backend Developer Intern — *Sycamore NG, Lagos (Hybrid)*

June 2023 – October 2023

- Built and tested RESTful APIs using Node.js.
- Improved database performance through schema optimization and query tuning.
- Conducted technical product research and reporting.

HONORS, AWARDS & SCHOLARSHIPS

- AfricaNLP Travel Support Award (USD \$2,000+)
- University of Rochester Master's Scholarship — Tuition-Based (unable to commence due to U.S. visa restrictions)
- Interspeech 2026 Travel Grant Award (AUD \$800)

PROFESSIONAL MEMBERSHIPS

- Masakhane Research Foundation
- Machine Learning Collective

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, TypeScript, SQL

Machine Learning: PyTorch, TensorFlow, Hugging Face Transformers, Scikit-learn

Speech AI: Whisper, Faster-Whisper, MMS-TTS, VITS, Orpheus-TTS, F5-TTS, Qwen3-TTS, ASR, TTS

NLP / GenAI: NLLB, M2M-100, mT5, Prompt Engineering, RAG, Summarization, Machine Translation

Cloud & DevOps: AWS SageMaker, EC2, Lambda, S3, Docker, Git, GitHub Actions

CERTIFICATIONS

- Deep Learning Specialization, DeepLearning.AI — 2025
- Machine Learning, IBM — 2023
- Introduction to Software Engineering, IBM — 2023
- Python Programming, University of Michigan (Coursera) — 2023