

[AFAAI Open Source project](https://afaai.org) 1M word to ethnic oriented GPT <https://afaai.org>



[AFAAI Open Source project](https://afaai.org) 1M word to Low Resource Language Model (ethnic oriented GPT)
Ethnic languages are suffering heavily with back translation which is rife in LLM. To solve this problem we create this open source project. Today we call on you to bring your talent and build a future for yourself and those to come.

Volunteer Opportunity: Backend Engineer (Azure / Linux OS)

About the Opportunity

Join the AFAAI Open Source Initiative in an ambitious mission to create innovative AI solutions for underrepresented languages. We are developing a 1M-word dictionary and NLP pipelines to preserve and elevate ethnic languages like Igbo, Yoruba, Hausa, and AAVE. As a Backend Engineer (Azure Expert), you'll play a critical role in designing scalable backend systems and advancing inclusivity in AI.

Responsibilities

Design and optimize backend systems to support the 1M-word dictionary and NLP pipelines on Azure Cloud Services.

Develop and integrate bit-level tokenization and normalization pipelines tailored for ethnic languages. Collaborate with linguists and researchers to ensure seamless API deployment and smooth end-to-end system functionality.

Requirements

Advanced expertise in Azure Functions, Azure Storage, and Azure Cognitive Services.

Proficiency in Python and backend frameworks such as FastAPI or Flask.

Experience with efficient encoding and optimization techniques for language models.

Strong collaboration skills to work effectively with linguists, researchers, and developers.

What's in It for You?

Skill Development: Gain hands-on experience with cutting-edge AI and backend systems tailored for low-resource languages.

Professional Growth: Enhance your expertise in Azure and backend development while contributing to a high-impact, innovative project.

Networking: Collaborate with a global team of linguists, researchers, and AI experts.

Training Opportunities: Access online study groups and training programs, including Cybersecurity SOC Level I Analyst courses, to diversify your skillset.

Portfolio Boost: Showcase your work in a groundbreaking initiative that advances inclusivity in AI and language technology.

Meaningful Impact: Contribute to preserving and empowering ethnic languages, leaving a lasting legacy in the AI field.

How to Apply

Are you ready to make a difference and expand your professional horizons? Send your resume and a brief note about your interest in this role to [Volunteer Application Form](#) with the subject line:

Volunteer Backend Engineer Application – AFAAI Initiative. When you are filling the form be very specific with your skills.



Join us to build innovative solutions that celebrate diversity in language and technology!

Volunteer Opportunity: Linguist About the Opportunity

Join the **AFAAI Open source project Initiative** as we work to revolutionize AI by creating inclusive solutions for low-resource languages. A key part of our mission is developing a **1M-word dictionary** to represent and preserve ethnic languages, including **Igbo, Yoruba, Hausa, and AAVE**. We are seeking dedicated volunteers to contribute their linguistic expertise to this groundbreaking project.

Responsibilities

- Curate high-quality corpora of ethnic languages for the **1M-word dictionary project**.
 - Ensure datasets fully represent alphabets, orthographic nuances, and comply with **Unicode standards**.
 - Validate dictionary entries and collaborate with native speakers to ensure linguistic accuracy and cultural sensitivity.
-

Requirements

- Fluency in one or more of the listed languages (**Igbo, Yoruba, Hausa, or AAVE**).
 - Experience in **linguistic research, lexicography, or dataset curation** for natural language processing (NLP) projects.
 - Passion for language preservation and inclusivity in technology.
-

Additional Benefits for Volunteers

- **Study Group Access:** Join our online study group, where you'll engage with like-minded professionals and enthusiasts.
 - **Cybersecurity Training:** Volunteers can access training programs for **SOC Level I Analyst roles**, gaining valuable cybersecurity skills and certifications to advance their careers.
 - **Skill Development:** Learn cutting-edge techniques in dataset curation, language preservation, and AI model development.
 - **Networking:** Collaborate with a global team of experts in linguistics, AI, and technology.
-



What's in It for You?

- Be part of an impactful initiative advancing representation in AI.
- Gain hands-on experience in linguistic dataset creation for cutting-edge NLP applications.
- Enhance your professional portfolio with valuable skills in both linguistics and cybersecurity.
- Contribute to a project that will shape the future of language inclusivity in technology.

To apply go to [Volunteer Application Form](#) Fill out what your skills are and we will call if you are fit.

Volunteer Opportunity: Full-Stack Developer

About the Opportunity

Be a part of the **AFAAI Open Source Project Initiative**, where we're pioneering AI-driven solutions for underrepresented languages. As a **Full-Stack Developer**, you will play a critical role in building the systems powering our **1M-word dictionary** and advancing natural language processing (NLP) tools for ethnic languages like **Igbo, Yoruba, Hausa, and AAVE**.

Responsibilities

- Build robust, end-to-end systems to support the **1M-word dictionary** and NLP pipelines using **BitsNet** and other NLP frameworks.
 - Develop APIs and backend integrations to enable **real-time word lookups** and **model inference**.
 - Collaborate with linguists and backend engineers to create scalable, efficient systems that preserve the unique characteristics of ethnic languages.
-

Requirements

- Experience with **PyTorch, NLTK**, and managing large-scale datasets.
 - Proficiency in backend frameworks like **Django** or **Node.js** and strong front-end development skills (e.g., React, Angular, or Vue.js).
 - Knowledge of **bit-level encoding, efficient API design**, and working with large datasets.
 - Ability to work collaboratively in a multi-disciplinary environment.
-

What's in It for You?

- **Skill Growth:** Enhance your expertise in **full-stack development, NLP, and AI applications** for low-resource languages.



- **Professional Opportunities:** Build your portfolio with hands-on experience in a high-impact AI project.
 - **Networking:** Collaborate with a diverse team of linguists, AI researchers, and developers from around the globe.
 - **Training Programs:** Access to **Cybersecurity SOC Level I Analyst** training and study groups to broaden your technical skill set.
 - **Impactful Work:** Contribute to the preservation and advancement of ethnic languages in AI, making a meaningful difference in the field of technology.
-

How to Apply

If you're passionate about technology, language, and inclusivity, we want to hear from you! Send your resume and a brief statement about your interest in this role to [Volunteer Application Form](#) with the subject line: **Volunteer Full-Stack Developer Application. Try to be specific with your skills.**

Join us in building innovative solutions that bridge the gap between technology and linguistic diversity!

Volunteer Opportunity: Cloud Infrastructure Engineer

About the Opportunity

Join the **AFAAI Open Source Project Initiative**, a groundbreaking project focused on creating AI tools and resources for underrepresented languages. As a **Cloud Infrastructure Engineer**, you'll design and maintain cloud-based solutions that power our **1M-word dictionary** and NLP pipelines for ethnic languages like **Igbo, Yoruba, Hausa, and AAVE**.

Responsibilities

- Design scalable cloud infrastructures to support the **1M-word dictionary** and NLP pipelines.
 - Deploy and manage large-scale datasets and AI models on platforms such as **Azure, AWS, or Google Cloud**.
 - Optimize storage and processing systems for **real-time dictionary lookups** and **model inference**.
 - Collaborate with linguists, full-stack developers, and backend engineers to ensure smooth integration of all components.
-



Requirements

- Expertise in cloud platforms, with a strong preference for **Azure**, and practical experience in **AI model deployment**.
 - Proficiency in **containerization** tools like **Docker** and **orchestration** platforms such as **Kubernetes**.
 - Familiarity with optimizing cloud resources for large-scale NLP and dictionary-related applications.
 - Strong problem-solving and teamwork skills to thrive in a multi-disciplinary environment.
-

What's in It for You?

- **Skill Development:** Advance your expertise in cloud infrastructure, AI, and NLP technologies.
- **Networking Opportunities:** Collaborate with a diverse team of AI researchers, linguists, and tech experts.
- **Training Access:** Join online study groups and training programs, including **Cybersecurity SOC Level I Analyst** courses, to diversify your skill set.
- **Meaningful Impact:** Contribute to a project that promotes inclusivity in AI by preserving and advancing underrepresented languages.
- **Professional Recognition:** Build your portfolio with hands-on experience in a cutting-edge AI initiative.

How to Apply

If you're passionate about technology, language, and inclusivity, we want to hear from you! Send your resume and a brief statement about your interest in this role to [Volunteer Application Form](#) with the subject line: **Volunteer Cloud Infrastructure Engineer Application. Try to be specific with your skills.**

Join us in building innovative solutions that bridge the gap between technology and linguistic diversity!

Volunteer Opportunity: AI Researcher (Python + NLP)

About the Opportunity

Join the **AFAAI Open Source Project**, a pioneering initiative aimed at advancing AI for underrepresented languages. As an **AI Researcher**, you'll contribute to fine-tuning **BitNet models**, developing embeddings, and creating tokenization techniques designed for ethnic languages such as **Igbo, Yoruba, Hausa, and AAVE**. Your efforts will support the creation of the **1M-word dictionary** and drive inclusivity in AI for low-resource languages.



Responsibilities

- Fine-tune **BitNet models** using ethnic language datasets to support the **1M-word dictionary**.
 - Develop custom embeddings and tokenization techniques to handle the unique characteristics of ethnic languages.
 - Apply **transfer learning** and **knowledge distillation** to optimize models for low-resource environments.
 - Work closely with linguists, engineers, and other researchers to integrate AI solutions effectively into the project's pipeline.
-

Requirements

- Proficiency in **Python**, with hands-on experience in NLP libraries such as **PyTorch**, **TensorFlow**, or **NLTK**.
 - Strong understanding of multilingual model development, tokenization strategies, and low-resource language challenges.
 - Familiarity with transfer learning, knowledge distillation, and other model optimization techniques.
 - A collaborative mindset to work within an interdisciplinary team focused on innovation and inclusivity.
-

What's in It for You?

- **Skill Development:** Expand your expertise in model fine-tuning, tokenization, and multilingual NLP.
 - **Collaboration:** Join a passionate community of linguists, developers, and AI researchers working on groundbreaking technologies.
 - **Training Access:** Participate in online study groups and access training programs such as **Cybersecurity SOC Level I Analyst** courses.
 - **Portfolio Building:** Gain hands-on experience with an impactful, open-source AI initiative.
 - **Meaningful Impact:** Contribute to a project that bridges linguistic gaps and promotes representation in AI.
-

How to Apply

Ready to make a difference? Apply now by filling out this form: [Volunteer Application Form](#).

Be part of the **AFAAI Open Source Project**, where your skills will help shape the future of AI for underrepresented languages!