

# Ahmad Shamail

(520)-609-1202 | [shamail@arizona.edu](mailto:shamail@arizona.edu) | [Website](#) | [LinkedIn](#)

## Education

---

### Master of Science in Computer Science

University of Arizona

GPA: 4.00 / 4.00

2024 - 2026

📍 Tucson, Arizona

### Bachelor of Science in Computer Science

Lahore University of Management Sciences (LUMS)

GPA: 3.89 / 4.00

2020 - 2024

📍 Lahore, Pakistan

**Awards:** Dean's Honor List (2021 – 2024), Merit Scholarship (2020 – 2024),

NMF Gold Medal - Sportsman of the Year (2024) – Male

Relevant coursework: Machine Learning, Data Science, Network Security, Topics in Internet and Network Security, Computer Graphics, Advanced Programming, Software Engineering

## Technical Skills

---

**Experienced:** Python, LangChain/LangGraph, C/C++, Bash, Docker, Wireshark, Linux, React, TypeScript, JavaScript, SQL

**Familiar:** React Native, MongoDB, Express.js, Node.js, R, Haskell, Selenium, TensorFlow

**Other:** Unity, Git, OpenCV, PyTorch, LaTeX, LLMs/Transformers, HPC

## Experience

---

**AI-Engineer Intern** | [StruxHub](#) (USA) | *June 2025 – August 2025* |

- Collaborated with the CEO to integrate AI-driven workflows into the platform using LangChain and LangGraph, streamlining construction form creation, recreation, and requests for on-site workers
- Developed AI-powered digitization of paper-based forms from images and PDFs, enabling construction site digitalization to improve conflict avoidance and increase planning speed by 70%
- Designed AI agent based solutions to generate daily and weekly site reports, saving teams 5-10 hours weekly while providing stakeholders with actionable insights into progress, setbacks, and improvement areas

**Protein-Specific LLM Researcher** | [University of Arizona](#) | *August 2024 – Ongoing* |

- Built data pipelines and GPU-optimized workflows to analyze neuron activations across millions of protein sequences, enabling large-scale study of protein representations.
- Developed a sparse autoencoder in PyTorch to enhance neuron specificity, with the goal of isolating signals tied to protein motifs, binding sites, and structural features.
- Designed evaluation and visualization tools to validate neuron behavior, supporting more accurate protein function annotation and structure prediction.

**Systems Reproducibility and Vulnerability Researcher** | [SRI International](#) | *May 2024 – August 2024* |

- Evaluated 66 datasets from eight provenance-based intrusion detection systems using deep learning, reproducing and benchmarking reported results
- Achieved exact match with published results in 6 evaluations and within two percent variance in 29 evaluations, revealing significant reproducibility gaps
- Published [Paper](#) in [ACM-REP 2025](#) showing none of the systems were fully reproducible end to end, and delivered targeted recommendations on code completeness, dataset standardization, automated evaluation, and deterministic training

## Projects

---

**LLM-Powered Fitness App** [[GitHub](#)]

- Built an LLM-powered system generating dynamic, personalized workout plans that automatically adapt based on user performance and feedback. Using LangGraph, React, and Gemini.
- Developed progress tracking with analytics dashboard, community features, and AI recommendations for optimized fitness outcomes

**AI Protein-Card Generator** [[GitHub](#)]

- Developed AI-powered protein intelligence system aggregating data from 6 biological databases using LangGraph workflow orchestration and Google Gemini API, creating comprehensive "Protein Cards" with interactive 3D visualization