Statement of Purpose

I want to create the next generation of **mega-scale distributed and embedded systems**, ensuring they also strengthen and uplift underserved communities. Growing up in Multan, a city in underdeveloped southern Punjab, meant seeing inequity: in disaster preparedness, education and even in healthcare. The haves and have-nots lived two different lives, and seeing myself as part of the haves meant that I felt a sense of responsibility. My work in LUMS would soon make me realize the potential of distributed and embedded systems to bridge regional and class divides.

As a sophomore RA under Dr. Malik Jahan, I collaborated with the Centre for Water Informatics and Technology at LUMS, deploying AIoT solutions for disasters. I conducted a field study using deployed IoT sensors and cameras to analyze environmental features like river proximity and identify key contributors to flood severity. I immediately saw how such technology could apply in flood-prone Multan. The research taught me how to navigate IoT devices and apply data-driven insights to address local environmental challenges, and to begin seeing how these tools could truly help my overarching goals, I knew I had a lot more to learn. When I finally found myself ready to work in distributed and embedded systems, an opportunity presented itself.

In regulated fields like healthcare, ensuring secure, user-friendly, and auditable access to sensitive data and AI models is critical. However, strict privacy regulations and the need for transparency make it difficult for authorities to trust the public with their datasets, and for healthcare professionals to access them securely. To address these challenges, I collaborated with Dr. Basit Shafiq (LUMS) and Dr. Jaideep Vaidya (Rutgers) on the "**Blockchain-Based AI Model Governance and Auditable Policy Compliance System**". Inspired by healthcare inequities in Multan, I developed a platform integrating an LLM-based architecture that converts natural language queries into composite BPMN (Business Process Model and Notation) workflows, with smart contract approvals ensuring compliance (a paper is currently in writing). I converted XACML-based policies into smart contracts using INFURA APIs over the Sepolia Testnet and created an interactive workflow visualization combining the Cox model and Kaplan-Meier curves for survival analysis of 400,000+ patients. The React-based frontend, which included BPMN Modeler, improved navigation and workflow efficiency. Addressing LLM-generated BPMN inaccuracies, which could lead to potential misdiagnosis, I employed chain-of-thought prompting with examples of correct XML code, improving interpretability. A RAG-based implementation is in progress to ensure workflows incorporate context-specific knowledge. By ensuring compliance, transparency, and accessible LLM-based natural language interactions, this system motivates authorities to share data and simplifies access for healthcare providers.

Among the first to link LLMs with auditable workflows in healthcare, this research demonstrated how this multilayered approach can democratize access to healthcare assets while ensuring compliance. The project taught me to manage complex codebases, refine AI workflows, and, most importantly, innovate solutions when refining the workings of a complex distributed system.

Motivated by a desire to use large-scale systems for disaster preparedness, I pursued a research internship at the Centre for Water Informatics and Technology. Tasked with creating a cost-effective, low-complexity edge solution for real-time flood detection in remote areas, I collaborated with Dr. Talha Manzoor to develop an **Edge AIoT system** using the ESP-32 Cam. Its accessible development environment and integrated camera reduced hardware costs to a tenth of traditional devices, making it ideal for TinyML. I employed YOLOv5-based TinyML (Edge Impulse) for accurate detection and supplemented it with LLMs for broad generalizability. I also integrated a serverless Google Drive for reliable data transfer without on-site servers. To address connectivity challenges, a GSM module was incorporated for long-distance communication, and sleep modes were introduced for energy efficiency and resilience. This adaptable architecture is now being refined for deployment in Pakistan's Northern Areas, paving the way for applications such as soil moisture and wildlife detection.

Building on this experience, I joined a WWF initiative with Dr. Murtaza Taj to address forest fires in Pakistan's remote regions. Using Hikvision IP cameras integrated with CLIP-based LLM and YOLO models, I localized fire hotspots through image analysis. By optimizing Pan, Tilt, and Zoom with the Hikvision API, I captured overlapping images for recursive analysis, pinpointing fire locations. With OpenCV, I created GIFs of capture events to aid interventions. The integrated LLM-YOLO system achieved 99% true positive accuracy and is now being deployed across five cameras to enhance disaster prevention. These projects showed me how the use of embedded technology with a problem-solving mindset can yield scalable IoT solutions that address real-world challenges.

Pursuing a thesis-based MS in Computer Science will equip me with the expertise and versatility to translate my passion for distributed systems into truly impactful solutions. The focused research approach with rigorous courses will help me quickly develop expertise, build a strong professional network, and position myself at the forefront of cutting-edge distributed systems research. All in all, an MS at a reputed university would provide me with an excellent faculty coupled with state-of-the-art research labs, primarily focused on creating an impact by leveraging distributed systems and a vibrant community of like-minded scholars.