I want to become a biomedical scientist because I am driven by the goal of contributing to better healthcare equity. My motivation comes from seeing how biomedical research has been a positive driving force in the development of new methods and technologies that have led to better healthcare outcomes and standards of living worldwide. I see research as a direct path towards contributing to such efforts, with my recent and ongoing works helping me further refine my commitment in the research and development of clinically effective, impactful and accessible computational methods. I hope to contribute to a future where advancements in biomedical engineering are widely accessible, and not limited by geographic and or socioeconomic status. My ultimate goal is to be part of a generation of scientists that embarks on transformational research addressing real-world global healthcare challenges.

I am applying to the Mayo Clinic College of Medicine and Science, Biomedical Engineering and Physiology PhD track, with a special interest in Biomedical Imaging. Having graduated with a Bachelor of Science in Biomedical Engineering, and currently underway in a Master's program in Data Science, I have had the opportunity to work on using data and computational methods for biomedical research. I believe the Mayo Clinic Biomedical Engineering and Physiology PhD program is the best environment to continue to learn and contribute in this field, given its diverse faculty in both patient-centered clinical research, and computational analysis. I believe this combination gives me the best opportunity for working on translational research with the right guidance. I have gained interest and worked on various areas of research including medical image analysis across modalities (neuroimaging, ultrasound, MRI), trustworthy AI methods, and multi-omic data integration and fusion for disease characterization and prognostication.

My current research interests lie at the intersection of artificial intelligence, biomedical data science, and computational medicine. I focus on developing trustworthy and interpretable methods for analyzing clinical and biomedical data, ranging from medical images to multi-modal datasets that integrate demographic and biomarker information. My work explores three complementary directions:

(A) Foundational algorithmic methods for building reliable, trustworthy automated machine learning systems such as incorporating uncertainty quantification to flag low-confidence predictions.

(B) Representation learning across diverse biomedical data modalities including ultrasound, magnetic resonance images, and computed tomography for predictive and diagnostic modeling.

(C) Ultimately working towards accessible and reliable methods for clinical applications, particularly in low-resource healthcare settings.

My overarching goal is to bridge the technical and translational dimensions of artificial intelligence in medicine, developing systems that are computationally efficient, clinically meaningful, and deployable in real-world environments. In the short term, I aim to advance my expertise in biomedical data science with emphasis on model generalization, interpretability, and equitable performance across populations. In the long term, I aspire to lead research that drives data-informed healthcare innovations, translating computational and artificial intelligence innovations into deployable healthcare tools that improve diagnostic accuracy and expand global access to quality care. Beyond research, I am equally committed to mentorship and capacity building, contributing to the development of institutions in underserved communities capable of achieving globally recognized scientific impact.

A very satisfying research project I recently worked on was the development of an end-to-end framework for breast tumor segmentation and pathological complete response (pCR) assessment in dynamic contrast-enhanced MRI [1]. The work, titled "*Large Kernel MedNeXt for Breast Tumor Segmentation and Self-Normalizing Network for pCR Classification in Magnetic Resonance Images*," produced mixed but insightful results, and taught me a lot about balanced approaches for reporting findings. For instance, subgroup analysis showed considerable variability: balanced accuracy reached up to 75% among patients aged 51–60, but declined to around 30% in those aged 71 and above, exhibiting notable performance differences across subpopulations. The method involved a two-stage training strategy of a ConvNeXt-UNet for segmentation, and a self-normalizing network for assessing pCR using radiomic biomarker features from the segmentation step. Through this work, I learned how to identify and extract meaningful information from complex medical data for computational analysis; particularly radiomic biomarker features. What made this project especially rewarding was that I pursued it independently, finding a right balance between computational applications and translational potential. Future iterations of this work will focus on integrating clinical and demographic factors with radiomic biomarkers to improve robustness and subgroup generalizability.

Another noteworthy project involved developing a context-specific large language model (LLM) application to support expectant mothers with gestational diabetes in disease management. This work required bridging multiple challenges: adapting general-purpose language models to provide evidence-based guidance in both English and local languages; and designing a system that could function reliably within Ghana's digital infrastructure constraints. Working closely with clinicians and maternal health specialists, we iteratively refined the system based on clinical feedback and piloted it in a controlled environment. What made this project particularly rewarding was witnessing the direct clinical utility, and its potential to help standardize patient education, freeing up consultation time for more complex cases.

My research experience spans roles at the Responsible Artificial Intelligence Lab (RAIL-KNUST), and at the Kumasi Center for Collaborative Research in Tropical Medicine (KCCR) as a Research Assistant and Engineer, where I have worked in developing AI-driven solutions for real-world healthcare challenges. At KCCR, a project I am proud of is the previously highlighted work on an LLM application designed to assist expectant mothers with gestational diabetes in disease management. Some other research oriented works I have undertaken include; developing a trustworthy breast tumor segmentation network with Monte Carlo dropout and deep ensembles for epistemic uncertainty estimation to improve reliability in breast ultrasound analysis [2], an automated ischemic stroke lesion detection framework for non-contrast CT images to support early diagnosis and improve treatment outcomes in low-resource settings [3], and a segmentation-aware data augmentation and deep ensembling approach that achieved state-of-the-art results for brain tumor segmentation in the Sub-Saharan African context as part of the BraTS 2025 challenge [4].

In service to community, I have had the honour of contributing in roles that allow me to give back and learn from those around me. Most recently, I have been involved in the Sprint AI Training Program for Medical Imaging Knowledge Translation (SPARK Academy) [1], an initiative dedicated to advancing research capacity in medical imaging and deep learning, where I help coordinate the program and develop instructional materials for up and coming researchers and clinicians. **A noteworthy outcome of this was supervising the Ghanaian team to winning first place in the Brain Tumor Segmentation Challenge (BraTS-SSA) at MICCAI 2025, marking the first time an African team has achieved this distinction** [4]. My community engagements also include facilitating *SheCodes*, a girl-child empowerment program in STEM, where I teach students from high-school to university levels in Python programming, microcontrollers, machine learning, and data science. In the open-source community, I have contributed to the TorchIO library for medical image processing, extended the nnU-Net framework to support additional use cases such as uncertainty estimation, and developed a Python library that simplifies access to medical image segmentation datasets for learning and benchmarking [2]. In all these efforts, my goal is consistent: to help build communities and tools that make advanced biomedical technologies more accessible, equitable, and impactful.

These experiences collectively inform my decision to pursue graduate study at Mayo Clinic, demonstrating my preparedness through both independent efforts [1], and effective teamwork and leadership evident in collaborative projects [2, 4], and coordination at the SPARK AI Training Program. I am especially inspired by the works of **Dr. Hamid R. Tizhoosh** on AI in medical imaging (recently in pathology), clinical data integration, **Dr. Richard D. White** on AI and radiomics for non-invasive image-based biopsies, and **Dr. Mostafa Fatemi** on advancing ultrasound image analysis, recognizing its accessibility as a diagnostic tool. I look forward to the opportunity of contributing my experience and perspective to the Mayo Clinic community, and learning from its culture of collaboration, where clinicians and scientists work together to transform research into tangible improvements in healthcare.

References

[1] **Toufiq Musah**. Large kernel mednext for breast tumor segmentation and self-normalizing network for pcr classification in magnetic resonance images. In *Deep Breast Workshop on AI and Imaging for Diagnostic and Treatment Challenges in Breast Care*, pages 72–80. Springer, 2025.

[2] **Toufiq Musah**, Chinasa Kalaiwo, Maimoona Akram, Ubaida Napari Abdulai, Maruf Adewole, Farouk Dako, Adaobi Chiazor Emegoakor, Udunna C Anazodo, Prince Ebenezer Adjei, and Confidence Raymond. Towards trustworthy breast tumor segmentation in ultrasound using monte carlo dropout and deep ensembles for epistemic uncertainty estimation. *arXiv preprint arXiv:2508.17768*, 2025.

[3] **Toufiq Musah**, Prince Ebenezer Adjei, and Kojo Obed Otoo. Automated segmentation of ischemic stroke lesions in non-contrast computed tomography images for enhanced treatment and prognosis, 2024.

[4] Claudia Takyi Ankomah, Livingstone Eli Ayivor, Ireneaus Nyame, Leslie Wambo, Patrick Yeboah Bonsu, Aondona Moses Iorumbur, Raymond Confidence, and **Toufiq Musah**. How we won brats-ssa 2025: Brain tumor segmentation in the sub-saharan african population using segmentation-aware data augmentation and model ensembling. *arXiv preprint arXiv:2510.03568*, 2025.

[1] https://event.fourwaves.com/spark/pages
[2] https://github.com/toufiqmusah/MedSegMNIST