Statement of Purpose

Growing up in Nigeria, I was always expected to become a civil engineer or a doctor, with no alternative. But during the four years of working after high school to save up for college, I began to read more about artificial intelligence (AI) and its potential applications for societal impacts. As a result, I bought my first computer to teach myself computer science and AI. After learning more about the field, I chose to pursue a Bachelor's degree in computer science at the Federal University of Agriculture Abeokuta (FUNAAB).

During my studies at FUNAAB, I learned that sub-Saharan Africa has the highest population of people with hearing disabilities but the least resources to support sign language training and other tools for communication. I began working independently on an end-to-end, lightweight, sign-to-speech application to address this issue. Specifically, I sought to develop a prototype that works in real-time to mitigate the strains of this communication barrier by detecting signs from live video feeds and returning the equivalent texts and speech. In the process, I created a pioneer dataset for the Nigerian Sign Language with help from a local TV broadcaster and two special education schools in Nigeria. I then annotated my dataset in YOLO and XML formats. I then performed classification experiments using several modeling techniques, including data augmentation, image classification, and object detection using YOLO and single-shot-detector architectures. I deployed the best-performing model using DeepStack, and my work resulted in a publication to appear at the ML4D workshop at NeurIPS 2021.

While working on my sign-to-speech project, I realized I needed to develop fundamental research skills and get regular feedback from experienced researchers. This led me to join ML Collective (MLC). In this open science lab, self-motivated researchers can proactively utilize the existing resources in MLC, such as project scoping and feedback process, to support their research. I make use of the various opportunities available in MLC to grow as a researcher in many ways, including (i) presenting my sign-to-speech and other projects regularly in the research lab meetings to get feedback from diverse, world-class researchers, (ii) joining the Vision and Language interest groups to review top literature weekly and to study publicly-available graduate courses from Carnegie Mellon and Stanford, and (iii) collaborating with a diverse range of researchers through the open-access research model at MLC and getting compute resources support, both of which were of great help for my sign-to-speech project.

Looking to continue developing my resource-efficient machine learning skills but lacking suitable advisors at FUNAAB, I independently sought opportunities to work remotely with research labs abroad. As a result of that effort, I now work as a volunteer research assistant for a PhD researcher, Nils Rethmeier, at the German Research Center for AI. Under his supervision, we try to address the universal need in research for huge compute power and massive datasets to build large, self-supervised language models (e.g. RoBERTa). To do so, the lab has been

exploring continual learning using a small-sized model that extends on the T5 model's approach via data-efficient contrastive self-supervision and performing text-to-text prediction in the embedding space instead of in the token space. I am currently studying how well this compact model can perform downstream tasks (e.g. ultra-fine entity typing) compared to current state-of-the-art models.

In addition to my research experiences, I have participated in community outreaches within and outside my academic environment. For instance, in my role as a Data Science Mentor across two cohorts of She Code Africa, a community that empowers young women in tech across Africa, I prepared immersive curriculums for female data science enthusiasts, mentored them actively, and carefully measured their progress until they reached industry-standard. Within my school, one of my more prominent roles was serving as the Campus Lead for AI+ FUNAAB, where I was the driving force behind our initiatives to democratize AI and help students, especially non-CS majors, acquire the necessary AI skills to build innovative solutions to problems in their own diverse fields. And until recently, I was the Community Lead for the Google Developers Student Club at FUNAAB, where I have initiated and now spearhead our programs to help student developers acquire industry-relevant technical skills and soft skills, connect them with local recruiters, and improve their visibility to attract job opportunities.

Given my background and experiences, I am interested in pursuing a PhD focused on resource-efficient ML. The University of Washington's Allen School presents an ideal environment to pursue my studies as it grants me proximity to AI2 and its other industry partners to apply my research directions for real-world impact, and its cross-disciplinary research culture gives me the chance to explore the extent to which my research could facilitate progress in other fields. In addition, I am quite interested in Dr. Ranjay Krishna's focus on resource-efficient machine learning for vision-language models and his study on how to make self-supervised learning work with less data, less compute and fewer parameters. I am also intrigued by Prof. Ali Farhadi's body of work in the language-vision intersection to create real-world impact - his work on YOLO is an ideal example of the research I aim to do - as well as by Prof. Hannaneh Hajishirzi's work on efficient ML algorithms for representing textual and visual data.

I have witnessed firsthand how the developing world and low-tech institutions rarely have sufficient compute resources or quality data in bulk for their ML workflows. To become an academic or industrial research scientist one day, I aim to become grounded in resource-constrained AI to make AI more accessible for people and institutions who don't have access to large-scale resources. I strongly believe that my prior research experiences, constant professional development, and my record of being proactive in everything I do, demonstrate that I am fully prepared to pursue my PhD at UW to advance state of the art in resource-constrained AI.